PSL
UNIVERSITÉ PARIS

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à MINES ParisTech

Development of geostatistical models using Stochastic Partial Differential Equations

Soutenue par

Ricardo Carrizo Vergara

Le 18 décembre 2018

École doctorale n°432

Sciences des Métiers de l'Ingénieur

Spécialité

Géostatistique et Probabilités Appliquées

Composition du jury :

Hermine BIERMÉ Professeur, Université de Poitiers	*Présidente*
Annika LANG Professeur associée, Chalmers University of Technology	*Rapporteuse*
Josselin GARNIER Professeur, École Polytechnique	*Rapporteur*
Finn LINDGREN Chaire de statistique, University of Edinburgh	*Examinateur*
Nicolas DESASSIS Chargé de recherche, MINES ParisTech	*Encadrant*
Denis ALLARD Directeur de recherche, INRA	*Co-directeur de thèse*
Hans WACKERNAGEL Directeur de recherche, MINES ParisTech	*Directeur de thèse*

MINES ParisTech | PSL

"*He*[1] *who loves practice without theory is like the sailor who boards ship without a rudder and compass and never knows where he may cast.*"

Leonardo da Vinci

[1] It also applies for a "*she*"... and for any *other* human being...

Abstract

This dissertation presents theoretical advances in the application of the Stochastic Partial Differential Equation (SPDE) approach in Geostatistics. This recently developed approach consists in interpreting a regionalised data-set as a realisation of a Random Field satisfying a SPDE. Within the theoretical framework of Generalized Random Fields with a mean-square analysis, we are able to describe with a great generality the influence of a linear SPDE over the covariance structure of its potential solutions. A criterion of existence and uniqueness of stationary solutions for a wide-class of conveniently defined linear SPDEs has been obtained, together with an expression for the related spectral measures. This result allows to encompass a great variety of already known relationships between stationary covariance models and SPDEs. It also allows us to obtain new stationary covariance models that are easily related to SPDEs, and to propose SPDEs for some already known covariance models such as the Stein model and the $J-$Bessel model. We apply these results to construct spatio-temporal covariance models having non-trivial properties. By analysing evolution equations presenting an arbitrary fractional temporal derivative order, we have been able to develop non-separable models with controllable non-symmetric conditions and separate regularity over space and time. We present results concerning stationary solutions for physically inspired SPDEs such as the advection-diffusion equation, the Heat equation, some Langevin equations and the Wave equation. We also present developments on the resolution of a first order evolution equation with initial condition. We then study a method of non-conditional simulation of stationary models within the SPDE approach, following the resolution of the associated SPDE through a convenient PDE numerical solver. This simulation method, whose practical applications are already present in the literature, can be catalogued as a spectral method. It consists in obtaining an approximation of the Fourier Transform of the stationary Random Field, using a procedure related to the classical development on Fourier basis, and for which the computations can be efficiently obtained through the use of the Fast Fourier Transform. We have theoretically proved the convergence of this method in suitable weak and strong senses. We show how to apply it to numerically solve SPDEs relating the stationary models developed in this work, and we present a qualitative error analysis in the case of the Matérn model. Illustrations of models presenting non-trivial properties and related to physically driven equations are then given.

Résumé

Ces travaux présentent des avancées théoriques pour l'application de l'approche EDPS (Équation aux Dérivées Partielles Stochastique) en Géostatistique. On considère dans cette approche récente que les données régionalisées proviennent de la réalisation d'un Champ Aléatoire satisfaisant une EDPS. Dans le cadre théorique des Champs Aléatoires Généralisés avec une analyse en moyenne-quadratique, nous avons décrit avec une grande généralité l'influence d'une EDPS linéaire sur la structure de covariance de ses éventuelles solutions. Un critère d'existence et d'unicité des solutions stationnaires pour une classe assez large d'EDPSs linéaires a été obtenu, ainsi que des expressions pour les mesures spectrales reliées. Ce résultat nous permet de rassembler dans un cadre unifié un grand nombre de liens déjà connus entre modèles de covariance stationnaires et EDPSs. Il nous permet en outre d'obtenir de nouveaux modèles de covariance stationnaires immédiatement reliés à des EDPSs, et de proposer des EDPSs pour des modèles de covariance déjà connus comme le modèle de Stein et le modèle $J-$Bessel. Nous appliquons ces résultats à la construction de modèles de covariance spatio-temporels présentant des propriétés intéressantes. À travers l'analyse des équations d'évolution comprenant un opérateur différentiel temporel d'ordre fractionnaire arbitraire, nous avons développé des modèles non-séparables ayant des conditions d'asymétrie et de régularités spatiale et temporelle séparées contrôlables. Nous présentons des résultats concernant des solutions stationnaires pour des EDPSs issues de la physique, telle que l'équation d'advection-diffusion, l'équation de la chaleur, quelques équations de Langevin, et l'équation d'onde. Nous présentons aussi des développements pour la résolution des modèles d'évolution de première ordre ayant une condition initiale. Puis, nous étudions une méthode de simulation non-conditionnelle pour des modèles stationnaires dans le cadre de l'approche EDPS. Pour cela, nous nous inspirons de la résolution de l'EDPS associée moyennant une méthode de résolution numérique des EDP choisie de manière appropriée. Cette méthode de simulation, dont son application pratique est déjà présente dans la littérature, peut être considérée comme une méthode spectrale. Elle consiste à obtenir une approximation de la Transformée de Fourier du Champ Aléatoire stationnaire par une procédure intimement reliée au développement classique en base de Fourier, et pour laquelle nous pouvons obtenir des méthodes de calcul efficaces grâce à la Transformée de Fourier Rapide. Nous avons démontré théoriquement la convergence de cette méthode dans aux sens faible et forte dans des conditions appropriées. Nous montrons comment appliquer cette méthode pour la résolution numérique des EDPSs reliant les modèles stationnaires

développés dans ces travaux, et nous présentons une analyse qualitative de l'erreur pour le cas du modèle Matérn. Des illustrations de modèles présentant des propriétés non-triviales et reliés à des équations de la physique sont alors présentées.

Acknowledgements

This acknowledgement section can be also considered as a *dédicace* section, since this work is also dedicated to every person mentioned here. It will be written in a quite particular way: I will use three different languages, depending on which language do I share in common with the people I mention (if there is no common language, I use English). I hope this section ne sera pas trop difficile de entender.

Quiero comenzar agradeciendo a la gente más importante de mi vida: a mi familia, a la cual le debo todo y aún más. Les debo sobretodo unas disculpas por aislarme durante mis períodos de mayor trabajo en vez de contar con ellos, estar más presente y permitirles darme los ánimos necesarios. Sólo en retrospectiva uno puede darse cuenta de cuánto hace falta la gente que amas en el momento correcto. Quiero agradecer el apoyo incondicional de mi Madre, de mi Padre y de mi Hermano. Los amo mucho y que no se les olvide que *a esto vine*. Agradezco también a esa familia con la que no tengo la suerte de compartir más seguido: a mi cuñada la Sole, a mis *primos* Cristóbal, Vale, Nacho y (en rigor prima también) Cahui, a toda la familia de Iquique, a mis padrinos, y a la familia Contzen Acevedo. Todos ustedes, *siempre* estuvieron presentes.

I would like to thank professors Annika Lang, Josselin Garnier, Finn Lindgren and Hermine Biermé, for having the kindness of having accepted to be part of my defense's jury, and for their time and patience used to read this work and for contributing to it with their critics, comments and suggestions. Being read by you has been an honour.

Je voudrais remercier avec un grand enthousiasme les deux personnes qui m'ont encadré tout au long de ces trois années (et plus!). À Denis Allard pour sa patience, pour la pression qu'il a mis au moment où il fallait, et pour ses tentatives, souvent (mais pas toujours) fructueuses de me faire comprendre et accepter la *réalité sociale*. À Nicolas Desassis pour avoir été beaucoup plus qu'un encadrant, un ami, un grand frère et un collègue qui était toujours là pour m'écouter et pour travailler avec moi, et qui m'a aussi respecté plus qu'il ne devait le faire. Je le remercie pour m'avoir supporté surtout aux pires moments. Je tiens à lui promettre que plus jamais personne ne devra s'endormir à 3h00 du matin à cause de moi...

Je voudrais remercier tout le personnel du Centre de Géostatistique de Fontainebleau qui m'a accueilli pendant (plus de!) trois ans. En particulier, je voudrais remercier mon directeur de thèse Hans Wackernagel pour son soutien, son encouragement, et sa capacité à me faire revenir à la réalité lorsque je le devais.

À Christian Lantuejoul, qui a été toujours disponible pour échanger avec moi sur des sujets divers, mais surtout des maths. Je suis très ravi de toutes les fois où il a compté sur moi pour partager ses propres recherches et questions: il a été une aide indispensable, morale et professionnellement parlant. Je remercie Jacques Rivoirard pour son bon travail de chef du centre, spécialement lorsqu'il s'agissait de me donner des bonnes leçons de *diplomatie*. À Christian Lajaunie pour ses contributions et suggestions. À Didier Renard pour créer la bonne ambiance tous les jours. À Thomas Romary pour toutes les discussions sur des sujets matheux et pas trop matheux. À Émilie Chautru pour sa gentillesse et sa disponibilité à m'aider dans des sujets professionnels ou pas. À Xavier Freulon pour avoir toujours été intéressé par mon travail. À Fabien Ors pour sa patience incroyable au moment de m'aider avec mes problèmes informatiques. À Gaëlle Le Loch pour ses leçons au moment de faire des TPs pour les CFSG et pour supporter les fois où je chantais plus fort que je ne devais près de son bureau. À Nathalie Dietrich et à Isabelle Schmitt pour m'avoir aidé à avoir toujours *tout en ordre*.

Je voudrais remercier le professeur Olivier Laffitte pour avoir donné le cours que j'ai le plus apprécié pendant mes années du cycle ingénieur des Mines. Je le remercie aussi pour m'avoir encouragé et guidé dans ma recherche de thèse, pour avoir toujours été attentif à mes problématiques pendant le doctorat, pour s'intéresser à ma soutenance à laquelle il a invité tout le cours de Théorie des Distributions, et pour son aide dans ma recherche d'emploi après thèse. C'est toujours un plaisir d'échanger avec vous!

Je voudrais remercier Angela Taddei et Michel Morange pour avoir organisé le cours de sensibilisation à l'Histoire, l'épistémologie et l'éthique des Sciences, qui est peut-être le cours que j'ai le plus apprécié de toute ma vie. Quelle émotion de faire autant de polémiques...

Quiero agradecer a varios de mis profesores de la Pontificia Universidad Católica de Chile, mi *verdadera y única* alma máter. Quiero agradecer a Rolando Rebolledo por todo el apoyo que me dió durante mis cursos de Integración y Probabilidades, además de su ayuda en las postulaciones a doble título y a doctorado en París. Agradezco a Luis Disset por su primer curso de matemáticas abstractas, en donde, con mucho dolor, aprendí a arreglar mi mente. A Carlos Jerez por sus fructuosos intentos de mejorar la especialidad de Ingeniería Matemática en la PUC, y a quien aprovecho de prometerle que *todas* las deudas se pagarán tarde o temprano. A Wolfgang Rivera por acompañarme durante todo mi desarrollo en la PUC, desde mis primeros pasos como novicio en las matemáticas (gracias por hacerme popular... $\neg\neg$), hasta como ayudante de sus propios cursos. Finalmente, quisiera agradecer a Álvaro Alarcón por su no-tan-formal curso de modelos estocásticos en donde aprendí quizás una de las lecciones más importantes que he recibido en mi vida académica: lo importante de un modelo matemático no es que sea realista, lo importante es que sea *útil*.

Je tiens à remercier maintenant mes collègues et amis avec qui l'on a partagé une partie de nos vies au centre de Fontainebleau. Je remercie Jihane Belhadj et Anna Bubnova pour m'avoir accueilli chaleureusement dans notre bureau (et donc pour avoir supporté tous mes délires et mes chansons lorsque je ne contrôlais pas mes envies de faire du bruit...). À Robin Alais pour ses magnifiques soirées chez lui, où l'on a profité

en particulier de ses très intéressants instruments de musique. À ma presque-colocatrice Marine Demangeot pour m'avoir sauvé la vie un million de fois en m'accueillant chez elle à Fontainebleau lorsque j'avais du mal à rentrer à Paris le soir. À mon canard d'air Laure Pizzella pour animer le bureau avec de nombreux commentaires marrants et des sourires contagieux. À Jean Langanay pour s'être toujours présent dans les rencontres sociales et pour être toujours intéressé par mes travaux. A los chilenos *cul...* Héctor y Esteban que me hicieron explotar de risa muchas veces en un centro de investigación que generalmente es un tanto serio. À MariAngela Donati pour les bonnes conversations autour de bières. To Eirini Skylogianni, who sadly did not stay long in the Center, so I miss her. Au reste des doctorants du Centre, en particulier à Léa, Mathieu, Alan, Aurelien, Albane, Angelique, Eric, Pierre, Nico, Leonardo, etc. Je remercie aussi Soufiane qui a rendu plus intéressantes mes visites en Avignon. Et pour finir, je voudrais remercier de façon spéciale mon pote, collègue du cycle ingé, collègue de thèse en Géostatistique, collègue des maths et collègue de la vie, Mike Pereira; partager avec toi ces moments, ces travaux, et en général, le plaisir de faire des mathématiques (appliquées...(?)), a été l'une des plus belles expériences de ma vie.

Je voudrais remercier toutes les personnes qui m'ont accompagné physiquement pendant mon séjour à Paris et qui ont allumé ma vie lors des dernières années. Quiero agradecer a Martín Costa por aparecer en mi vida de modo sorpresivo y por tener siempre algo que hacer para distraerme desde entonces. Quiero agradecer con un cariño especial a Juan Pablo Vigneaux, por todas esas inquietantes, a veces desesperantes pero siempre fructíferas y nutritivas discusiones sobre temas científicos, musicales, filosóficos y políticos: tu presencia ha sido indispensable. Quiero agradecer a Marta Jordana por todas esas relajantes *serbesas* que nos tomamos juntos; a Madeleine van Wassenhove (lo escribí bien??) por su cariño y amistad incondicional; à la mignonne Fanny Zimmermann qui s'est toujours inquiétée pour moi, spécialement lors des périodes stressantes d'isolement de rédaction; a la adorable Florence Dubart a quien le agradezco su presencia sobre todo al final y que me hizo notar que era una persona con la que siempre podría contar a pesar de cualquier lejanía; à François Pacaud pour partager avec moi une bonne tonne de goûts et de sujets d'intérêt et pour croire à chaque bêtise que je dis (arrête de le faire!); au Peingch de l'école des Mines qui m'a toujours considéré dans ses activités; a los chilenos que conocí en París y que animaron mi estadia, Paula Jouannet, Javier León y los hermanos Neumann. Quiero agradecer especialmente a Jorge Clarke por haberme invitado a trabajar con él a Valparaíso, y por su no-medible (medida interior infinita) ayuda al momento de redactar, por quien aprovecho de decir que esta tesis no habría podido escribirse sin su presencia e insistente presión. À Carlo Bellingeri pour montrer son intérêt à mes travaux et pour m'introduire peu à peu dans les sujets d'intérêt de la communauté *SPDE*-iste. Gracias a los mexicanotes Caroline Rivera, Eduardo Güemes y al Ariel cul... que nunca responde los mensajes. Je voudrais aussi remercier ceux qui ont fait partie de ma vie en France mais qui habitent maintenant ailleurs. Merci à l'adorable Tabea Wäfler pour illuminer mes journées chaque fois que j'ai la chance de la voir. Gracias a mis queridos españoles Alejandro Villar, Alejandro Suárez y Pilar Barbero quienes han tenido la gentileza de invitarme a compartir con ellos en Zaragoza y Madrid, y quienes han mostrado una paciencia enorme soportando mis salidas. Un último agradecimiento

para MaryCarmen Domínguez quien siempre estuvo preocupada por mí desde la lejanía del Perú.

I would also like to thank all the people with whom I have shared one of the main pleasures of my life: Music. Merci aux Conejitas en Reproducción, Louis Guitton, Thomas Brehon et Lionel Sergent pour nos bons moments partagés avec la musique. Je voudrais dire merci à Mattia Cavicchi pour ses inoubliables soirées et pour toutes les fois que l'on a joué (et que l'on jouera) Radiohead ensemble. Je voudrais remercier tous les membres de l'équipe organisateur des concerts des résidents à la Cité Universitaire de Paris (des trois dernières années): merci pour l'impeccable organisation et pour la capacité de créer une belle communauté de musiciens dans l'un des environnements les plus beaux de Paris. I acknowledge every musician with whom I have played in this context: Liburn Jupolli and the members of The Teko Sacke Orchestra; Michel Melia who in addition had the kindness of receiving me at his home in New York; Aude Nyadanu, avec qui j'ai appris plus d'une bonne leçon de vie; Mathilde Elmgaard Blanc et Guru Pope. But above all, I would like to acknowledge the three persons who have touched deeper my Parisian musical experience: Johanna Kuperminc, Dominykas Gustas and Nicolas Judalet. Playing with you guys has been more than a pleasure: it has been an honour.

Y no me puedo olvidar del parráfo más largo de todos: el agradecimiento a todos los *weones* que abandoné en Chile. Quiero comenzar por agradecer a los eternos, a los que han estado, están y estarán siempre en mi memoria, cerca o lejos, desde el momento en que los conocí hace ya tantos años: Matías Puelma, Maximiano Coronel, Tomás Yáñez y Fernanda Kossack. Quiero agredecer también a todos mis amigos ingenieros matemáticos con quienes siempre es un placer, e incluso una necesidad el compartir la investigación y la pasión por nuestros campos de estudio. Gracias a Matías Carvajal, Álvaro Gajardo, José Poblete, Felipe Ossa y David Cozmar, quienes siempre se mostraron interesados en mis trabajos y a quienes aprovecho de confesarles que esta tesis fue escrita en inglés para que ustedes la entendieran (a quién le importa el jurado de tesis no franco-parlante??!!... y el franco-parlente... a nadie tampoco!). En general, gracias a todos los ingenieros matemáticos de la PUC (aproximadamente 10 por generación) con quienes disfrutamos mientras nos hacían pedazos en las prubeas, pero con los que siempre compartimos un especial placer por lo que hacíamos. Quiero agradecer a todos los chilenos del ambiente musical chileno a los cuales aprovecho de decir que los extraño y no poco. A Javier Andina, a Bruno Castillo, a Cristián Ayala, a Esteban Olivares y a Delia Valdebenito por su hermosa música que me acompañó estos años. Quiero darle las gracias a mis amigos miembros de The LeoLea Experience, Tomás Barrios, Ignacio Carrillo, Ignacio Flores y Leonardo Olea, con quienes la nostalgia es siempre más grande que la distancia, pero jamás más grande que el placer que compartimos juntos. Gracias también a Sofía Gallego a quien tengo la suerte de ver en Europa a menudo. Quiero agradecer a los chilenos que conocí en Francia durante los primero años, y que no tienen idea de cuánta falta han hecho en los últimos: gracias a Rodrigo Lara, Esteban Iglesias, Joaquín Muñoz y a Daisy Castillo. Quiero agradecer a todo chileno que haya venido a *pecharme casa* en París: sepa usted que yo lo pasé mejor! Recibir chilenos y sacarlos a pasear haciéndose pasar por guía turístico en París es lejos uno de los mayores placeres de la vida Parisina! Quiero agradecer a todos mis amigos que conocí en la PUC,

que me dieron los mejores años de mi vida en un ambiente enriquecedor. Menciono especialmente a Vicente Cartagena, a Jorge Sandoval, a Olivia Bordeu, a Felipe Vera, a Omar Seguel, a Francisca Rojas, a Simón Ríos (seeeeco monito!), a Ixsy Valdés, a Naked, a Florencia Obilinovic, a Luis Bravo (qué lata tener que escribir tu nombre acá...), a Francisco Larrondo y a Sebastián Otárola quien además se preocupó especialmente por mí en los momentos de mi defensa de tesis. Quiero también agradecer a los miembros del CAi en los años en los que trabajé, especialmente aquellos con los que organizábamos el Woodstock UC y otras actividades culturales; pienso particularmente en Felipe Salfeeld, Pablo Torres y Roberto Flores. Quiero dar un agradecimiento especial a tres grandes amigos quienes, gracias al destino, también decidieron venir a Europa y gracias a lo cual tengo la posibilidad de verlos más seguido para alegrarme la vida: Rodolfo del Valle, Sebastián Flores y Fabián Salguero, gracias por estar *siempre* presentes. No lo olviden nunca: *la vida es una lenteja*. Para terminar, quiero agredecer a todos y cada uno de los miembros de la Changa del Friday (varios ya se repitieron el plato...); este magnífico grupo me permitió hacer durante todos estos años aquello que más le hace falta a un chileno viviendo en Francia: *reírse a carcajadas por puras weas*.

To finish, I would like to thank some people that I do not know personally (or not so much...), but who have been at any state indispensable for this work.

À la mémoire de deux grands mathématiciens qui ont rendu possibles tous les développements de cette thèse: George Matheron et Laurent Schwartz. Merci à Monsieur Matheron pour donner les bases théoriques de la pratique de la Géostatistique et pour présenter un cadre épistémologique inspirateur pour la compréhension des phénomènes (ou plutôt des *modèles*...) aléatoires. Merci à Monsieur *Papillon* Schwartz de nous permettre de regarder beaucoup de problèmes mathématiques *en hauteur*, grâce à une magnifique *généralisation pertinente*.

À Claudine Tiercelin et à Stanislas Dehaene pour leur magnifiques cours au Collège de France.

À tous les travailleurs et mécènes de la Bibliothèque Publique d'Information du Centre Pompidou et de la Bibliothèque Sainte-Geneviève. Une bonne partie de cette thèse a été rédigée dans ces beaux endroits, en profitant de la gratuité et de l'ouverture jusqu'à 22h00.

À tous les travailleurs du meilleur resto-kebab du monde: Manhattan. J'exagère à peine lorsque je dis que cette thèse n'aurait jamais vu la lumière sans un bon menu triple cheese toutes les semaines.

À la mémoire de Frédéric Chopin. To Radiohead, to Sigur Rós, and to Pink Floyd (all of its *five* members, with special dedication to the memory of Syd Barret and Richard Wright). Thanks for so much.

Finally, I would like to acknowledge two special historical persons. Even if these two persons worked in subjects *a little bit* far away from the one I work in, it must be said that their *opus*, their ways of thinking and, above all, their *ways of being*, are and will always be a deep source of inspiration and guidance for my work, my own personal development, and for any possible achievement I could obtain in this life. I dedicate this work to the memory of Friedrich Nietzsche and Ludwig Wittgenstein. I thank the first one for liberating

me from the intellectual biases produced by my owns desires and convictions. I thank the second one for liberating me from the intellectual biases produced by a misunderstanding of how language's creation works.

Anyone who at any time uses or get influenced by this work or by any other future work I may do, should also acknowledge each one of the people mentioned in this section. Nothing could be possible without them.

Contents

Chapter 1

Introduction

1.1 Geostatistics and Stochastic Calculus

The stochastic modelling of natural phenomena can be done following methodologies grounded on many different approaches[1]. Let us consider two approaches in particular: the approach of *Stochastic Analysis* or *Stochastic Calculus*, and the approach of *Geostatistics*. Within the first approach, the modelling of a particular phenomenon is often done by following a physically-based model imposed to the phenomenon, usually expressed through a differential equation. This differential equation is then *stochastized* in some sense, that is, some *random mathematical objects* are added in the equation or the deterministic objects are interpreted as random mathematical objects. These objects may describe, for example, a *noise* acting as a source term or as a force vector field, the structure of irregular media or geometries, an initial or boundary condition which we do not know in detail, etc. The resulting differential equation is then called a *Stochastic Differential Equation* (SDE), and when the problem is treated in a spatial context with dimension higher than 1, it is called a **Stochastic Partial Differential Equation** (SPDE). The branch of mathematics which rigorously formalizes these notions and studies its properties and rules of use is called *Stochastic Calculus* or *Stochastic Analysis*, and it is a sub-branch of Probability Theory. Its main mathematical tool is the *Stochastic*

[1]We understand by *stochastic modelling* of a natural phenomenon any *mathematical* modelling of a natural phenomenon grounded on the consideration that *we do not know* how this phenomenon behaves and that *we do not know how, or we do not want to describe it with full precision*. The objective is then to describe *grosso-modo* the behaviour of the variables involved, to recover their main *general* characteristics and to describe roughly their variability. The modelling is done by describing the behaviour of the unknown quantities through random mathematical objects defined in *Probability Theory*, more precisely, *random variables*, whose behaviour is determined by *probability laws*. The *precision level* with which the natural phenomenon is studied in a stochastic context depends on the needs and objectives of the user of the model. The criterion to select such a precision level is then, left to the freewill of who makes the model and why does this person wants it. We understand by a *random phenomenon* any phenomenon for which *we* do not exactly know how it behaves. In the last rigorous sense, every natural phenomenon is a random phenomenon, but in practice, we call *deterministic* phenomenon every natural phenomenon which we know and understand how it behaves up to some precision level which is sufficient for our objectives and needs. Hence, the *determinism* or *stochasticism* of a natural phenomenon is not really a property of the phenomenon but of our knowledge about it.

Process. Stochastic Calculus provides then a rigorous framework to treat mathematically the intuition of a SPDE and to apply it to model natural phenomena.

On the other hand, we have the *geostatistical approach*[2]. In principle, this approach does neither consider a physical model that the studied variable must follow, nor some kind of differential equation it must satisfy. This approach is based on *Data Analysis*, studying data-values of variables changing over the space and/or time. A typical methodology consists in interpreting the obtained data-values as the evaluation of a realisation of a *Random Function*. The variability of the studied variable is then described statistically through a selected *covariance function*, *variogram* or another mathematical tool determining the Random Function. The *selected* model must, in some sense, *fit the data* obtained during the study of a particular case of a natural phenomenon. Once the model is selected, it is then used to treat the unknown quantities of the phenomenon: prediction at a future time or at a non-sampled location, or the general behaviour of the variability of the variable along the space-time, etc.

These two approaches differ in their inspiration but coincide in the mathematical tools used. A *Stochastic Process* and a *Random Function* are *exactly* the same thing: both are a family of random variables indexed by a non-empty set[3]. The difference relies rather on the *way to describe it or to determine it*. Stochastic Analysis does it by imposing a SPDE the Random Function must satisfy. Geostatistics does it by imposing a covariance structure that the Stochastic Process must follow. Under suitable mathematical conditions, both approaches determine completely this mathematical object. From these considerations, a question arises somewhat *naturally*: are these approaches related? Is it equivalent to fix a covariance structure the Random Function must follow or to fix a SPDE that it must satisfy?

The answer is, roughly speaking, *yes*. In this dissertation we will get into the details of this answer and we will address other questions arising from this issue. In an intuitive way, we can remark the following fact: Geostatistics has always described the *increments* of the studied variable with respect to changes in the spatial or temporal components. These increments are modelled as *random quantities* described *statistically* through the specification of their laws, moments or mutual dependences structures. On the other hand, in a typical deterministic modelling context the increments of a studied variable are described infinitesimally through a differential equation. Hence it also describes the increments with respect to changes on the spatial or temporal components, or with respect to others variables of interest. It is not very surprising then that the *stochastized* version of such a differential equation would describe the variability of the variable in a similar

[2]In what concerns this paragraph, we can use as synonyms, Geostatistics, Spatial Statistics, Spatio-temporal Statistics, Time-Series Analysis, etc. In general, any branch of Statistics for which its methodology of study considers the *place where* and/or the *moment when* the data-value was obtained as an important and determinant information, in addition to the data-value itself.

[3]Some authors restrain the term *stochastic process* to the cases where the indexation set represents a time-interval or an ordered set. The concept of Random Function is then more general, usually used when the indexation set is the space or the space-time. Another typical terminology which involves both concepts is the term *Random Field*. We have decided not to make a strong distinction between these terms, since, mathematically, they present no difference in principle: what changes is only the indexation set, which will be always specified.

way as Geostatistics does: through an interpretation of the increments as *random quantities*, but controlling the statistical behaviour of these increments through a differential equation imposing a particular behaviour to the infinitesimal increments. Hence, it is reasonable to think that both methodologies are connected and may be equivalent under some suitable hypotheses.

In the last decade a new geostatistical modelling paradigm based on these considerations (either explicitly or implicitly) has been developed. It is called the **SPDE Approach**. It has arisen from the needs of the *statistical community* and not from the probabilist community. It consists in interpreting the studied variable as the realisation of a Random Function which satisfies some SPDE. Although this kind of modelling has always been done in Stochastic Analysis, it has not necessarily been grounded on the need of conveniently fitting a stochastic model to a data-set, nor by the need of interpreting statistical techniques in an *analyst* way. This approach has allowed many theoretical and practical developments. From the practical point of view, it allows the analysis of geostatistical models through the use of numerical tools used in the analysis of Partial Differential Equations (PDEs). PDE numerical solvers such as the Finite Element Method (FEM) or spectral methods can now be used to inspire new simulation and statistical inference methods of geostatistical models. All the imaginable benefits of the world of Numerical Analysis are then applicable in Geostatistics. In particular, the computing time for simulations and inference methods has been notably reduced thanks to the fast computing performance of PDE numerical solvers in some contexts. From the theoretical viewpoint, this approach has allowed the introduction of new geostatistical models related to SPDEs which can be added to the already known valid covariance models. In some cases, these models can present a traditional physical meaning, and hence, the parameters of classical geostatistical covariance models can carry a traditional physical interpretation. A classical geostatistical parameter such as the *scale*, which describes roughly the spatial or spatio-temporal *range*, defined as the distance below which the correlation is significant enough, can be interpreted as a *damping parameter*. Other parameters, now considered as *parameters of the associated SPDE* rather than of the covariance model itself, can be also physically interpreted. This is the case for example of a *velocity vector*, a *diffusivity coefficient* or an *anisotropic diffusivity matrix*, a *curvature coefficient*, or a *wave propagation velocity*.

In the next sections we will enter into the details and precisions of such a paradigm, giving the adequate bibliographical sources of the statements claimed in this section.

1.2 The SPDE Approach in Geostatistics: state-of-the-art

The term *SPDE Approach* was first used in the seminal discussion paper by Lindgren et al. (2011). The selection of such a terminology arises from the point of view of a statistical community which did not necessarily face their problems using concepts and methodologies associated to Stochastic Calculus. Hence, it was worth being called a *new approach* in geostatistical analysis. In this paper, the authors considered a

SPDE over the Euclidean space $\mathbb{R}^d$ of the form

$$(\kappa^2 - \Delta)^{\frac{\alpha}{2}} U = W, \tag{1.1}$$

where $\kappa > 0$, $\alpha > \frac{d}{2}$, Δ is the Laplacian operator and W is a Gaussian White Noise. The operator $(\kappa^2 - \Delta)^{\frac{\alpha}{2}}$ is fractional *pseudo-differential* operator which can be defined through the Fourier Transform. Then, they exploited the theoretical link between this equation and the stationary covariance function

$$\rho(h) = \frac{1}{(2\pi)^{d/2} 2^{\alpha-1} \kappa^{2\alpha-d} \Gamma(\alpha)} (\kappa|h|)^{\alpha-d/2} K_{\alpha-d/2}(\kappa|h|), \quad h \in \mathbb{R}^d, \tag{1.2}$$

where $K_{\alpha-\frac{d}{2}}$ denotes the modified Bessel function of the second kind of order $\alpha - \frac{d}{2} > 0$. The members of this class of covariance models are called Matérn models or $K-$Bessel models. The function (1.2) is actually the only possible covariance function which a stationary solution to Eq. (1.1) can follow. This theoretical result was obtained in Whittle (1963). Before entering into the details of the exploitation of this theoretical link done in Lindgren et al. (2011) and in reasons as to why this idea has been so beneficial and fertile, let us make a little historical analysis of similar theoretical relationships between covariance models and SPDEs.

The idea of obtaining covariance models arising from solutions of SPDEs is actually quite old. Rigorously speaking, it can be said that the probabilist community has always done this. Since the very beginning of the theory of Stochastic Processes and Stochastic Calculus, the covariance function has been an important mathematical tool which characterises roughly the stochastic processes involved[4]. However, they do not necessarily use this mathematical tool as a *basis* of model construction or modelling methodology. From the statisticians standpoint, most techniques are grounded on this mathematical tool, such as kriging, simulation methods, conditional simulations and inference methods based on the analysis of the second moments of the random variables involved such as variographic analysis or likelihood methods in square-integrable contexts. Within the statistical objectives, many authors have obtained and described covariance models from the resolution of stochastic differential equations. The earlier works we have found in this aim are those of Heine and Whittle. Heine (1955) presents formulas of stationary covariance functions describing solutions to some SPDEs involving hyperbolic, parabolic, and other type of second order differential operators in spatial dimension 2. In Whittle (1954) the author is inspired by a typical time-series interpretation, analysing the increments of a random process with respect to constant gaps in the temporal domain, adding independent innovation terms. The analogue idea is then applied to the spatial case with symmetric second-order gaps in two dimensions. It is then concluded that such a model follows a stochastic Laplace equation with

[4]Kolmogorov and Prokhorov referred to the *correlation function* (Kolmogorov & Prokhorov, 1992). Without using a particular name for it, Doob presents this concept in the framework of real Gaussian Processes where it plays a determinant role since, together with the mean function, it determines completely the Random Function (Doob, 1953, Chapter II, Section 3). Itô also refers to the *Khintchine's covariance function* referring to Khintchine (1934), and he also refers to the *covariance distribution* in the case of Generalized Random Processes (Itô, 1954).

damping (Eq. (1.1) with $\alpha = 2$). In Whittle (1963), the author generalizes these results. He develops a general framework where stationary Random Functions are related to some SPDEs, showing the link between Eq. (1.1) and the Matérn model and presenting other examples such as spatio-temporal models related to diffusion equations with damping. Later, Vecchia (1985) proposed models obtained from SPDEs involving operators which are compositions of operators exposed in Heine (1955) and in Whittle (1963), defining spectral densities following products of diverse spectral densities. Gay & Heyde (1990) proposed models based on solutions of SPDEs involving fractional Laplacian operators without damping parameters ($\kappa = 0$ in Eq. (1.1)). These models are said to have a long-range dependence covariance structure, and they have been worked out in great detail in Kelbert et al. (2005) and in the works of members of the Granada school (Anh et al., 1999; Angulo et al., 2000; M. Ruiz-Medina et al., 2003). In Anh et al. (1999) examples of such fractional models are exposed, and their regularity is analysed through the use of the Reproducing Kernel Hilbert Space associated to the covariance structure, presenting also a SPDE these models must satisfy. In Kelbert et al. (2005), the authors obtain models associated to fractional forms of the stochastic Heat equation. A summary exposition of this kinds of models can be found in M. Ruiz-Medina et al. (2003). Fontainebleau's school of Geostatistics also provided advances in this framework. In the doctoral thesis Dong (1990), different covariance structures of univariate and multivariate geostatistical models are obtained from the analysis of PDEs, particularly the cases of the Poisson equation and other equations arising from Hydrogeology. In Jones & Zhang (1997) stationary covariance spatio-temporal models issued from some SPDEs are developed, which allow the construction of non-separable models. Examples of the stochastic diffusion equation with damping and associated generalizations in spatial dimension 2 are presented. More recently, Lim & Teo (2009) and Bolin & Lindgren (2011) proposed covariance models which are associated to more general forms of the SPDE (1.1), and hence presented as generalizations of the Matérn Model.

It is then concluded that the idea of obtaining new covariance models from SPDEs is not a new idea but rather an already well established practice in the statistical community. However, the *exploitation* of these models *taking advantage* of this *explicit link* between their covariance structures and a SPDE is a quite new practice. Bibliographical sources can be found where this link is exploited in the *backwards sense* as it has been done in Lindgren et al. (2011). That is, *a SPDE is fitted to data using classical statistical techniques*. The earlier work we have found considering the explicit notion of fitting a SPDE to a data-set is in Jones (1989). In this work, the author proposes to fit a SPDE to aquifer data by considering the example of Eq. (1.1) with $\alpha = 2$. He presents the SPDE, makes reference to the relation with a Matérn model following the results obtained in Whittle (1963), and then fits the model to a data-set using classical likelihood-based statistical methods. The relation with the SPDE is almost anecdotal. Jones fitted a SPDE without never really considering the SPDE itself, but using its related covariance model and applying typical statistical techniques. In the conclusion of its work, Jones also states that this method was somewhat a "brute force" method due to the high computational cost of likelihood based algorithms.

From these considerations we can identify the real *major contribution* of the work Lindgren et al. (2011).

The approach in this paper is on the opposite direction as the one in Jones (1989). The authors begin considering the Matérn covariance model, which is particularly popular among the statistical community. They then present the connection with the SPDE (1.1) obtained in Whittle (1963), and then they *forget the related covariance function and work with discretized versions of the SPDE*, which are provided by the FEM. Hence, the approach consists in considering that the Random Function follows a particular SPDE which determines its covariance structure and then work with *that SPDE*, rather than with the covariance function itself. It turns out that this method presents real advantages in the case of the Matérn model (1.2) with integer values of α. The authors show that when applying the FEM to discretize the linear SPDE (1.1), the matrix involved in the numerical method can be identified as the precision matrix (that is, the inverse of the covariance matrix) of the approximated Random Function evaluated at the nodes of the triangulation mesh, which in this case is sparse. The sparsity of the matrices involved in the FEM approximation is theoretically justified by the Markovian behaviour of the Matérn model for integer values of α. This consideration allows to immediately work with a sparse precision matrix completely determined by the FEM triangulation, hence avoiding inverting the covariance matrices as it is done in usual geostatistical techniques. This has allowed to reduce considerably the computational time of geostatistical techniques which require the precision matrix, such as non-conditional and conditional simulations, Bayesian inference methods, kriging, etc. Precisely, the authors show that the complexity of the computations are reduced from $\mathcal{O}(N^3)$ to $\mathcal{O}(N^{\frac{3}{2}})$ in $2D$, N being the number of sampling points. This method provides then a methodology for handling large to very large ($> 10^6$) data sets. As stated in Jones (1989), such an amount of data sets could not be treated satisfactorily with classical likelihood based methods.

The SPDE approach has then set a new paradigm for geostatistical modelling. The Matérn model, which enjoyed a considerable popularity within the statistical community even before the introduction of the SPDE approach (see the conclusive expression "*use the Matérn model*" in Stein (1999, page 14)), is equipped now with new treatment techniques which makes it even more attractive for practical applications. Thanks to the fast computation treatment provided by this technique, the SPDE approach has been widely used for analysing large data sets, in particular in environment or climate science (Bolin & Lindgren, 2011; Cameletti et al., 2013; Huang et al., 2017; Mena & Pfurtscheller, 2017). Some authors even consider that it is no longer really necessary to explicitly use the covariance function when analysing some geostatistical models right now, since we can now count on a SPDE which implicitly imposes a covariance structure and whose numerical resolution provides more practical treatment techniques. We refer to Simpson et al. (2012) for such a discussion, together with a comparison of the computational benefits when using SPDE approach techniques with respect to other classical geostatistical techniques. In addition, since the positive-definiteness restriction on a covariance function makes the construction of new models intricate, the SPDE approach allows to implicitly construct models through the specification of SPDEs. This has allowed, for example, the development of non-stationary models (Fuglstad et al., 2013). Generalizations of the Matérn model to more general manifolds as in the case of the sphere representing the planet Earth can now easily be obtained through the

resolution of suitable SPDEs defined over the sphere; see the application section in Lindgren et al. (2011) for the details on this approach, and see Lang et al. (2015) for a theoretical analysis of Gaussian Random Fields related to SPDEs over the sphere, together with practical simulation methods. It is interesting to contrast this SPDE-based construction of covariance models with other more classical techniques of constructing positive-definite functions over the sphere, framework which presents special theoretical issues. We refer to Porcu et al. (2016) and White & Porcu (2018) for the difficulties and advances within this *classical* geostatistical approach, considering covariance functions over the sphere cross time. The SPDE approach has also inspired the development of other PDE-solver based methods with efficient performances for a wider class of models. We refer to Sigrist et al. (2015) for the study of a stochastic form of the advection-diffusion equation with damping, using Fourier Analysis methods in space and a strict resolution of the SPDE in time to perform efficient simulation techniques based on the Fast Fourier Transform. See Liu et al. (2016) for the case of approximations of Matérn models over the space using bivariate splines, approach which allows, in particular, to easily consider extensions to non-stationary models. Finally, we refer to Bolin & Kirchner (2017) for adaptations of the FEM to the cases of Matérn models with non-integer parameter α.

We finally remark that the SPDE approach has allowed to consider particular physical meanings for some parameters of the developed models. We consider for instance the already mentioned case of the advection-diffusion equation worked out in (Sigrist et al., 2015), where the SPDE involves parameters defining a damping number, a velocity vector, and an anisotropic diffusion matrix. We remark also the work M. D. Ruiz-Medina et al. (2016) which allows to obtain new spatio-temporal covariance models related to SPDEs with fractional regularities on time and defined over bounded sets in space. The approach consists in the resolution of a deterministic version of a fractional PDE (without random source terms) considering a random initial condition. The solution Random Field and its covariance are expressed through their developments in a convenient orthonormal basis of functions, and the models are treated using wavelet-based methods. Such models involve, for example, fractional versions of the Heat equation with fractional temporal derivatives and a fractional spatial Laplacian operator.

1.3 Objectives

This new approach relating Geostatistics and the analysis of SPDEs open many doors in both theoretical and practical aspects. When this PhD project was conceived, the main general questions that were aimed to be worked out were the following ones:

1. How can we obtain new covariance models from the analysis of SPDEs, in order to add them to the *catalogue* of valid available models? How can we describe the main properties of these models such as its variance, range and regularity from the analysis of the parameters of the associated SPDE? Which of these new models are related to classical physically driven PDEs, and hence with parameters

describing for example transport, diffusions and wave propagation phenomena, among other possibilities? How can we apply conveniently this approach in order to obtain models in a spatio-temporal context with non-trivial properties?

2. Once the link between a geostatistical model and a SPDE is established, how can we exploit this link in order to obtain ad-hoc treatment techniques of the geostatistical model? In particular, which PDE numerical solver approach is more convenient to use for solving a SPDE in order to obtain an adapted framework for simulations and statistical inference methods?

3. How can we relate well-known geostatistical models to particular classes of SPDEs, and hence allowing to treat these models with techniques issued from the SPDE approach?

These three questions are very generic and they lead to many different research works involving more specific questions. In this dissertation we have mainly worked out the *theoretical* issues which appear when facing these questions, mainly for questions 1 and 2. Let us present the issues and objectives of this work in this aim. Since all the chapters in this dissertation present suitable introductory and discussion sections with plenty of details and bibliographical sources, in this section we will not give many bibliographical sources for our statements. All of them are treated more deeply further.

Question 1 proposes the challenge of relating explicitly the resolution of SPDEs with the construction of covariance models. Hence, here we have to study and exploit the connection between the framework of Stochastic Analysis and Geostatistics introduced roughly in Section 1.1. The question is rather *how does the SPDE impose a particular behaviour to the covariance of a stochastic process*. This requires us to enter into the details of Stochastic Calculus and the resolution and well-posedness of SPDEs. Hence, the technical details which appear in the theory of stochastic processes and in Stochastic Analysis are present. In particular, the well-posedness of a SPDE is a crucial question. The definition of a differential operator acting on a stochastic process is one of the first basic issues in Stochastic Calculus, since many of the most important stochastic processes do not have a regular behaviour. Even if such operation is well-defined, the questions about the existence and uniqueness of solutions to some SPDEs, including in which sense we interpret these potential solutions, are determinant. In some situations there are no solutions to a SPDE, hence there are no covariance models at all to be concerned about. In other cases there exist many different solutions, so there may be many possible covariance models whose associated Random Fields satisfy the equation, hence the covariance structure is not completely determined by the equation. If we are in the case of uniqueness of a solution, we still have to verify, in general, if such a solution has a covariance structure: it could be a stochastic process with no square-integrable evaluations. Only after all these issues have been tackled, we can really consider *the* covariance model which *the* solution to the SPDE follows. Finally, by imposing a particular square-integrable multi-dimensional law which the process must follow (Gaussian, for instance), we can really say that in such a context posing the SPDE and fixing the corresponding covariance

model are equivalent methodologies.

It is then necessary to select a convenient framework between the possible ones already present in Stochastic Analysis to work with. This is not an immediate selection to do and a naive choice may not be very adapted to the objectives of this dissertation. We could try, as a first tentative, to use a typical *mainstream* framework in Stochastic Calculus as the one based on a particular stochastic process: the Brownian motion (or the White Noise, which is its *derivative*). Such a framework is the basis of Itô Calculus and other similar approaches of Stochastic Calculus involving stochastic integrals. This framework often deals with the analysis of filtrations, martingales and Markovian behaviours, and it is usually inspired by a strictly *temporal* framework. However, a geostatistician who just starts entering into the technical details of Stochastic Analysis may be a little bit surprised by the general need of basing the whole theory on this particular stochastic process or on the already mentioned concepts. In principle, a geostatistician is not particularly interested in specific technical conditions such as a martingale behaviour or measurability along a particular direction of the axes using filtrations. Such kind of properties are rather characteristics of particular stochastic processes which the geostatistician has no reason to impose to a model in a first sight. Even Markovian models are not theoretically preferred in principle. In practice, they are quite popular for practical reasons such as allowing fast computations through the specification of sparse precision matrices, as already mentioned in Section 1.2. Although this motivation is quite important, in the general sense there is no other reason to restrict our work to a Markovian model[5]. At the end of the day, the data-set and the simplicity of the model are the main criteria determining which kind of model is preferable for a particular situation.

Concerning the definition of Stochastic Differential Equations, an interesting methodological question arises. If the aim of Stochastic Calculus is *to do Calculus with Random Functions*, where does the need of fixing a *basis* stochastic process, such as Brownian motion or White Noise, come from? If classical deterministic calculus and analysis of PDEs are not *based* on a particular function but rather on the concepts of continuity, differentiability, integration and other related concepts, why does Stochastic Calculus need to be based on a particular stochastic process? The answer is simple: *it does not*. Indeed, all we need is a good definition of differential operators acting on a stochastic process. This includes the correct specification of these operations, the class of stochastic processes that they can be applied to, and to which class of stochastic processes belongs the result of these operations. The necessity of fixing a particular stochastic process as basis, or to restrict the work to processes presenting martingale or Markovian behaviour is not really present. The construction of a stochastic integral, which is a tool often used to solve SPDEs, does not really require to be based on Brownian Motion, nor does it necessitate a process presenting a Markovian behaviour or being a martingale over the time. All we really need to define a stochastic integral is a *Random Measure* and a precise *Integration Theory* with respect to this measure.

It is then necessary to focus in other, maybe less mainstream but also simpler theories of stochastic calcu-

[5]For instance, geostatistical techniques do not require the concept of *causality*, even when working in a temporal context. Hence, the notion of *the future depending on the past* in a particular manner is not really necessary.

lus. First of all, the typical geostatistical framework works with the covariance structure, and hence a square-integrability condition must be imposed to the processes. The so-called *mean-square theory* (Sobczyk, 1991) seems to be the most adapted to geostatistical purposes. Here the convergence and equalities of the random variables are all considered in a mean-square sense. Hence, the connection with the covariance structure framework in Geostatistics is immediate. In addition, one can construct Stochastic Differential Equations without imposing the right side to be a typical model such as a White Noise. Instead, one only interprets the equation as a PDE with Random Functions involved. Hence, this framework is not *based* on any particular stochastic process, for instance, not on White Noise[6]. Finally, the definition of a differential operator acting on an arbitrary stochastic process poses more sophisticated theoretical issues. These are often worked out by interpreting the differential equation as an integral equation. However, another option which is *simpler*, is to use the more sophisticated theory of *Generalized Random Fields*, that is, Random Distributions, the stochastic version of the Theory of Distributions. In this theory many operations such as differential operators and the Fourier Transform can be applied freely, and even some pseudo-differential operators interpreted as fractional forms of classical differential operators can be applied under suitable conditions. We will thus see that this framework is perfectly adapted to the analysis and treatment of linear SPDEs over the space or space-time and for the description of the covariance structure of Random Fields. This framework also allows to put in the same bag Random Functions, Random Measures and Random Distributions and to work with them in a unified context. For instance, for the analysis of stationary Generalized Random Fields, the freedom we gain when considering the Fourier Transform of a stationary Random Field as an orthogonal Random Measure allows us to obtain a quite special treatment in this context and to well-define, analyse and solve a wide-class of linear SPDEs, with a simple description of the covariance model following immediately. This is thus, the framework which we decided to choose in this dissertation: Generalized Random Fields in a mean-square analysis context. The mathematical tools needed to develop this framework are exposed in Chapters 2 and 3. Its application to analyse stationary solutions for a wide-class of linear SPDEs is presented in Chapter 4 and in Chapter 5 we apply it to develop spatio-temporal covariance models presenting non-trivial properties and being related to physically driven SPDEs.

Concerning Question 2, in this dissertation we deal with the problem of choosing a suitable PDE solver numerical method which can be easily adapted to the developments presented in this dissertation. In the literature the main methods are the Finite Element Method and spectral methods. Both kinds of methods present advantages and disadvantages, mainly considering its versatility to treat wide classes of SPDEs, and hence wide classes of possible covariance models in a geostatistical framework. Considering the developments in this dissertation, we concluded that the most adapted method for non-conditional simulation of stationary models related to SPDEs is the one based on a spectral method based on the approximation of the Fourier Transform of the Random Field. This method is closely related to the development on the Fourier basis,

[6] Along this dissertation we will see that the White Noise, while not a cornerstone, has many special properties which makes it meritorious of our attention, either for theoretical or practical purposes.

although it is not exactly *the same* method. This method is not new (Pardo-Iguzquiza & Chica-Olmo, 1993; Lang & Potthoff, 2011), although in the literature there is no theoretical proof for its performance. Within our framework we have been able to prove theoretically the convergence of the numerical method to the theoretical solution of the associated SPDE in convenient weak and strong senses. We have been able then, to apply it to illustrate approximations of the models presented in Chapters 4 and 5, specially those presenting non-trivial properties and related to physically driven SPDEs. These developments are presented in Chapter 6. We do not enter on the problem of developing adapted inference methods and conditional simulations for the SPDE Approach.

Regarding Question 3, on the relation of known geostatistical models with particular SPDEs, we do not present general explicit advances in this dissertation. Implicitly, in Chapter 4 we develop a framework where it is easy to relate a stationary covariance model to a SPDE when we know the spectral measure of the Random Field and if it has the form of a density with respect to another spectral measure. This has allowed us to obtain new relationships between some known covariance models and some type of SPDE which we will specify further. However, the problem of relating a general geostatistical model to a convenient SPDE has not been tackled. In the conclusive Chapter 7 we show indices for advances in this aim, which are also embedded in our Generalized Random Fields framework with a mean-square approach.

1.4 Organisation of this dissertation

This dissertation is organised in five main chapters. The first two chapters are mainly expositions of mathematical tools used in this dissertation. The other three chapters are devoted to new results and applications with geostatistical purposes.

In Chapter 2 we introduce the main *deterministic* mathematical tools which are required on the formalism of the SPDE approach. It consists of an exposition of Measure Theory over the Euclidean space and of Distribution Theory. We present the notion of a complex locally finite random measure over the Euclidean space. We remark the special cases of slow-growing, compactly supported and finite measures. We recall the classical Riesz Representation Theorem which allows to characterize locally finite and finite measures as continuous linear functionals over convenient spaces of continuous functions, and we present analogue results for the cases of slow-growing and compactly supported measures. We remark also the important case of measures concentrated on subsets of the Euclidean space. We present the most important definitions and results of Distribution Theory in a tempered framework. We then show how differential operators and the Fourier Transform are applicable in this context. We present the concepts of tensor products of distributions and operators, and we recall important results such as the Exchange Formula for the Fourier Transform and the Nuclear Theorem.

In Chapter 3 we present in detail the stochastic tools used both in classical geostatistical frameworks and

in this dissertation. We present the classical framework of Random Functions, with all necessary notions to relate geostatistical analysis with basic theories of PDEs, such as the definitions of continuity, differentiability and integrability in a mean-square context. We then introduce the *stochastized* version of the deterministic tools presented in Chapter 3, that is, Random Measures and Generalized Random Fields or Random Distributions. We define those tools in the context of the *mean-square* theory, where the main characteristics of these objects are determined by the characteristics of the covariance structure. For instance, a Random Measure is defined as being determined by a covariance measure. The cases of finite, slow-growing and compactly supported Random Measures are related to analogue properties for the covariance measure. We present the important case of orthogonal Random Measures. We show how to define integrals of deterministic functions with respect to Random Measures within this framework. We then present the theory of Generalized Random Fields. We show the main important aspects which allow to properly deal with differential operators and the Fourier Transform on Random Fields with a huge generality. We present the definition of a stationary Generalized Random Field and we recall the important result relating them with slow-growing orthogonal Random Measures through the Fourier Transform. This result is widely used. We then present our definition of a SPDE, and we show how linear SPDEs impose deterministic PDEs to the covariance structures of the involved fields. We also give a brief but enlightening way to construct bivariate geostatistical models through the SPDE approach. The final two sections of this chapter are devoted to explain the theoretical issues that arise in Stochastic Analysis and the differences and similarities between the framework used in this dissertation and other typical approaches to Stochastic Calculus. We present the differences between *mean-square theories* and *sample-paths* theories, which are determinant on the cases of Random Functions and Random Measures. We also present the theoretical issues that arise when trying to define multiplications between GeRFs and hence when trying to define non-linear SPDEs or SPDEs involving a multiplicative noise. We show that this issue is also related to the classical problem of the non-canonical definition of a stochastic integral of general stochastic processes with respect to Random Measures. These theoretical issues push us to restrict our work to the cases of linear SPDEs involving deterministic operators.

In Chapter 4 we generalize the results in Whittle (1963) regarding stationary solutions for a wide class of linear SPDEs. Within the framework of stationary Generalized Random Fields, we obtain conditions under which there exist strict stationary solutions and under which there is a unique solution. The criteria consists in a suitable integrability condition between the symbol function defining the operator and the spectral measure of the source term. We present the particular case of a White Noise source term and we show that it can be considered as a fundamental case since, under suitbale conditions, the covariances of the solutions with more general source term can be related to the one of the White Noise source term case through convolution. These results allow to recover and encompass stationary models already present in the literature. They also give a direct enlightening on the problem of relating stationary models to SPDEs. We recall the Matérn model, the Matérn model without scale parameter, and Markovian stationary models. We present examples of SPDEs related to the J-Bessel covariance model and the Stein model (Stein, 2005). We end this chapter

with a remark concerning the associated deterministic PDEs. This framework can be used to analyse the means of the Random Fields involved in a SPDE without necessarily supposing a constant or zero mean as it is done in a stationary framework.

In Chapter 5 we present new spatio-temporal covariance models which we obtained within this SPDE approach. We begin by recalling the special issues and difficulties in the framework of spatio-temporal Geostatistics. We recall the concepts of separability, symmetry, spatial and temporal traces and margins in a classical geostatistical framework, and we show how to properly define these notions in the framework of Generalized Random Fields. We present then new stationary spatio-temporal covariance models associated to evolution equations presenting a fractional derivative order in time and an arbitrary operator defined through a symbol in space. We present sufficient conditions for existence and uniqueness of a stationary solution regardless of the source term and the imaginary part of the spatial symbol function. We then describe the spectral measures associated to these models and we show how we can easily control the separability, the symmetry, and the separate spatial and temporal regularity of the model. We give more details in the cases of first order and second order evolution models. We describe the covariance of the spatial traces of these models in symmetric cases. We point out already existent particular cases of these models which are present in the literature, such as the case of the advection-diffusion equation (Sigrist et al., 2015) and some Langevin's equations (Hristopulos & Tsantili, 2016). We then introduce the *Evolving Matérn* models, which are spatio-temporal stationary solutions to these evolution equations which follow a Matérn spatial covariance model. We also obtain interesting results on the existence and uniqueness of stochastic forms of the Heat and Wave equations. In the case of the Wave equation, we show that there exists a great variety of stationary models which satisfy spatio-temporally its homogeneous form, and which can be chosen to follow an arbitrary spatial covariance behaviour. We call these kinds of models *Waving models*. The last section of the chapter is devoted to the study of first order evolution models satisfying an initial condition. The results, presented informally, generalize well-known results on the analysis of such type of spatio-temporal PDEs and SPDEs, which involve for instance the advection-diffusion equation, Langevin's equations and the Heat equation with a fractional Laplacian operator. The problem is solved considering solutions in a suitable space of tempered distributions for which an *initial condition* makes sense. Under suitable assumptions, it is claimed that the solution of this initial value problem converges asymptomatically spatio-temporally as the time flows to the spatio-temporal stationary solution (in the geostatistical sense) associated to a first order evolution model, already studied in this chapter.

In Chapter 6 we present a method of simulation of approximations of the models developed in this dissertation. It is a well-known simulation model based on the spectral representation of stationary Random Fields, taking advantage of the computational benefits of the Fast Fourier Transform. This method has already been introduced by Pardo-Iguzquiza & Chica-Olmo (1993) in a geostatistical context and by Lang & Potthoff (2011) as an efficient numerical method for solving suitable SPDEs. This method turned out to be perfectly adapted to the approach presented in this dissertation. We present the theoretical basis of

the methods and we have been able to theoretically prove the convergence of the method when increasing the approximation order. The convergence is considered in suitable mean-square weak and strong senses. We show how to apply this method to the resolution of SPDEs of the form presented in Chapter 4, to numerically solve the initial value problem related to first order evolution models, and to simulate Waving models. We implement this method in a particular convenient setting. We show a qualitative error analysis of the simulated approximations by comparing the average over 50 independent realisations with the theoretical variogram in the case of the Matérn model. We then illustrate simulations of different type of models, presenting advection effects, different regularities along different directions, and non-symmetric behaviour inspired by the developments done in Chapter 5. We also present illustrations of first order evolution models with random initial condition and of Waving models. We finish with final words on the advantages and disadvantages of this simulation method.

We finish with the conclusive Chapter 7 where we summarize the obtained results and we present possible future courses of research within the SPDE approach, which are closely related to the issues exposed in this dissertation.

Chapter 2

Theoretical Framework: Deterministic Tools

SUMMARY

In this chapter we present the main non-stochastic mathematical tools that will be used in this dissertation. It is basically a recall on Measure Theory for the Euclidean space and Distribution Theory. It can be considered as a special chapter of this dissertation which does not deal with strictly speaking geostatistical concepts. Hence, some geostatisticians who do not often use these theories may find this exposition useful. Some of the notions and terminologies we use here are not broadly used in classical treaties of these theories or they are presented in a different way as we do. Hence, even if the reader knows well these theories, we suggest anyway to make at least a fast reading of this chapter.

Section 2.1 deals with Measure Theory for the Euclidean space. Here the concept of locally finite or Radon complex measure over $\mathbb{R}^d$ is presented and exploited. We present the definition of complex measures as set functions. The vector space of complex measures over $\mathbb{R}^d$ is described. The space of finite complex measures is also presented, together with the concept of measure of total variation. We recall the construction of the Lebesgue integral with respect to positive and complex measures. We introduce the space of slow-growing complex measures which will be of great importance in this dissertation. We recall the classical Riesz Representation Theorem which states that any continuous linear functional over the space of compactly supported continuous functions is a complex Radon measure. Variants of the Riesz Representation Theorem are also presented, including the cases of compactly supported measures, finite measures and slow-growing measures as continuous linear functionals over the spaces of continuous functions, continuous functions vanishing at infinity and fast-decreasing continuous

functions respectively. We finish this section with an exposition of measures concentrated on subsets of the Euclidean space, where we remark measures concentrated on the Sphere, on the Hyperplane $\{y = x\}$ and on the spatio-temporal cone.

Section 2.2 deals with Schwartz's Distribution Theory. We restrain ourselves to the case of tempered distributions. We introduce the Schwartz space and its dual space of tempered distributions. Some examples of such distributions are given. We present examples of operations which can be defined for tempered distributions, all of them defined through an adjoint. We give the most important examples: differential operators, multiplication with multiplicators of the Schwartz space, convolution with fast-decreasing distributions and the Fourier Transform, together with its famous multiplication-convolution exchange formula. We recall the concept of tensor product for the cases of functions, measures, tempered distributions and linear operators over tempered distributions. Finally, we recall the important Nuclear Theorem. Some comments on other spaces of distributions are also given.

Since we have used some notions and terminologies which are different to standard ones, many claims presented in this section are not easily findable in the literature in the way we state them, even if they could seem obvious for some specialists. In these cases, we always give a proof in Appendix A or a convenient reference.

2.1 Measures over the Euclidean space

In this section we recall some concepts and results of Measure Theory for the Euclidean space $\mathbb{R}^d$, with $d \in \mathbb{N}_*$. We will always work with **Borel** measures, that is, our measurable space will always be $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, with $\mathcal{B}(\mathbb{R}^d)$ being the Borel $\sigma-$algebra of $\mathbb{R}^d$.

Some definitions and terminologies that we have chosen to use in this work differ with classical terminologies that can be found in most bibliographical sources. This choice is done mainly for practical reasons. We will sometimes make reference to some treaties or articles using the same mathematical objects as we do but with different *names* for those. We will specify the details when necessary. Some of the results presented in this section are not easily findable in the literature, mainly because of this different usage of terminology. For some of them, as Theorems 2.1.5 and 2.1.6, we are not aware about if they are new or not, but we did not find a source where they are stated *in the way we needed*. We give, of course, proofs of those and to any other result which is not immediately easy to find in the literature. We think, however, that these results are, if not evident, at least intuitive for a Measure Theoretician or for an Analyst.

We refer to Knapp (2005, Chapter 6) for a general description of positive Borel measures over the Euclidean space, and to Rudin (1987, Chapter 6) for a general theory of complex (finite) measures over abstract measurable spaces. Here we make a general compendium of the main ideas on those and other bibliograph-

ical sources, with no restriction to finite complex measures. Sources based on *the other way* of defining measures over $\mathbb{R}^d$, that is, by using the Riesz Representation Theorem as basis are also recommendable, for which we remark the student-oriented Demengel & Demengel (2000) or the classical (and polemical) Bourbaki (1965). This approach is also discussed in Section 2.1.4.

2.1.1 Locally finite complex measures as set functions

We begin by giving our definition of a *Radon complex measure*, or a *locally finite complex measure* over $\mathbb{R}^d$. The term *Radon measure* is mainly used for measures defined over more abstract measurable spaces with extra properties required for the measure, namely, inner and outer regularity. However, in the case of the Euclidean space with its Borel $\sigma-$algebra, the property of being locally finite is equivalent to being Radon. See Knapp (2005, Theorem 6.2), the author using the term *Borel* measure for what we would call here a *positive locally finite measure*. Hence, in this context it is not necessary to make a distinction between the adjectives *Radon* and *locally finite*. We will set-up the next convention: *all measures defined over the Euclidean space used in this work are supposed to be Radon, and so locally finite, unless explicitly stated otherwise*. We will then drop the adjectives Radon or locally finite unless it is useful to recall them.

We denote by $\mathcal{B}_B(\mathbb{R}^d)$ the collection of all **bounded** Borel subsets of $\mathbb{R}^d$.

Definition 2.1.1. *A **locally finite complex measure** (from now on, a **complex measure**, or simply a **measure**) over $\mathbb{R}^d$ is a function $\mu : \mathcal{B}_B(\mathbb{R}^d) \to \mathbb{C}$ such that for every countable collection of mutually disjoint bounded Borel sets $(A_n)_{n\in\mathbb{N}} \subset \mathcal{B}_B(\mathbb{R}^d)$ such that $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{B}_B(\mathbb{R}^d)$, it holds that*

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n). \tag{2.1}$$

This definition is not a traditional one. Most bibliographical sources require the measure to be defined over the whole $\sigma-$algebra of Borel sets and not just over the bounded Borel sets. Some authors use the term *pre-measure* for this mathematical object, at least in the positive case (see C. Rogers, 1970, Definition 5 in Chapter 1). Indeed, that name is often used when the function μ is not defined over the whole $\sigma-$algebra of subsets of the space but rather over a *ring* or over a $\delta-$*ring* of subsets of the space[1]. However, the stronger requirement that μ must be defined over the whole $\sigma-$algebra $\mathcal{B}(\mathbb{R}^d)$ produces problems when trying to define a complex measure over unbounded sets, since in those cases expressions of the form $\infty - \infty + i(\infty - \infty)$ may arise, even for very basic and useful measures (the Lebesgue measure, for instance). Actually, it can be proven that if we define a complex set function satisfying (2.1) for every arbitrary countable partition

[1] If X is a non-empty set, a *ring* of subsets of X is a collection of subsets of X stable under finite unions and under set differences. A $\delta-$*ring* of subsets of X is a ring of subsets of X stable under countable intersections. Every $\sigma-$algebra is a $\delta-$ring, but the converse is of course not true in general. The collection $\mathcal{B}_B(\mathbb{R}^d)$ forms a $\delta-$ring but not a $\sigma-$ring. See Rao (2012, Chapter 1) for an introduction to measure theory using these notions.

of any Borel set, this measure is necessarily a finite measure, notion that will be explained later (see Rudin, 1987, chapter 6). Definition 2.1.1 allows thus to bypass this problem since we are not concerced on what happens over unbounded Borel sets. Some authors use this notion of complex measure, often inspired by the manipulation of complex measures in Distribution Theory; see for example Schwartz (1966, Chapter I, §1).

The property related to Eq. (2.1) is called the σ**-additivity** property. We remark that the series in (2.1) must be absolutely convergent, since $\bigcup_{n\in\mathbb{N}} A_n$ is still the same set for every rearrangement of the family $(A_n)_{n\in\mathbb{N}}$. We denote by $\mathscr{M}(\mathbb{R}^d)$ the space of locally finite complex measures over $\mathbb{R}^d$. This space is a complex vector space with the sum $(\mu+\nu)(A) := \mu(A)+\nu(A)$ and with the scalar multiplication $(\alpha\mu)(A) := \alpha\mu(A)$, for all $\mu,\nu \in \mathscr{M}(\mathbb{R}^d)$, $\alpha\in\mathbb{C}$, and $A\in\mathcal{B}_B(\mathbb{R}^d)$. If a measure μ satisfies $\mu(A)\in\mathbb{R}$ for every $A\in\mathcal{B}_B(\mathbb{R}^d)$, it is said to be **real**. If a measure μ satisfies $\mu(A)\geqslant 0$ for every $A\in\mathcal{B}_B(\mathbb{R}^d)$, it is said to be **positive**. We denote by $\mathscr{M}^+(\mathbb{R}^d)$ the space of all positive measures over $\mathbb{R}^d$.

If $\mu\in\mathscr{M}(\mathbb{R}^d)$, its **reflection** measure $\check{\mu}$ is defined as $\check{\mu}(A) := \mu(-A)$ for every $A\in\mathcal{B}_B(\mathbb{R}^d)$, where $-A := \{x\in\mathbb{R}^d : -x\in A\}$. It is straightforward that $\check{\mu}$ is a well-defined measure. A measure μ is said to be **even** if $\mu=\check{\mu}$ and **odd** if $\check{\mu}=-\mu$. Its **conjugate** measure $\overline{\mu}$ is defined as $\overline{\mu}(A) := \overline{\mu(A)}$ for every $A\in\mathcal{B}_B(\mathbb{R}^d)$, and it is a well-defined measure. The **real part** of μ is the real measure $\mu_R := \frac{\mu+\overline{\mu}}{2}$, and the **imaginary part** of μ is the real measure $\mu_I := \frac{\mu-\overline{\mu}}{2i}$, satisfying $\mu=\mu_R+i\mu_I$. A measure μ is said to be **Hermitian** if $\mu=\check{\overline{\mu}}$, that is, if its real part is even and its imaginary part is odd.

We introduce the next important definition.

Definition 2.1.2. *Let $\mu\in\mathscr{M}(\mathbb{R}^d)$. Its **measure of total variation** is defined as the measure $|\mu|\in\mathscr{M}^+(\mathbb{R}^d)$ defined for every $A\in\mathcal{B}_B(\mathbb{R}^d)$ by*

$$|\mu|(A) := \sup\left\{\sum_{n\in\mathbb{N}} |\mu(E_n)| \;\middle|\; (E_n)_{n\in\mathbb{N}}\subset\mathcal{B}(\mathbb{R}^d) \text{ partition of } A\right\}. \tag{2.2}$$

The measure of total variation $|\mu|$ is actually, as its name states, a measure, which is in addition positive. It is also the *smallest* positive measure satisfying $|\mu|(A)\geqslant|\mu(A)|$ for all $A\in\mathcal{B}_B(\mathbb{R}^d)$. Those claims can be proven following Rudin (1987, Theorems 6.2 and 6.4). It can thus be concluded that $|\mu|(A)<\infty$ for every $A\in\mathcal{B}_B(\mathbb{R}^d)$. It is clear that if $\mu\in\mathscr{M}^+(\mathbb{R}^d)$, then $|\mu|=\mu$. If $\mu\in\mathscr{M}(\mathbb{R}^d)$ is a real measure, we define its **positive part** as the positive measure $\mu^+=\frac{|\mu|+\mu}{2}$, and its **negative part** as the positive measure $\mu^-=\frac{|\mu|-\mu}{2}$, satisfying then $\mu=\mu^+-\mu^-$. This decomposition of real measures is also called the *Jordan decomposition* (Rudin, 1987, Section 6.6). Using the positive and negative parts of the real and imaginary parts of a complex measure μ, it is easy to see that μ can be decomposed in four positive measures $\mu_R^+,\mu_R^-,\mu_I^+,\mu_I^-\in\mathscr{M}^+(\mathbb{R}^d)$ through $\mu=\mu_R^+-\mu_R^-+i(\mu_I^+-\mu_I^-)$. If $\mu\in\mathscr{M}(\mathbb{R}^d)$ and $A\in\mathcal{B}_B(\mathbb{R}^d)$, we say that A is a **null set** of μ, or a $\mu-$null set, if $|\mu|(A)=0$. This definition can also be extended to unbounded Borel sets.

We give now our definition of a measure of finite total mass.

Definition 2.1.3. *A measure $\mu \in \mathscr{M}(\mathbb{R}^d)$ is said to be a **finite measure**, or to have a **finite total mass** if its domain can be extended to $\mathcal{B}(\mathbb{R}^d)$ maintaining the $\sigma-$additivity property:*

- *$\mu(A) \in \mathbb{C}$ for all $A \in \mathcal{B}(\mathbb{R}^d)$,*
- *for every countable collection of mutually disjoint Borel sets $(A_n)_{n\in\mathbb{N}} \subset \mathcal{B}(\mathbb{R}^d)$, the $\sigma-$ additivity property (2.1) holds.*

We denote by $\mathscr{M}_F(\mathbb{R}^d)$ the space of complex finite measures over $\mathbb{R}^d$, and $\mathscr{M}_F^+(\mathbb{R}^d)$ the set of all positive finite measures over $\mathbb{R}^d$. It is immediate that $\mathscr{M}_F(\mathbb{R}^d)$ is a complex vector space. It is also clear that for a positive measure $\mu \in \mathscr{M}^+(\mathbb{R}^d)$, being a finite measure is equivalent to having $\mu(\mathbb{R}^d) < \infty$. For a complex measure, an analogue condition is required to the measure of total variation, as the next Proposition states.

Proposition 2.1.1. *Let $\mu \in \mathscr{M}(\mathbb{R}^d)$. Then, $\mu \in \mathscr{M}_F(\mathbb{R}^d)$ if and only if $|\mu| \in \mathscr{M}_F^+(\mathbb{R}^d)$.*

Since we have used non-traditional definitions of measure and finite measures, this result is not easily findable in the literature in the exact way we state it. We then give a proof of Proposition 2.1.1 in Appendix A.1. From Proposition 2.1.1, it is straightforward that the reflection, the conjugate, the imaginary and real parts, and the positive and negative parts (in the real case) of a finite measure are also finite measures. We can also conclude that the Jordan decomposition of a finite measure μ consists of four positive finite measures $\mu_R^+, \mu_R^-, \mu_I^+, \mu_I^- \in \mathscr{M}_F^+(\mathbb{R}^d)$, having $\mu = \mu_R^+ - \mu_R^- + i(\mu_I^+ - \mu_I^-)$. If μ is a finite measure, the positive real value $|\mu|(\mathbb{R}^d)$ is called the **total variation** of μ.

We finish this section with two basic but essential examples of measures over $\mathbb{R}^d$:

- The **Lebesgue measure**, denoted by Leb, which satisfy to be the unique measure in $\mathscr{M}(\mathbb{R}^d)$ that gives to every set of the form $[a_1, b_1] \times ... \times [a_d, b_d]$, with $-\infty < a_j \leqslant b_j < \infty$ for all $j \in \{1, ..., d\}$, the value $\mathrm{Leb}([a_1, b_1] \times ... \times [a_d, b_d]) = \prod_{j=1}^d |b_j - a_j|$. It is also the unique measure in $\mathscr{M}^+(\mathbb{R}^d)$ which is invariant under translations[2] and which gives the value 1 to the hyper-cube $[0, 1]^d$. Hence, the Lebesgue measure is the formalisation of the intuitive notion of area in $\mathbb{R}^2$ or the volume in $\mathbb{R}^3$. It is not a finite measure.
- The **Dirac measure** at $x \in \mathbb{R}^d$, denoted by δ_x, which is the measure that for every $A \in \mathcal{B}_B(\mathbb{R}^d)$ gives the value $\delta_x(A) = 1$ if $x \in A$, and $\delta_x(A) = 0$ if $x \notin A$. It is a positive finite measure. If $x = 0$, the Dirac measure at x is simply denoted by δ.

2.1.2 Reminders on Lebesgue integrability

In this section we recall some notions of *Lebesgue integrability* over the Euclidean space. As the reader probably knows, the Lebesgue integral can be defined on quite abstract measure spaces. The recall made

[2]That is, that for every $A \in \mathcal{B}_B(\mathbb{R}^d)$, $\mathrm{Leb}(A + h) = \mathrm{Leb}(A)$ for all $h \in \mathbb{R}^d$, where $A + h$ denotes the set $\{x \in \mathbb{R}^d \mid x - h \in A\}$.

here is restricted to integration over the Euclidean space, and it is presented mainly to clarify the language, notions and results that will be used further in this work, with special emphasis on the construction of the Lebesgue integral. We do not give any proof of the claims presented in this section, since they are either broadly known, or straightforward from well-known results. There are many classical bibliographical sources about this subject. We suggest for example Rudin (1987). Fast and effective introductions can be found in Donoghue (1969, Chapter 5) and in L. Rogers & Williams (2000, Chapter 2, §1), the latter with a *probabilistic* approach. This section can be skipped by a reader already familiar with this theory.

We consider the measurable space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. A **not-necessarily Radon *positive* measure** over this space, is a set function $\mu : \mathcal{B}(\mathbb{R}^d) \to [0, \infty]$ satisfying the $\sigma-$additivity property (2.1) for every countable mutually disjoint family of Borel sets. Such a measure can take infinite (positive) values over bounded Borel sets, and it has necessarily the property $\mu(A) \leqslant \mu(B)$ if $A \subset B$. If $A \in \mathcal{B}(\mathbb{R}^d)$ is such that $\mu(A) = 0$, we say that A is a **null set** of μ, or a $\mu-$*null set*.

We will first focus on integration of positive extended-real valued functions. A positive extended-real valued function $f : \mathbb{R}^d \to [0, \infty]$ is said to be **measurable** if $f^{-1}(\mathcal{B}([0, \infty])) \subset \mathcal{B}(\mathbb{R}^d)$, that is, if the pre-image of every Borel subset of $[0, \infty]$ is a Borel subset of $\mathbb{R}^d$. The set of positive extended-real functions is a cone stable under multiplication, maximum and minimum, and point-wise monotone convergence. A particular class of measurable positive extended-real valued functions are the so-called **simple functions**, which are measurable functions taking a finite number of values. A simple function $f : \mathbb{R}^d \to [0, \infty]$ can be expressed as a finite linear combination of indicators functions of Borel sets:

$$f = \sum_{j \in J} a_j \mathbf{1}_{A_j}, \tag{2.3}$$

with $A_j \in \mathcal{B}(\mathbb{R}^d)$ and $a_j \in [0, \infty]$ for all $j \in J$, with $\#(J) < \infty$, where $\#(J)$ denotes the cardinality of the index set J. If f is a function of this form, its Lebesgue integral with respect to a positive not-necessarily Radon measure μ is defined as

$$\int_{\mathbb{R}^d} f(x) d\mu(x) = \sum_{j \in J} a_j \mu(A_j). \tag{2.4}$$

This expression can take infinite positive values, even when μ is in $\mathscr{M}^+(\mathbb{R}^d)$ and f takes finite values, since one of the involved Borel sets could be not bounded. For a positive extended-real measurable function f, it is known that we can always construct a sequence of positive simple functions that converges monotonically (increasing) point-wise to f. The **Lebesgue integral** of a positive extended-real measurable function f is then defined as follows: if $(f_n)_{n \in \mathbb{N}}$ is a sequence of positive simple functions monotonically point-wise convergent to f, we define

$$\int_{\mathbb{R}^d} f(x) d\mu(x) := \lim_{n \to \infty} \int_{\mathbb{R}^d} f_n(x) d\mu(x). \tag{2.5}$$

This limit always exists (it can be infinite). This limit does not depend on the choice of the sequence of

simple functions converging to f. The Lebesgue integral with respect to μ defines a linear mapping over the cone of positive measurable functions. When the limit (2.5) is a finite positive real number we say that the function f is **integrable** with respect to μ. The Lebesgue integral is monotonic in the following sense: if f, g are two positive extended-real measurable functions such that $f \leqslant g$, and μ is a positive not-necessarily Radon measure, then

$$\int_{\mathbb{R}^d} f(x)d\mu(x) \leqslant \int_{\mathbb{R}^d} g(x)d\mu(x). \tag{2.6}$$

Let us right now consider the case of a complex function $f : \mathbb{R}^d \to \mathbb{C}$. We say that f is measurable if $f^{-1}(\mathcal{B}(\mathbb{C})) \subset \mathcal{B}(\mathbb{R}^d)$. Every continuous complex function is a measurable function. The space of all complex measurable functions is a complex vector space stable under complex conjugation, multiplication, maximum and minimum in the case of real functions, and under point-wise limits. The space of measurable complex functions is actually the sequential completition of the space of continuous complex functions equipped with the topology of point-wise convergence. Every complex measurable function f can be decomposed in four positive measurable functions $f_R^+, f_R^-, f_I^+, f_I^-$ through $f = f_R^+ - f_R^- + i(f_I^+ - f_I^-)$. We can take, for instance, $f_R^+ = \max\{\frac{f+\overline{f}}{2}, 0\}, f_R^- = -\min\{\frac{f+\overline{f}}{2}, 0\}, f_I^+ = \max\{\frac{f-\overline{f}}{2i}, 0\}, f_I^- = -\min\{\frac{f-\overline{f}}{2i}, 0\}$, with $\overline{f}$ being the complex conjugate of f. We remark that with this decomposition, $|f| = f_R^+ + f_R^- + f_I^+ + f_I^-$. If μ is a positive not-necessarily Radon measure, we say that f is integrable with respect to μ if $|f|$ is integrable with respect to μ, which holds if and only if all the four positive functions $f_R^+, f_R^-, f_I^+, f_I^-$ are integrable with respect to μ. In such a case, the Lebesgue integral of f with respect to μ is defined as

$$\int_{\mathbb{R}^d} f(x)d\mu(x) := \int_{\mathbb{R}^d} f_R^+(x)d\mu(x) - \int_{\mathbb{R}^d} f_R^-(x)d\mu(x) + i\left(\int_{\mathbb{R}^d} f_I^+(x)d\mu(x) - \int_{\mathbb{R}^d} f_I^-(x)d\mu(x)\right). \tag{2.7}$$

We finally consider the case where $\mu \in \mathscr{M}(\mathbb{R}^d)$ and f is a measurable complex function. In such a case, we say that f is **integrable** with respect to μ if $|f|$ is integrable with respect to $|\mu|$. Using the decomposition of μ in four positive measures, $\mu = \mu_R^+ - \mu_R^- + i(\mu_I^+ - \mu_I^-)$, it is immediate that f is integrable with respect to μ if and only if all the positive measurable functions $f_R^+, f_R^-, f_I^+, f_I^-$ are integrable with respect to each one of the positive locally finite measures $\mu_R^+, \mu_R^-, \mu_I^+, \mu_I^-$. In such a case, the **Lebesgue integral** of f with respect to μ is defined as

$$\int_{\mathbb{R}^d} f(x)d\mu(x) := \int_{\mathbb{R}^d} f(x)d\mu_R^+(x) - \int_{\mathbb{R}^d} f(x)d\mu_R^-(x) + i\left(\int_{\mathbb{R}^d} f(x)d\mu_I^+(x) - \int_{\mathbb{R}^d} f(x)d\mu_I^-(x)\right). \tag{2.8}$$

In an analogous way, a positive extended-real measurable function f is said to be integrable with respect to a complex measure $\mu \in \mathscr{M}(\mathbb{R}^d)$ if it is integrable with respect to $|\mu|$, which is equivalent to require that f is integrable with respect to each one of the four measures $\mu_R^+, \mu_R^-, \mu_I^+, \mu_I^-$. In such a case, the Lebesgue integral of f with respect to μ is defined through the expression (2.8). When the measure μ is the Lebesgue measure, the integral of a function f in any of the aforementioned cases, if well-defined, is denoted by

$\int_{\mathbb{R}^d} f(x)dx$.

In all the mentioned cases for f and μ, if the Lebesgue integral of f with respect to μ is well-defined, the following inequality holds:

$$\left|\int_{\mathbb{R}^d} f(x)d\mu(x)\right| \leqslant \int_{\mathbb{R}^d} |f(x)|d|\mu|(x). \tag{2.9}$$

Let $A \in \mathcal{B}(\mathbb{R}^d)$, f be a function and μ a measure in any of the aforementioned cases. If the function $\mathbf{1}_A f$ is integrable with respect to μ, the Lebesgue integral of f with respect to μ over A is defined by

$$\int_A f(x)d\mu(x) := \int_{\mathbb{R}^d} \mathbf{1}_A(x)f(x)d\mu(x). \tag{2.10}$$

If f is a measurable positive extended-real function and μ is a positive not-necessarily Radon measure, the expression (2.10) is always well-defined independently of the integrability condition (it can be infinite). In such a case, the application $A \in \mathcal{B}(\mathbb{R}^d) \mapsto \int_A f(x)d\mu(x)$ is actually a positive not-necessarily Radon measure, called the **multiplication** between f and μ, and it is denoted by $f\mu$.

If μ is a measure in any of the aforementioned cases, and f is a measurable function, complex or positive extended-real, such that for every compact subset K of $\mathbb{R}^d$ the function $\mathbf{1}_K f$ is integrable with respect to μ, we say that f is **locally integrable** with respect to μ. In such a case, the Lebesgue integral (2.10) is well-defined for every bounded Borel set A, and the application $A \in \mathcal{B}_B(\mathbb{R}^d) \mapsto \int_A f(x)d\mu(x)$ is a measure in $\mathscr{M}(\mathbb{R}^d)$, which is also called the **multiplication** between f and μ and it is denoted by $f\mu$. Every locally bounded measurable function f (that is, such that $\|f\|_{\infty,A} := \sup_{x\in A} |f(x)| < \infty$ for every bounded set $A \subset \mathbb{R}^d$) is locally integrable with respect to every measure $\mu \in \mathscr{M}(\mathbb{R}^d)$, and thus $f\mu \in \mathscr{M}(\mathbb{R}^d)$. This holds in particular if f is a continuous function. We also remark that if f is any measurable function and μ is any measure of the aforementioned cases, if f is integrable with respect to μ, the multiplication $f\mu$ is a finite measure. In particular, every measurable and bounded complex function f is integrable with respect to every finite measure μ, and thus the multiplication $f\mu$ is a finite measure for which it holds that $|(f\mu)(\mathbb{R}^d)| = |\int_{\mathbb{R}^d} f(x)d\mu(x)| \leqslant \|f\|_\infty |\mu|(\mathbb{R}^d)$, where $\|f\|_\infty$ denotes de supremum norm of f, $\|f\|_\infty := \sup_{x\in\mathbb{R}^d} |f(x)|$. In order to fix some notation and terminology, when $\nu = f\mu$ is the multiplication between a function f and a measure μ, we note this fact also as $d\nu(x) = f(x)d\mu(x)$, and we say that ν **has a density** f with respect to μ. In the case where μ is the Lebesgue measure, we rather denote this by $d\nu(x) = f(x)dx$, and we simply say that ν *has a density*, without necessarily specifying that the density is with respect to the Lebesgue measure.

We recall two important results from Integration Theory.

Theorem 2.1.1 (Monotone Convergence Theorem). *Let μ be a positive measure over $\mathbb{R}^d$, Radon or not. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of positive extended-real measurable functions which is monotonically increasing.*

Let f be the point-wise limit $f(x) = \lim_{n\to\infty} f_n(x)$ (which is also measurable). Then,

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} f_n(x) d\mu(x) = \int_{\mathbb{R}^d} f(x) d\mu(x). \tag{2.11}$$

Theorem 2.1.2 (Lebesgue's Dominated Convergence Theorem). *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of complex measurable functions over $\mathbb{R}^d$ which is point-wise convergent to a measurable complex function f. Let μ be a complex measure over $\mathbb{R}^d$ and suppose that there exists a positive function g integrable with respect to μ such that $|f_n| \leqslant g$ for all $n \in \mathbb{N}$. Then, f is integrable with respect to μ and*

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} f_n d\mu(x) = \int_{\mathbb{R}^d} f(x) d\mu(x). \tag{2.12}$$

This Theorem is also applicable for point-wise converging sequences of positive extended-real valued functions which are bounded by a positive extended-real measurable function integrable with respect to μ. The condition of point-wise convergence can also be relaxed to have point-wise convergence outside a $\mu-$null set.

We finally recall the definitions and notations of the **Lebesgue spaces**. Given a measure μ over $\mathbb{R}^d$, Radon or not, we denote by $\mathscr{L}^1(\mathbb{R}^d, \mu)$ or $\mathscr{L}(\mathbb{R}^d, \mu)$ the set of all *complex*[3] measurable functions which are integrable with respect to μ. For $p \in [1, \infty)$ the set $\mathscr{L}^p(\mathbb{R}^d, \mu)$ denotes the space of all complex measurable functions f such that $|f|^p$ is integrable with respect to μ. We denote by $\mathscr{L}^\infty(\mathbb{R}^d, \mu)$ the space of all complex measurable functions f such that there exists $C \geqslant 0$ such that $|f(x)| \leqslant C$ for all $x \in \mathbb{R}^d \backslash D$ where $D \in \mathcal{B}(\mathbb{R}^d)$ is a $\mu-$null set. The spaces $\mathscr{L}^p(\mathbb{R}^d, \mu)$ with $p \in [0, \infty]$ are complex vector spaces. The associated quotient spaces of those spaces with respect to the equivalence relation of equality outside a $\mu-$null set are denoted by $L^p(\mathbb{R}^d, \mu)$. Hence the spaces $L^p(\mathbb{R}^d, \mu)$ are not spaces of functions but rather of equivalence classes of measurable functions. For $p \in [1, \infty)$, the spaces $L^p(\mathbb{R}^d, \mu)$ are endowed with the norm $\|f\|_{L^p(\mathbb{R}^d,\mu)} := \left(\int_{\mathbb{R}^d} |f(x)|^p d|\mu|(x)\right)^{\frac{1}{p}}$ for any f representing its equivalence class of functions. For $p = \infty$, $L^p(\mathbb{R}^d, \mu)$ is endowed with the norm of the *essential supremum*, $\|f\|_{L^\infty(\mathbb{R}^d,\mu)} := \inf\{C \geqslant 0 \mid |f(x)| \leqslant C \text{ for all } x \text{ outside a } \mu-\text{null set}\}$. For every $p \in [1, \infty]$, the so-constructed topological vector space $L^p(\mathbb{R}^d, \mu)$ is a Banach space, that is, a complete normed space. For $p = 2$, it is a Hilbert space, with the inner product defined through $(f, g)_{L^2(\mathbb{R}^d,\mu)} := \int_{\mathbb{R}^d} f(x)\overline{g}(x) d|\mu|(x)$, being f, g two any representatives of their equivalence classes. When the measure μ is the Lebesgue measure, we drop the "μ" in the notation of the associated Lebesgue spaces and their norms, denoting them simply as $\mathscr{L}^p(\mathbb{R}^d)$, $L^p(\mathbb{R}^d)$ and $\|\cdot\|_{L^p(\mathbb{R}^d)}$.

[3] Some authors also include the extended-real measurable functions. We do not.

2.1.3 Slow-growing measures

We introduce now another space of measures which plays a particular role in the theory of stationary Random Fields. In contrast to the case of locally finite measures and finite measures, this definition cannot be immediately extended to more abstract measurable spaces.

Definition 2.1.4. *Let $\mu \in \mathscr{M}(\mathbb{R}^d)$. We say that μ is a **slow-growing measure** if there exists a strictly positive polynomial $p : \mathbb{R}^d \to \mathbb{R}_*^+$ such that the measure $\frac{1}{p}\mu$ is finite, or equivalently, if there exists $N \in \mathbb{N}$ such that*

$$\int_{\mathbb{R}^d} \frac{d|\mu|(x)}{(1+|x|^2)^N} < \infty. \tag{2.13}$$

The equivalence stated in Definition 2.1.4 comes from the fact that for every polynomial $p : \mathbb{R}^d \to \mathbb{C}$ there exist $N \in \mathbb{N}$ and $C > 0$ such that $|p(x)| \leqslant C(1+|x|^2)^N$ for all $x \in \mathbb{R}^d$. We denote by $\mathscr{M}_{SG}(\mathbb{R}^d)$ the set of slow-growing complex measures over $\mathbb{R}^d$. It is immediate that it is a complex vector space, and that the inclusion $\mathscr{M}_F(\mathbb{R}^d) \subset \mathscr{M}_{SG}(\mathbb{R}^d)$ holds, since for every finite measure it suffices to set $N = 0$ in (2.13). If μ is a slow-growing measure, its reflection, its conjugate, its imaginary and real parts, and its positive and negative parts in the real case are also slow-growing. We denote by $\mathscr{M}_{SG}^+(\mathbb{R}^d)$ the set of all positive slow-growing measures. The Jordan decomposition of a slow-growing measure μ consists then of four positive slow-growing measures $\mu_R^+, \mu_R^-, \mu_I^+, \mu_I^- \in \mathscr{M}_{SG}^+(\mathbb{R}^d)$, having $\mu = \mu_R^+ - \mu_R^- + i(\mu_I^+ - \mu_I^-)$.

Let $f : \mathbb{R}^d \to \mathbb{C}$ be a polynomially bounded measurable function and let $\mu \in \mathscr{M}_{SG}(\mathbb{R}^d)$. Then, the multiplication $f\mu$ is a slow-growing measure. To see this, consider $N_f \in \mathbb{N}$ such that $(1+|x|^2)^{-N_f} f$ is bounded and $N_\mu \in \mathbb{N}$ such that $(1+|x|^2)^{-N_\mu}\mu$ is finite. As every bounded measurable function is integrable with respect to any finite measure, we obtain

$$\int_{\mathbb{R}^d} \frac{d|f\mu|(x)}{(1+|x|^2)^{N_f+N_\mu}} \leqslant \int_{\mathbb{R}^d} \frac{|f(x)|}{(1+|x|^2)^{N_f}} \frac{d|\mu|(x)}{(1+|x|^2)^{N_\mu}} < \infty, \tag{2.14}$$

from which we conclude that $f\mu$ is a slow-growing measure, as it can be seen by setting $N = N_f + N_\mu$ in Eq. (2.13).

2.1.4 Measures as linear functionals over spaces of continuous functions

In this section we recall some classical results which identify spaces of complex measures as members of the dual of some vector spaces of continuous functions. Hence, we present the Riesz Representation Theorem and some of its variants.

We denote by $C(\mathbb{R}^d)$ the space of all **complex continuous** functions over $\mathbb{R}^d$ and $C_B(\mathbb{R}^d)$ the space of all **bounded complex continuous** functions over $\mathbb{R}^d$. Let us first recall the definition of support of a function

and of a measure.

Definition 2.1.5. *Let $f : \mathbb{R}^d \to \mathbb{C}$ be a function. Its **support** is defined as the closure of the set where f is not null:*

$$\operatorname{supp}(f) := \overline{\{x \in \mathbb{R}^d \ : \ f(x) \neq 0\}}. \tag{2.15}$$

*If $\mu \in \mathscr{M}(\mathbb{R}^d)$, its **support** is defined as the complementary of the largest open set where the total variation is null:*

$$\operatorname{supp}(\mu) = \left(\bigcup \{O \subset \mathbb{R}^d \ : \ O \text{ is open and } |\mu|(O) = 0\}\right)^c. \tag{2.16}$$

The support of a function or a measure is always a closed set. It is, roughly speaking, the set where the function or the measure is not null. A function or a measure is said to be **compactly supported** if its support is a compact set. This definition is also applicable to extended-real functions and to positive not-necessarily Radon measures. It is immediate that $|\mu|(\operatorname{supp}(\mu)^c) = 0$ and thus $|\mu|(\operatorname{supp}(\mu)) = |\mu|(\mathbb{R}^d)$. It is also easy to see that if f is a measurable complex (or positive extended-real) function, its Lebesgue integral with respect to any measure μ, Radon or not, satisfies, when the integral is well-defined,

$$\int_{\mathbb{R}^d} f(x)d\mu(x) = \int_{\operatorname{supp}(f)} f(x)d\mu(x) = \int_{\operatorname{supp}(\mu)} f(x)d\mu(x) = \int_{\operatorname{supp}(f)\cap\operatorname{supp}(\mu)} f(x)d\mu(x). \tag{2.17}$$

It follows that in the case where $\operatorname{supp}(f) \cap \operatorname{supp}(\mu) = \varnothing$, then $f\mu = 0$. The same applies if $|\mu|(\operatorname{supp}(f)) = 0$ or if f is null over $\operatorname{supp}(\mu)$.

Let us introduce the next spaces of continuous functions:

- $C_c(\mathbb{R}^d)$, the space of **compactly supported continuous functions**:

$$C_c(\mathbb{R}^d) = \{\varphi \in C(\mathbb{R}^d) \ \big| \ \operatorname{supp}(\varphi) \text{ is compact}\}. \tag{2.18}$$

- $C_0(\mathbb{R}^d)$, the space of **continuous functions vanishing at infinity**:

$$C_0(\mathbb{R}^d) = \{\varphi \in C(\mathbb{R}^d) \ \big| \ \lim_{|x|\to\infty} \varphi(x) = 0\}. \tag{2.19}$$

- $C_{FD}(\mathbb{R}^d)$, the space of **fast-decreasing continuous functions**, that is, functions that decrease faster than any polynomial:

$$C_{FD}(\mathbb{R}^d) = \{\varphi \in C(\mathbb{R}^d) \ \big| \ \|(1+|x|^2)^N\varphi\|_\infty = \sup_{x\in\mathbb{R}^d} |(1+|x|^2)^N\varphi(x)| < \infty \quad \forall N \in \mathbb{N}\}. \tag{2.20}$$

We remark that $C_c(\mathbb{R}^d) \subset C_{FD}(\mathbb{R}^d) \subset C_0(\mathbb{R}^d) \subset C(\mathbb{R}^d)$. Each one of these sets are complex vector spaces, and each one of them will be endowed with a particular topology which makes them complete

locally convex topological vector spaces (see Appendix D). Their dual spaces will be identified with spaces of measures. We recall that for a general complex topological vector space E, its **dual** is the space of all continuous linear functionals over E, and it is denoted by E'. If $T \in E'$, its action over an element $x \in E$, that is $T(x)$, (T is a function from E to $\mathbb{C}$) is more comfortably denoted by $\langle T, x\rangle$, to highlight the fact that T is linear.

The topologies these spaces will be endowed with are not all trivial. The space $C_0(\mathbb{R}^d)$ will be endowed with a norm which makes it a Banach space. The spaces $C_{FD}(\mathbb{R}^d)$ and $C(\mathbb{R}^d)$ will be endowed with a metric topology which makes them Fréchet spaces. The space $C_c(\mathbb{R}^d)$ will be endowed with a more technical topology which makes it a Hausdorff locally convex topological vector space.

We start with the space $C_c(\mathbb{R}^d)$. The topology defined in this space is rather technical and it will be explained in detail in a footnote[4]. Most authors do not make it explicit, but rather describe it roughly through the description of the convergent *sequences* on this space[5]. A sequence $(\varphi_n)_{n\in\mathbb{N}} \subset C_c(\mathbb{R}^d)$ converges to 0 in $C_c(\mathbb{R}^d)$, denoted by $\varphi_n \overset{C_c}{\to} 0$, if (and only if) $\|\varphi_n\|_\infty \to 0$ and if there exists a compact $K \subset \mathbb{R}^d$ such that $\mathrm{supp}(\varphi_n) \subset K$ for all $n \in \mathbb{N}$. If $\varphi \in C_c(\mathbb{R}^d)$, a sequence $(\varphi_n)_{n\in\mathbb{N}} \subset C_c(\mathbb{R}^d)$ is said to converge to φ in $C_c(\mathbb{R}^d)$, denoted by $\varphi_n \overset{C_c}{\to} \varphi$ if $\varphi_n - \varphi \overset{C_c}{\to} 0$. Another way of describing the topology of $C_c(\mathbb{R}^d)$ is through the characterisation of the continuous linear functionals defined over it. A linear functional $T : C_c(\mathbb{R}^d) \to \mathbb{C}$ is continuous if (and only if) for all compact $K \subset \mathbb{R}^d$ there exists $C_K > 0$ such that

$$|\langle T, \varphi\rangle| \leqslant C_K \|\varphi\|_\infty, \quad \forall \varphi \in C_c(\mathbb{R}^d) \text{ such that } \mathrm{supp}(\varphi) \subset K. \tag{2.22}$$

[4]The topology of $C_c(\mathbb{R}^d)$ is defined in order to make it a Hausdorff complete locally convex topological vector space. Its topology can be fully determined through the specification of the associated family of semi-norms. This family will be indexed by the set of all decreasing to zero sequences of strictly positive real numbers. Let $(\epsilon_n)_{n\in\mathbb{N}} \subset \mathbb{R}_*^+$ be such a sequence. The associated semi-norm is defined as

$$p_{(\epsilon_n)_{n\in\mathbb{N}}}(\varphi) = \sup_{n\in\mathbb{N}}\{\sup_{|x|\geqslant n} \frac{|\varphi(x)|}{\epsilon_n}\}, \quad \forall \varphi \in C_c(\mathbb{R}^d). \tag{2.21}$$

It is difficult to find authors presenting this topology in this way. What we have done here is just a copy-paste of the description of the topology of the classical space of smooth and compactly supported functions in Distribution Theory, $\mathscr{D}(\mathbb{R}^d)$, as done in Schwartz (1966, Chapter III), and restrain the definition of the semi-norms to the case of non-differentiable functions. It can be proven that $C_c(\mathbb{R}^d)$ equipped with this topology is, as expected, complete. Theorem I in Schwartz (1966, Chapter III, §1) states that $\mathscr{D}(\mathbb{R}^d)$ is complete, and the same arguments can be used to prove that $C_c(\mathbb{R}^d)$ with this topology is complete. Another approach to prove the completeness of $C_c(\mathbb{R}^d)$ is by considering that it is the *strict inductive limit of Banach spaces*; see Reed & Simon (1980, Section V.4) for an introduction of this concept and its properties. In this same source and section, the authors present in Example 1 the analogue to our space $C_c(\mathbb{R}^d)$, in the case $d = 1$, there denoted by $\kappa(\mathbb{R})$.

[5]In a general topological space, the description of the convergent sequences in the space does not suffice to define the topology. This would hold, for example, over a metrizable topological space. The space $C_c(\mathbb{R}^d)$ is an example of a non-metrizable topological vector space, whose topology cannot be completely determined by its convergent sequences. A generalization of the concept of sequence is the concept of *net*, which can be used to describe completely the continuous functions over a topological space, and hence to describe topologies defined from a family of functions desired to be continuous. See (Reed & Simon, 1980, Section IV.2).

We remark that any function in $C_c(\mathbb{R}^d)$ is integrable with respect to any measure $\mu \in \mathscr{M}(\mathbb{R}^d)$, since

$$\left|\int_{\mathbb{R}^d} \varphi(x)d\mu(x)\right| = \left|\int_{\mathrm{supp}(\varphi)} \varphi(x)d\mu(x)\right| \leqslant \|\varphi\|_\infty |\mu|(\mathrm{supp}(\varphi)) < \infty. \tag{2.23}$$

We conclude that the integral with respect to any $\mu \in \mathscr{M}(\mathbb{R}^d)$ defines a linear functional over $C_c(\mathbb{R}^d)$, which can be seen by setting $C_K = |\mu|(K)$ for a corresponding compact set $K \subset \mathbb{R}^d$ in Eq. (2.22). We present the famous Riesz Representation Theorem for Radon measures, which states the converse: any continuous linear function over $C_c(\mathbb{R}^d)$ can be represented by a measure in $\mathscr{M}(\mathbb{R}^d)$.

Theorem 2.1.3 (Riesz Representation for locally finite complex measures). *$\mathscr{M}(\mathbb{R}^d) = C_c'(\mathbb{R}^d)$, that is, every measure $\mu \in \mathscr{M}(\mathbb{R}^d)$ defines a continuous linear functional T over $C_c(\mathbb{R}^d)$ through the integral*

$$\langle T, \varphi\rangle = \int_{\mathbb{R}^d} \varphi(x)d\mu(x), \quad \forall \varphi \in C_c(\mathbb{R}^d). \tag{2.24}$$

Conversely, for every continuous linear functional $T : C_c(\mathbb{R}^d) \to \mathbb{C}$ there exists a unique $\mu \in \mathscr{M}(\mathbb{R}^d)$ such that (2.24) holds.

This Theorem is quite remarkable and powerful. It allows to completely describe a measure by its action over continuous functions with compact support rather than over sets, which in some cases simplifies the analysis. It also gives a criterion for discriminating when a set function or a linear functional over some vector spaces of functions actually defines a measure, allowing to use freely the properties and operations well-defined for measures. But maybe the most important consequence is that this Theorem provides a framework where an Integration Theory can be constructed using tools of topological vector spaces and relate them to Distribution Theory. Indeed, some authors *define* a Radon Measure as a continuous linear functional over $C_c(\mathbb{R}^d)$ and then construct the Integration Theory over Borel sets. This is the approach described in Bourbaki (1965). A student-oriented exposition of this approach can be found in Demengel & Demengel (2000). Other sources which take advantage of this vector space oriented theory are Schwartz (1966) and Trèves (1967, Chapter 21).

Theorem 2.1.3 is usually presented in a more general setting than the measure space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Indeed, an analogue result holds for more abstract measurable spaces, namely, when $\mathbb{R}^d$ is replaced by a locally compact Hausdorff topological space, and it is endowed with its Borel $\sigma-$algebra. The representation is done using Radon measures in the strict sense: locally finite, inner and outer regular measures. A proof of the *positive* version of this Theorem, that is, that every positive linear functional can be represented by a positive Radon measure, dropping the continuity condition, can be found in a general form in Donoghue (1969, Chapter 5) or in Reed & Simon (1980, Theorem IV.18). A proof of the case of complex measures over $\mathbb{R}$, whose arguments also hold for the case over $\mathbb{R}^d$ stated as in Theorem 2.1.3, can be found in (Reed & Simon, 1980, Section V.4, Example 1), with a reference to other developments done in the book. There, a simple

argument considering that $C_c(\mathbb{R}^d)$ is an inductive limit of Banach spaces is presented. The authors then conclude that its dual is the space of locally finite complex measures (called *Baire measures* in this source). See also the comments on (Reed & Simon, 1980, Theorem IV.18) and its related results.

Consider right now a case with a simpler topology. Consider the space $C_0(\mathbb{R}^d)$. We endow it with the topology induced by the supremum norm $\|\cdot\|_\infty$. $C_0(\mathbb{R}^d)$ with this norm is a Banach space. Since $C_0(\mathbb{R}^d) \subset C_B(\mathbb{R}^d)$, any function in $C_0(\mathbb{R}^d)$ is Lebesgue integrable with respect to any finite measure $\mu \in \mathscr{M}_F(\mathbb{R}^d)$, and it is immediate that the integral defines a bounded (and hence continuous) linear functional over $C_0(\mathbb{R}^d)$.

Theorem 2.1.4 (Riesz Representation for finite measures). *$\mathscr{M}_F(\mathbb{R}^d) = C_0'(\mathbb{R}^d)$, that is, every finite measure $\mu \in \mathscr{M}_F(\mathbb{R}^d)$ defines a continuous linear functional T over $C_0(\mathbb{R}^d)$ through the integral*

$$\langle T, \varphi \rangle = \int_{\mathbb{R}^d} \varphi(x) d\mu(x), \quad \forall \varphi \in C_0(\mathbb{R}^d). \tag{2.25}$$

Conversely, for every continuous linear functional $T : C_0(\mathbb{R}^d) \to \mathbb{C}$ there exists a unique $\mu \in \mathscr{M}_F(\mathbb{R}^d)$ such that (2.25) holds.

For a proof, we suggest the one presented in Rudin (1987, Theorem 6.19). This Theorem is also usually presented in the more general setting of a locally compact Hausdorff topological space. This Theorem is also often presented before the generic Riesz Representation Theorem 2.1.3 since the topology over $C_0(\mathbb{R}^d)$ is easier to describe than the one of $C_c(\mathbb{R}^d)$.

We present now two variants of the Riesz Representation Theorem, for which we have not found proofs in the literature. They are almost a direct application of the previous Theorems, in an adequate way.

Let us consider the space $C(\mathbb{R}^d)$. We endow this complex vector space with the topology of uniform convergence over compact sets, that is, a sequence $(\varphi_n)_n \in C(\mathbb{R}^d)$ is said to converge to 0 in $C(\mathbb{R}^d)$, noted $\varphi_n \xrightarrow{C} 0$ if for every $K \subset \mathbb{R}^d$ compact, $\|\varphi_n\|_{\infty,K} := \sup_{x\in K} |\varphi_n(x)| \to 0$ as $n \to \infty$. This topology is equivalent to the one induced by the metric:

$$(\varphi, \phi) \mapsto \sum_{N\in\mathbb{N}_*} \frac{1}{2^N} \frac{\|\varphi - \phi\|_{\infty,\overline{B_N(0)}}}{1 + \|\varphi - \phi\|_{\infty,\overline{B_N(0)}}}, \tag{2.26}$$

where $B_r(x) \subset \mathbb{R}^d$ denotes the open ball of radius $r \geqslant 0$ centred at x. Using standard methods of basic analysis, it is easy to prove that $C(\mathbb{R}^d)$ endowed with this metric is a complete metric space (actually a Fréchet space).

Let us right now describe a particular class of complex measures. We denote by $\mathscr{M}_c(\mathbb{R}^d)$ the space of all **complex compactly supported measures** in $\mathscr{M}(\mathbb{R}^d)$. It is a complex vector subspace of $\mathscr{M}(\mathbb{R}^d)$. The local finiteness of the measures in $\mathscr{M}(\mathbb{R}^d)$ guarantees that $\mathscr{M}_c(\mathbb{R}^d) \subset \mathscr{M}_F(\mathbb{R}^d)$. We denote by $\mathscr{M}_c^+(\mathbb{R}^d)$ the

space of all positive compactly supported measures over $\mathbb{R}^d$. It is straightforward that the measure of total variation, the reflection, the conjugate, the imaginary and real parts, and the positive and negative parts (in the real case) of a compactly supported measure are also compactly supported measures. We also conclude that the Jordan decomposition of a compactly supported measure μ consists of four positive compactly supported measures $\mu_R^+, \mu_R^-, \mu_I^+, \mu_I^- \in \mathscr{M}_c^+(\mathbb{R}^d)$, having $\mu = \mu_R^+ - \mu_R^- + i(\mu_I^+ - \mu_I^-)$.

Every function $\varphi \in C(\mathbb{R}^d)$ is locally bounded. This implies that φ is integrable with respect to any measure $\mu \in \mathscr{M}_c(\mathbb{R}^d)$, since

$$\left|\int_{\mathbb{R}^d} \varphi(x)d\mu(x)\right| = \left|\int_{\operatorname{supp}(\mu)} \varphi(x)d\mu(x)\right| \leqslant \|\varphi\|_{\infty,\operatorname{supp}(\mu)}|\mu|(\operatorname{supp}(\mu)) < \infty. \tag{2.27}$$

This suggests that we can obtain an analogue to Riesz Representation Theorem for the dual of the space $C(\mathbb{R}^d)$.

Theorem 2.1.5 (Representation for compactly supported measures). *$\mathscr{M}_c(\mathbb{R}^d) = C'(\mathbb{R}^d)$, that is, every compactly supported measure $\mu \in \mathscr{M}_c(\mathbb{R}^d)$ defines a continuous linear functional T over $C(\mathbb{R}^d)$ through the integral*

$$\langle T, \varphi\rangle = \int_{\mathbb{R}^d} \varphi(x)d\mu(x), \quad \forall \varphi \in C(\mathbb{R}^d). \tag{2.28}$$

Conversely, for every continuous linear functional $T : C(\mathbb{R}^d) \to \mathbb{C}$ there exists a unique $\mu \in \mathscr{M}_c(\mathbb{R}^d)$ such that (2.28) holds.

We give a proof of this Theorem in Appendix A.2.1. We remark that this Theorem can be generalized to more abstract measurable spaces, provided that the arguments used to prove it hold also for those spaces: analogues to Lemmas A.2.2 and A.2.3 must be verified. This holds for example, over every separable locally compact metric space, using Radon measures in the strict sense of the term.

We finally consider the case of the space $C_{FD}(\mathbb{R}^d)$. We equip this space with the following topology: a sequence of functions $(\varphi_n)_{n\in\mathbb{N}} \subset C_{FD}(\mathbb{R}^d)$ converges to 0, denoted by $\varphi_n \overset{C_{FD}}{\to} 0$, if for all $N \in \mathbb{N}$ we have that $\|(1 + |x|^2)^N\varphi_n\|_\infty \to 0$. This is equivalent to require that the sequence $(\varphi_n)_{n\in\mathbb{N}}$ is such that $(p\varphi_n)_{n\in\mathbb{N}}$ converges uniformly to 0 for every polynomial $p : \mathbb{R}^d \to \mathbb{C}$. We say that a sequence $(\varphi_n)_{n\in\mathbb{N}} \subset C_{FD}(\mathbb{R}^d)$ converges to $\varphi \in C_{FD}(\mathbb{R}^d)$, denoted by $\varphi_n \overset{C_{FD}}{\to} \varphi$, if $\varphi_n - \varphi \overset{C_{FD}}{\to} 0$. This topology over $C_{FD}(\mathbb{R}^d)$ is induced by the metric

$$(\varphi, \phi) \mapsto \sum_{N\in\mathbb{N}} \frac{1}{2^N} \frac{\|(1 + |x|^2)^N(\varphi - \phi)\|_\infty}{1 + \|(1 + |x|^2)^N(\varphi - \phi)\|_\infty}. \tag{2.29}$$

The space $C_{FD}(\mathbb{R}^d)$ equipped with this metric is a complete metric space, which in addition is Fréchet[6].

[6]This is not complicated to conclude using standard arguments. For instance, the same arguments used to prove the completeness of the Schwartz space $\mathscr{S}(\mathbb{R}^d)$ can be used to prove the completeness of $C_{FD}(\mathbb{R}^d)$. See Definition 2.2.1 in Section 2.2.1 and the references therein.

We remark that every function $\varphi \in C_{FD}(\mathbb{R}^d)$ is Lebesgue integrable with respect to any slow-growing measure $\mu \in \mathscr{M}_{SG}(\mathbb{R}^d)$. Indeed, let μ be a slow-growing measure and let $N \in \mathbb{N}$ such that $(1+|x|^2)^{-N}|\mu|$ is a finite measure. Then $(1+|x|^2)^N\varphi \in C_B(\mathbb{R}^d)$ (it is actually in $C_{FD}(\mathbb{R}^d)$ too), from which we obtain

$$\left|\int_{\mathbb{R}^d} \varphi(x)d\mu(x)\right| = \left|\int_{\mathbb{R}^d} \frac{(1+|x|^2)^N\varphi}{(1+|x|^2)^N}d\mu(x)\right| \leqslant \|(1+|x|^2)^N\varphi\|_\infty \left|(1+|x|^2)^{-N}\mu\right|(\mathbb{R}^d) < \infty. \tag{2.30}$$

We have then an inspiration for an analogous to Riesz Representation Theorem for the case of the space $C_{FD}(\mathbb{R}^d)$.

Theorem 2.1.6 (Representation for slow-growing measures). *$\mathscr{M}_{SG}(\mathbb{R}^d) = C'_{FD}(\mathbb{R}^d)$, that is, every slow-growing measure $\mu \in \mathscr{M}_{SG}(\mathbb{R}^d)$ defines a continuous linear functional T over $C_{FD}(\mathbb{R}^d)$ through the integral*

$$\langle T, \varphi\rangle = \int_{\mathbb{R}^d} \varphi(x)d\mu(x), \quad \forall\varphi \in C_{FD}(\mathbb{R}^d). \tag{2.31}$$

Conversely, for every continuous linear functional $T : C_{FD}(\mathbb{R}^d) \to \mathbb{C}$ there exists a unique $\mu \in \mathscr{M}_{SG}(\mathbb{R}^d)$ such that (2.31) holds.

We give a proof of Theorem 2.1.6 in Appendix A.2.2. Since this Theorem uses the multiplicative structure of the components of a vector in $\mathbb{R}^d$, used to define polynomials, it cannot be generalized to more abstract measure spaces without a suitable adaptation, in contrast to the cases of Theorems 2.1.3, 2.1.4 and 2.1.6.

We summarize the duality and inclusion relationships. We put the dual of every space below itself:

$$\begin{aligned} &C_c(\mathbb{R}^d) \subset C_{FD}(\mathbb{R}^d) \subset C_0(\mathbb{R}^d) \subset C(\mathbb{R}^d) \\ &\mathscr{M}(\mathbb{R}^d) \supset \mathscr{M}_{SG}(\mathbb{R}^d) \supset \mathscr{M}_F(\mathbb{R}^d) \supset \mathscr{M}_c(\mathbb{R}^d). \end{aligned} \tag{2.32}$$

Taking advantage of the new interpretation of a measure μ as a continuous linear functional, we will often use the notation

$$\langle \mu, f\rangle := \int_{\mathbb{R}^d} f(x)d\mu(x), \tag{2.33}$$

when f is a measurable complex function integrable with respect to μ. We will use conveniently both the *linear functional* notation or the fully integral notation, depending on which one is more convenient to write or more explicit for communicating the desired message. In the same spirit, along this work we will use conveniently both interpretations of a measure μ as a linear functional over a space of continuous function or as a set function.

2.1.5 Measures concentrated on subsets

In this section we focus on measures which are, in some sense, null outside some particular Borel subset and thus it is not necessary to treat them outside of it.

Definition 2.1.6. *Let $\mu \in \mathscr{M}(\mathbb{R}^d)$, and let $A \in \mathcal{B}(\mathbb{R}^d)$. We say that μ is **concentrated on the set** A if for all $B \in \mathcal{B}_B(\mathbb{R}^d)$, $A \cap B = \varnothing \Rightarrow \mu(B) = 0$.*

The $\sigma-$additivity allows to easily conclude that μ is concentrated on A if and only if $\mu(A \cap B) = \mu(B)$ for all $B \in \mathcal{B}_B(\mathbb{R}^d)$. It is also true that μ is concentrated on A if and only if $|\mu|$ is concentrated on A, claim which can be concluded by analysing the definition of the total variation measure (2.2). The relationship between a set on which μ is concentrated and the support of μ is not immediate to describe. Of course, any measure is concentrated on its support. It is also immediate that a measure μ is concentrated on any Borel set A such that $\operatorname{supp}(\mu) \subset A$, and if A is a closed set, then μ is concentrated on A if and only if $\operatorname{supp}(\mu) \subset A$. Nevertheless, there are examples of measures which are concentrated on sets strictly included in their supports. Take for instance, $\mu = \sum_{n\in\mathbb{N}_*} \frac{1}{n^2}\delta_{\frac{1}{n}}$, which is a finite measure. Then, μ is concentrated on $\bigcup_{n\in\mathbb{N}_*}\{\frac{1}{n}\}$, but $\operatorname{supp}(\mu) = \{0\} \cup \bigcup_{n\in\mathbb{N}_*}\{\frac{1}{n}\}$.

The main interest of this definition is that now we are able to consider measures concentrated on subsets of $\mathbb{R}^d$ which can have null Lebesgue measure. The main example is the Dirac measure at a point $x \in \mathbb{R}^d$, δ_x, for which $\operatorname{supp}(\delta_x) = \{x\}$. It is particularly interesting to define measures which are concentrated on sub-manifolds of $\mathbb{R}^d$ which have dimension smaller than d. When working on the one dimensional Euclidean space $\mathbb{R}$, sub-manifolds of dimension 0 would be for example point-sets, and hence measures concentrated on these sub-manifolds are linear combinations (possibly countable, if local finiteness is provided) of Dirac measures. In higher dimensions we can still use Dirac measures, but other more interesting measures can appear since there exist sub-manifolds of higher dimensions, like for example curves.

We give some examples of these kinds of measures. We will extensively use the Riesz Representation Theorem 2.1.3, since describing a measure concentrated in a sub-manifold is usually easier by describing its action on continuous functions with compact support rather than its action over Borel sets.

Example 2.1.1. Consider the two dimensional Euclidean space $\mathbb{R}^2$. Consider the sphere of radius $R > 0$ centred at the origin, which we will denote by $\partial B_R^{(2)}(0)$. We can travel across this set with a typical parametrization of the curve, using the mapping $\gamma : [0, 2\pi) \to \mathbb{R}^2$ defined through $\gamma(\theta) = R\,(\cos(\theta), \sin(\theta))$. A measure $\mu_\nu \in \mathscr{M}(\mathbb{R}^d)$ **concentrated on the sphere** $\partial B_R^{(2)}(0)$ can then be defined through a measure over $\mathbb{R}$, $\nu \in \mathscr{M}(\mathbb{R})$, by

$$\langle \mu_\nu, \varphi \rangle = \int_{[0,2\pi)} \varphi(\gamma(\theta)) d\nu(\theta), \quad \forall \varphi \in C_c(\mathbb{R}^d). \tag{2.34}$$

Since $\left|\int_{[0,2\pi)} \varphi(\gamma(\theta))d\nu(\theta)\right| \leqslant \|\varphi\|_{\infty,\partial B_R^{(2)}(0)} |\nu|([0,2\pi)) < \infty$, μ is a continuous linear functional over $C(\mathbb{R}^d)$ (cf. Eq. (A.6)) and thus μ_ν is a well-defined compactly supported measure (Theorem 2.1.5). We

remark that this definition depends on the measure ν and on the selected parametrization γ. When using the already specified parametrization γ and when the measure ν is of the form $d\nu(\theta) = (2\pi R)^{-1}d\theta$, the measure μ_ν defined through (2.34) is called the **uniform measure** over the sphere $\partial B_R^{(2)}(0)$. It is a positive finite measure with total mass equal to 1, that is, a probability measure. This concept is naturally generalized to the $d-$dimensional case using polar coordinates and the surface of the $d-1$-sphere, $2R^{d-1}\frac{\pi^{d/2}}{\Gamma(d/2)}$. We will denote by $\mu_{unif}^{\partial B_R^{(d)}(0)}$ the uniform measure supported on $\partial B_R^{(d)}(0) \subset \mathbb{R}^d$. ▫

More generally, consider $A \subset \mathbb{R}^d$ a Borel sub-manifold of dimension $m < d$. Let us suppose, for simplicity, that A is homeomorphic to a subset D of $\mathbb{R}^m$, and hence we can use a parametrization of A given by a continuous mapping $\gamma : D \subset \mathbb{R}^m \to A \subset \mathbb{R}^d$ that defines an homeomorphism between D and A (γ is bijective with continuous inverse). Since γ is continuous, D is a Borel subset of $\mathbb{R}^m$. We then consider a measure $\nu \in \mathscr{M}(\mathbb{R}^m)$ and we define

$$\langle \mu_\nu, \varphi \rangle := \int_D \varphi(\gamma(\theta))d\nu(\theta), \quad \forall \varphi \in C_c(\mathbb{R}^d). \tag{2.35}$$

This defines a continuous linear functional over $C_c(\mathbb{R}^d)$. Indeed, let $K \subset \mathbb{R}^d$ be a compact set. Since γ is an homeomorphism, γ^{-1} is continuous. Hence, the set $\gamma^{-1}(A \cap K)$ is bounded since $\gamma^{-1}(A \cap K) = \gamma^{-1}(K)$ and $\gamma^{-1}(K)$ is compact. We conclude that for all $\varphi \in C_c(\mathbb{R}^d)$ such that $\mathrm{supp}(\varphi) \subset \gamma^{-1}(A \cap K)$ it holds that

$$|\langle \mu, \varphi \rangle| = \left| \int_D \varphi(\gamma(\theta))d\nu(\theta) \right| = \left| \int_{\gamma^{-1}(A\cap K)} \varphi(\gamma(\theta))d\nu(\theta) \right| \leqslant \|\varphi\|_\infty \underbrace{|\nu|(\gamma^{-1}(A \cap K))}_{<\infty}. \tag{2.36}$$

Hence, μ_ν defines a continuous linear functional over $C_c(\mathbb{R}^d)$ (Eq. (2.22)). From Riesz Representation Theorem 2.1.3, we obtain that μ_ν is a well-defined measure in $\mathscr{M}(\mathbb{R}^d)$, and it is concentrated on A.

Example 2.1.2. This is the **MOST IMPORTANT** example of this section. Consider the space of *doubled dimension* $\mathbb{R}^d \times \mathbb{R}^d$ ($= \mathbb{R}^{2d}$). Consider the hyperplane $\{y = x\} := \{(x, y) \in \mathbb{R}^d \times \mathbb{R}^d \mid y = x\}$, which is a sub-manifold of dimension d. Let $\mu \in \mathscr{M}(\mathbb{R}^d)$. We then define a measure over $\mathbb{R}^d \times \mathbb{R}^d$, denoted by $\mu\delta^{\{y=x\}}$, as

$$\langle \mu\delta^{\{y=x\}}, \varphi \rangle = \int_{\mathbb{R}^d} \varphi(x, x)d\mu(x), \quad \forall \varphi \in C_c(\mathbb{R}^d \times \mathbb{R}^d). \tag{2.37}$$

The measure $\mu\delta^{\{y=x\}}$ defined in this way is then **concentrated on the hyperplane** $\{y = x\}$. For this type of measures, a two-dimensional integral is reduced to a one-dimensional one, since we have

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \varphi(x, y)d(\mu\delta^{\{y=x\}})(x, y) = \int_{\mathbb{R}^d} \varphi(x, x)d\mu(x), \tag{2.38}$$

for every $\varphi \in C_c(\mathbb{R}^d \times \mathbb{R}^d)$. This kind of measure can also be described in a quite simple way through its

action to (some) Borel subsets of $\mathbb{R}^d \times \mathbb{R}^d$. Indeed, if we consider subsets of $\mathbb{R}^d \times \mathbb{R}^d$ of the form $A \times B$ with $A, B \in \mathcal{B}_B(\mathbb{R}^d)$, then

$$\mu\delta^{\{y=x\}}(A \times B) = \mu(A \cap B). \tag{2.39}$$

This can be shown by approaching the indicator function $\mathbf{1}_{A\times B}(x, y) = \mathbf{1}_A(x)\mathbf{1}_B(y)$ by a suitable sequence of functions in $C_c(\mathbb{R}^d \times \mathbb{R}^d)$. See Lemma A.2.2 in Appendix A.2, applicable when A and B are open, which is enough to completely characterize the measure. This also allows to conclude, using typical results of Measure Theory, that Eq. (2.38) also holds for any measurable φ such that $x \mapsto \varphi(x, x)$ is integrable with respect to μ.

We can actually prove the stronger condition that *every* measure in $\mathscr{M}(\mathbb{R}^d \times \mathbb{R}^d)$ which is concentrated on $\{y = x\}$ can be expressed in the form (2.39) for some measure $\mu \in \mathscr{M}(\mathbb{R}^d)$. Indeed, if $\nu \in \mathscr{M}(\mathbb{R}^d \times \mathbb{R}^d)$ is concentrated on $\{y = x\}$, we can define $\mu(A) := \nu(A \times \mathbb{R}^d)$ for all $A \in \mathcal{B}_B(\mathbb{R}^d)$. Since ν is concentrated on $\{y = x\}$, it follows that

$$\mu(A) = \nu(A \times \mathbb{R}^d) = \nu\left((A \times \mathbb{R}^d) \cap \{y = x\}\right) = \nu\left((A \times A) \cap \{y = x\}\right) = \nu(A \times A),$$

hence $\mu(A) \in \mathbb{C}$ since A is bounded. The $\sigma-$additivity of μ follows immediately from the $\sigma-$additivity of ν. μ is then a well-defined measure. If $A, B \in \mathcal{B}_B(\mathbb{R}^d)$, then

$$\mu(A\cap B) = \nu\left((A \cap B) \times (A \cap B)\right) = \nu\left(\{y = x\} \cap ((A \cap B) \times (A \cap B))\right) = \nu\left(\{y = x\} \cap (A \times B)\right) = \nu(A\times B). \tag{2.40}$$

We consider the next result which relates the characteristics of μ to those of $\mu\delta^{\{y=x\}}$. The proof of this Proposition is presented in Appendix A.3.1.

Proposition 2.1.2. *Let j designating "c", "F" or "SG". Then, $\mu\delta^{\{y=x\}} \in \mathscr{M}_j(\mathbb{R}^d \times \mathbb{R}^d)$ if and only if $\mu \in \mathscr{M}_j(\mathbb{R}^d)$.*

The kind of measure exposed in this example will be used to describe orthogonal Random Measures, a stochastic tool which is a key concept in the study of Random Fields, specially in a stationary framework. We finally remark that if μ is the Lebesgue measure, the associated measure $\mu\delta^{\{y=x\}}$ is more often denoted by $\delta(y-x)$ (or $\delta(x-y)$). This measure of two variables plays a central role in the theory of Partial Differential Equations since it is used to define *Green's functions*. ▫

Example 2.1.3. Let $c > 0$. Consider the subset of $\mathbb{R}^d \times \mathbb{R}$, $\mathcal{C}^c := \{(x, t) \in \mathbb{R}^d \times \mathbb{R} \;\big|\; |t| = c|x|\}$, which we call a **spatio-temporal cone**. Even if $\mathcal{C}^c$ is not strictly speaking a manifold, we can still define measures concentrated on $\mathcal{C}^c$ using the same principle as in (2.35). Consider two measures $\mu_1, \mu_2 \in \mathscr{M}(\mathbb{R}^d)$. We define the measure $\mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)} \in \mathscr{M}(\mathbb{R}^d \times \mathbb{R})$ as

$$\langle \mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}, \psi\rangle := \int_{\mathbb{R}^d} \psi(x, c|x|)d\mu_1(x) + \int_{\mathbb{R}^d\setminus\{0\}} \psi(x, -c|x|)d\mu_2(x), \quad \forall \psi \in C_c(\mathbb{R}^d \times \mathbb{R}). \tag{2.41}$$

Let $K \subset \mathbb{R}^d \times \mathbb{R}$ be a compact set. Let $K_d \subset \mathbb{R}^d$ and $K_1 \subset \mathbb{R}$ compacts such that $K \subset K_d \times K$. If $\psi \in C_c(\mathbb{R}^d \times \mathbb{R})$ is such that $\text{supp}(\varphi) \subset K$, then

$$\left|\int_{\mathbb{R}^d} \psi(x, c|x|)d\mu_1(x)\right| = \left|\int_{K_d} \psi(x, c|x|)d\mu_1(x)\right| \leqslant |\mu_1|(K_d) \sup_{x\in K_d} |\psi(x, c|x|)| \leqslant |\mu_1|(K_d)\|\psi\|_\infty. \quad (2.42)$$

The same principle is applied to the integral with respect to μ_2 in (2.41). This proves that $\mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}$ is a continuous linear functional over $C_c(\mathbb{R}^d \times \mathbb{R})$ and hence a well-defined measure. Integration with respect to such a measure transforms a $d+1-$ dimensional integral into two $d-$dimensional integrals, having

$$\int_{\mathbb{R}^d\times\mathbb{R}} \psi(x,t)d\mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}(x,t) = \int_{\mathbb{R}^d} \psi(x, c|x|)d\mu_1(x) + \int_{\mathbb{R}^d\setminus\{0\}} \psi(x, -c|x|)d\mu_2(x), \quad (2.43)$$

for every $\psi \in C_c(\mathbb{R}^d \times \mathbb{R})$. This kind of measure can also be described through its action over some Borel subsets of $\mathbb{R}^d \times \mathbb{R}$. Consider the sets of the form $A \times B$ with $A \in \mathcal{B}_B(\mathbb{R}^d)$ and $B \in \mathcal{B}_B(\mathbb{R})$. Then, it holds that

$$\mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}(A \times B) = \int_A \delta_{c|x|}(B)d\mu_1(x) + \int_{A\setminus\{0\}} \delta_{-c|x|}(B)d\mu_2(x). \quad (2.44)$$

The same arguments used in Example 2.1.2 to prove the analogue relation (2.39) can be used to prove (2.44). Similarly, we can prove that *every* measure in $\mathscr{M}(\mathbb{R}^d \times \mathbb{R})$ concentrated on $\mathcal{C}^c$ can be expressed in the form (2.44) for some pair of measures $\mu_1, \mu_2 \in \mathscr{M}(\mathbb{R}^d)$. Indeed, if $\nu \in \mathscr{M}(\mathbb{R}^d \times \mathbb{R})$ is concentrated on the cone, we take $\mu_1(A) = \nu(A \times \mathbb{R}^+)$ and $\mu_2(A) = \nu(A \times \mathbb{R}^-_*)$ for every $A \in \mathcal{B}_B(\mathbb{R}^d)$. With the same arguments as in Example 2.1.2, it can be proven that μ_1 and μ_2 are well-defined measures in $\mathscr{M}(\mathbb{R}^d)$ and that they satisfy (2.44) for $\mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)} = \nu$. It also follows that expression (2.43) is still valid for every measurable ψ such that the mappings $x \mapsto \psi(x, c|x|)$ and $x \mapsto \psi(x, -c|x|)$ are integrable with respect to μ_1 and μ_2 respectively.

We have, in addition, an analogous to Proposition 2.1.2. The proof of this Proposition is presented in Appendix A.3.2.

Proposition 2.1.3. *Let j designating "c", "F" or "SG". Then, $\mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)} \in \mathscr{M}_j(\mathbb{R}^d \times \mathbb{R})$ if and only if both μ_1 and μ_2 are in $\mathscr{M}_j(\mathbb{R}^d)$.*

The name *spatio-temporal cone* makes reference to the use of the set $\mathcal{C}^c$ in the study of some physical phenomena in a spatio-temporal context. As it will be seen in Chapter 5, this set plays an important role when analysing solutions to the homogeneous Wave equation. ▫

2.2 Distributions

The *Theory of Distributions* is a mathematical theory developed by Laurent Schwartz in the middle of the 20th century whose main aim is to rigorously define the derivative of a large class of objects, such as any ar-

bitrary continuous function or Radon measures. In this framework the main objects are not functions which can be evaluated at points of the space, but rather objects that act over test-functions belonging to convenient functional spaces, describing how the variable acts on a region of the space rather than in a singular point. Such objects are called *Distributions* or sometimes also *Generalized Functions.* The development of this theory has allowed enormous advances in many branches of theoretical and applied mathematics. Borrowing words from (Demengel & Demengel, 2000, Preface): *"...this theory is so revolutionary in its concept as Einstein's Relativity Theory in Physics"*. The *freedom* in the manipulation of the new objects has allowed to well-pose and analyse equations arising from Physics, Engineering, Signal Analysis and other fields. Regarding Probability Theory and Statistics, it provides a framework where calculus with Stochastic Processes or Random Functions can be rigorously treated, and hence the analysis of SPDEs is possible. We will present this last application of Distribution Theory in Section 3.4 of Chapter 3.

There are plenty of treaties and bibliographic sources concerning the Theory of Distributions and its associated issues. The *Bible* of this theory is the classical Schwartz's treaty (Schwartz, 1966). A source with a little less depth but with an effective and clear exposition is Donoghue (1969). The exposition presented in Reed & Simon (1980) is also remarkable. Here the authors present the basis of modern Functional Analysis, together with the main concepts of Topology, which allows them to compile a coherent and clear presentation passing through the most important details of the theory of locally convex topological vector spaces. The exposition of the so-called *tempered distributions*, done in Reed & Simon (1980, Section V.3 and Appendix), is presented before the exposition of general distributions, and it stands out for its compactness, clarity, and also because of the inclusion of quite deep and important results such as the Regularity Theorem, the Nuclear Theorem and the development on the Hermite basis. Another introductory source on Theory of Distribution, which is non-specialist oriented and less technical can be found in Richards & Youn (1995). This source is recommendable for an easier understanding of the theory and its basic results, but it is also remarkable because of its Chapter 7, where a very interesting *symmetric* definition of the multiplication and convolution of distributions in a more general framework than the classical ones is exposed with an astonishing simplicity. We finally suggest a special last bibliographic source, which gives a fast and effective introduction to Distribution Theory with *geostatistician objectives*: the Appendix A in the Geostatistics's classical opus Matheron (1965).

In this work we focus on *tempered distributions*, which is a framework in which differential operators and the Fourier Transform can be used freely. In section 2.2.5 we make some comments on more general spaces of distributions. The definitions and results presented in this section will be used extensively in this dissertation.

All along this work we will extensively use the convenient *multi-index notation* for differential operators and vector powers. If $\alpha \in \mathbb{N}^d$, we denote by $|\alpha| = \alpha_1 + ... + \alpha_d$. For differential operators over $\mathbb{R}^d$, we use

the notation:

$$D^{\alpha} := \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} ... \partial x_d^{\alpha_d}}. \tag{2.45}$$

For a vector $x = (x_1, ..., x_d) \in \mathbb{R}^d$, the symbol x^{α} denotes the real number determined by

$$x^{\alpha} := x_1^{\alpha_1} x_2^{\alpha_2} ... \, x_d^{\alpha_d}. \tag{2.46}$$

We will sometimes work with spaces of *double dimensions*, that is, over spaces of the form $\mathbb{R}^d \times \mathbb{R}^m = \mathbb{R}^{d+m}$ with $d, m \in \mathbb{N}_*$, where the *first variables* in $\mathbb{R}^d$ may play a different role that the *second variables* in $\mathbb{R}^m$. In such a case the *first components* will be denoted by the letter x and the *second components* by the letter y. If $\alpha \in \mathbb{N}^d$ and $\beta \in \mathbb{N}^m$, its concatenation multi-index is denoted by $(\alpha, \beta) := (\alpha_1, ..., \alpha_d, \beta_1, ..., \beta_m) \in \mathbb{N}^d \times \mathbb{N}^m = \mathbb{N}^{d+m}$. Hence, in such a case $D^{(\alpha,\beta)}$ denotes the differential operator over $\mathbb{R}^d \times \mathbb{R}^m$: $\frac{\partial^{|\alpha|+|\beta|}}{\partial x_1^{\alpha_1} ... \partial x_d^{\alpha_d} \partial y_1^{\beta_1} ... \partial y_d^{\beta_d}}$. For the power, we will have, of course, $(x, y)^{(\alpha,\beta)} = x^{\alpha} y^{\beta} = x_1^{\alpha_1} x_2^{\alpha_2} ... \, x_d^{\alpha_d} y_1^{\beta_1} y_2^{\beta_2} ... \, y_d^{\beta_d}$, for $(x, y) \in \mathbb{R}^d \times \mathbb{R}^m$. Sometimes, we will also write $\alpha \leqslant \beta$ for $\alpha, \beta \in \mathbb{N}^d$, meaning $\alpha_j \leqslant \beta_j$ for all $j \in \{1, ..., d\}$.

We make explicit a terminological convention which is quite necessary in this work since we are going to use concepts from Distribution Theory and from Probability Theory: *all along this dissertation, we will always use the word "distribution" referring to a Generalized Function, and NEVER to the "probability distribution" of a random variable, term widely used in Statistics and Probability. For the latter mathematical object, we use the word "law".*

2.2.1 Tempered distributions

We denote by $C^{\infty}(\mathbb{R}^d)$ the space of smooth complex functions over $\mathbb{R}^d$. We introduce the following space of smooth functions.

Definition 2.2.1. *The **Schwartz space**, denoted by $\mathscr{S}(\mathbb{R}^d)$, is defined as the space of smooth fast-decreasing complex functions over $\mathbb{R}^d$. Explicitly,*

$$\mathscr{S}(\mathbb{R}^d) := \{\varphi \in C^{\infty}(\mathbb{R}^d) \;\big|\; \|x^{\alpha} D^{\beta} \varphi\|_{\infty} < \infty, \; \forall \alpha, \beta \in \mathbb{N}^d\}. \tag{2.47}$$

Equivalently, the Schwartz space can be defined as the space of complex smooth functions such that all of its derivatives are in $C_{FD}(\mathbb{R}^d)$. Typical examples of members of this space are Gaussian functions, that is, functions of the form $\varphi(x) = be^{-a|x|^2}$ with $a > 0$ and $b \in \mathbb{C}$. A function in $\mathscr{S}(\mathbb{R}^d)$ will be often called a **test-function**.

The Schwartz space is equipped with a topology determined by the metric

$$(\varphi, \phi) \mapsto \sum_{\alpha,\beta \in \mathbb{N}^d} \frac{1}{2^{|\alpha|+|\beta|+2}} \frac{\|x^\alpha D^\beta(\varphi - \phi)\|_\infty}{1 + \|x^\alpha D^\beta(\varphi - \phi)\|_\infty}, \quad \varphi, \phi \in \mathscr{S}(\mathbb{R}^d). \tag{2.48}$$

This topology can be described equivalently using semi-norms of the form $\varphi \mapsto \|(1 + |x|^2)^N D^\beta \varphi\|_\infty$ with $N \in \mathbb{N}$ and $\beta \in \mathbb{N}^d$, similarly to the case of the space $C_{FD}(\mathbb{R}^d)$ (see Eq. (2.29)) but with differentiable functions. It is known that $\mathscr{S}(\mathbb{R}^d)$ equipped with this topology is a complete locally convex metric space, hence it is a Fréchet space; see Reed & Simon (1980, Theorem V.9) or the arguments in Donoghue (1969, Chapter 28). We can thus speak about continuous linear functionals over $\mathscr{S}(\mathbb{R}^d)$.

Definition 2.2.2. *A **tempered distribution** is a continuous linear functional over the Schwartz space, $T : \mathscr{S}(\mathbb{R}^d) \to \mathbb{C}$, that is, a member of the dual space of $\mathscr{S}(\mathbb{R}^d)$. This dual space, denoted by $\mathscr{S}'(\mathbb{R}^d)$ is called the **space of tempered distributions**.*

We present a general criterion for verifying if a linear functional is a tempered distribution, which comes immediately from Theorem D.0.1. Let $T : \mathscr{S}(\mathbb{R}^d) \to \mathbb{C}$ be a linear functional. Then, $T \in \mathscr{S}'(\mathbb{R}^d)$ if and only if there exists $C > 0$ and $N \in \mathbb{N}$ such that

$$|\langle T, \varphi \rangle| \leqslant C \sum_{\substack{\alpha,\beta \in \mathbb{N}^d \\ |\alpha|,|\beta| \leqslant N}} \|x^\alpha D^\beta \varphi\|_\infty, \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^d). \tag{2.49}$$

Let us present some important examples of tempered distributions.

Example 2.2.1. Let $f : \mathbb{R}^d \to \mathbb{C}$ be a **polynomially bounded measurable function**. Then, the linear functional defined through the integral

$$\langle f, \varphi \rangle := \int_{\mathbb{R}^d} f(x)\varphi(x)dx, \quad \varphi \in \mathscr{S}(\mathbb{R}^d), \tag{2.50}$$

defines a tempered distribution. This is easy to verify using the criterion (2.49). This holds in particular for polynomially bounded continuous functions. It can also be proven that any **function in the Lebesgue spaces** $f \in L^p(\mathbb{R}^d)$, with $p \in [1, \infty]$, defines a tempered distribution through the integral (2.50). Hence, in all of this cases the function f can be identified with a temperate distribution. In this work we will use the common abuse of language of saying that the function f *is* a tempered distribution. ▫

Example 2.2.2. Let μ be a **slow-growing measure**. The linear functional defined through the integral

$$\langle \mu, \varphi \rangle = \int_{\mathbb{R}^d} \varphi(x)d\mu(x), \quad \varphi \in \mathscr{S}(\mathbb{R}^d), \tag{2.51}$$

defines a tempered distribution, which can be seen directly from Eq. (2.30). Hence, every measure $\mu \in \mathscr{M}_{SG}(\mathbb{R}^d)$ determines a tempered distribution. As in the case of functions, we will simply say that μ *is* a tempered distribution. It follows that every **finite measure** and every **compactly supported measure** is also a tempered distribution. It can be proven that a positive measure $\mu \in \mathscr{M}^+(\mathbb{R}^d)$ defines a tempered distribution if and only if it is slow-growing. Some non-positive non-slow-growing measures in $\mathscr{M}(\mathbb{R}^d)$ can also determine tempered distributions but not exactly through the Lebesgue integral but rather by taking a limit of Lebesgue integrals (See Schwartz, 1966, Theorem VII, Chapter VII and the comments therein). ▫

The space of tempered distributions can then be seen as an abstract space where usual objects such as some kinds of functions and measures are included. Of course, there are *many* other types of objects in this class. The main motivation of this space is to have a unified class where typical operations of calculus and Fourier Analysis can be applied without formal problems.

Let $T \in \mathscr{S}'(\mathbb{R}^d)$. If $\langle T, \varphi\rangle \in \mathbb{R}$ for every *real* function $\varphi \in \mathscr{S}(\mathbb{R}^d)$, T is said to be **real**. The complex **conjugate** of T, denoted by $\overline{T}$, is defined as the distribution determined by $\langle \overline{T}, \varphi\rangle := \overline{\langle T, \overline{\varphi}\rangle}$ for all $\varphi \in \mathscr{S}(\mathbb{R}^d)$. Following criterion (2.49) it is immediate that $\overline{T} \in \mathscr{S}'(\mathbb{R}^d)$. The **real part** of T is defined as $T_R := \frac{T+\overline{T}}{2}$, and its **imaginary part** is defined as $T_I = \frac{T-\overline{T}}{2i}$. Both T_R and T_I are real tempered distributions, and it holds that $T = T_R + iT_I$. If $\varphi : \mathbb{R}^d \to \mathbb{C}$ is any function, its reflection, denoted by $\check{\varphi}$ is the function defined as $\check{\varphi}(x) = \varphi(-x)$ for every $x \in \mathbb{R}^d$. The **reflection** of the tempered distribution T, denoted by $\check{T}$, is defined through $\langle \check{T}, \varphi\rangle = \langle T, \check{\varphi}\rangle$ for every $\varphi \in \mathscr{S}(\mathbb{R}^d)$ and we can also conclude that $\check{T} \in \mathscr{S}'(\mathbb{R}^d)$. If T satisfies $\check{T} = T$, it is said to be **even**. If it satisfies $\check{T} = -T$, it is said to be **odd**. If it satisfies $\overline{\check{T}} = T$, that is, if its real part is even and its imaginary part is odd, it is said to be **Hermitian**. The reader can verify that all the definitions given in this paragraph coincide with the corresponding classical definitions when T is a function or a measure.

Let us present a topology that we will sometimes use for the space of tempered distribution $\mathscr{S}'(\mathbb{R}^d)$. Since this space is a dual space, there are many possible topologies that it can be endowed with (see Trèves, 1967, Chapter 19). We are going to use the simplest one: the so-called **weak-star**, or **weak-$*$**, topology, which is no other but the topology of point-wise sequential convergence[7]. A *sequence* of tempered distributions $(T_n)_{n\in\mathbb{N}}$ is said to converge to $T \in \mathscr{S}'(\mathbb{R}^d)$ in the weak-$*$ topology, denoted by $T_n \overset{\mathscr{S}'}{\to} T$, if for every $\varphi \in \mathscr{S}(\mathbb{R}^d)$, $\langle T_n, \varphi\rangle \to \langle T, \varphi\rangle$.

[7]To be more precise, the weak-$*$ topology on $\mathscr{S}'(\mathbb{R}^d)$ is the one determined by the family of semi-norms:

$$p_{\varphi_1,\dots,\varphi_N}(T) := \sup_{j\in\{1,\dots,N\}} |\langle T, \varphi_j\rangle|, \quad T \in \mathscr{S}'(\mathbb{R}^d), \tag{2.52}$$

for every finite family of test-functions $(\varphi_j)_{j\in\{1,\dots,N\}} \subset \mathscr{S}(\mathbb{R}^d)$. This topology is not metric, and the description of convergence sequences is not sufficient to completely describe the topology.

2.2.2 Operations on Tempered distributions

In this section we will define some linear operations that are defined for tempered distributions and which generalize classical operations in calculus and Analysis.

We begin by giving a criterion to determine when a linear operator over the Schwartz space $\mathcal{L} : \mathscr{S}(\mathbb{R}^d) \to \mathscr{S}(\mathbb{R}^d)$ is continuous. Such a linear operator is continuous if and only if for every $\alpha, \beta \in \mathbb{N}^d$ there exists $C > 0$ and $N \in \mathbb{N}$ such that

$$\|x^\alpha D^\beta \mathcal{L}(\varphi)\|_\infty \leqslant C \sum_{\substack{\alpha',\beta' \in \mathbb{N}^d \\ |\alpha'|,|\beta'| \leqslant N}} \|x^{\alpha'} D^{\beta'} \varphi\|_\infty, \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^d). \tag{2.53}$$

Similarly to criterion (2.49), this criterion is obtained directly from Theorem D.0.1. We define now a class of linear operators over the space of tempered distributions. The members of this class are said to be **defined through an adjoint**.

Definition 2.2.3. *Let $\mathcal{L} : \mathscr{S}(\mathbb{R}^d) \to \mathscr{S}(\mathbb{R}^d)$ be a continuous and linear operator over the Schwartz space. Its **adjoint** operator, denoted by $\mathcal{L}^*$, is the linear operator over the space of tempered distributions $\mathcal{L}^* : \mathscr{S}'(\mathbb{R}^d) \to \mathscr{S}'(\mathbb{R}^d)$ defined through,*

$$\langle \mathcal{L}^* T, \varphi \rangle := \langle T, \mathcal{L}\varphi \rangle, \quad \forall T \in \mathscr{S}'(\mathbb{R}^d), \varphi \in \mathscr{S}(\mathbb{R}^d). \tag{2.54}$$

In other words, the adjoint operator simply does $\mathcal{L}^* T = T \circ \mathcal{L}$ for every $T \in \mathscr{S}'(\mathbb{R}^d)$. It is immediate that $\mathcal{L}^* T$ is in $\mathscr{S}'(\mathbb{R}^d)$ since it is a linear functional which is the composition of continuous linear mappings, hence it is continuous. It is also straightforward that $\mathcal{L}^*$ is sequentially[8] continuous with the weak-$*$ topology.

Most linear operators defined over $\mathscr{S}'(\mathbb{R}^d)$ used in practice (and theory) are defined in this way. Those operators are very convenient since they are completely defined through an action on test-functions, for which many traditional linear operators are defined.

Let us remark two simple examples of operators defined in this way. The first is the reflection operator $T \in \mathscr{S}'(\mathbb{R}^d) \mapsto \check{T}$, defined as in Section 2.2.1. By definition, it is clear that the reflection operator on $\mathscr{S}'(\mathbb{R}^d)$ is the adjoint of the reflection operator over $\mathscr{S}(\mathbb{R}^d)$, $\varphi \in \mathscr{S}(\mathbb{R}^d) \mapsto \check{\varphi}$. Another example is the **translation operator**. Let $h \in \mathbb{R}^d$. The translation by h of a function $\varphi : \mathbb{R}^d \to \mathbb{C}$, denoted by $\tau_h \varphi$, is the function $\tau_h \varphi : \mathbb{R}^d \to \mathbb{C}$ defined as $\tau_h \varphi(x) := \varphi(x - h)$ for all $x \in \mathbb{R}^d$. The translation by h defines a linear operator from $\mathscr{S}(\mathbb{R}^d)$ to $\mathscr{S}(\mathbb{R}^d)$ and it is also continuous[9]. The translation by h for tempered distributions is defined as the adjoint of the translation by $-h$ for test-functions. Explicitly, if $T \in \mathscr{S}'(\mathbb{R}^d)$, we define

[8] It is, actually, continuous. See the discussion after Example 1 in Reed & Simon (1980, Section V.3).

[9] A few technical arguments using Taylor's expansion and convexity, as it is done in the case of Lemma A.4.3 in Appendix A.4.2 can be used to prove this claim.

its translation by h, denoted by $\tau_h T$, as the distribution defined through $\langle \tau_h T, \varphi \rangle = \langle T, \tau_{-h}\varphi \rangle$ for every $\varphi \in \mathscr{S}(\mathbb{R}^d)$. Then, $\tau_h : \mathscr{S}'(\mathbb{R}^d) \to \mathscr{S}'(\mathbb{R}^d)$ is a continuous linear operator. Using a change of variable in the integral defining the linear functional, the reader can verify that this definition coincides with the classical definition of translation when T is a function or a measure.

We remark that $\mathscr{S}(\mathbb{R}^d) \subset \mathscr{S}'(\mathbb{R}^d)$ in the sense of distributions since every test-function can also determine a tempered distribution following Example 2.2.1. Hence, the adjoint operator $\mathcal{L}^*$ is also defined for test-functions. It must not be mistaken, however, with the initial operator $\mathcal{L}$. The case where $\mathcal{L}$ can be identified with $\mathcal{L}^*$ is the case of so-called **self-adjoint operators**, which are operators satisfying $\langle \phi, \mathcal{L}\varphi \rangle = \langle \mathcal{L}\phi, \varphi \rangle$ for all couple of test-functions $\varphi, \phi \in \mathscr{S}(\mathbb{R}^d)$. For example, the reflection operator is self-adjoint, while the translation operator is not.

We present now the most important examples of linear operators over the space of tempered distributions which are defined in this way.

Differentiation

Let $\alpha \in \mathbb{N}^d$ and let D^α its associated differential operator. $D^\alpha : \mathscr{S}(\mathbb{R}^d) \to \mathscr{S}(\mathbb{R}^d)$ is linear and continuous, which is easy to obtain following criterion (2.53). Its adjoint can then be defined through the expression (2.54). Nevertheless, it is more important to directly define D^α over the space of tempered distributions and then to identify its *pre-adjoint*, which is not necessarily the operator D^α for test-functions.

Definition 2.2.4. *The **differential operator** over the space of tempered distributions $D^\alpha : \mathscr{S}'(\mathbb{R}^d) \to \mathscr{S}'(\mathbb{R}^d)$ is defined through*

$$\langle D^\alpha T, \varphi \rangle := (-1)^{|\alpha|} \langle T, D^\alpha \varphi \rangle, \quad \forall T \in \mathscr{S}'(\mathbb{R}^d), \varphi \in \mathscr{S}(\mathbb{R}^d). \tag{2.55}$$

In other words, D^α is the adjoint of the operator $(-1)^{|\alpha|} D^\alpha$ defined over the Schwartz space. This definition is inspired by the *integration by parts formula*. The reader can verify, for example, that if $f \in C^{|\alpha|}(\mathbb{R}^d)$ is a polynomially bounded function with polynomially bounded derivatives, then the function $D^\alpha f$ in the classical sense satisfies $\langle D^\alpha f, \varphi \rangle = (-1)^{|\alpha|} \langle f, D^\alpha \varphi \rangle$ for all $\varphi \in \mathscr{S}(\mathbb{R}^d)$. Hence, the derivative D^α is a generalization of the classical notion of the derivative for enough regular functions, and it can be applied to any tempered distributions any number of times. This allows to formally differentiate complicated objects. For instance, any function as presented in the Example 2.2.1 can be differentiated any number of times. Slow-growing measures can also be differentiated any number of times. The objects which are obtained after applying this operator are not necessarily, of course, functions or measures, but they are well-defined tempered distributions. Actually, it is true that *every* tempered distribution is the derivative of large enough order of a polynomially bounded continuous function (Reed & Simon, 1980, Theorem V.10).

We give an example of application: the classical example of the *Heaviside function* on $\mathbb{R}$, which is the indicator function of the positive real line, $\mathbf{1}_{\mathbb{R}^+}$. This function is of course not differentiable at 0. However, its derivative in distributional sense is the Dirac measure at 0, $\frac{d}{dx}(\mathbf{1}_{\mathbb{R}^+}) = \delta$. The Dirac measure can also be differentiated, obtaining the tempered distribution $\varphi \in \mathscr{S}(\mathbb{R}) \mapsto -\frac{d\varphi}{dx}(0)$.

We finally remark that differential operators of even order are self-adjoint, while differential operators of odd order are not.

Multiplication with $\mathcal{O}_M(\mathbb{R}^d)$

Let us introduce the next space of functions:

Definition 2.2.5. *The space of* ***multiplicators of the Schwartz space****, denoted by $\mathcal{O}_M(\mathbb{R}^d)$ is defined as the space of all complex smooth functions such that all of their derivatives of all orders are polynomially bounded. Explicitly,*

$$\mathcal{O}_M(\mathbb{R}^d) := \{f \in C^\infty(\mathbb{R}^d) \;\big|\; \forall \alpha \in \mathbb{N}^d \; \exists C > 0 \; \exists N \in \mathbb{N} \text{ such that } |D^\alpha f(x)| \leqslant C(1 + |x|^2)^N \; \forall x \in \mathbb{R}^d\}. \tag{2.56}$$

If $f \in \mathcal{O}_M(\mathbb{R}^d)$ and $\varphi \in \mathscr{S}(\mathbb{R}^d)$, then $f\varphi \in \mathscr{S}(\mathbb{R}^d)$. Moreover, the application $\varphi \mapsto f\varphi$ is a continuous linear operator from $\mathscr{S}(\mathbb{R}^d)$ to $\mathscr{S}(\mathbb{R}^d)$. This can be seen by applying criterion (2.53) together with the polynomials bounding the derivatives of f. It can actually be proven that if f is a measurable function, then the multiplication by f, $\varphi \mapsto f\varphi$ over $\mathscr{S}(\mathbb{R}^d)$, is a continuous linear operator from $\mathscr{S}(\mathbb{R}^d)$ to $\mathscr{S}(\mathbb{R}^d)$ if and only if $f \in \mathcal{O}_M(\mathbb{R}^d)$ (Reed & Simon, 1980, Problem 23 in Chapter V). We remark, in addition, that $\mathcal{O}_M(\mathbb{R}^d) \subset \mathscr{M}_{SG}(\mathbb{R}^d) \subset \mathscr{S}'(\mathbb{R}^d)$.

The **multiplication** with $f \in \mathcal{O}_M(\mathbb{R}^d)$ over the space of tempered distributions $\mathscr{S}'(\mathbb{R}^d)$ is defined as the adjoint operator of the multiplication with f over the Schwartz space $\mathscr{S}(\mathbb{R}^d)$. Explicitly, if $T \in \mathscr{S}'(\mathbb{R}^d)$ and $f \in \mathcal{O}_M(\mathbb{R}^d)$, the multiplication $fT \in \mathscr{S}'(\mathbb{R}^d)$ is defined as the distribution which satisfies

$$\langle fT, \varphi \rangle := \langle T, f\varphi \rangle, \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^d). \tag{2.57}$$

It follows that the multiplication with f is a self-adjoint operator.

It is not immediate to generalize the notion of multiplication, that is, to define ST for two arbitrary tempered distributions T and S, while maintaining good topological properties and the analogies to the case of continuous functions. It is actually, a kind of *weakness* of the Theory of Distributions: the lack of a multiplicative algebra of spaces of distributions. See (Schwartz, 1954) for a discussion in the case generic distributions (not necessarily tempered). The particularity of the space $\mathcal{O}_M(\mathbb{R}^d)$ is that members of this space can be multiplied with *any* tempered distribution. However, it is possible to define products between

less regular functions with members in more restricted subspaces of $\mathscr{S}'(\mathbb{R}^d)$. For instance, we know (cf. Section 2.1.3) that every polynomially bounded measurable function can be multiplied with any distribution in $\mathscr{M}_{SG}(\mathbb{R}^d)$, obtaining as a result a member of $\mathscr{M}_{SG}(\mathbb{R}^d) \subset \mathscr{S}'(\mathbb{R}^d)$. However, some "nice properties" of the multiplication which we would like to hold are lost[10]. Some generalizations can be found in Shiraishi & Itano (1964) and the references therein. Other ways of describing multiplicative products between tempered distributions which will be sharply pointed out in Section 2.2.3 can be found in Richards & Youn (1995, Chapter 7).

Convolution with $\mathcal{O}'_c(\mathbb{R}^d)$

In this section we restrain ourselves to the convolution between a tempered distribution and a distribution of *fast decreasing behaviour*. This concept is entirely described in Schwartz (1966, Chapter VII, §5), but a simpler exposition, which is the one we will follow, can be found in Trèves (1967, Chapter 30).

We recall that for two complex valued functions over $\mathbb{R}^d$, φ, ϕ, which by simplicity are supposed to be in $\mathscr{S}(\mathbb{R}^d)$, their **convolution** product is defined as

$$(\varphi * \phi)(x) := \int_{\mathbb{R}^d} \varphi(x-y)\phi(y)dy. \tag{2.58}$$

The convolution product is commutative. It is also true that for functions φ, ϕ in the Schwartz space, its convolution is also in $\mathscr{S}(\mathbb{R}^d)$ (Donoghue, 1969, Chapter 29). If D^α is a differential operator, it is well-known that $D^\alpha(\varphi * \phi) = D^\alpha\varphi * \phi = \varphi * D^\alpha\phi$.

If $T \in \mathscr{S}'(\mathbb{R}^d)$, its **convolution with a test-function** $\varphi \in \mathscr{S}(\mathbb{R}^d)$ is defined as the *function*

$$x \in \mathbb{R}^d \mapsto (T * \varphi)(x) := \langle T, \tau_x\check{\varphi}\rangle. \tag{2.59}$$

It can be proven that this function is in $C^\infty(\mathbb{R}^d)$. Moreover, it is actually in $\mathcal{O}_M(\mathbb{R}^d)$ (Trèves, 1967, Theorem 30.2). In particular, $T * \varphi$ is a tempered distribution. If D^α is a differential operator, then it holds that $D^\alpha(T * \varphi) = D^\alpha T * \varphi = T * D^\alpha\varphi$. The next definition is the one stated in Trèves (1967, Definition 30.1).

Definition 2.2.6. *Let $T \in \mathscr{S}'(\mathbb{R}^d)$. We say that T is of **fast decreasing behaviour** if for all $N \in \mathbb{N}$ there exists $M_N \in \mathbb{N}$ and there exists a finite family of continuous functions $(f_\alpha)_{\alpha\in\mathbb{N}^d, |\alpha|\leqslant M_N} \subset C(\mathbb{R}^d)$ such that $(1+|x|^2)^N f_\alpha \in C_0(\mathbb{R}^d)$ for all $\alpha \in \mathbb{N}^d$ with $|\alpha| \leqslant M_N$, and such that*

$$T = \sum_{|\alpha|\leqslant M_N} D^\alpha f_\alpha, \tag{2.60}$$

where the derivatives are taken in the distributional sense.

[10] For instance, the Exchange Formula of the Fourier Transform, which will be presented in Section 2.2.2 could fail.

The space of distributions of fast decreasing behaviour over $\mathbb{R}^d$ is denoted by $\mathcal{O}'_c(\mathbb{R}^d)$ and it is also often called the **space of convoluters of tempered distributions**. If $S \in \mathcal{O}'_c(\mathbb{R}^d)$ and $\varphi \in \mathscr{S}(\mathbb{R}^d)$, then $S * \varphi \in \mathscr{S}(\mathbb{R}^d)$, and it can be proven that the application $\varphi \mapsto S * \varphi$ is a continuous linear operator from $\mathscr{S}(\mathbb{R}^d)$ to $\mathscr{S}(\mathbb{R}^d)$ (Trèves, 1967, Theorem 30.1).

Let $S \in \mathcal{O}'_c(\mathbb{R}^d)$. The **convolution with** S is the linear operator defined over $\mathscr{S}'(\mathbb{R}^d)$ as the adjoint of the convolution with $\check{S}$ over $\mathscr{S}(\mathbb{R}^d)$. Explicitly,

$$\langle T * S, \varphi \rangle := \langle T, \check{S} * \varphi \rangle, \quad \forall T \in \mathscr{S}'(\mathbb{R}^d), \varphi \in \mathscr{S}(\mathbb{R}^d). \tag{2.61}$$

The convolution $T * S$ is then a well-defined tempered distribution. It also holds that if D^α is a differential operator, then $D^\alpha(T * S) = D^\alpha T * S = T * D^\alpha S$. The reader can verify that the definition of convolution following Eq. (2.61) coincides with the classical one when T and S are convolable functions.

Fourier Transform

The Fourier Transform is the *raison d'être* of the Schwartz space and of tempered distributions: these spaces are created in order to apply the Fourier Transform to more general objects than in the *classical* framework of functions in $L^2(\mathbb{R}^d)$, while maintaining all of its interesting properties. In this section we make precise the convention of the Fourier Transform we use in this work and we recall its main properties.

Let $\varphi \in \mathscr{S}(\mathbb{R}^d)$. Its **Fourier Transform** is defined as the function

$$\mathscr{F}(\varphi)(\xi) := \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{-i\xi^T x} \varphi(x) dx, \quad \xi \in \mathbb{R}^d. \tag{2.62}$$

The Fourier Transform is a linear and continuous bijective operator from $\mathscr{S}(\mathbb{R}^d)$ to $\mathscr{S}(\mathbb{R}^d)$. Its inverse operator is the **Inverse Fourier Transform** and it is determined by

$$\mathscr{F}^{-1}(\varphi)(\xi) := \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{i\xi^T x} \varphi(x) dx, \quad \xi \in \mathbb{R}^d. \tag{2.63}$$

Of course, $\mathscr{F}^{-1}$ also defines a continuous linear operator from $\mathscr{S}(\mathbb{R}^d)$ to $\mathscr{S}(\mathbb{R}^d)$. See Donoghue (1969, Chapter 29). We will mainly use the letter ξ to describe the variables in the space $\mathbb{R}^d$ *after applying a Fourier Transform*. This space is called the *frequency space*.

Definition 2.2.7. *The **Fourier Transform** over $\mathscr{S}'(\mathbb{R}^d)$ is defined as the adjoint operator of the Fourier Transform over $\mathscr{S}(\mathbb{R}^d)$. Explicitly,*

$$\langle \mathscr{F}(T), \varphi \rangle := \langle T, \mathscr{F}(\varphi) \rangle, \quad \forall T \in \mathscr{S}'(\mathbb{R}^d), \varphi \in \mathscr{S}(\mathbb{R}^d). \tag{2.64}$$

This definition is inspired by the *transfer formula*, which states that for $\varphi, \phi \in L^2(\mathbb{R}^d)$ it holds that $\langle \mathscr{F}(\varphi), \phi \rangle = \langle \varphi, \mathscr{F}(\phi) \rangle$. This allows then to define the Fourier Transform for a quite large class of objects, such as polynomially bounded functions, not necessarily integrable or in $L^2(\mathbb{R}^d)$, and for slow-growing measures, not necessarily finite. We give as typical examples, $\mathscr{F}(\mathbf{1}) = (2\pi)^{\frac{d}{2}}\delta$ and, with more generality, $\mathscr{F}(x^\alpha) = (2\pi)^{\frac{d}{2}} i^{|\alpha|} D^\alpha \delta$, with $\alpha \in \mathbb{N}^d$. We remark that if $T \in \mathscr{S}'(\mathbb{R}^d)$, then $\mathscr{F}(T)$ is real if and only if T is Hermitian.

The classical property of the Fourier Transform of a derivative also holds for tempered distributions, and with more interest than in the classical case since now all objects are differentiable. Thus, $\mathscr{F}(D^\alpha T) = (i\xi)^\alpha \mathscr{F}(T)$ for all $T \in \mathscr{S}'(\mathbb{R}^d)$ and for all $\alpha \in \mathbb{N}^d$. Here we have denoted by $(i\xi)^\alpha$ the function $\xi \in \mathscr{S}(\mathbb{R}^d) \mapsto (i\xi)^\alpha$ which is in $\mathcal{O}_M(\mathbb{R}^d)$.

Another important property of the Fourier Transform is the **Exchange Formula** between the convolution and the multiplication. For two test-functions $\varphi, \phi \in \mathscr{S}(\mathbb{R}^d)$, it holds that $\mathscr{F}(\varphi * \phi) = (2\pi)^{\frac{d}{2}} \mathscr{F}(\varphi)\mathscr{F}(\phi)$ (Donoghue, 1969, Chapter 29). For distributions it is a little bit more restrictive since we cannot multiply or convolute arbitrary tempered distributions at will, but it holds for the cases seen in this section. It can be proven that the Fourier Transform is a bijective linear operator from $\mathcal{O}_M(\mathbb{R}^d)$ to $\mathcal{O}'_c(\mathbb{R}^d)$: the Fourier Transform exchanges the space of multiplicators with the space of convoluters (Trèves, 1967, Theorem 30.3). If $T \in \mathscr{S}'(\mathbb{R}^d)$, $S \in \mathcal{O}'_c(\mathbb{R}^d)$ and $f \in \mathcal{O}_M(\mathbb{R}^d)$, then

$$\mathscr{F}(T * S) = (2\pi)^{\frac{d}{2}} \mathscr{F}(T)\mathscr{F}(S) \quad ; \quad \mathscr{F}(fT) = (2\pi)^{-\frac{d}{2}} \mathscr{F}(T) * \mathscr{F}(f). \tag{2.65}$$

Some generalizations of this formula will be pointed-out in section 2.2.3.

We finally recall an important result of Fourier Analysis in the classical case of functions in $L^1(\mathbb{R}^d)$.

Theorem 2.2.1 (Riemann-Lebesgue Lemma). *Let $f : \mathbb{R}^d \to \mathbb{C}$ be a function in $L^1(\mathbb{R}^d)$. Then, $\mathscr{F}(f) \in C_0(\mathbb{R}^d)$.*

See Donoghue (1969, Chapter 30) for a proof. Obviously, this also applies when considering the Inverse Fourier Transform $\mathscr{F}^{-1}$.

2.2.3 Tensor products

In this section we recall the definitions of tensor products of functions, measures, distributions and linear operators over $\mathscr{S}(\mathbb{R}^d)$ and $\mathscr{S}'(\mathbb{R}^d)$. The latter case is the only case which is not easily findable in the literature in the way we state it here. For the rest, we just follow usual terminologies and results.

Tensor product of Functions and Measures

Let $f : \mathbb{R}^d \to \mathbb{C}$ and $g : \mathbb{R}^m \to \mathbb{C}$ be two functions. The **tensor product between the functions** f and g is the function $(f \otimes g) : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{C}$ defined through $(f \otimes g)(x, y) = f(x)g(y)$ for all $x \in \mathbb{R}^d$ and $y \in \mathbb{R}^m$.

If $\mu \in \mathscr{M}(\mathbb{R}^d)$ and $\nu \in \mathscr{M}(\mathbb{R}^m)$, the **tensor product between the measures** μ and ν is a measure over $\mathbb{R}^d \times \mathbb{R}^m$, denoted by $\mu \otimes \nu$, which satisfies that $(\mu \otimes \nu)(A \times B) = \mu(A)\nu(B)$ for every $A \in \mathcal{B}_B(\mathbb{R}^d)$ and $B \in \mathcal{B}_B(\mathbb{R}^m)$. A typical result from Measure Theory guarantees that there is a unique measure $\mu \otimes \nu \in \mathscr{M}(\mathbb{R}^d \times \mathbb{R}^m)$ satisfying this condition. The next Proposition also holds.

Proposition 2.2.1. *If $\mu \in \mathscr{M}(\mathbb{R}^d)$ and $\nu \in \mathscr{M}(\mathbb{R}^m)$, then $|\mu \otimes \nu| = |\mu| \otimes |\nu|$.*

A proof of this Proposition is presented in Appendix A.4.1. From this Proposition it is straightforward that the tensor product between finite (respectively, slow-growing, respectively compactly supported) measures is a finite (respectively, slow-growing, respectively compactly supported) measure.

Tensor product of distributions

Let $T \in \mathscr{S}'(\mathbb{R}^d)$ and $S \in \mathscr{S}'(\mathbb{R}^m)$. The **tensor product between the distributions** T and S is the tempered distribution $T \otimes S \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^m)$ defined through

$$\langle T \otimes S, \psi \rangle := \langle T, x \mapsto \langle S, \psi(x, \cdot)\rangle\rangle, \quad \forall \psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m). \tag{2.66}$$

It can be proven that this definition determines a unique tempered distribution in $\mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^m)$ (Richards & Youn, 1995, Theorem A in Chapter 7). In addition, the following "*Fubini's Theorem*" holds:

$$\langle S, y \mapsto \langle T, \psi(\cdot, y)\rangle\rangle = \langle T \otimes S, \psi\rangle = \langle T, x \mapsto \langle S, \psi(x, \cdot)\rangle\rangle. \tag{2.67}$$

In Schwartz (1966, Chapter IV) this result is presented in the case of generic distributions. We remark that if $\psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m)$ is of the form $\psi = \varphi \otimes \phi$, with $\varphi \in \mathscr{S}(\mathbb{R}^d)$ and $\phi \in \mathscr{S}(\mathbb{R}^m)$, then the tensor product satisfies $\langle T \otimes S, \varphi \otimes \phi\rangle = \langle T, \varphi\rangle\langle S, \phi\rangle$.

Let us make a *special* comment about this product: it can be used to give a more general definition of the convolution and the multiplicative product of distributions than the one presented in Section 2.2.2. We will detail the case of the convolution since it is not necessary to introduce new definitions to do it. The definition of the convolution between two generic distributions, as presented in Schwartz (1966, Chapter VI) is based on the idea of tensor products. For the specific case of multiplications between tempered distributions, we suggest the more recent and didactic Richards & Youn (1995, Chapter 7).

In the case of functions, one can show that, if $f \in L^1(\mathbb{R}^d)$ and $g \in L^1(\mathbb{R}^d)$ for instance, then

$$\int_{\mathbb{R}^d} (f * g)(x)\varphi(x)dx = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(x)g(y)\varphi(x+y)dxdy, \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^d). \tag{2.68}$$

This inspires a criterion of convolvability of two arbitrary tempered distributions. Two tempered distributions $S, T \in \mathscr{S}'(\mathbb{R}^d)$ are said to be $\mathscr{S}'-$**convolvable** (Dierolf & Voigt, 1978) if the application

$$\varphi \in \mathscr{S}(\mathbb{R}^d) \mapsto \langle T \otimes S, (x,y) \mapsto \varphi(x+y)\rangle \tag{2.69}$$

defines a tempered distribution. In that case, this distribution is called the **convolution** between T and S. In Richards & Youn (1995, Chapter 7) a general description about the convolvability of two tempered distributions is done. In this same source, the tensor product is used to define a multiplication criterion between two tempered distributions, based on the idea of *localization*. This concept uses similar notions to those considered in the Example 2.1.2 of measures supported on the hyperplane $\{y = x\}$, applied to the case of distributions. Within this framework, the authors prove that the *multiplicability* of two tempered distributions is a necessary and sufficient condition for the convolvability of their Fourier Transforms, and an Exchange Formula for the Fourier Transform, similar to (2.65), is stated in the case where convolvability and *multiplicativity* are satisfied (Richards & Youn, 1995, Theorem 7.6). See Richards & Youn (2000) for a briefer exposition about this theory of localization, multiplication and convolution of distributions, and the associated Exchange Formula for the Fourier Transform.

Tensor product of linear operators

We consider now two linear and continuous operators defined through an adjoint, $\mathcal{L}_1^* : \mathscr{S}'(\mathbb{R}^d) \to \mathscr{S}'(\mathbb{R}^d)$ and $\mathcal{L}_2^* : \mathscr{S}'(\mathbb{R}^m) \to \mathscr{S}'(\mathbb{R}^m)$. The objective of this section is to present the definition of the tensor product between $\mathcal{L}_1^*$ and $\mathcal{L}_2^*$, denoted by $\mathcal{L}_1^* \otimes \mathcal{L}_2^*$, which is a linear and continuous mapping from $\mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^m)$ to $\mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^m)$. The intuitive meaning of this operator is that it applies the operator $\mathcal{L}_1^*$ to the *first component* of its argument and the operator $\mathcal{L}_2^*$ to the *second* one.

We will first consider the definition of the tensor product of linear operators on the space of test-functions. We denote by $\mathcal{I}_d$ the identity operator from $\mathscr{S}(\mathbb{R}^d)$ to $\mathscr{S}(\mathbb{R}^d)$ and $\mathcal{I}_m$ its analogue for $\mathbb{R}^m$.

Definition 2.2.8. *Let $\mathcal{L}_1 : \mathscr{S}(\mathbb{R}^d) \to \mathscr{S}(\mathbb{R}^d)$ be linear and continuous. We define the mapping $\mathcal{L}_1 \otimes \mathcal{I}_m : \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m) \to \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m)$ as*

$$\mathcal{L}_1 \otimes \mathcal{I}_m(\psi) := (x,y) \mapsto \mathcal{L}_1(\psi(\cdot, y))(x), \quad \forall \psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m). \tag{2.70}$$

Let us clarify how this operator acts on a test-function $\psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m)$. For a function ψ and for a fixed

$y \in \mathbb{R}^m$, the function $\psi(\cdot, y)$ is clearly in $\mathscr{S}(\mathbb{R}^d)$. The operator $\mathcal{L}_1$ is applied to that function, and the result is evaluated at x. Hence, this operator follows the intuition of applying $\mathcal{L}_1$ to the *first* component of ψ, while doing nothing to the *second* one. It is not obvious that this procedure defines a function in $\mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m)$, nor that the operation is linear and continuous.

Proposition 2.2.2. *The operator $\mathcal{L}_1 \otimes \mathcal{I}_m : \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m) \to \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m)$ is well-defined and it is linear and continuous.*

Although Proposition 2.2.2 may seem intuitive, its proof is not easily findable in the literature without entering into deep and sophisticated theories of tensor products and Kernels. We give in Appendix A.4.2 a proof of this Proposition which is more accessible given the notions introduced in this work, but also probably longer than what it could be if we were using more sophisticated theories. We remark that the main difficulty is rather topological than algebraical. The definition of the operator $\mathcal{L}_1 \otimes \mathcal{I}_m$ over the space $\mathscr{S}(\mathbb{R}^d) \otimes \mathscr{S}(\mathbb{R}^m)$ of finite linear combinations of tensor products between functions in $\mathscr{S}(\mathbb{R}^d)$ and in $\mathscr{S}(\mathbb{R}^m)$ is straightforward. A density argument could guarantee that the definition can be extended to the whole space $\mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m)$, but for this the continuity of the operator must be verified.

If $\mathcal{L}_2 : \mathscr{S}(\mathbb{R}^m) \to \mathscr{S}(\mathbb{R}^m)$ is linear and continuous, the construction of the operator $\mathcal{I}_d \otimes \mathcal{L}_2 : \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m) \to \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m)$ is done similarly to Definition 2.2.8. We thus define $\mathcal{L}_1 \otimes \mathcal{L}_2$ easily.

Definition 2.2.9. *Let $\mathcal{L}_1 : \mathscr{S}(\mathbb{R}^d) \to \mathscr{S}(\mathbb{R}^d)$ and $\mathcal{L}_2 : \mathscr{S}(\mathbb{R}^m) \to \mathscr{S}(\mathbb{R}^m)$ be linear and continuous operators. We define the* ***tensor product*** *between $\mathcal{L}_1$ and $\mathcal{L}_2$ as the linear and continuous operator $\mathcal{L}_1 \otimes \mathcal{L}_2 : \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m) \to \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m)$ defined through*

$$(\mathcal{L}_1 \otimes \mathcal{L}_2)(\psi) = (\mathcal{I}_d \otimes \mathcal{L}_2)\left((\mathcal{L}_1 \otimes \mathcal{I}_m)(\psi)\right), \quad \psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m). \tag{2.71}$$

Hence, $\mathcal{L}_1 \otimes \mathcal{L}_2$ is just the composition between $\mathcal{I}_d \otimes \mathcal{L}_2$ and $\mathcal{L}_1 \otimes \mathcal{I}_m$. It is clear that for two test-functions $\varphi \in \mathscr{S}(\mathbb{R}^d)$ and $\phi \in \mathscr{S}(\mathbb{R}^m)$ it holds that $(\mathcal{L}_1 \otimes \mathcal{L}_2)(\varphi \otimes \phi) = \mathcal{L}_1\varphi \otimes \mathcal{L}_2\phi$. An argument using the density of $\mathscr{S}(\mathbb{R}^d) \otimes \mathscr{S}(\mathbb{R}^m)$ in $\mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m)$ guarantees[11] that Definition 2.2.9 does not depend on the order in which the composition is done, that is, $\mathcal{L}_1 \otimes \mathcal{L}_2 = (\mathcal{L}_1 \otimes \mathcal{I}_m) \circ (\mathcal{I}_d \otimes \mathcal{L}_2) = (\mathcal{I}_d \otimes \mathcal{L}_2) \circ (\mathcal{L}_1 \otimes \mathcal{I}_m)$.

From this the way of defining the tensor product of two operators defined through an adjoint, $\mathcal{L}_1^* : \mathscr{S}'(\mathbb{R}^d) \to \mathscr{S}'(\mathbb{R}^d)$ and $\mathcal{L}_2^* : \mathscr{S}'(\mathbb{R}^m) \to \mathscr{S}'(\mathbb{R}^m)$, is quite intuitive.

Definition 2.2.10. *Let $\mathcal{L}_1 : \mathscr{S}(\mathbb{R}^d) \to \mathscr{S}(\mathbb{R}^d)$ and $\mathcal{L}_2 : \mathscr{S}(\mathbb{R}^m) \to \mathscr{S}(\mathbb{R}^m)$ be two continuous linear operators. Let $\mathcal{L}_1^*$ and $\mathcal{L}_2^*$ be the adjoint operators of $\mathcal{L}_1$ and $\mathcal{L}_2$ respectively. The tensor product between*

[11] An argument using the development of functions in $\mathscr{S}(\mathbb{R}^d)$ on the base of Hermite functions can be used to conclude the density of $\mathscr{S}(\mathbb{R}^d) \otimes \mathscr{S}(\mathbb{R}^m)$ in $\mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m)$. See Reed & Simon (1980, Appendix to V.3). An analogue result also holds for the space $\mathscr{S}'(\mathbb{R}^d) \otimes \mathscr{S}'(\mathbb{R}^m)$, which is sequentially dense in the space $\mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^m)$ with the weak-$*$ topology.

$\mathcal{L}_1^*$ and $\mathcal{L}_2^*$ is defined as the adjoint of the operator $\mathcal{L}_1 \otimes \mathcal{L}_2$:

$$\mathcal{L}_1^* \otimes \mathcal{L}_2^* := (\mathcal{L}_1 \otimes \mathcal{L}_2)^*. \tag{2.72}$$

This definition, although intuitive, is a little bit artificial since we have avoided the notion of defining the operators acting on each component of its argument, as we did in Definition 2.2.8. Nevertheless, it can be proven that such a definition would be equivalent. This can be seen by considering that for tempered distributions in $\mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^m)$ of the form $T \otimes S$, with $T \in \mathscr{S}'(\mathbb{R}^d)$ and $S \in \mathscr{S}'(\mathbb{R}^m)$, the operator $\mathcal{L}_1^* \otimes \mathcal{L}_2^*$ does simply $(\mathcal{L}_1^* \otimes \mathcal{L}_2^*)(T \otimes S) = \mathcal{L}_1^* T \otimes \mathcal{L}_2^* S$. The equivalence can thus be concluded using density arguments[12]. We will omit those details and we will just work with Definition 2.2.10.

2.2.4 Nuclear Theorem

We present here a particular Theorem which is valid for the space of tempered distributions and which has important consequences on the theory of Generalized Stochastic Processes.

Let $K : \mathscr{S}(\mathbb{R}^d) \times \mathscr{S}(\mathbb{R}^m) \to \mathbb{C}$ be a bilinear form. We say that K is **separately continuous** if for every $\varphi \in \mathscr{S}(\mathbb{R}^d)$, the mapping $\phi \in \mathscr{S}(\mathbb{R}^m) \mapsto K(\varphi, \phi)$ is continuous and hence an element of $\mathscr{S}'(\mathbb{R}^m)$, and in a symmetric way, if for every $\phi \in \mathscr{S}(\mathbb{R}^m)$ the mapping $\varphi \in \mathscr{S}(\mathbb{R}^d) \mapsto K(\varphi, \phi)$ is in $\mathscr{S}'(\mathbb{R}^d)$.

Theorem 2.2.2 (Nuclear Theorem for tempered distributions). *Let $K : \mathscr{S}(\mathbb{R}^d) \times \mathscr{S}(\mathbb{R}^m) \to \mathbb{C}$ be a separately continuous bilinear form. Then, there exists a unique tempered distribution $T \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^m)$ such that*

$$\langle T, \varphi \otimes \phi \rangle = K(\varphi, \phi), \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^d), \phi \in \mathscr{S}(\mathbb{R}^m). \tag{2.73}$$

We suggest Reed & Simon (1980, Theorem V.12) and the comments in Appendix to section V.3 of this source for a proof. A general description of spaces which satisfy a similar result, namely, *Nuclear spaces*, can be found in Trèves (1967, Part III).

2.2.5 Comments on other spaces of distributions

The Theory of Distributions has been developed using a space of test-functions more restrictive than the Schwartz space. It is the space of compactly supported smooth functions over $\mathbb{R}^d$, denoted by $\mathscr{D}(\mathbb{R}^d)$. This space is endowed with a suitable topology which makes it to be a complete Hausdorff locally convex topological vector space, similarly to the case of the space $C_c(\mathbb{R}^d)$. Its dual $\mathscr{D}'(\mathbb{R}^d)$ is the space of *distributions*, in the generic sense, and the Theory of Distributions in the broadest sense uses this space. Here, every

[12] Cf. footnote 11.

continuous function is identified with a member of $\mathscr{D}'(\mathbb{R}^d)$, not just the polynomially bounded ones. Analogously, every measure in $\mathscr{M}(\mathbb{R}^d)$, being or not slow-growing, determines a member in $\mathscr{D}'(\mathbb{R}^d)$. Differential operators as defined in Section 2.2.2 are also defined for any distribution in $\mathscr{D}'(\mathbb{R}^d)$ and hence, any continuous function and any complex measure over $\mathbb{R}^d$ can be differentiated any number of times. Multiplication with any smooth function, not necessarily a member of the space $\mathcal{O}_M(\mathbb{R}^d)$ is also possible. The convolution is more restricted, being the space of so-called *distributions with compact support*[13] the only space of distributions whose elements can be convoluted with *any* distribution in $\mathscr{D}'(\mathbb{R}^d)$. Tensor products and thus convolutions in more general cases, as presented in section 2.2.3 are also immediately extended to this space. The Nuclear Theorem 2.2.2 also holds for this space, result which is known as the *Schwartz's Kernel Theorem* (Trèves, 1967, Theorem 51.7). The Fourier Transform, however, is just defined for tempered distributions, and hence all Fourier Analysis techniques to solve and treat PDEs are restricted, in principle, to the tempered framework.

Another commonly used space is the space of smooth functions denoted by $\mathscr{E}(\mathbb{R}^d)$, which is actually nothing but the space $C^\infty(\mathbb{R}^d)$. The notation $\mathscr{E}(\mathbb{R}^d)$ is used when it is understood that the space is endowed with the topology of uniform convergence on compact sets of the derivatives, similarly to the case of the space $C(\mathbb{R}^d)$ described in Section 2.1.4 but with smooth functions. Its dual $\mathscr{E}'(\mathbb{R}^d)$ is actually the space of distributions with compact support, and it is a subspace of the space of tempered distributions. A Nuclear Theorem analogue to Theorem 2.2.2 can also be obtained (Trèves, 1967, Theorem 51.6 and Corollary). We summarize the inclusion relationships between spaces of test-functions and distributions:

$$\mathscr{D}(\mathbb{R}^d) \subset \mathscr{S}(\mathbb{R}^d) \subset \mathscr{E}(\mathbb{R}^d) \text{“} \subset \text{”} \mathscr{E}'(\mathbb{R}^d) \subset \mathscr{S}'(\mathbb{R}^d) \subset \mathscr{D}'(\mathbb{R}^d). \tag{2.74}$$

Other inclusions which are useful to retain are the ones related to spaces of measures:

$$\begin{aligned} \mathscr{D}(\mathbb{R}^d) \subset C_c(\mathbb{R}^d) \text{“} \subset \text{”} C_c'(\mathbb{R}^d) = \mathscr{M}(\mathbb{R}^d) \subset \mathscr{D}'(\mathbb{R}^d), \\ \mathscr{S}(\mathbb{R}^d) \subset C_{FD}(\mathbb{R}^d) \text{“} \subset \text{”} C_{FD}'(\mathbb{R}^d) = \mathscr{M}_{SG}(\mathbb{R}^d) \subset \mathscr{S}'(\mathbb{R}^d). \end{aligned} \tag{2.75}$$

And the simplest but maybe the most important one for many applications:

$$\mathscr{S}(\mathbb{R}^d) \subset L^2(\mathbb{R}^d) \subset \mathscr{S}'(\mathbb{R}^d). \tag{2.76}$$

We claim that all of these inclusions are *dense* when the corresponding spaces are endowed with a suitable topology (see Proposition C.1.1 in Appendix C for the case of $\mathscr{S}(\mathbb{R}^d) \subset C_{FD}(\mathbb{R}^d)$, the other cases are well-known in Distribution Theory).

[13] The support of a distribution can be defined analogously to the definition of support of a measure (2.16), using test-functions supported on open sets. See Donoghue (1969, Chapter 29).

Chapter 3

Theoretical Framework: Stochastic Tools

SUMMARY

In this chapter we present the stochastic tools that will be used in the next chapters. It is basically a framework where Random Functions, Random Measures and Random Distributions can be defined starting from specified mean and covariance structures. We explain how to do Stochastic Calculus with these objects and how to pose and analyse some linear SPDEs. All this framework is embedded in the so-called mean-square analysis, where the random variables are supposed to be square-integrable and the convergences are considered in a mean-square sense.

In Section 3.2 we recall the classical geostatistical framework of square-integrable Random Functions. We recall the concepts of mean and covariance functions. We recall the concepts of stationarity, positive-definite functions and spectral measures. We study the mean-square regularity of Random Functions, and we present the definitions of continuity, differentiability and integrability with respect to deterministic measures.

In Section 3.3 we present our concept of Random Measure. We define it as a stochastic process indexed by the bounded Borel sets being determined by mean and covariance measures. We present the construction of the stochastic integral of deterministic functions with respect to Random Measures. We give our definitions of finite, slow-growing and compactly supported Random Measures. We present the interpretation of Random Measures as linear functionals over spaces of continuous functions, following an analogy to Riesz Representation Theorems in the deterministic case. We also present the notion of a Random Measure concentrated on a subset. We finally introduce the class of orthogonal Random Measures and we recall its relationship with stationary Random Functions.

In Section 3.4 we present the theory of Generalized Random Fields or Random Distributions,

which is the stochastic analogue of the Theory of Distributions. We define them as stochastic processes indexed by the Schwartz space determined by mean and covariance distributions. We show how to apply linear operators defined through an adjoint over Generalized Random Fields, and we explain how this operators modify the mean and covariance structures. We recall the definition of a stationary Generalized Random Field, its main properties and its relationship with slow-growing orthogonal Random Measures.

In Section 3.5 we present the class of SPDEs which will be considered in the next chapters. We always consider SPDEs defined through a deterministic operator. We make the distinction between satisfying a SPDE strictly, in law, and in a second-order sense. We specify how a linear SPDE determines the mean and covariance structure of a model, translated on linear PDEs to be satisfied by the mean and the covariance.

In Section 3.6 we give comments about how the SPDE framework can be used in multivariate Geostatistics, particularly in the case of bivariate models. We give a brief overview of the concept of a bivariate model in a generalized sense, introducing the cross-covariance. We show the relationships that two processes must satisfy in order to be equal in a convenient sense, and we present how to apply this condition when the variables in the model are related through a SPDE. We give a necessary and sufficient condition for a general linear SPDE to be satisfied strictly, which is stated through PDEs that the means, the covariances and the cross-covariance must satisfy.

In Section 3.7 we make some comments about other frameworks of Stochastic Analysis beyond the mean-square theory. We present the issues involved when trying to define Random Functions and Random Measures in a stricter sense than the mean-square sense. We remark the impossibility of defining convenient orthogonal Random Measures. We remark that, contrarily to the case of Random Functions and Random Measures, the case of Random Distributions can be worked out in a stricter framework without problems, due to the Bochner-Minlos Theorem, applicable to the case of tempered distributions since the Schwartz space is Nuclear.

We end in Section 3.8 with some comments about stochastic integrals of Random Functions with respect to Random Measures and how some non-linear SPDEs can be defined through them. This framework is not used in the rest of this dissertation. We present the classical Itô Integral and we explain the typical issue of the non-canonical way of defining a stochastic integral. We show that this issue is related to the structure of the cross-covariance Kernel between the Random Function to be integrated and the reference Random Measure. We show why these notions are important in order to pose non-linear SPDEs or SPDEs with multiplicative noise and we give some examples of such SPDEs, together with a brief explanation of their issues and some related theories already developed in the literature in order to treat them.

3.1 General introduction

In this chapter we introduce the stochastic objects which we will work with along this dissertation. We keep in mind the following *almost correct* idea: what has been defined in the deterministic world, has its counterpart in the stochastic world. Hence, we will be able to work with Random Functions, Random Measures and Random Distributions, all of them defined through a mean and a covariance structure. The intuition in mind is that a SPDE is just a *PDE with random objects involved*, and doing Stochastic Calculus is simply *doing calculus with Random Functions* and their *generalizations*. Reasons about why we have used the "*almost correct*" expression in this paragraph are commented in Section 3.7.

We keep our loyalty to the tradition of the Fontainbleau school. Hence, *we do not specify the laws of the random variables involved*[1]. Indeed, the developments in this chapter are done, in principle, without any regard to the laws of the random variables involved, besides the fact that we require them to be square-integrable. This *tradition* is based on the historical development of Geostatistics, which has been developed as a framework to deal with *unique phenomena*. We will sometimes, however, make references to the framework of Gaussian processes.

From now on, we will always work with a fixed arbitrary probability space $(\Omega, \mathcal{A}, \mathbb{P})$. All random variables we use are supposed to be defined over this space. We will work with complex stochastic processes, that is, families of complex random variables indexed by a non-empty set T, $(X_t)_{t\in T}$ with some characteristics to be described. The existence of such a mathematical entity is guaranteed by Kolmogorov's Theorem, which provides an enough general mathematical framework to work with. We refer to Appendix B for a statement of this Theorem and its application to construct the objects introduced in this chapter.

None of the results presented in this chapter is essentially new. We give proofs for some of them which may be difficult to find in the literature in the way we state them. Which may be considered as *new*, is the *compendium* of different applications of the *mean-square* approach to Stochastic Calculus that can be found in the literature. Here we focus our exposition in a mean-square based framework in order to simplify its potential application in geostatistical analysis.

3.2 Random Functions: the classical geostatistical framework

A real **Random Function** over $\mathbb{R}^d$, also called a *Stochastic Process* indexed by $\mathbb{R}^d$, is a family of real random variables indexed by the Euclidean space $(Z(x))_{x\in\mathbb{R}^d}$. We suppose that all of the random variables are square-integrable: $Z(x) \in L^2(\Omega, \mathcal{A}, \mathbb{P})$ for all $x \in \mathbb{R}^d$. In such a case, the mean and covariance structures of Z can be studied. The **mean function** is the function $m_Z : \mathbb{R}^d \to \mathbb{R}$ defined by $m_Z(x) = \mathbb{E}(Z(x))$. The **covariance function** is the function $C_Z : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ defined by $C_Z(x, y) = \mathbb{C}ov(Z(x), Z(y))$. The

[1] **Note of the author:** I am not only an apprentice of the *Bellifontaine* school, I am also a *Chilean*: I *cannot* care about *the law*.

covariance function must be a **positive-definite Kernel**: a function (of two-variables) $C : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is said to be a positive-definite Kernel if for every finite collections of complex numbers and points in the space, $(\lambda_1, ..., \lambda_N) \in \mathbb{C}^N$ and $(x_1, ..., x_N) \subset (\mathbb{R}^d)^N$, with $N \in \mathbb{N}_*$, it holds that

$$\sum_{j,k=1}^{N} \lambda_j C(x_j, x_k)\overline{\lambda_k} \geqslant 0. \tag{3.1}$$

Every positive-definite Kernel can be used as a covariance function of a square-integrable stochastic process. This is a consequence of Kolmogorov's Theorem; see Appendix B. The definition of C_Z implies that $C_Z(x, x) \geqslant 0$ for all $x \in \mathbb{R}^d$ and that it is symmetric: $C_Z(x, y) = C_Z(y, x)$. The Cauchy-Schwarz inequality implies that C_Z satisfies $|C_Z(x, y)| \leqslant \sqrt{C_Z(x, x)C_Z(y, y)}$.

The typical methodology in geostatistical analysis of unique phenomena consists in supposing that a variable varying spatially or spatio-temporally is the realisation of a Random Function for which its covariance and mean functions are to be selected to model the phenomenon. The mean is often selected to be null or constant, although any arbitrary function can be used as a mean function of a Random Function. The choice of the covariance function is more determinant and intricate since it has to satisfy the Kernel positive-definiteness condition (3.1). Geostatisticians usually work with a sort of *catalogue* of well-known positive-definite Kernel functions that can be used as covariance functions and for which their properties are well studied and mastered. The selection of the covariance function, usually refereed in a broader sense as *selection of the model*, is done by diverse techniques of statistical inference such as variographic analysis or likelihood maximisation. Typical treaties on this practice are Chilès & Delfiner (1999) and Wackernagel (2003).

3.2.1 Stationarity

Let $Z = (Z(x))_{x \in \mathbb{R}^d}$ be a real square integrable Random Function. We say that Z is **second order stationary** (from now on, simply **stationary**) if its mean function is constant and if its covariance function depends only on the gap $x - y$ of the variables. Hence, $m_Z(x) = m_Z \in \mathbb{R}$, and there exists a function $\rho_Z : \mathbb{R}^d \to \mathbb{R}$ such that $C_Z(x, y) = \rho_Z(x - y)$ for all $x, y \in \mathbb{R}^d$. The stationarity condition is equivalent to require that the mean function m_Z is invariant under translations and that the covariance function C_Z is invariant under equal translations on both of its components. Explicitly, $m_Z(x) = m_Z(x + h)$ and $C_Z(x, y) = C_Z(x + h, y + h)$ for all $x, y, h \in \mathbb{R}^d$.[2] When Z is stationary, its associated function ρ_Z is called the **stationary covariance function** or simply its *covariance function* if stationarity is clear in context and it is not mistaken with C_Z.

[2]This is not the typical and strict definition of *stationarity* of Random Functions. In a more traditional terminology, a Random Function is said to be stationary or *strictly stationary*, if its finite-dimensional laws are invariant under translations. Hence, it is not only the first and second order structures that are invariant under translations but the whole probability law. Since we work with the covariance as the main tool, we will always use the concept of stationarity only referring to its *second order* sense.

The function ρ_Z must be a **positive-definite function**[3], also called a **function of positive-type**. This means that ρ_Z satisfies that for every finite family of complex numbers and points in the space, $(\lambda_1, ..., \lambda_N) \in \mathbb{C}^N$ and $(x_1, ..., x_N) \in (\mathbb{R}^d)^N$ with $N \in \mathbb{N}_*$, it holds that[4]

$$\sum_{j,k=1}^{N} \lambda_j \rho_Z(x_j - x_k)\overline{\lambda_k} \geqslant 0. \tag{3.2}$$

This implies that $\rho_Z(0) \geqslant 0$ and that ρ_Z is even: $\rho_Z(h) = \rho_Z(-h)$ for all $h \in \mathbb{R}^d$. The Cauchy-Schwarz inequality implies that $|\rho(h)| \leqslant \rho(0)$ for all $h \in \mathbb{R}^d$.

A practical description of real and continuous positive-definite functions is given by the well-known **Bochner's Theorem:** *a real*[5] *and continuous function $\rho : \mathbb{R}^d \to \mathbb{R}$ is positive-definite if and only if it is the Fourier Transform of a positive even and finite measure $\mu \in \mathscr{M}_F^+(\mathbb{R}^d)$:*

$$\rho(h) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{-ih^T\xi} d\mu(\xi). \tag{3.3}$$

If Z is a stationary Random Function and if ρ_Z is its covariance distribution, the measure μ_Z that satisfies $\rho_Z = \mathscr{F}(\mu_Z)$ is called the **spectral measure** of Z. Bochner's Theorem has a crucial importance in both Probability Theory and Geostatistics. Probabilists use it (in a more general version) to show the existence of convenient probability measures over abstract spaces, starting from well-defined positive-definite functions over those spaces. Geostatisticians use it mainly to obtain new covariance models to add to their *catalogue*, since defining a finite measure over the Euclidean space is not a complicated task to do, while the direct construction of different classes of positive-definite functions may be intricate. In this work we focus on the geostatistician application of this theorem. See Donoghue (1969, Chapter 37) for an exposition and proof of this theorem.

Stationary Random Functions are basic tools for the geostatistical modelling since they provide a suitable statistical methodology for unique phenomena: even if we suppose that we analyse a single realisation of the Random Function, stationarity guarantees that the behaviour of the variable will be similar in every part of the analysed region, and thus it provides, intuitively, an analogue to the *independent and identically*

[3]The reader must not confuse a *positive-definite function* with a *positive-definite Kernel*. This terminological distinction is crucial when remarking the difference between a stationary covariance model and a non-stationary one. The term *positive-definite Kernel* designates always a *function of two variables*, like the covariance function C_Z. The term *positive-definite*, without the mention "*Kernel*", always refers to a function of *one variable*, like ρ_Z, used to describe stationary models. This terminology selection is the most appropriated given its current use in the literature. We will use an analogue terminology for the cases of Random Measures and Generalized Random Fields

[4]Some authors require the function ρ_Z to satisfy that the quadratic form (3.2) is strictly positive, unless the complex numbers involved are all null. We do not require that stronger condition. If the function ρ_Z satisfies this stronger requirement, it will be said to be *strictly positive-definite*.

[5]Bochner's Theorem does not really require the function to be real but complex, and thus the associated measure needs not to be even. Of course, it must remains positive and finite.

distributed random variables framework of classical statistical methodologies.

We give four important examples of stationary covariance models over $\mathbb{R}^d$ which are widely used in the practice of Geostatistics. We present them together with their associated spectral measures. All of them are presented in their *isotropic version*, that is, the covariance function ρ_Z depends only on the Euclidean norm of its argument.

- The **Gaussian model**:

$$\rho_Z(h) = \sigma^2 e^{-|\frac{h}{a}|^2} \quad ; \quad d\mu_Z(\xi) = \frac{\sigma^2 a^d}{2^{\frac{d}{2}}} e^{-\frac{a^2}{4}|\xi|^2} d\xi, \tag{3.4}$$

  with $\sigma^2, a > 0$.

- The **Exponential model**:

$$\rho_Z(h) = \sigma^2 e^{-|\frac{h}{a}|} \quad ; \quad d\mu_Z(\xi) = \frac{\sigma^2 2^{\frac{d}{2}} \Gamma(\frac{d+1}{2})}{a\sqrt{\pi}} \frac{d\xi}{(\frac{1}{a^2} + |\xi|^2)^{\frac{d+1}{2}}}, \tag{3.5}$$

  with $\sigma^2, a > 0$. Γ denotes the Gamma function.

- The **Matérn model** (also called K-*Bessel model*):

$$\rho_Z(h) = \frac{1}{(2\pi)^{d/2} 2^{\alpha-1} \kappa^{2\alpha-d} \Gamma(\alpha) a} (\kappa|h|)^{\alpha-d/2} K_{\alpha-d/2}(\kappa|h|) \quad ; \quad d\mu_Z(\xi) = \frac{d\xi}{(2\pi)^{\frac{d}{2}} a(\kappa^2 + |\xi|^2)^{\alpha}}, \tag{3.6}$$

  with $\kappa, a > 0$ and $\alpha > \frac{d}{2}$. $K_{\alpha-\frac{d}{2}}$ denotes the modified Bessel function of the second kind of order $\alpha - \frac{d}{2} > 0$. This model is actually a generalization of the Exponential model.

- The $J-$**Bessel model**:

$$\rho_Z(h) = \frac{\sigma^2}{\kappa|h|^{\frac{d}{2}-1}} J_{\frac{d}{2}-1}(\kappa|h|) \quad ; \quad \mu_Z = \sigma^2 \mu_{unif}^{\partial B_\kappa^{(d)}(0)}, \tag{3.7}$$

  where $\sigma^2, \kappa > 0$ and $J_{d/2-1}$ denotes the Bessel function of the first kind of order $d/2 - 1$. We recall that $\mu_{unif}^{\partial B_\kappa^{(d)}(0)}$ denotes the uniform measure on the $d-1$-sphere of radius κ (see Example 2.1.1).

3.2.2 Regularity of Random Functions and operations

Within this framework of Random Functions it is already possible to enter into the domain of Stochastic Calculus and to consider some kinds of SPDEs. To do this, the stochastic analogues of continuity, integrals and derivatives must be specified. All of these concepts are defined in the deterministic world through

limits of things. In the stochastic framework, since there are many possible ways in which a sequence of random variables may converge, we need to specify the sense in which the limits are considered. Since the main tool of geostatistical analysis is the covariance function, the most adapted framework is the so-called *mean-square analysis*, where the limits are taken in the sense of $L^2(\Omega, \mathcal{A}, \mathbb{P})$. Here we are going to present the basic notions and concepts of this framework. This is not exactly the most traditional way of exposing Stochastic Calculus: usually a stronger or different mode of convergence is required, considering for example an almost-surely continuity for Random Functions, the existence of a *continuous modification* (typical requirement for Brownian Motion for instance, see Øksendal (2003, Definition 2.2.2)), or simply that the Random Functions have sample paths which *are always*, strictly speaking, continuous functions. We give comments on other approaches of Stochastic Calculus in Section 3.7.

A sufficient and simple exposition about the *mean-square analysis* can be found in Sobczyk (1991, Part II, Chapter 14 and Part III, Section 21.1). The exposition presented here follows similar principles. Other approaches to Stochastic Calculus can be also found in this same treaty.

Continuity and Integrals

Let us begin with the concept of **continuity**. Let $Z = (Z(x))_{x \in \mathbb{R}^d}$ be a Random Function. Z is said to be **mean-square continuous** or, more simply from now on, **continuous**, if both its mean and covariance functions are continuous. This definition is equivalent to the following one: Z is a continuous Random Function if for every convergent sequence of points in the space $(x_n)_{n \in \mathbb{N}} \subset \mathbb{R}^d$, $x_n \to x \in \mathbb{R}^d$, it holds that $Z(x_n) \overset{L^2(\Omega)}{\to} Z(x)$ (Sobczyk, 1991, Definition 2.2 and Theorem 2.2).

For a real continuous Random Function many mathematical operations are possible. Let us remark one: the classical Riemann Integral. We are going to present it in detail since the concepts presented here will be used further in this dissertation. Let $K \subset \mathbb{R}^d$ be a compact set. The Riemann Integral of a continuous function in the deterministic case is constructed using a limit of Riemann sums. We are going to make explicit the same procedure. Consider a family of non-empty Borel subsets of $\mathbb{R}^d$, $(V_j^N)_{j \in \{1,...,N\}, N \in \mathbb{N}_*}$, with the following properties:

- $\bigcup_{j=1}^N V_j^N = K$ for all $N \in \mathbb{N}_*$.
- $V_j^N \cap V_k^N = \varnothing$ for all $N \in \mathbb{N}_*$ and for all $j, k \in \{1, ..., N\}$ such that $j \neq k$.
- $\max_{j \in \{1,...,N\}} \operatorname{diam}(V_j^N) \to 0$ as $N \to \infty$.

Here $\operatorname{diam}(A)$ denotes the diameter of the set A. Hence, $(V_j^N)_{j \in \{1,...,N\}, N \in \mathbb{N}_*}$ is a class of subsets such that for a fixed N the collection $(V_j^N)_{j \in \{1,...,N\}}$ is a partition of K, and such that the *size* of every set in the partition converges to 0 as N grows. For simplicity, we will give a name to this kind of class of sets: it will

be called a **Riemann sequence of partitions of** K. The typical example when defining Riemann sums is by taking rectangles of smaller and smaller size covering K. Let us consider in addition, for every set V_j^N, an arbitrary point $x_j^N \in V_j^N$. The collection $(x_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$ will be said to be a collection of **tag points** of $(V_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$. We define the **integral of** Z **over** K as the limit in the sense of $L^2(\Omega,\mathcal{A},\mathbb{P})$:

$$\int_K Z(x)dx := \lim_{N\to\infty}\sum_{j=1}^{N}|V_j^N|Z(x_j^N), \tag{3.8}$$

where $|V_j^N| = \mathrm{Leb}(V_j^N)$ denotes the Lebesgue measure of the set V_j^N.

Rather than verifying if the integral (3.8) is well-defined, we will use this idea to define a more general integral of continuous Random Functions. Note that the definition of a Riemann sequence of partitions of K can be extended immediately to every $K \in \mathcal{B}_B(\mathbb{R}^d)$.

Definition 3.2.1. *Let Z be a real continuous Random Function over $\mathbb{R}^d$. Let $\mu \in \mathscr{M}(\mathbb{R}^d)$ and let $A \in \mathcal{B}_B(\mathbb{R}^d)$. The integral of Z with respect to μ over A is defined as the limit in $L^2(\Omega,\mathcal{A},\mathbb{P})$:*

$$\int_A Z(x)d\mu(x) := \lim_{N\to\infty}\sum_{j=1}^{N}\mu(V_j^N)Z(x_j^N), \tag{3.9}$$

where $(V_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$ is a Riemann sequence of partitions of A and $(x_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$ is a collection of tag points of $(V_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$.

The following result guarantees that the previous definition is consistent, and it describes some properties of the integral.

Proposition 3.2.1. *Let Z be a real continuous Random Function over $\mathbb{R}^d$, $\mu \in \mathscr{M}(\mathbb{R}^d)$ and $A \in \mathcal{B}_B(\mathbb{R}^d)$. Then, the integral $\int_A Z(x)d\mu(x)$ is well-defined as a random variable in $L^2(\Omega,\mathcal{A},\mathbb{P})$ and it does neither depend on the choice of the Riemann sequence of partitions of A nor on the choice of its tag points. Moreover, if $A, B \in \mathcal{B}_B(\mathbb{R}^d)$, and $\mu,\nu \in \mathscr{M}(\mathbb{R}^d)$, the following expressions for the mean and the covariance of the integral hold:*

$$\mathbb{E}\left(\int_A Z(x)d\mu(x)\right) = \int_A m_Z(x)d\mu(x), \tag{3.10}$$

$$\mathbb{C}ov\left(\int_A Z(x)d\mu(x), \int_B Z(x)d\nu(x)\right) = \int_A\int_B C_Z(x,y)d\mu(x)d\overline{\nu}(y). \tag{3.11}$$

A proof of this Proposition is presented in Appendix A.5. Notice that Definition 3.2.1, which is inspired by the classical construction of Riemann sums, is also related to the construction of the Lebesgue integral for a continuous function. Indeed, the sum in (3.9) can be interpreted as the Lebesgue integral of the *Random Simple Function* $Z_N(x) = \sum_{j=1}^{N} Z(x_j^N)\mathbf{1}_{V_j^N}(x)$ with respect to μ over A, and one can prove that $Z_N(x) \to$

$Z(x)$ in the sense of $L^2(\Omega, \mathcal{A}, \mathbb{P})$ for every $x \in A$. Hence, the integral can be interpreted also as a limit of integrals of simple functions converging, in a *mean-square-point-wise* sense to Z. Another, maybe more interesting way of interpreting this integral is, rather than approaching the function Z by simple functions, is by *approaching the measure* μ by a linear combination of punctual masses: $\mu_N = \sum_{j=1}^{N} \mu(V_j^N)\delta_{x_j^N}$. The sequence of measures $(\mu_N)_{N\in\mathbb{N}_*}$ converges *in some sense* to the measure μ. Therefore, we can expect that the limit of the integrals is the integral of the limit[6].

Extensions of Definition 3.2.1 to the case of not bounded sets are possible, thereby obtaining a notion of a continuous Random Function *integrable* with respect to μ. We will make precise this idea in Section 3.3.

Derivatives

Let Z be a real Random Function over $\mathbb{R}^d$. A broadly used notion of differentiability of Z is the next one: Z is **mean-square continuously differentiable** if its mean function is continuously differentiable and its covariance function is two-times continuously differentiable. It can be proven that this definition is equivalent to the next one: $Z = (Z(x))_{x\in\mathbb{R}^d}$ is mean-square continuously differentiable if for every $x \in \mathbb{R}^d$, for every normal vector $v \in \mathbb{R}^d$, and for every converging-to-zero sequence of positive real numbers $(h_n)_{n\in\mathbb{N}} \in (0, \infty)$, the limits of the form $\lim_{n\to\infty} \frac{Z(x+vh_n)-Z(x)}{h_n}$ converge in the sense of $L^2(\Omega, \mathcal{A}, \mathbb{P})$ to a random variable $Y_v(x)$, and the so defined Random Functions $(Y_v(x))_{x\in\mathbb{R}^d}$ are continuous. This can be concluded from the developments in Sobczyk (1991, Section 14.3).

If Z is mean-square continuously differentiable and if $\frac{\partial Z}{\partial x_j}$ is the partial derivative with respect to the $j-$th component on $\mathbb{R}^d$, we have the expressions for the mean and the covariance:

$$m_{\frac{\partial Z}{\partial x_j}} = \frac{\partial m_Z}{\partial x_j} \quad ; \quad C_{\frac{\partial Z}{\partial x_j}} = \frac{\partial^2 C_Z}{\partial x_j \partial y_j}. \tag{3.12}$$

Hence, the derivative of the mean is the mean of the derivative, and the covariance of the derivative is the corresponding double derivative of the covariance (we recall that the covariance C_Z is a function of two vectorial variables). The definition of the derivative of arbitrary order $N \in \mathbb{N}$ is done analogously, with the mean m_Z required to be in $C^N(\mathbb{R}^d)$ and with the covariance C_Z required to be in $C^{2N}(\mathbb{R}^d \times \mathbb{R}^d)$. Considering the multi-index notation, the generalization of Eq. (3.12) is in such a case

$$m_{D^\alpha Z} = D^\alpha m_Z \quad ; \quad C_{D^\alpha Z} = D^{(\alpha,\alpha)} C_Z, \tag{3.13}$$

for $\alpha \in \mathbb{N}^d$ with $|\alpha| \leqslant N$.

When Z is stationary with zero-mean, the differentiability condition is equivalent to require that the sta-

[6]This holds, for example, when approaching measures of compact support in the sense of the weak-$*$ topology in the space $\mathscr{M}_c(\mathbb{R}^d)$ which, as we have seen, is the dual of the space of continuous functions $C(\mathbb{R}^d)$ (Theorem 2.1.5).

tionary covariance function ρ_Z is twice differentiable, with the same analogue for higher orders of derivation. In such case, we have

$$\rho_{D^\alpha Z} = (-1)^{|\alpha|} D^{2\alpha} \rho_Z. \tag{3.14}$$

An important point to notice in the stationary case is the relation between the mean-square regularity of Z and the growing behaviour of its spectral measure μ_Z. Using the relation between ρ_Z and the spectral measure μ_Z, $\rho_Z = \mathscr{F}(\mu_Z)$ (Eq. (3.3)), it is not hard to conclude using a typical application of Dominated Convergence Theorem that ρ_Z is $N-$times continuously differentiable if and only if the multiplication measure $|\xi|^N d\mu_Z(\xi)$ is finite. Hence, the slower μ_Z *increases* at infinity, the more regular ρ_Z (and thus the Random Function Z) is. For example, the Gaussian model (3.4) is infinitely differentiable since its spectral measure has a density which decreases faster than any polynomial. The $J-$Bessel model (3.7) is also infinitely differentiable since its spectral measure is compactly supported. The Matérn model (3.6) has the particularity that the parameter α controls its regularity: using polar coordinates, one verifies that the covariance of the Matérn model is $2N-$times differentiable, and thus Z is $N-$times mean-square differentiable, if and only if $\alpha > \frac{d+N}{2}$. This *control of regularity* available for the Matérn model is one of the reasons of its popularity; see for instance the comment *"use the Matérn model"* in Stein (1999, page 14) and an example of application in this spirit in Minasny & McBratney (2005). We finally remark that the Exponential model (3.5) is not mean-square differentiable. This can be concluded by considering that the Exponential model is a particular case of the Matérn model with $\alpha = \frac{d+1}{2}$. However, it is, of course, continuous, and it is one of the most used models in practice.

The notions exposed in this section are actually sufficient to develop a rich enough Stochastic Calculus and to well-pose and solve some SPDEs. See the examples discussed in Chilès & Delfiner (1999, Section 8.3), and the developments in Sobczyk (1991, Chapter III). Nevertheless, the strong restriction of differentiability constrains us to a rather limited framework both in the theory of Stochastic Analysis and in the practice of Geostatistics. Indeed, many models used in practice are non-differentiable. Hence it is not clear how they could be related to a differential equation with differential operators acting on them. In addition, most of the well-known SPDE-based models (mainly the Matérn model) are related to a SPDE which *is not properly defined* in the framework of Random Functions: they involve a differential operator (or something *like that*) applied over an insufficiently differentiable Random Function. This issue is also present in the domain of Stochastic Calculus based on Brownian Motion: as the reader probably knows, the sample paths of Brownian Motion are continuous but nowhere differentiable, and many of the most interesting processes used in this theory present also this condition. There is thus a need for working with things that are more *general* than a function, where the objects can be differentiated freely. Having in mind the theory presented in Section 2.2, we know in which direction we should be heading to. We will describe this theory in Section 3.4. We remark the discussion on this subject proposed in Chilès & Delfiner (1999, Chapter 8), where all of these issues are commented, presenting also interesting applications of the *regular* framework presented on

this section in Hydrogeology[7].

3.3 Random Measures

A Random Measure can be defined in many different ways. The definition that we are going to use is inspired in a *mean-square analysis* as we did in the case of Random Functions. Hence, we interpret the $\sigma-$additivity condition in a mean-square sense. We will also restrict our work to the case where the covariance Kernel determines a covariance measure. The bibliographical source which exposes a theory of Random Measures in the most similar way as we do here is Rao (2012). We will not follow strictly its terminology, but the *main idea* is the same. We give some comments on other ways of approaching the Random Measure Theory in Section 3.7.

3.3.1 Formal definitions

We begin with a first generic definition. We will not actually use *this* concept as main tool of analysis due to some issues that will be exposed later. Nevertheless, it is worth being presented and discussed. The name selected for this object is strictly restricted to this dissertation and it is not inspired by the literature[8].

Definition 3.3.1. *A L^2-**generic complex random measure** (simply, a **Generic Random Measure**) over $\mathbb{R}^d$ is a complex stochastic process indexed by the bounded Borel sets $M := (M(A))_{A\in\mathcal{B}_B(\mathbb{R}^d)}$ such that for every countable collection of mutually disjoint bounded Borel sets $(A_n)_{n\in\mathbb{N}} \subset \mathcal{B}_B(\mathbb{R}^d)$ such that $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{B}_B(\mathbb{R}^d)$, it holds that*

$$M\left(\bigcup_{n\in\mathbb{N}} A_n\right) \overset{a.s.}{=} \sum_{n\in\mathbb{N}} M(A_n), \tag{3.15}$$

where the series in (3.15) is taken as a limit in $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

We will call the property (3.15) the $L^2-\ \sigma-$**additivity condition**. We remark that this additivity condition is not *strict* as in the sense of the deterministic Definition 2.1.1.

Let M be a Generic Random Measure over $\mathbb{R}^d$. Consider the set function $m_M : \mathcal{B}_B(\mathbb{R}^d) \to \mathbb{C}$, defined as $m_M(A) = \mathbb{E}(m_M(A))$. An immediate implication of the $L^2-\ \sigma-$additivity condition (3.15) is that the function m_M is in $\mathscr{M}(\mathbb{R}^d)$. We call m_M the **mean measure** of M. Consider now the function $K_M : \mathcal{B}_B(\mathbb{R}^d) \times \mathcal{B}_B(\mathbb{R}^d) \to \mathbb{C}$, defined through $K_M(A, B) = \mathbb{C}ov(M(A), M(B))$. K_M is called the **covariance bi-measure Kernel** of M. The choice of the name is important. Using the $L^2-\ \sigma-$additivity,

[7] Chapter 8 in Chilès & Delfiner (1999) has been, unfortunately, dropped out of the latest editions of this book. The reference is then, particularly, to the 1999's edition.

[8] We do not *suggest it* neither.

one can verify that K_M is what some authors call a *bi-measure* (Horowitz, 1977; Morse, 1955; Rao, 2012): for any fixed $B \in \mathcal{B}_B(\mathbb{R}^d)$, the set application $A \in \mathcal{B}_B(\mathbb{R}^d) \mapsto K_M(A,B)$ is in $\mathscr{M}(\mathbb{R}^d)$, and correspondingly, for any fixed $A \in \mathcal{B}_B(\mathbb{R}^d)$, the set application $B \in \mathcal{B}_B(\mathbb{R}^d) \mapsto K_M(A,B)$ is also in $\mathscr{M}(\mathbb{R}^d)$. This Kernel must be a **positive-definite bi-measure Kernel**, that is, for every finite collection of bounded Borel sets $(A_j)_{j\in\{1,...,N\}} \subset \mathcal{B}_B(\mathbb{R}^d)$ and for every finite collection (of the same size) of complex numbers $(\lambda_1, ..., \lambda_N) \in \mathbb{C}^N$, $N \in \mathbb{N}_*$, it must hold that

$$\sum_{j,k=1}^{N} \lambda_j K_M(A_j, A_k)\overline{\lambda_k} \geqslant 0. \tag{3.16}$$

Let us explain now why we have added the adjective "*generic*" in Definition 3.3.1. It would be expected that this Kernel determines a sort of *covariance measure*, $C_M \in \mathscr{M}(\mathbb{R}^d \times \mathbb{R}^d)$, which satisfies $C(A \times B) = K_M(A,B)$ for $A, B \in \mathcal{B}_B(\mathbb{R}^d)$. The theoretical problem that arises is that this does not hold at all: *the covariance bi-measure Kernel does not necessarily determine a measure in* $\mathscr{M}(\mathbb{R}^d \times \mathbb{R}^d)$.[9] This is a remarkable consideration which holds particularly in the case of measures. As we shall see later, in the case of Random Distributions this issue is not present thanks to the Nuclear Theorem.

We remark that this is not actually a real *theoretical* problem: even if the bi-measure Kernel does not determine a measure, an Integration Theory with respect to the associated Generic Random Measure can be anyway developed, with remarkable results. The construction of the associated integral is grounded on the concept of integral with respect to a *vector valued measure*, that is, a measure taking values in a Banach space. Indeed, the Generic Random Measure M according to definition 3.3.1 is nothing but a $\sigma-$additive function taking values in the Hilbert space $L^2(\Omega, \mathcal{A}, \mathbb{P})$. The integral with respect to such a measure is called the *Dunford-Schwartz Integral*. The related theory can be found in its general deterministic form in Dunford & Schwartz (1958, Chapter III and Section IV.10). We refer to (Rao, 2012, Chapter 2) for the associated stochastic theory. We thus conclude that even if a Generic Random Measure does not have a strictly speaking *covariance measure* describing its second order structure, it can be anyway manipulated as a measure in a satisfactory manner. The real issue is rather *practical* or methodological, and it concerns for example the practice of geostatsticians: it is far simpler and intuitive to define a Random Measure starting from a covariance *measure*, than using the more generic but complicated concept of *covariance bi-measure Kernel*.

We will thus avoid all of these issues and we will work with the following definition of *Random Measure*.

[9] In Rao (2012, Chapter 2, Example 2) the author presents sophisticated arguments which implicitly prove that the positive-definite bi-measure Kernel of the form

$$K_M(A,B) = \sum_{n\in\mathbb{N}} \nu_n(A)\overline{\nu_n}(B), \tag{3.17}$$

with $\nu_n(A) = \int_{(0,2\pi)} e^{inx}\mathbf{1}_A(x)dx$, does not determine a measure in $\mathscr{M}(\mathbb{R}^d \times \mathbb{R}^d)$. We say *implicitly* because the author does not show explicitly the Kernel in the form (3.17), but he rather works with the associated Generic Random Measure.

Definition 3.3.2. *A L^2-**Complex Random Measure** over $\mathbb{R}^d$ (from now on simply a **Random Measure**) is a Generic Random Measure M such that there exists a measure $C_M \in \mathscr{M}(\mathbb{R}^d \times \mathbb{R}^d)$ such that*

$$C_M(A \times B) = \mathbb{C}ov(M(A), M(B)), \quad \forall A, B \in \mathcal{B}_B(\mathbb{R}^d). \tag{3.18}$$

It is possible to give necessary and sufficient conditions for a covariance bi-measure Kernel to actually determine a measure in the way of Definition 3.3.2. The extra requirement is that the associated *potential* measure must have finite total variation over bounded Borel subsets of $\mathbb{R}^d \times \mathbb{R}^d$. See Horowitz (1977) for the associated result in the case of finite measures in more general measure spaces. The case of non-finite measures can be proven following similar arguments restricting the analysis to bounded sets. Another approach may be based on convenient extension theorems, which can be used to prove the existence and uniqueness of a complex measure which *extends* the domain of a complex measure of local bounded variation defined over the ring of sets of the form $A \times B$, $A, B \in \mathcal{B}_B(\mathbb{R}^d)$. See for example Takahashi (1966) for such an extension theorem. Kupka (1978) also provides an extension theorem which may be useful in this aim. We omit the details since this issue is beyond the scope of this dissertation. We also claim, without proof, that if K_M is positive, that is, $K_M(A, B) \geqslant 0$ for all $A, B \in \mathcal{B}_B(\mathbb{R}^d)$, then it does determine a covariance measure on $\mathscr{M}^+(\mathbb{R}^d \times \mathbb{R}^d)$[10], although such a requirement excludes some useful covariance models such as the $J-$Bessel model.

We focus hence, once and for all, on Random Measures as in Definition 3.3.2. Let M be a Random Measure over $\mathbb{R}^d$. The measure $C_M \in \mathscr{M}(\mathbb{R}^d \times \mathbb{R}^d)$ satisfying (3.18) is called the **covariance measure** of M. Analogously to covariance functions, this measure must satisfy a positive-definiteness condition associated to condition (3.16). We will make it explicit. We say that $C_M \in \mathscr{M}(\mathbb{R}^d \times \mathbb{R}^d)$ **defines a positive-definite Kernel** if for every finite collection of bounded Borel sets $(A_j)_{j\in\{1,...,N\}} \subset \mathcal{B}_B(\mathbb{R}^d)$ and for every finite collection (of the same size) of complex numbers $(\lambda_1, ..., \lambda_N) \in \mathbb{C}^N$, $N \in \mathbb{N}_*$, it holds that

$$\sum_{j,k=1}^{N} \lambda_j C_M(A_j \times A_k)\overline{\lambda_k} \geqslant 0. \tag{3.19}$$

Hence, if a Random Measure model is needed for describing some phenomenon, a geostatistician needs to fix a mean measure and a covariance measure defining a positive-definite Kernel, analogously to the case of Random Functions. Then, classical geostatistical tools such as Kriging and simulations follow immediately.

Let M be a Random Measure over $\mathbb{R}^d$. Its **complex conjugate** is defined as the Random Measure $\overline{M}(A) := \overline{M(A)}$ for every $A \in \mathcal{B}_B(\mathbb{R}^d)$. M can be decomposed into its **real** and **imaginary parts**, denoted

[10]This can be concluded using the same arguments presented in Dellacherie & Meyer (1978, Chapter III, N° 74) for the case of probability measures, restricting the analysis over bounded sets and then extending. We remark that this holds only when working with Radon measures, which satisfy inner and outer regular conditions and are defined over Borel sets, which is our case. For instance, Rao (2012, Chapter 3, Example 2) provides a counterexample considering non-Borel sets.

by M_R and M_I respectively, determined by $M_R(A) = \frac{M(A)+\overline{M}(A)}{2}$ and $M_I(A) = \frac{M(A)-\overline{M}(A)}{2i}$ for every $A \in \mathcal{B}_B(\mathbb{R}^d)$. M is said to be **real** if $M_I(A) \overset{a.s.}{=} 0$ for all $A \in \mathcal{B}_B(\mathbb{R}^d)$, or equivalently, if for every $A \in \mathcal{B}_B(\mathbb{R}^d)$, the random variable $M(A)$ is almost surely a real random variable. The **reflection measure** of M is the Random Measure defined as $\check{M}(A) := M(-A)$ for every $A \in \mathcal{B}_B(\mathbb{R}^d)$. If for every $A \in \mathcal{B}_B(\mathbb{R}^d)$, it holds that $M(A) \overset{a.s.}{=} \check{M}(A)$, we said that M is **even**, and if it holds that $M(A) \overset{a.s.}{=} -\check{M}(A)$, we say that it is **odd**. Finally, if M satisfies that for every $A \in \mathcal{B}_B(\mathbb{R}^d)$, $M(A) \overset{a.s.}{=} \check{\overline{M}}(A)$, M is said to be an **Hermitian** Random Measure.

In Section 3.2 we have restricted our framework to *real* Random Functions, while in the case of Random Measures we have allowed the use of complex random variables. We have done this because along this dissertation many of the Random Measures used are complex, although with a special condition: they are obtained as a result of a *complex operation* applied to a *real* stochastic process. In a general framework of complex square-integrable stochastic processes, the covariance Kernel *is not sufficient to fully-characterise the second-order structure of the process:* the *cross-covariance* structure between the real and imaginary parts must also be specified (see Section 3.6). Hence, the theoretical model of a geostatistician is not complete by only setting the covariance Kernel. Nevertheless, if the complex stochastic process is actually the result of a complex operation applied to a real stochastic process, the covariance Kernel paradigm still works, being the real and imaginary parts of the complex process determined implicitly by the complex operation and the real covariance Kernel of the real stochastic process. This will be clarified later when analysing the spectral behaviour of stationary Random Functions. In this dissertation we always work with *real* stochastic processes as basis, but we will sometimes apply complex operations on them such as, for example, the Fourier Transform.

We make a last remark. We have not defined the stochastic analogous to the *total variation measure*, nor the Jordan decomposition of the Random Measure (see Section 2.1.1). In more generality, we have not defined *positive Random Measures*. We are not going to do that. Reasons for this choice will be commented in Section 3.7. We can always, however, work with the total variation measure and with the Jordan decomposition of the mean and covariance measures.

3.3.2 Examples

The first two examples of Random Measures presented here will be widely used along this dissertation. The third one is rather presented to show to the reader the kinds of distinct areas of Stochastic Analysis that can be included within this framework.

Example 3.3.1 (White Noise). This is **the most important example of a Random Measure** used in this work[11]. Let $W = (W(A))_{A \in \mathcal{B}_B(\mathbb{R}^d)}$ be a Random Measure. We say that W is a **White Noise** if its mean

[11] It would not be very bold to say that it is the most important example of all times... until now at least...

measure m_W is null and if its covariance measure C_W is determined by

$$C_W(A \times B) = \text{Leb}(A \cap B). \tag{3.20}$$

Hence C_W is the measure supported on the hyperplane $\{y = x\}$, $C_W = \delta(x - y)$ (Example 2.1.2). It is easy to prove that C_W defines a positive-definite Kernel. A particularity of this measure is that if A and B are two disjoint Borel sets with equal Lebesgue measure (for example, take A a bounded rectangle and B be an enough distant translation of A), then the random variables $W(A)$ and $W(B)$ are non-correlated with the same variance. Hence, it is a particularly useful tool for modelling variables with stationary erratic behaviour. It is widely used, for example, in Signal Analysis and it is a *basis* of all standard developments in Stochastic Calculus and its applications[12]. Other interesting properties of White Noise will be presented further in this work. ▫

Example 3.3.2. Let Z be a real continuous Random Function over $\mathbb{R}^d$, and let $\mu \in \mathscr{M}(\mathbb{R}^d)$. We define for every $A \in \mathcal{B}_B(\mathbb{R}^d)$, the random variable

$$(Z\mu)(A) := \int_A Z(x)d\mu(x), \tag{3.21}$$

where the integral is defined as in Definition 3.2.1. The so-defined stochastic process $((Z\mu)(A))_{A\in\mathcal{B}_B(\mathbb{R}^d)}$ will be called the **multiplication** between Z and μ. From Proposition 3.2.1, it follows immediately that the multiplication $Z\mu$ is a Random Measure over $\mathbb{R}^d$, with mean measure given by $m_{Z\mu} = m_Z\mu \in \mathscr{M}(\mathbb{R}^d)$ and covariance measure given by $C_{Z\mu} = C_Z(\mu \otimes \overline{\mu}) \in \mathscr{M}(\mathbb{R}^d \times \mathbb{R}^d)$. Hence, we are able to define a stochastic analogue of the multiplication of a continuous function with a measure. When the measure μ is the Lebesgue measure, we say simply that the multiplication measure $Z(x)dx$ *is* Z. Hence, *every continuous Random Function defines a Random Measure*. ▫

Example 3.3.3. We remark the example of **Point Processes**, which we will describe roughly. Consider $(X_n)_{n\in\mathbb{N}}$ a sequence of $\mathbb{R}^d-$valued random variables. We suppose this sequence is such that almost-surely for every bounded set $A \subset \mathbb{R}^d$ the quantity of random variables in the family $(X_n)_{n\in\mathbb{N}}$ that belongs to A is finite. We define then the random variables

$$P(A) = \sum_{n\in\mathbb{N}} \delta_{X_n}(A), \quad A \in \mathcal{B}_B(\mathbb{R}^d). \tag{3.22}$$

An almost-surely $\sigma-$additivity condition follows immediately from the definition of P, and the local finite-

[12]Maybe the reader knows that the most popular branches of Stochastic Calculus are not exactly based on the White Noise but rather on Brownian Motion. Actually, both can be equivalently used as basis for the same developments: the White Noise is just the *derivative* of Brownian motion, in a generalized sense. See Section 3.8. See for example Holden et al. (2009), where a Stochastic Calculus is developed starting from White Noise rather than Brownian motion, and the same concepts arise from any of the two starting concepts.

ness guarantees that the random variable (3.22) is well-defined. This kind of process is called a *Point Process*, and it is of crucial importance in Probability Theory and its applications. Although we are not going to focus on this concept in this work, it is worth emphasizing that it can be included in our analysis without major technical difficulties[13]. If we suppose that for every set $A \in \mathcal{B}_B(\mathbb{R}^d)$ the random variable $P(A)$ is in $L^2(\Omega, \mathcal{A}, \mathbb{P})$, Eq. (3.22) defines a Random Measure according to Definition 3.3.2, and we can describe its mean and covariance measures. These measures depend of course on the dependence structure of the family $(X_n)_{n\in\mathbb{N}}$. For instance, when the point process is a homogeneous Poisson process, its covariance measure is the same as the covariance measure of the White Noise. Although in the framework of Point Processes the covariance bi-measure Kernel and the mean measure are not the most important or determinant tools, they can at any rate be used to characterise some properties of the behaviour of the process. For instance, some authors develop inference and extrapolation methods for Point Processes using the structure of their first two moments. See for example, Gabriel (2014); Gabriel et al. (2017).▫

3.3.3 Integrals

Analogously to the deterministic framework, integrals of (deterministic) measurable functions with respect to Random Measures can be defined without major difficulties. The approach is the same the one of the Lebesgue integral described in Section 2.1.2. Here we follow the exposition in Rao (2012, Section 2.2).

Let M be a Random Measure. Let $f : \mathbb{R}^d \to \mathbb{C}$ be a simple function of the form $f = \sum_{j=1}^N a_j \mathbf{1}_{A_j}$, with $(a_j)_{j\in\{1,\dots,N\}} \subset \mathbb{C}$ and $(A_j)_{j\in\{1,\dots,N\}} \subset \mathcal{B}_B(\mathbb{R}^d)$. Then, its *Integral with respect to* M is defined as

$$\int_A f(x)dM(x) = \sum_{j=1}^N a_j M(A_j). \tag{3.23}$$

Let $f : \mathbb{R}^d \to \mathbb{C}$ be measurable and let $(f_n)_{n\in\mathbb{N}}$ be a sequence of simple functions such that $f_n \to f$ point-wise and $|f_n| \to |f|$ point-wise monotonically increasing. If the sequence of random variables $\left(\int_{\mathbb{R}^d} f_n(x)dM(x)\right)_{n\in\mathbb{N}} \in L^2(\Omega, \mathcal{A}, \mathbb{P})$ is a Cauchy sequence, we say that f is **integrable** with respect to M. We define in such a case the **integral of f with respect to** M as

$$\int_{\mathbb{R}^d} f(x)dM(x) := \lim_{n\to\infty} \int_{\mathbb{R}^d} f_n(x)dM(x), \tag{3.24}$$

where the limit is taken in the sense of $L^2(\Omega, \mathcal{A}, \mathbb{P})$. In such a case, standard arguments prove that the limit does not depend on the sequence of simple functions $(f_n)_{n\in\mathbb{N}}$ approaching f. It can also be proven that this integral acts linearly on f, and that analogue results to Dominated Convergence Theorem and Monotone Convergence Theorem hold. See the comments in Rao (2012, Definition 2.2).

[13] We can, for example, study SPDEs having a point process as a *source term*.

Proposition 3.3.1. *Let M be a Random Measure over $\mathbb{R}^d$ with mean measure m_M and covariance measure C_M. Let f be a measurable complex function over $\mathbb{R}^d$. If $f \in \mathscr{L}^1(\mathbb{R}^d, m_M)$ and $f \otimes f \in \mathscr{L}^1(\mathbb{R}^d \times \mathbb{R}^d, C_M)$, then f is integrable with respect to M. If f and g are two complex measurable functions satisfying these requirements, then*

$$\mathbb{E}\left(\int_{\mathbb{R}^d} f(x)dM(x)\right) = \int_{\mathbb{R}^d} f(x)dm_M(x), \tag{3.25}$$

$$\mathbb{C}ov\left(\int_{\mathbb{R}^d} f(x)dM(x)\,,\int_{\mathbb{R}^d} g(x)dM(x)\right) = \int_{\mathbb{R}^d\times\mathbb{R}^d} f(x)\overline{g(y)}dC_M(x,y). \tag{3.26}$$

See Appendix A.6 for a proof.

A measurable function $f : \mathbb{R}^d \to \mathbb{C}$ is said to be **locally integrable** with respect to M if for every compact set $K \subset \mathbb{R}^d$ the function $\mathbf{1}_K f$ is integrable with respect to M. It is immediate from Proposition 3.3.1 that this holds when f is locally integrable with respect to m_M and $f \otimes f$ is locally integrable with respect to C_M. In such a case we define the **multiplication** between the function f and the Random Measure M, denoted by fM, as the Random Measure defined by

$$(fM)(A) := \int_{\mathbb{R}^d} \mathbf{1}_A(x)f(x)dM(x), \quad \forall A \in \mathcal{B}_B(\mathbb{R}^d). \tag{3.27}$$

From Proposition 3.3.1 it is immediate that

$$m_{fM} = fm_M \quad ; \quad C_{fM} = (f\otimes\overline{f})C_M. \tag{3.28}$$

3.3.4 Finite, slow-growing and compactly supported Random Measures

As seen in Section 2.1.1, the concept of a finite measure can be defined equivalently through an extension-domain to the whole $\sigma-$algebra of Borel sets requirement, or through a finite total variation measure. Here we are going to take an *easy way out* and we will just determine the finiteness of a Random Measure through a finiteness condition of its mean and covariance measures.

Definition 3.3.3. *Let M be a Random Measure over $\mathbb{R}^d$ with mean measure m_M and covariance measure C_M. We say that M is a **finite Random Measure** if $m_M \in \mathscr{M}_F(\mathbb{R}^d)$ and $C_M \in \mathscr{M}_F(\mathbb{R}^d \times \mathbb{R}^d)$.*

Let M be a finite Random Measure. It is immediate from Proposition 3.3.1 that every measurable and bounded function is integrable with respect to M. It follows also that a characterisation analogue to Definition 2.1.3 can be done: the indexation set of the stochastic process $(M(A))_{A\in\mathcal{B}_B(\mathbb{R}^d)}$ can be extended almost-surely uniquely to the set of all Borel sets, obtaining a square-integrable process $(M(A))_{A\in\mathcal{B}(\mathbb{R}^d)}$ for which the mean-square-$\sigma-$additivity condition (3.15) holds for every arbitrary countable partition of Borel sets. Indeed, this follows immediately from the integrability with respect to M of the functions of the form

$\mathbf{1}_A$ with $A \in \mathcal{B}(\mathbb{R}^d)$.[14]

Let us consider the case of a continuous Random Function Z. As we have seen in Example 3.3.2, it defines a Random Measure through the application $A \in \mathcal{B}_B(\mathbb{R}^d) \mapsto \int_A Z(x)dx$. We say that Z is an **integrable Random Function** if this measure is a finite Random Measure, condition obtained when its mean function m_Z is integrable with respect to the Lebesgue measure over $\mathbb{R}^d$ and its covariance function C_Z is integrable with respect to the Lebesgue measure over $\mathbb{R}^d \times \mathbb{R}^d$. Following the same spirit, we can define a continuous Random Function integrable with respect to an arbitrary deterministic measure $\mu \in \mathscr{M}(\mathbb{R}^d)$.

We introduce in a natural way the definition of a slow-growing Random Measure.

Definition 3.3.4. *Let M be a Random Measure over $\mathbb{R}^d$. We say that M is a **slow-growing Random Measure** if there exists a strictly positive polynomial $p : \mathbb{R}^d \to \mathbb{R}_*^+$ such that $\frac{1}{p}M$ is a finite Random Measure, or equivalently, if there exists $N \in \mathbb{N}$ such that $(1 + |x|^2)^{-N}M$ is a finite Random Measure.*

It is immediate from the definition of a finite Random Measure 3.3.3 that M is slow-growing if and only if $m_M \in \mathscr{M}_{SG}(\mathbb{R}^d)$ and $C_M \in \mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^d)$. Indeed, the finiteness of $\frac{1}{p}M$, with $p : \mathbb{R}^d \to \mathbb{R}_*^+$ being a strictly positive polynomial implies the finiteness of $\frac{1}{p}m_M$ and $(\frac{1}{p} \otimes \frac{1}{p})C_M$. It is also immediate that any polynomially bounded measurable function is locally integrable with respect to a slow-growing Random Measure M.

The definition of the support of a Random Measure will be quite intuitive, although the typical subtleties of *almost-surely* defined properties are involved.

Definition 3.3.5. *Let M be a Random Measure over $\mathbb{R}^d$. Its **support** is defined to be the complement of the largest open set where the measure M has L^2-norm equal to zero:*

$$\operatorname{supp}(M) := \left(\bigcup \{O \subset \mathbb{R}^d \ : \ O \text{ is open and } \mathbb{E}(|M(O)|^2) = 0\}\right)^c. \tag{3.29}$$

A Random Measure M is said to be **compactly supported** if its support is a compact set. Using standard arguments, one proves that a Random Measure M is compactly supported if and only if $m_M \in \mathscr{M}_c(\mathbb{R}^d)$ and $C_M \in \mathscr{M}_c(\mathbb{R}^d \times \mathbb{R}^d)$. It can also be concluded that $\operatorname{supp}(M)$ is always contained inside the union between $\operatorname{supp}(m_M)$ and the complementary of the largest open set $O \subset \mathbb{R}^d$ such that $|C_M|(O \times O) = 0$. It is immediate that any locally bounded measurable function is integrable with respect to any compactly supported Random Measure M.

[14] We remark that we have catalogued Definition 3.3.3 as an *easy way out* since we have avoided the statement of an equivalence between the finiteness of m_M and C_M and the properties described in this paragraph. Indeed, it would have been more sophisticated to define a finite Random Measure as a Random Measure whose indexation domain can be extended to the whole system of Borel sets of $\mathbb{R}^d$, similarly to Definition 2.1.3, and then conclude that the mean and covariance measures are finite. We ignore if such an implication holds.

3.3.5 Interpretation as linear functional

Let us right now remark the interpretation of M as a linear functional over spaces of continuous functions, analogously to the case of deterministic measures explained in Section 2.1.4. From now on, we will sometimes use the notation $\langle M, f\rangle$ to denote the integral of a measurable function f with respect to a Random Measure M, when f is integrable with respect to M.

It is straightforward from the results on Section 3.3.4 that every function $\varphi \in C_c(\mathbb{R}^d)$ (respectively in $C_{FD}(\mathbb{R}^d)$, in $C_0(\mathbb{R}^d)$, in $C(\mathbb{R}^d)$) is integrable with respect to any Random Measure over $\mathbb{R}^d$ (respectively, with respect to any slow-growing Random Measure, with respect to any finite Random Measure, with respect to any compactly supported Random Measure). We will focus on the general case of the space $C_c(\mathbb{R}^d)$.

If M is a Random Measure over $\mathbb{R}^d$, we can then define a square-integrable stochastic process indexed by the set $C_c(\mathbb{R}^d)$ through the integrals, $(\langle M, \varphi\rangle)_{\varphi \in C_c(\mathbb{R}^d)}$. The mean and covariance of this process are then described through

$$\mathbb{E}(\langle M, \varphi\rangle) = \langle m_M, \varphi\rangle \quad ; \quad \mathbb{C}ov(\langle M, \varphi\rangle, \langle M, \phi\rangle) = \langle C_M, \varphi \otimes \overline{\phi}\rangle. \tag{3.30}$$

for very $\varphi, \phi \in C_c(\mathbb{R}^d)$. The following statement can be concluded.

Proposition 3.3.2. *The mapping $M : C_c(\mathbb{R}^d) \to L^2(\Omega, \mathcal{A}, \mathbb{P})$ defined through $M(\varphi) := \langle M, \varphi\rangle$ for every $\varphi \in C_c(\mathbb{R}^d)$ determines a continuous linear functional.*

The proof of this Proposition can be found in Appendix A.7. This proof is grounded on the continuity of m_M and C_M interpreted as continuous linear functionals over $C_c(\mathbb{R}^d)$ and $C_c(\mathbb{R}^d \times \mathbb{R}^d)$ respectively (Riesz Representation Theorem 2.1.3). We remark that the converse is not true: a continuous linear mapping from $C_c(\mathbb{R}^d) \to L^2(\Omega, \mathcal{A}, \mathbb{P})$ does not necessarily define a Random Measure in the sense of Definition 3.3.2. Instead, it would be expected to define a Generic Random Measure in the sense of Definition 3.3.1. We do not enter in those details[15].

An analogue to Proposition 3.3.2 can be concluded for slow-growing Random Measures, finite Measures and compactly supported Measures using the spaces $C_{FD}(\mathbb{R}^d)$, $C_0(\mathbb{R}^d)$ and $C(\mathbb{R}^d)$ respectively. A sketch of proof of this is presented in the proof of Proposition 3.3.2.

We conclude that we can interpret a Random Measure M both as a set-function or as a continuous linear functional over $C_c(\mathbb{R}^d)$. Passing from one to the other version is done using typical procedures of Measure Theory. If we start from the linear functional version, we obtain the set-function version by taking a pointwise limit of functions in $C_c(\mathbb{R}^d)$ to indicator functions. The convergence of the related random variables

[15] It is the lack of a Nuclear Theorem similar to Theorem 2.2.2 for the case of the space $C_c(\mathbb{R}^d)$ which does not allow us to conclude that M determines a Random Measure. The same problem arises in the cases of the spaces $C_{FD}(\mathbb{R}^d)$, $C_0(\mathbb{R}^d)$ and $C(\mathbb{R}^d)$.

is always interpreted in the sense of $L^2(\Omega, \mathcal{A}, \mathbb{P})$. The inverse is done through the definition of the integral presented in Section 3.3.3.

The last remark we do is that we can now, without difficulties, define a Random Measure concentrated on a subset of $\mathbb{R}^d$, and describe integrals through expressions analogue to Eq. (2.35).

Definition 3.3.6. *Let M be a Random Measure over $\mathbb{R}^d$ and let $A \subset \mathcal{B}(\mathbb{R}^d)$. We say that M is **concentrated** on A if for every set $B \in \mathcal{B}_B(\mathbb{R}^d)$ such that $A \cap B = \varnothing$, it holds that $M(B) \overset{a.s.}{=} 0$.*

It is immediate that a Random Measure is concentrated on a set $A \in \mathcal{B}(\mathbb{R}^d)$ if and only if m_M is concentrated on A and C_M is concentrated on $A \times A$. The procedures explained in Section 2.1.5 can be applied in this stochastic framework without technical difficulties. Hence, Random Measures concentrated on the sphere, on the hyperplane $\{y = x\}$ or on the spatio-temporal cone (Examples 2.1.1, 2.1.2 and 2.1.3) are constructed analogously to the deterministic case.

3.3.6 Orthogonal Random Measures and relationship with stationary Random Functions

In this Section we introduce an important class of Random Measures which is, actually, what justifies all the material presented up to now in this section.

Definition 3.3.7. *Let M be a Random Measure over $\mathbb{R}^d$. We say that M is an **orthogonal Random Measure** if its mean measure is null and if its covariance measure is concentrated on the hyperplane $\{y = x\}$.*

Let M be an orthogonal Random Measure and let $C_M \in \mathscr{M}(\mathbb{R}^d \times \mathbb{R}^d)$ be its covariance measure. Following Example 2.1.2, there exists a unique measure $\nu_M \in \mathscr{M}(\mathbb{R}^d)$ such that $C_M(A \times B) = \nu_M(A \cap B)$ for every $A, B \in \mathcal{B}_B(\mathbb{R}^d)$. We emphasize this fact using the notation introduced in Example 2.1.2, $C_M = \nu_M \delta^{\{y=x\}}$. Since C_M defines a positive-definite Kernel, $0 \leqslant C_M(A \times A) = \nu_M(A)$ for every $A \in \mathcal{B}_B(\mathbb{R}^d)$, and hence ν_M is a positive measure. This measure $\nu_M \in \mathscr{M}^+(\mathbb{R}^d)$ is called the **weight** of the orthogonal Random Measure M. It follows that

$$\mathbb{C}ov(M(A), M(B)) = \nu_M(A \cap B), \quad \forall A, B \in \mathcal{B}_B(\mathbb{R}^d). \tag{3.31}$$

Hence, an orthogonal Random Measure produces non-correlated values when evaluated at disjoint Borel sets. If we consider $\varphi, \phi \in C_c(\mathbb{R}^d)$, it follows from Eq. (2.38) and from Proposition 3.3.1 that

$$\mathbb{C}ov(\langle M, \varphi \rangle, \langle M, \phi \rangle) = \int_{\mathbb{R}^d} \varphi(x) \overline{\phi}(x) d\nu_M(x). \tag{3.32}$$

Hence, the covariance between the action of M over two functions in $C_c(\mathbb{R}^d)$ is the inner product on $L^2(\mathbb{R}^d, \nu_M)$ of the two involved functions. If the functions φ and ϕ have disjoint supports, it follows that

$\langle M, \varphi \rangle$ and $\langle M, \phi \rangle$ are not correlated. Following an approximation argument, we conclude from Eq. (3.32) that a measurable function $f : \mathbb{R}^d \to \mathbb{C}$ is integrable with respect to M if and only if $f \in \mathscr{L}^2(\mathbb{R}^d, \nu_M)$. We remark, finally, that if two functions $f, g \in \mathscr{L}^2(\mathbb{R}^d, \nu_M)$ are orthogonal with respect to the inner-product of $L^2(\mathbb{R}^d, \nu_M)$, then the random variables $\langle M, f \rangle$ and $\langle M, g \rangle$ are uncorrelated, regardless of their supports.

From Proposition 2.1.2, it follows that an orthogonal Random Measure is slow-growing if and only if $\nu_M \in \mathscr{M}^+_{SG}(\mathbb{R}^d)$, finite if and only if $\nu_M \in \mathscr{M}^+_F(\mathbb{R}^d)$, and compactly supported if and only if $\nu_M \in \mathscr{M}^+_c(\mathbb{R}^d)$.

The following result also holds for orthogonal Random Measures. Its proof is presented in Appendix A.8.

Proposition 3.3.3. *Let M be an orthogonal Random Measure over $\mathbb{R}^d$ with weight ν_M. Let us suppose that M is Hermitian. Then, ν_M is an even measure, and the real and imaginary parts of M are uncorrelated real Random Measures with the following covariance measures:*

$$C_{M_R}(A \times B) = \frac{\nu_M(A \cap B) + \nu_M(A \cap (-B))}{2} \quad ; \quad C_{M_I}(A \times B) = \frac{\nu_M(A \cap B) - \nu_M(A \cap (-B))}{2}, \tag{3.33}$$

for all $A, B \in \mathcal{B}_B(\mathbb{R}^d)$.

Example 3.3.4. The **White Noise** over $\mathbb{R}^d$, W, is an orthogonal Random Measure whose weight is the Lebesgue Measure, $d\nu_M(x) = dx$. Hence, every function integrable with respect to the White Noise is necessarily a function in $L^2(\mathbb{R}^d)$. Since the Lebesgue measure is slow-growing, the White Noise is a slow-growing orthogonal Random Measure. It is not a finite Random Measure. ▫

The main importance of orthogonal Random Measures is their relationship with stationary Random Functions. We recall that relationship in the following Proposition.

Proposition 3.3.4. *Let Z be a real stationary Random Function with zero mean, with spectral measure μ_Z. Then, Z is the Fourier Transform of a finite Hermitian orthogonal Random Measure M whose weight is $\nu_M = (2\pi)^{\frac{d}{2}} \mu_Z$:*

$$Z(x) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{-ix^T \xi} dM(\xi). \tag{3.34}$$

The Hermitian condition in this Proposition comes from the fact that Z is real. This is a well-known result which simplifies considerably the theoretical and practical treatment of stationary Random Functions. We will not give a proof of this result, since we will actually work with a more general version which will be presented in Section 3.4.3, where the finiteness condition can be replaced by a slow-growing condition. We remark that we could have used the Inverse Fourier Transform in this result without changing the covariance structure of M: we would obtain the reflection $\check{M}$[16].

[16]The reflection of an orthogonal Random Measure with even weight measure has the same covariance measure that the non-

3.4 Generalized Random Fields

In this section we present the stochastic version of the Theory of Distributions. The idea is analogue to the stochastic version of functions and measures: the *Random Distribution* is determined by a mean distribution and a covariance distribution, and the linearity and continuity conditions are always interpreted in a mean-square sense. We then show that this notion provides a generalization of the notion of Random Function and Random Measure where differential operators of arbitrary order can be freely applied. Hence, linear SPDEs in a quite general and useful framework can be well-posed and analysed.

This theory is actually quite old. The first important treaty on this subject is Gelfand & Vilenkin (1964, Chapter III). Due to its earlier work (Gelfand, 1955), Gelfand is often considered as the first author to introduce the concept of *Generalized Stochastic Process* and its related theory. Other authors also worked on this theory at the time. We note for instance the work by Itô in the case of stationary Generalized Stochastic Processes on the real line (Itô, 1954). Since these early developments, many authors have based their works on this theory. It is in particular the case of Y. A. Rozanov (1982), where the author develops a theory of Markov Random Fields based on the concept of *Generalized Random Field*. (Matheron, 1965, Chapter X) presents an interesting *geostatistically-oriented* exposition of Random Distributions. Although this theory is widely used in Probability Theory and Stochastic Analysis, not many authors from the statistical community deal with this concept and take advantage from this theory in order to construct new geostatistical models or develop adapted inference methods. We point out the exceptions of Kelbert et al. (2005) and Angulo et al. (2000), although these authors prefer to restrain their analysis to the case of Hilbert spaces, usually focusing on stochastic processes indexed by convenient Sobolev spaces.

We present here the theory in a tempered framework, always restricted to a *mean-square analysis*. We do not base our work on the remarkable Bochner-Minlos Theorem in this subject. Some comments on this theorem and related developments are given in Section 3.7. Our framework is based on the simple fact that we *can* define a stochastic process of square-integrable random variables indexed by the Schwartz space with linearity and continuity conditions, following desired mean and covariance structures. See Appendix B for the proof of this claim, based on the classical Kolmogorov Existence Theorem.

3.4.1 Formal definitions

Random Distributions, also referred as Generalized Random Fields, are stochastic processes indexed by a space of test-functions, satisfying some linearity and continuity conditions. We are going to use the Schwartz space $\mathscr{S}(\mathbb{R}^d)$ as space of test-functions, working thus in a tempered framework. We also restrain ourselves to the case of *real* Generalized Random Fields as basis. Any *complex* Generalized Random Field that could

reflected orthogonal Random Measure. This is the situation in Proposition 3.3.4 since the spectral measure μ_Z is supposed to be even.

appear in this work will be the result of a complex operation applied over a real Generalized Random Field.

Definition 3.4.1. *A real mean-square-**tempered Random Distribution** (also called Random Distribution, Generalized Stochastic Process, Generalized Random Function, or, to set a unique terminology, **Generalized Random Field**, and abbreviated **GeRF**) over $\mathbb{R}^d$ is a real and continuous linear operator from $\mathscr{S}(\mathbb{R}^d)$ to $L^2(\Omega, \mathcal{A}, \mathbb{P})$.*

Let us make this definition more explicit. A mapping $Z : \mathscr{S}(\mathbb{R}^d) \to L^2(\Omega, \mathcal{F}, \mathbb{P})$ is a real GeRF if

- $Z(\varphi + \phi) \overset{a.s.}{=} Z(\varphi) + Z(\phi)$ and $Z(\lambda\varphi) \overset{a.s.}{=} \lambda Z(\varphi)$ for all $\varphi, \phi \in \mathscr{S}(\mathbb{R}^d)$ and $\lambda \in \mathbb{C}$.
- If $\varphi \in \mathscr{S}(\mathbb{R}^d)$ is a real test-function, $Z(\varphi)$ is an almost-surely real random variable.
- If $(\varphi_n)_{n\in\mathbb{N}} \subset \mathscr{S}(\mathbb{R}^d)$ is a sequence of test-functions such that $\varphi_n \overset{\mathscr{S}}{\to} 0$, then $Z(\varphi_n) \overset{L^2(\Omega)}{\to} 0$.

Thus, Z can be interpreted as a stochastic process indexed by the Schwartz space, $(Z(\varphi))_{\varphi\in\mathscr{S}(\mathbb{R}^d)}$, with a real, linear and mean-square continuous behaviour. Since both $\mathscr{S}(\mathbb{R}^d)$ and $L^2(\Omega, \mathcal{A}, \mathbb{P})$ are metric spaces, the sequential continuity guarantees the continuity of Z. In order to emphasize that Z works as a continuous linear functional, we will from now on explicitly write

$$\langle Z, \varphi \rangle := Z(\varphi), \tag{3.35}$$

for every $\varphi \in \mathscr{S}(\mathbb{R}^d)$.

Since the random variables $(\langle Z, \varphi\rangle)_{\varphi\in\mathscr{S}(\mathbb{R}^d)}$ are all in $L^2(\Omega, \mathcal{A}, \mathbb{P})$, a mean and a covariance structure can be described. For instance, let us consider the mean function $m_Z : \mathscr{S}(\mathbb{R}^d) \to \mathbb{C}$ defined as

$$m_Z(\varphi) = \mathbb{E}(\langle Z, \varphi\rangle). \tag{3.36}$$

Since Z is real and linear, m_Z is also real and linear. If $(\varphi_n)_{n\in\mathbb{N}} \subset \mathscr{S}(\mathbb{R}^d)$ is a sequence such that $\varphi_n \overset{\mathscr{S}}{\to} 0$, we argue by Hölder inequality that

$$|m_Z(\varphi_n)| \leqslant \sqrt{\mathbb{E}(|\langle Z, \varphi_n\rangle|^2)} \to 0,$$

and thus m_Z is a continuous real linear functional, so it is a real tempered distribution: $m_Z \in \mathscr{S}'(\mathbb{R}^d)$. We call m_Z the **mean distribution** of Z. We write thus $\langle m_Z, \varphi\rangle := m_Z(\varphi)$ for all $\varphi \in \mathscr{S}(\mathbb{R}^d)$.

Let us now define the application $K_Z : \mathscr{S}(\mathbb{R}^d) \times \mathscr{S}(\mathbb{R}^d) \to \mathbb{C}$ through

$$K_Z(\varphi, \phi) = \mathbb{C}ov(\, \langle Z, \varphi\rangle \,,\, \langle Z, \phi\rangle \,). \tag{3.37}$$

We call K_Z the **covariance Kernel** of Z. By definition of the covariance, the application K_Z is a sesquilinear form. In addition, it is a **positive-definite Kernel**. Indeed, let $(\varphi_1, ..., \varphi_N) \in \mathscr{S}(\mathbb{R}^d)^N$ and $(\lambda_1, ..., \lambda_N) \in \mathbb{C}^N$, with $N \in \mathbb{N}_*$. By the sesquilinearity, we argue that

$$\begin{aligned}\sum_{j,k=1}^{N} \lambda_j K_Z(\varphi_j, \varphi_k)\overline{\lambda_k} &= \sum_{j,k=1}^{N} K_Z(\lambda_j\varphi_j, \varphi_k\lambda_k) \\ &= K_Z\left(\sum_{j=1}^{N} \lambda_j\varphi_j, \sum_{k=1}^{N} \varphi_k\lambda_k\right) \\ &= \mathbb{C}ov\left(\langle Z, \sum_{j=1}^{N} \lambda_j\varphi_j\rangle, \langle Z, \sum_{k=1}^{N} \lambda_k\varphi_k\rangle\right) \\ &= \mathbb{V}ar\left(\langle Z, \sum_{j=1}^{N} \lambda_j\varphi_j\rangle\right) \geqslant 0.\end{aligned} \tag{3.38}$$

We remark that the vector space structure of $\mathscr{S}(\mathbb{R}^d)$ allows us to describe the positive-definiteness of the covariance Kernel K_Z in a simpler way: a sesquilinear form $K : \mathscr{S}(\mathbb{R}^d) \times \mathscr{S}(\mathbb{R}^d) \to \mathbb{C}$ is a positive-definite Kernel if and only if $K(\varphi, \varphi) \geqslant 0$ for all $\varphi \in \mathscr{S}(\mathbb{R}^d)$.

Let us right now fix $\phi \in \mathscr{S}(\mathbb{R}^d)$, and consider the linear functional over $\mathscr{S}(\mathbb{R}^d)$, $\varphi \mapsto K_Z(\varphi, \phi)$. Let $(\varphi_n)_{n\in\mathbb{N}} \subset \mathscr{S}(\mathbb{R}^d)$ such that $\varphi_n \xrightarrow{\mathscr{S}} 0$. Since Z is continuous, $\langle Z, \varphi_n\rangle \xrightarrow{L^2} 0$, and using the Cauchy-Schwarz inequality we obtain that

$$|K_Z(\varphi_n, \phi)| \leqslant \sqrt{\mathbb{V}ar(\langle Z, \varphi_n\rangle)\mathbb{V}ar(\langle Z, \phi\rangle)} \to 0.$$

Hence, $\varphi \mapsto K_Z(\varphi, \phi)$ is a continuous linear functional, thus it is in $\mathscr{S}'(\mathbb{R}^d)$. Doing the same procedure with the linear functional $\phi \mapsto K_Z(\varphi, \overline{\phi})$, $\varphi \in \mathscr{S}(\mathbb{R}^d)$ being fixed, one concludes that it is also in $\mathscr{S}'(\mathbb{R}^d)$. Hence, the bilinear form

$$(\varphi, \phi) \mapsto K_Z(\varphi, \overline{\phi}) \tag{3.39}$$

defines a separately continuous form on $\mathscr{S}(\mathbb{R}^d) \times \mathscr{S}(\mathbb{R}^d)$. By the Nuclear Theorem 2.2.2, there exists a unique tempered distribution $C_Z \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^d)$ such that

$$\langle C_Z, \varphi \otimes \phi\rangle = K_Z(\varphi, \overline{\phi}), \quad \forall \varphi, \phi \in \mathscr{S}(\mathbb{R}^d). \tag{3.40}$$

This tempered distribution is called the **covariance distribution** of Z, and it satisfies, of course,

$$\langle C_Z, \varphi \otimes \overline{\phi}\rangle = \mathbb{C}ov(\langle Z, \varphi\rangle, \langle Z, \phi\rangle), \quad \forall \varphi, \phi \in \mathscr{S}(\mathbb{R}^d). \tag{3.41}$$

We say also that *C_Z defines a positive-definite Kernel*. In general, a distribution $C \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^d)$ is said to

define a positive-definite Kernel if

$$\langle C, \varphi \otimes \overline{\varphi} \rangle \geqslant 0, \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^d). \tag{3.42}$$

Since we have supposed that Z is a real GeRF, it also follows that C_Z is a real distribution in $\mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^d)$. With a few basic but tedious computations, one proves that C_Z is also *symmetric* in the sense that

$$\langle C_Z, \varphi \otimes \phi \rangle = \langle C_Z, \phi \otimes \varphi \rangle, \quad \forall \varphi, \phi \in \mathscr{S}(\mathbb{R}^d). \tag{3.43}$$

We remark that in this case, the Nuclear Theorem has allowed us to identify the covariance Kernel with a covariance distribution in its own right, contrarily to the case of Random Measures exposed in Section 3.3. More comments about this are given in Section 3.7. Hence, in order to model the variable of a phenomenon as a distribution, a mean and a covariance distribution (of two-variables) must be set, analogously to the case of Random Functions.

Example 3.4.1 (Random Functions with polynomially bounded mean and covariance). Let $m_Z : \mathbb{R}^d \to \mathbb{R}$ be a continuous polynomially bounded function and let $C_Z : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be a polynomially bounded continuous function which is a positive-definite Kernel. Let $(Z(x))_{x \in \mathbb{R}^d}$ be a real Random Function with covariance function C_Z and with mean function m_Z. We can thus define the generalized version of Z by defining for every $\varphi \in \mathscr{S}(\mathbb{R}^d)$:

$$\langle Z, \varphi \rangle = \int_{\mathbb{R}^d} Z(x)\varphi(x)dx. \tag{3.44}$$

Following Proposition 3.2.1 we conclude that for every $\varphi \in \mathscr{S}(\mathbb{R}^d)$ this integral is well-defined. Indeed, the random function $x \in \mathbb{R}^d \mapsto \varphi(x)Z(x)$ is integrable since both functions $x \in \mathbb{R}^d \mapsto \varphi(x)m_Z(x)$ and $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d \mapsto \varphi(x)\overline{\varphi}(y)C_Z(x, y)$ are integrable over their respective domains. We obtain thus a stochastic process indexed by the Schwartz space $(\langle Z, \varphi \rangle)_{\varphi \in \mathscr{S}(\mathbb{R}^d)}$, and it is not hard to prove that it defines a real GeRF. The covariance distribution of Z, also noted $C_Z \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^d)$ is determined by

$$\langle C_Z, \psi \rangle = \int_{\mathbb{R}^d \times \mathbb{R}^d} C_Z(x, y)\psi(x, y)dxdy, \quad \forall \psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^d). \tag{3.45}$$

An analogue result holds for the mean function. Thus, the *generalization* of the continuous Random Function Z, that is, its interpretation as a Random Distribution, is done analogously to the interpretation of continuous functions as a distribution in the deterministic case: using its integral. In addition, we can verify that the criteria of positive-definiteness of C_Z interpreted as a distribution is equivalent to the classical definition for continuous functions. We make the statement in the general continuous, not-necessarily polynomially bounded case.

Proposition 3.4.1. *Let $C_Z : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be a continuous function. Then, it is a positive-definite Kernel if*

and only if

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} C_Z(x,y)\varphi(x)\overline{\varphi}(y)d(x,y) \geqslant 0 \quad \forall\varphi \in \mathscr{D}(\mathbb{R}^d). \tag{3.46}$$

We give a sketch of proof of this Proposition in a footnote[17]. In general, we will say that a continuous Random Function with continuous polynomially bounded mean and covariance functions *is* a GeRF. Conversely, we will say that a GeRF over $\mathbb{R}^d$ *is* a continuous Random Function if both its mean and covariance distributions *are* continuous functions. ▫

Example 3.4.2 (Slow-growing Random Measures). Let M be a slow-growing Random Measure over $\mathbb{R}^d$, with mean measure $m_M \in \mathscr{M}_{SG}(\mathbb{R}^d)$ and covariance measure $C_M \in \mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^d)$. Since every function in the Schwartz space decreases faster than any polynomial, it follows that for every $\varphi \in \mathscr{S}(\mathbb{R}^d)$, the multiplication measures φm_M and $(\varphi \otimes \overline{\varphi})C_M$ are finite measures, and hence φ is integrable with respect to M (Proposition 3.3.1). We consider then

$$\langle M, \varphi\rangle = \int_{\mathbb{R}^d} \varphi(x)dM(x), \quad \forall\varphi \in \mathscr{S}(\mathbb{R}^d). \tag{3.47}$$

The so-defined process $(\langle M, \varphi\rangle)_{\varphi\in\mathscr{S}(\mathbb{R}^d)}$ is then a GeRF, due to the linearity of the integral and to the fact that $m_M \in \mathscr{M}_{SG}(\mathbb{R}^d) \subset \mathscr{S}'(\mathbb{R}^d)$ and $C_Z \in \mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^d) \subset \mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^d)$. Its covariance distribution thus C_M, interpreted as a tempered distribution:

$$\langle C_M, \psi\rangle = \int_{\mathbb{R}^d\times\mathbb{R}^d} \psi(x,y)dC_M(x,y), \quad \forall\psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^d). \tag{3.48}$$

An analogue result holds for the mean measure. Hence, here again the procedure is done analogously to the deterministic case. The interpretation of a slow-growing Random Measure as a GeRF is done through the associated integral. We keep thus this idea and we will simply say that a slow-growing Random Measure *is* a GeRF. Conversely, we say that a GeRF over $\mathbb{R}^d$, Z, *is* a slow-growing Random Measure if both its mean and covariance distributions are slow-growing measures. ▫

3.4.2 Operations with GeRFs

In this section we explain how to apply some continuous linear operators defined for tempered distributions to a GeRF. The definition is actually straightforward: for an operator defined through an adjoint, we can pass

[17] **Sketch of proof of Proposition 3.4.1:** for the necessity, express the integral as a limit of convenient Riemann sums (using partitions of $\mathrm{supp}(\varphi) \times \mathrm{supp}(\varphi)$ in rectangles and fixing the middle points as tag points, for instance). One verifies that the sums obtained have the form of a quadratic form which are always positive due to the positive-definiteness of C_Z, and thus the integral is positive as a limit of positive numbers. For the sufficiency, if we consider $(x_1, ..., x_N) \in (\mathbb{R}^d)^N$ and $(\lambda_1, ..., \lambda_N) \in \mathbb{C}^N$ arbitrary, we can consider N sequences of functions in $\mathscr{D}(\mathbb{R}^d)$, $(\varphi_n^1)_{n\in\mathbb{N}}, ..., (\varphi_n^N)_{n\in\mathbb{N}}$ such that $\varphi_n^j \to \delta_{x_j}$ in the sense of $C'(\mathbb{R}^d) = \mathscr{M}_c(\mathbb{R}^d)$ for every $j \in \{1, ..., N\}$. One verifies then that the sums $\sum_{j,k} \omega_j\langle C_Z, \varphi_n^j \otimes \overline{\varphi_n^k}\rangle\overline{\omega_k}$ are positive and that they converge to the associated quadratic form $\sum_{j,k} \omega_j C_Z(x_j, x_k)\omega_k$, which is then positive. ■

the action of the operator *to the test-function*, and hence apply it without problems.

We will slightly change the notations in a way that is not standard at all. Rather than using the notation presented in Section 2.2.2, where $\mathcal{L}$ denotes a continuous linear operator from $\mathscr{S}(\mathbb{R}^d)$ to $\mathscr{S}(\mathbb{R}^d)$ and $\mathcal{L}^*$ denotes its adjoint, we will use the following *inverted* notation: $\mathcal{L}$ is going to be a linear and continuous operator from $\mathscr{S}'(\mathbb{R}^d)$ to $\mathscr{S}'(\mathbb{R}^d)$ which will be always supposed to be defined through an adjoint, and we will note $\mathcal{L}_* : \mathscr{S}(\mathbb{R}^d) \to \mathscr{S}(\mathbb{R}^d)$ its *pre-adjoint*, that is, the continuous linear operator for which $(\mathcal{L}_*)^* = \mathcal{L}$. The motivation of this usage if that in this dissertation we will mainly work with operators applied to GeRFs and tempered distributions, and hence it is better to keep for them the letter $\mathcal{L}$ without any $*$ symbol on it.

Consider then $\mathcal{L} : \mathscr{S}'(\mathbb{R}^d) \to \mathscr{S}'(\mathbb{R}^d)$ be a continuous linear operator defined through an adjoint. Let $\mathcal{L}_* : \mathscr{S}(\mathbb{R}^d) \to \mathscr{S}(\mathbb{R}^d)$ be its pre-adjoint. Let Z be a real GeRF. We define then the GeRF $\mathcal{L}Z$ as

$$\langle \mathcal{L}Z, \varphi \rangle := \langle Z, \mathcal{L}_* \varphi \rangle, \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^d). \tag{3.49}$$

The operation is well-defined algebraically speaking since $\mathcal{L}_*\varphi$ is in $\mathscr{S}(\mathbb{R}^d)$ for every $\varphi \in \mathscr{S}(\mathbb{R}^d)$. The continuity of $\mathcal{L}Z$ as a linear mapping from $\mathscr{S}(\mathbb{R}^d)$ to $L^2(\Omega, \mathcal{A}, \mathbb{P})$ is guaranteed by the continuity of Z and $\mathcal{L}_*$. Hence, $(\langle \mathcal{L}Z, \varphi \rangle)_{\varphi \in \mathscr{S}(\mathbb{R}^d)}$ is a well-defined GeRF. We notice that it is not necessarily real since the operator $\mathcal{L}$ may be a complex operator.

Let us make explicit the mean and the covariance distributions of $\mathcal{L}Z$. For the mean, we obtain for every $\varphi \in \mathscr{S}(\mathbb{R}^d)$,

$$\langle m_{\mathcal{L}Z}, \varphi \rangle = \mathbb{E}(\langle \mathcal{L}Z, \varphi \rangle) = \mathbb{E}(\langle Z, \mathcal{L}_* \varphi \rangle) = \langle m_Z, \mathcal{L}_* \varphi \rangle = \langle \mathcal{L} m_Z, \varphi \rangle. \tag{3.50}$$

For the covariance, consider $\varphi, \phi \in \mathscr{S}(\mathbb{R}^d)$ arbitrary. Then,

$$\begin{aligned} \langle C_{\mathcal{L}Z}, \varphi \otimes \overline{\phi} \rangle &= \mathbb{C}ov(\langle \mathcal{L}Z, \varphi \rangle, \langle \mathcal{L}Z, \phi \rangle) = \mathbb{C}ov(\langle Z, \mathcal{L}_* \varphi \rangle, \langle Z, \mathcal{L}_* \phi \rangle) \\ &= \langle C_Z, \mathcal{L}_* \varphi \otimes \overline{\mathcal{L}_* \phi} \rangle = \langle C_Z, \mathcal{L}_* \varphi \otimes \overline{\mathcal{L}_*}(\overline{\phi}) \rangle \\ &= \langle C_Z, (\mathcal{L}_* \otimes \overline{\mathcal{L}_*})(\varphi \otimes \overline{\phi}) \rangle = \langle (\mathcal{L} \otimes \overline{\mathcal{L}}) C_Z, \varphi \otimes \overline{\phi} \rangle. \end{aligned} \tag{3.51}$$

Here we have used the definition of the tensor product operator $\mathcal{L} \otimes \overline{\mathcal{L}}$ and its pre-adjoint $\mathcal{L}_* \otimes \overline{\mathcal{L}_*}$, following Section 2.2.3. $\overline{\mathcal{L}}$ denotes the *complex conjugate* of $\mathcal{L}$, which does $\overline{\mathcal{L}}(T) = \overline{\mathcal{L}(\overline{T})}$ for every $T \in \mathscr{S}'(\mathbb{R}^d)$. $\overline{\mathcal{L}_*}$ is the complex conjugate of $\mathcal{L}_*$, with an analogous definition. We conclude that for $\mathcal{L}Z$ we have

$$m_{\mathcal{L}Z} = \mathcal{L} m_Z \quad ; \quad C_{\mathcal{L}Z} = (\mathcal{L} \otimes \overline{\mathcal{L}}) C_Z. \tag{3.52}$$

We will make explicit the application of the operators presented in Section 2.2.2 to stochastic objects. In these examples, Z denotes a real GeRF over $\mathbb{R}^d$ with mean distribution $m_Z \in \mathscr{S}'(\mathbb{R}^d)$ and covariance distribution $C_Z \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^d)$.

Example 3.4.3 (**Derivatives**). Let $\alpha \in \mathbb{N}^d$ be a multi-index and let $D^\alpha : \mathscr{S}'(\mathbb{R}^d) \to \mathscr{S}'(\mathbb{R}^d)$ be the differential operator associated. The derivative of Z is then simply the GeRF defined through

$$\langle D^\alpha Z, \varphi \rangle := (-1)^{|\alpha|} \langle Z, D^\alpha \varphi \rangle, \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^d). \tag{3.53}$$

Hence, a GeRF *can be differentiated any number of times*. It is thus the appropriate tool to well-pose linear SPDEs. The mean and covariance distributions of $D^\alpha Z$ satisfy

$$m_{D^\alpha Z} = D^\alpha m_Z \quad ; \quad C_{D^\alpha Z} = D^{(\alpha,\alpha)} C_Z. \tag{3.54}$$

Hence, we have obtained the generalized form of Eq. (3.14) of the case of Random Functions. ▫

Example 3.4.4 (**Multiplication with $\mathcal{O}_M(\mathbb{R}^d)$**). Let $f \in \mathcal{O}_M(\mathbb{R}^d)$ (deterministic). The multiplication between Z and f is the GeRF defined through

$$\langle fZ, \varphi \rangle := \langle Z, f\varphi \rangle, \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^d). \tag{3.55}$$

The mean and covariance distributions of fZ are simply

$$m_{fZ} = f m_Z \quad ; \quad C_{fZ} = (f \otimes \overline{f}) C_Z. \quad ▫ \tag{3.56}$$

Example 3.4.5 (**Convolution with $\mathcal{O}_c'(\mathbb{R}^d)$**). Let $S \in \mathcal{O}_c'(\mathbb{R}^d)$ (deterministic). The convolution between Z and S is the GeRF defined through

$$\langle Z * S, \varphi \rangle := \langle Z, \check{S} * \varphi \rangle, \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^d). \tag{3.57}$$

A few computations which will be omitted prove that the mean and covariance distributions of $Z * S$ are

$$m_{Z*S} = S * m_Z \quad ; \quad C_{Z*S} = (S \otimes \overline{S}) * C_Z. \quad ▫ \tag{3.58}$$

Example 3.4.6 (**Fourier Transform**). The Fourier Transform of Z is simply the GeRF defined through

$$\langle \mathscr{F}(Z), \varphi \rangle := \langle Z, \mathscr{F}(\varphi) \rangle, \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^d). \tag{3.59}$$

The mean and the covariance distributions of $\mathscr{F}(Z)$ are

$$m_{\mathscr{F}(Z)} = \mathscr{F}(m_Z) \quad ; \quad (\mathscr{F} \otimes \mathscr{F}^{-1})(C_Z). \tag{3.60}$$

Here we have used that $\overline{\mathscr{F}} = \mathscr{F}^{-1}$. An analogue definition and result holds for the Inverse Fourier Trans-

form. It also holds, of course, that $\mathscr{F}^{-1}(\mathscr{F}(Z)) = Z$. Finally, the Exchange formula

$$\mathscr{F}(Z * S) = (2\pi)^{\frac{d}{2}} \mathscr{F}(S)\mathscr{F}(Z) \tag{3.61}$$

holds for every deterministic $S \in \mathcal{O}'_c(\mathbb{R}^d)$, for which we recall that $\mathscr{F}(S) \in \mathcal{O}_M(\mathbb{R}^d)$. ▫

3.4.3 Stationary GeRFs

We now provide more details about **stationary GeRFs** with their main properties. In order to introduce this concept intuitively, consider first of all the case of a real continuous stationary Random Function $(Z(x))_{x\in\mathbb{R}^d}$ with covariance function (Kernel) C_Z and with stationary covariance function ρ_Z. From the inequality $|\rho_Z(h)| \leqslant \rho_Z(0)$, it follows that ρ_Z is continuous and bounded and hence it defines a tempered distribution in $\mathscr{S}'(\mathbb{R}^d)$. C_Z is also continuous and bounded hence it defines a tempered distribution in $\mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^d)$. The mean function of Z is a constant so it also defines a tempered distribution in $\mathscr{S}'(\mathbb{R}^d)$. It follows (Example 3.4.1) that Z defines a GeRF, and hence, *any stationary Random Function can be interpreted as a tempered Random Distribution*. Recalling that $C_Z(x, y) = \rho_Z(x - y)$, the covariance distribution of Z (also noted C_Z) satisfies

$$\langle C_Z, \varphi \otimes \phi\rangle = \int_{\mathbb{R}^d\times\mathbb{R}^d} C_Z(x, y)\varphi(x)\phi(y)d(x, y) = \int_{\mathbb{R}^d\times\mathbb{R}^d} \rho_Z(x - y)\varphi(x)\phi(y)d(x, y), \tag{3.62}$$

for every $\varphi, \phi \in \mathscr{S}(\mathbb{R}^d)$. With a change of variable and Fubini's Theorem one obtains

$$\langle C_Z, \varphi \otimes \phi\rangle = \int_{\mathbb{R}^d} \rho_Z(u) \int_{\mathbb{R}^d} \varphi(u - y)\check{\phi}(y)dy\, du = \int_{\mathbb{R}^d} \rho_Z(u)(\varphi * \check{\phi})(u)du. \tag{3.63}$$

Inspired by this, we give a more general definition of a stationary GeRF.

Definition 3.4.2. *Let Z be a real GeRF over $\mathbb{R}^d$ with mean distribution m_Z and covariance distribution C_Z. We say that Z is a **second order stationary GeRF** (from now on, **stationary GeRF**) if its mean distribution is a constant function and if there exists a tempered distribution $\rho_Z \in \mathscr{S}'(\mathbb{R}^d)$ such that*

$$\langle C_Z, \varphi \otimes \phi\rangle = \langle \rho_Z, \varphi * \check{\phi}\rangle. \tag{3.64}$$

The distribution ρ_Z stated in Definition 3.4.2 is called the **stationary covariance distribution** of Z. When the stationarity is clear in context and C_Z does not intervene in the exposition, we call it simply the *covariance distribution* of Z. Since C_Z is a real distribution, it follows that ρ_Z is also a real distribution. We remark that ρ_Z is an even distribution: $\check{\rho}_Z = \rho_Z$. This can be concluded from the *symmetry* of C_Z stated in Eq. (3.43), the commutativity of the convolution product, and from $\widetilde{\varphi * \check{\phi}} = \check{\varphi} * \phi$.

The stationarity condition stated in Definition 3.4.2 implies, and actually is implied by a translation invariant property of its second order structures, analogously to the case of Random Functions. Indeed, it can be proven that a GeRF Z is stationary if and only if for every $h \in \mathbb{R}^d$ it holds that $\tau_h m_Z = m_Z$ and $(\tau_h \otimes \tau_h) C_Z = C_Z$, with τ_h being the translation operator. See the arguments in Matheron (1965, Chapter X, Section 2).

The fact that C_Z defines a positive-definite Kernel also implies a positive-definiteness condition for the distribution ρ_Z that we will specify in general. A distribution $\rho \in \mathscr{D}'(\mathbb{R}^d)$, that is, not necessarily tempered, is said to be a **positive-definite distribution** or a **distribution of positive-type** if

$$\langle \rho, \varphi * \check{\overline{\varphi}} \rangle \geqslant 0, \quad \forall \varphi \in \mathscr{D}(\mathbb{R}^d). \tag{3.65}$$

In Definition 3.4.2 we have imposed ρ_Z to be in $\mathscr{S}'(\mathbb{R}^d)$. The following theorem, which is a generalization of Bochner's Theorem, known as the **Bochner-Schwartz Theorem**, guarantees that we have lost nothing with this restriction.

Theorem 3.4.1 (Bochner-Schwartz). *Let $\rho \in \mathscr{D}'(\mathbb{R}^d)$. Then, ρ is a real*[18] *positive-definite distribution if and only if $\rho \in \mathscr{S}'(\mathbb{R}^d)$ and ρ is the Fourier Transform of an even positive slow-growing measure $\mu \in \mathscr{M}^+_{SG}(\mathbb{R}^d)$: $\rho = \mathscr{F}(\mu)$.*

See Donoghue (1969, Chapter 42) for a proof. Hence, from this Theorem we conclude that if Z is a stationary GeRF with stationary covariance distribution ρ_Z, there exists a unique even measure $\mu_Z \in \mathscr{M}^+_{SG}(\mathbb{R}^d)$ such that

$$\rho_Z = \mathscr{F}(\mu_Z). \tag{3.66}$$

The measure μ_Z which satisfies (3.66) is called the **spectral measure of** Z. Since both the distribution ρ_Z and the measure μ_Z are even, it follows that $\rho_Z = \mathscr{F}(\mu_Z) = \mathscr{F}^{-1}(\mu_Z)$. We conclude quite easily the generalization of Proposition 3.3.4.

Theorem 3.4.2. *Let Z be a real, zero mean, stationary GeRF over $\mathbb{R}^d$ with spectral measure μ_Z. Then, Z is the Fourier Transform of an Hermitian slow-growing orthogonal Random Measure M_Z whose weight is proportional to its spectral measure, $\nu_{M_Z} = (2\pi)^{\frac{d}{2}} \mu_Z$. Conversely, the Fourier Transform of any Hermitian slow-growing orthogonal Random Measure M_Z over $\mathbb{R}^d$ with weight ν_{M_Z} is a real zero-mean stationary GeRF over $\mathbb{R}^d$ with spectral measure $\mu_Z = (2\pi)^{-\frac{d}{2}} \nu_{M_Z}$.*

This theorem is quite old and well-known. It can be found in Matheron (1965, Chapter X, Section 3), and in Itô (1954) for dimension $d = 1$. Anyway, with the framework exposed until now, the proof of this

[18] Again, this Theorem does not actually require the distribution to be real. The related measure must not necessarily be even, but it has to be positive and slow-growing.

theorem is quite straightforward: we have just to analyse the covariance distribution of $\mathscr{F}^{-1}(Z)$. We make explicit the calculation since it is simple and beautiful.

$$\begin{aligned}\mathbb{C}ov(\langle\mathscr{F}^{-1}(Z),\varphi\rangle,\langle\mathscr{F}^{-1}(Z),\phi\rangle) &= \mathbb{C}ov(\langle Z,\mathscr{F}^{-1}(\varphi)\rangle,\langle Z,\mathscr{F}^{-1}(\phi)\rangle)\\ &= \langle C_Z,\mathscr{F}^{-1}(\varphi)\otimes\overline{\mathscr{F}^{-1}(\phi)}\rangle\\ &= \langle\rho_Z,\mathscr{F}^{-1}(\varphi)*\check{\overline{\mathscr{F}^{-1}(\phi)}}\rangle\\ &= \langle\rho_Z,(2\pi)^{\frac{d}{2}}\mathscr{F}^{-1}(\varphi\overline{\phi})\rangle\\ &= (2\pi)^{\frac{d}{2}}\langle\mathscr{F}^{-1}(\rho_Z),\varphi\overline{\phi}\rangle\\ &= \langle(2\pi)^{\frac{d}{2}}\mu_Z,\varphi\overline{\phi}\rangle = \langle(2\pi)^{\frac{d}{2}}\mu_Z\delta^{\{y=x\}},\varphi\otimes\overline{\phi}\rangle.\end{aligned} \tag{3.67}$$

Here we have used the definition of $\mathscr{F}^{-1}$, the relation $\check{\overline{\mathscr{F}^{-1}(\phi)}} = \mathscr{F}^{-1}(\overline{\phi})$, the Exchange Formula for the Inverse Fourier Transform (it holds analogously to the case of the Fourier Transform), and the Bochner-Schwartz Theorem 3.4.1. Hence, we conclude that

$$C_{\mathscr{F}^{-1}(Z)} = (2\pi)^{\frac{d}{2}}\mu_Z\delta^{\{y=x\}}. \tag{3.68}$$

This proves that $\mathscr{F}^{-1}(Z)$ is a slow-growing orthogonal Random Measure with weight $(2\pi)^{\frac{d}{2}}\mu_Z$. Conversely, starting from an Hermitian slow-growing orthogonal Random Measure M_Z, the result is straightforward considering that M_Z defines a GeRF through the well-defined integrals $\langle M_Z,\varphi\rangle$ for all $\varphi\in\mathscr{S}(\mathbb{R}^d)$ (Example 3.4.2), and hence $\mathscr{F}(M_Z)$ is well-defined as a GeRF. The stationarity of $\mathscr{F}(M_Z)$ is proven following a computation similar to (3.67).

We note that the Hermitian condition in Theorem 3.4.2 is a consequence of Z being real. We could have used the Inverse Fourier Transform to state the Theorem without changes in the covariance structure of M_Z: we would obtain the reflection of the Random Measure in 3.4.2, which has the same covariance measure since μ_Z is even, analogously to the case of Proposition 3.3.4. We also note that we could have stated a similar theorem without the assumption of Z being zero-mean. In this case, Z would be the Fourier Transform of a slow-growing Random Measure whose mean measure is proportional to Dirac measure, and its covariance measure still, of course, concentrated on the hyperplane $\{y=x\}$.

Example 3.4.7 (White Noise). Consider the White Noise over $\mathbb{R}^d$, W. As we have seen in Example 3.3.1, the covariance measure of W is $C_W = \delta^{\{y=x\}} = \delta(y-x)$. Hence,

$$\langle C_W,\varphi\otimes\overline{\phi}\rangle = \int_{\mathbb{R}^d}\varphi(x)\overline{\phi}(x)dx = \int_{\mathbb{R}^d}\varphi(0-x)\overline{\phi}(-x)dx = \langle\delta,\varphi*\check{\overline{\phi}}\rangle. \tag{3.69}$$

Hence, W is a stationary GeRF with stationary covariance distribution $\rho_W=\delta$. Since δ is not a function, the White Noise cannot be interpreted as a Random Function, but only as a Random Measure or as a GeRF.

From the relation $\mu_W = \mathscr{F}(\rho_W)$, it follows that the spectral measure of the White Noise is proportional to the Lebesgue measure: $d\mu_W(\xi) = (2\pi)^{-\frac{d}{2}} d\xi$. Let us describe the orthogonal Random Measure associated to the White Noise $M_W = \mathscr{F}(W)$. Using formula (3.68) (which also holds for the Fourier Transform), one concludes

$$C_{\mathscr{F}(W)} = (2\pi)^{\frac{d}{2}} \mu_W \delta^{\{y=x\}} = (2\pi)^{\frac{d}{2}} ((2\pi)^{-\frac{d}{2}} \operatorname{Leb}) \delta^{\{y=x\}} = \delta^{\{y=x\}} = C_W. \tag{3.70}$$

Hence, *the Fourier Transform of a White Noise is a White Noise*. This is a particularly interesting property of the White Noise arising from the relation between its covariance measure and the inner-product on $L^2(\mathbb{R}^d)$. Another property related with this is that the White Noise is, up to a multiplicative constant, *the only stationary orthogonal Random Measure*. This can be concluded considering that every locally finite measure over $\mathbb{R}^d$ which is invariant under translations is proportional to the Lebesgue measure. Hence, the weight measure of a stationary orthogonal Random Measure must be proportional to the Lebesgue measure to satisfy the invariance under translation imposed by the stationarity. From this it also follows that *the White Noise is the only slow-growing orthogonal Random Measure such that its Fourier Transform is also an orthogonal Random Measure*, up to a multiplicative constant. ▫

3.5 Stochastic Partial Differential Equations

The notion of SPDE presented here is not the most general conception of such a concept. The main difference between our definition and others that can be found in Stochastic Analysis and its applications, is that we require the involved operator to be a deterministic operator. Other branches of Stochastic Analysis do not require that, and the cases where the operator is also a random object are numerous. We mention for instance the concept of *stochastic homogenezation*, which arises when dealing for example in problems of diffusion in random media; see Armstrong et al. (2017) for an exposition of this theory and its applications. SPDEs determined by a *multiplicative noise* are also excluded. We restrain our work to deterministic operators and hence a SPDE will be simply a *PDE with Generalized Random Fields involved*.

We call a **SPDE** over $\mathbb{R}^d$ an equation of the form

$$\mathcal{L}U = X, \tag{3.71}$$

where both U and X are GeRFs over $\mathbb{R}^d$, and $\mathcal{L} : D \subset \mathscr{S}'(\mathbb{R}^d) \mapsto \mathscr{S}'(\mathbb{R}^d)$ is a mapping defined over a subset of the space of tempered distributions such that its action over GeRFs is well-defined[19]. We do not

[19] **Important remark:** here we have done an explicit and *shameless* abuse of language. Since the operator $\mathcal{L}$ is not necessarily a *differential* operator, this equation is not a *SPDE, stricto sensus*. We will nevertheless maintain this abuse of language anyway. The main motivation of doing this is (like almost every motivation on terminology selection), a *social one*: the term *SPDE Approach* was forged when working in a framework where the stochastic equations used did not involve strictly speaking differential operators (namely, in the case of the SPDE related to the Matérn model, Eq. (1.1)), but it was anyway popularised with the term *SPDE*. We decided to maintain this *popular* name. If the reader is not satisfied with this usage, we suggest to consider that the letter *P* in *SPDE*

suppose, for instance, that $\mathcal{L}$ is linear, nor continuous, but we will assume that it is well-defined for GeRFs without entering at this stage in further details. The GeRF X in equation (3.71) is called the **source term**. As it may be expected, we will fix the source term and try to find a GeRF U which satisfies (3.71).

We need more precisions about this concept, namely, in which sense do we interpret the equality (3.71). We will use the following terminology: for a fixed source term X, we say that a GeRF U satisfies (3.71) **strictly** if

$$\langle \mathcal{L}U, \varphi \rangle \overset{a.s.}{=} \langle X, \varphi \rangle, \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^d). \tag{3.72}$$

In the language of stochastic processes, this is equivalent to require that the process $(\langle \mathcal{L}U, \varphi \rangle)_{\varphi \in \mathscr{S}(\mathbb{R}^d)}$ is a *modification* of the stochastic process $(\langle X, \varphi \rangle)_{\varphi \in \mathscr{S}(\mathbb{R}^d)}$. When working with Random Functions or Random Measures, we use an analogue definition replacing φ with points in the space (or more precisely, with Dirac measures at points in the space) or with indicator functions of bounded Borel sets, respectively.

The strict sense (3.72) is the strongest notion of SPDE we will use in this work. However, in some cases we will work with weaker conditions. We say that a GeRF U satisfies (3.71) **in law** if for every finite vector of test-functions $(\varphi_1, ..., \varphi_N) \in \mathscr{S}(\mathbb{R}^d)^N$, $N \in \mathbb{N}_*$, we have

$$(\langle \mathcal{L}U, \varphi_1 \rangle, ..., \langle \mathcal{L}U, \varphi_N \rangle) \overset{law}{=} (\langle X, \varphi_1 \rangle, ..., \langle X, \varphi_N \rangle), \tag{3.73}$$

where the equality $\overset{law}{=}$ means an equality in law of the random vectors involved, that is, that they have the same probability law over $\mathbb{R}^N$. In such a case, we write

$$\mathcal{L}U \overset{law}{=} X. \tag{3.74}$$

This notion of SPDE is only useful to describe the behaviour of the law of the stochastic process U when it is supposed to satisfy Eq. (3.71). It does not describe any kind of *equality* between the random variables involved as in equation (3.72). Indeed, if we suppose U to satisfy (3.71) in law, we do not even impose any equality conditions between the random variables obtained when evaluating U and X over test-functions: both processes can even be independent and still satisfy (3.73).

An even weaker but useful condition is the following one: a GeRF U satisfies (3.71) in the **second-order sense**, if both $\mathcal{L}U$ and X *have the same mean and covariance distributions*, explicitly if

$$m_{\mathcal{L}U} = m_X \quad ; \quad C_{\mathcal{L}U} = C_X. \tag{3.75}$$

In such a case, we write

$$\mathcal{L}U \overset{2nd\ o.}{=} X. \tag{3.76}$$

---

stands for *"Pseudo-"*, and $\mathcal{L}$ can be called a *pseudo-differential operator*.

This notion is even weaker than the solution in law since it only involves the first two moments. In some cases both conditions are equivalent, for example for Gaussian processes.

Let us consider right now the well-defined and specified case where $\mathcal{L}$ is a linear operator defined through an adjoint. If U satisfies Eq. (3.76), then from equations (3.52) and (3.75), we conclude that m_U and C_U must satisfy the following deterministic PDEs:

$$\mathcal{L}m_U = m_X \quad ; \quad (\mathcal{L} \otimes \overline{\mathcal{L}})C_U = C_X. \tag{3.77}$$

The lesson of this section lies on this equation: in the second order framework, the SPDE relating X to U leads to usual PDEs relating the means and the covariance structures. Hence, the problem of describing covariance models through a SPDE is related to a problem involving deterministic PDEs.

We recall that if condition (3.77) holds, this does not imply a strict equality between $\mathcal{L}U$ and X. If this stronger equality is desired, one must analyse the *cross-covariance relationship* between $\mathcal{L}U$ and X. We will specify this notion in Section 3.6.

3.6 SPDEs and bivariate models

We now relate the SPDE approach to an important branch of Geostatistics: multivariate Geostatistics. Although we are not going to explicitly enter in the framework of multivariate models in this work, we will show in this section that the SPDE Approach can be used as an inspiration to describe cross-covariances models in a quite simple way. We will introduce the notion of a bivariate geostatistical model in the context of GeRFs. The classical case of Random Functions follows immediately.

Let X and Y be two real GeRFs over $\mathbb{R}^d$. We call the **cross-covariance Kernel** between X and Y the application $K_{X,Y} : \mathscr{S}(\mathbb{R}^d) \times \mathscr{S}(\mathbb{R}^d) \to \mathbb{C}$ defined through

$$K_{X,Y}(\varphi, \phi) = \mathbb{C}ov(\langle X, \varphi \rangle, \langle Y, \phi \rangle), \quad \varphi, \phi \in \mathscr{S}(\mathbb{R}^d). \tag{3.78}$$

Following the same arguments as in Section 3.4.1, one concludes that $K_{X,Y}$ defines a separately continuous sesquilinear form. From the Nuclear Theorem 2.2.2 there exists a unique distribution $C_{X,Y} \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^d)$ such that

$$\langle C_{X,Y}, \varphi \otimes \overline{\phi} \rangle = K_{X,Y}(\varphi, \phi) \quad \forall \varphi, \phi \in \mathscr{S}(\mathbb{R}^d). \tag{3.79}$$

The distribution $C_{X,Y}$ is called the **cross-covariance distribution between X and Y**. This distribution is real since both X and Y are supposed to be real GeRFs. An analogue definition holds for the covariance Kernel and distribution *in the reverse sense* Y, X, that is, $K_{Y,X}$ and $C_{Y,X}$. By definition of the covariance

and the reality of $C_{Y,X}$, it holds that

$$\langle C_{X,Y}, \varphi \otimes \overline{\phi} \rangle = \overline{\langle C_{Y,X}, \phi \otimes \overline{\varphi} \rangle} = \langle C_{Y,X}, \overline{\phi} \otimes \varphi \rangle, \quad \forall \varphi, \phi \in \mathscr{S}(\mathbb{R}^d). \tag{3.80}$$

In this context there is also a positive-definiteness condition that $C_{X,Y}$ must satisfy, but it is not a condition in $C_{X,Y}$ alone but on the whole system of bi-variate distributions $C_X, C_{X,Y}, C_{Y,X}, C_Y$. Specifically, any random vector of the form $(\langle X, \varphi \rangle, \langle Y, \phi \rangle)$, with $\varphi, \phi \in \mathscr{S}(\mathbb{R}^d)$, must have a positive-definite covariance matrix, and hence it must hold that

$$\langle C_X, \varphi \otimes \overline{\varphi} \rangle + \langle C_{X,Y}, \varphi \otimes \overline{\phi} \rangle + \langle C_{Y,X}, \phi \otimes \overline{\varphi} \rangle + \langle C_Y, \phi \otimes \overline{\phi} \rangle \geqslant 0, \quad \forall \varphi, \phi \in \mathscr{S}(\mathbb{R}^d). \tag{3.81}$$

The linearity of the distributions involved allows us to conclude that an analogue equation holds for every possible (finite-dimensional) random vector consistent in evaluations of the GeRFs X and Y over arbitrary test-functions[20]. A system of four distributions $C_X, C_{X,Y}, C_{Y,X}, C_Y \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^d)$ that satisfies (3.81) and (3.80) is called a **valid system of cross-covariances**. This is the required condition that a geostatistician must have in mind when selecting a model (that is, selecting covariances and cross-covariances distributions) in a bivariate framework. If we consider for any $\varphi, \phi \in \mathscr{S}(\mathbb{R}^d)$ the complex number $\lambda_{\varphi,\phi} \in \mathbb{C}$ to be such that $|\lambda_{\varphi,\phi}| = 1$ and such that $\lambda_{\varphi,\phi} \langle C_{X,Y}, \varphi \otimes \overline{\phi} \rangle = |\langle C_{X,Y}, \varphi \otimes \overline{\phi} \rangle|$, one proves by using $-\lambda_{\varphi,\phi}\varphi$ instead of φ in Eq. (3.81) and using the relationship (3.80), that a system of distributions $C_X, C_{X,Y}, C_{Y,X}, C_Y \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^d)$ is a valid system of cross-covariances if and only if

$$|\langle C_{X,Y}, \varphi \otimes \overline{\phi} \rangle| \leqslant \frac{\langle C_X, \varphi \otimes \overline{\varphi} \rangle + \langle C_Y, \phi \otimes \overline{\phi} \rangle}{2}, \quad \forall \varphi, \phi \in \mathscr{S}(\mathbb{R}^d). \tag{3.82}$$

This criterion does not use $C_{Y,X}$ since it is completely determined by $C_{X,Y}$ from (3.80).

The cross-covariance distribution $C_{X,Y}$ describes thus the interactions between the processes X and Y beyond their own covariance structures. For example, $C_{X,Y} = 0$ implies that the GeRFs X and Y are uncorrelated, and hence independent if we assume them to be real Gaussian processes. If we choose $C_{X,Y}$ appropriately, we can guarantee that both GeRFs are *equal*, in the sense that one is a modification of the other. More precisely, two real GeRFs X and Y satisfy $\langle X, \varphi \rangle \overset{a.s.}{=} \langle Y, \varphi \rangle$ for all $\varphi \in \mathscr{S}(\mathbb{R}^d)$ if and only if

$$m_X = m_Y \quad \text{and} \quad C_X = C_Y = C_{X,Y}. \tag{3.83}$$

The necessity is straightforward. The sufficiency arises immediately from the analysis of $\mathbb{E}(|\langle X, \varphi \rangle - \langle Y, \varphi \rangle|^2)$. With this fact in mind, it is straightforward to obtain a necessary and sufficient condition for

[20] More precisely, set $\varphi = \sum_{j=1}^N \lambda_j \varphi_j$ and $\phi = \sum_{k=1}^M \omega_k \phi_k$, with $(\varphi_1, ..., \varphi_N) \in \mathscr{S}(\mathbb{R}^d)^N$, $(\phi_1, ..., \phi_M) \in \mathscr{S}(\mathbb{R}^d)^M$, $(\lambda_1, ..., \lambda_N) \in \mathbb{C}^N$ and $(\omega_1, ..., \omega_M) \in \mathbb{C}^M$. Eq. (3.81) gives then the variance of $\sum_{j=1}^N \lambda_j \langle X, \varphi_j \rangle + \sum_{k=1}^M \omega_k \langle Y, \phi_k \rangle$, and hence the possible quadratic forms of the associated random vector $(\langle X, \varphi_1 \rangle, ..., \langle X, \varphi_N \rangle, \langle Y, \phi_1 \rangle, ..., \langle Y, \phi_M \rangle)$.

a GeRF U to satisfy *strictly* the SPDE (3.71) expressed through the first and second order structures of U and the source term X. A GeRF U satisfies Eq. (3.71) strictly if and only if

$$m_{\mathcal{L}U} = m_X \quad \text{and} \quad C_{\mathcal{L}U} = C_X = C_{\mathcal{L}U,X}. \tag{3.84}$$

Hence, when $\mathcal{L}$ is linear and defined through an adjoint, we obtain that U satisfies (3.71) strictly if and only if the following PDEs hold:

$$\mathcal{L}m_U = m_X \quad ; \quad (\mathcal{L} \otimes \overline{\mathcal{L}})C_U = C_X = (\mathcal{L} \otimes \mathcal{I})C_{U,X}, \tag{3.85}$$

where $\mathcal{I} : \mathscr{S}'(\mathbb{R}^d) \to \mathscr{S}'(\mathbb{R}^d)$ denotes the identity operator. This expression can be obtained following the same procedure as in (3.51). Hence, the SPDE is fully described by the determination of the covariances and cross-covariances of the GeRFs U and X, which follow suitable PDEs.

These results show the intimate relationship between the SPDE Approach and multivariate Geostatistics. Indeed, we can describe a bivariate model either *classically* by setting a system of cross-covariances distributions, or instead with the SPDE Approach, by setting a model for one variable and then choosing a SPDE relating the two GeRFs. From this fact arises an idea worth to be discussed: *the real interest of the SPDE Approach is in multivariate Geostatistics*. Indeed, more interesting than having a particular SPDE for a given model, is to have a SPDE *relating two different variables of interest directly*. This approach allows to discriminate between different bivariate covariance models based for example on traditional physical models, where the variables involved are related through a PDE. Cokriging techniques can then be adapted by taking advantage of the relationship defined by the specified SPDE. This approach has already been worked out in Dong (1990), where estimation methods for variables related through PDEs are developed, with applications to the Poisson equation and to PDEs from Hydrogeology.

3.7 A general comment on Random Functions, Measures and Distributions

In this section we make some general comments regarding the stochastic framework presented above and we compare it to other approaches that can be found in the literature.

In this chapter we have defined Random Functions, Measures and Distributions in order to well-define classical operations of PDE analysis on them. We have restrained ourselves to a *mean-square analysis*, where all random variables are square-integrable and all convergences involved are interpreted in the sense of $L^2(\Omega, \mathcal{A}, \mathbb{P})$. This has an important implication: we have never worked with, strictly speaking, continuous functions, measures or distributions. Indeed, the sample-paths of the processes involved are not required to satisfy themselves any particular regularity property. Let us be explicit.

Consider our probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Consider, first of all, a real and continuous Random Function

$(Z(x))_{x\in\mathbb{R}^d}$ according to the terminology fixed in Section 3.2.2. As every stochastic process, Z can be interpreted as a function $Z : \Omega \times \mathbb{R}^d \to \mathbb{R}$, for which the functions $x \in \mathbb{R}^d \mapsto Z(\omega, x)$ for a fixed $\omega \in \Omega$ are called the *sample-paths* of Z. Although we have supposed the mean and covariance functions to be continuous, *this does not imply that the sample-paths are continuous functions*. Hence, the interpretation that we use of this process as a *continuous random function* is rather special: we do not guarantee that our function will have continuous realisations in general, but we continue to *manipulate it* as a continuous function. This works quite well in order to define integrals, limits, derivatives in a regular case, and with a little more generality, to define linear operations. However, some non-linear operations that we do with deterministic continuous functions are not necessarily well-defined in this mean-square framework. For instance, calculating the maxima of a Random Function over a bounded subset of $\mathbb{R}^d$ is not necessarily a well-defined operation without supposing extra regularity conditions on the sample-paths of the process. The typical approach in Stochastic Calculus is to use processes which have a *modification* with almost-surely continuous sample-paths. In this stricter framework, properties of boundedness of the Random Functions, for example, are usually better described than in a mean-square analysis. We refer to Sobczyk (1991, Chapter II) for a general exposition of both the mean-square approach and the almost-surely continuous sample-path approach. We refer to (Øksendal, 2003, Chapter 2) for the concept of modification with almost-surely continuous sample-paths.

The case of Random Measures is not better. Actually, a theory of mean-square-Random Measures, where the $\sigma-$additivity is considered in the sense of $L^2(\Omega, \mathcal{A}, \mathbb{P})$ is not a quite standard framework. Although the concepts necessary to describe it and to establish an associated Integration Theory are an immediate application of the Dunford-Schwartz integral (Dunford & Schwartz, 1958), the only big treaty we could find which works completely in this framework is the recent Rao (2012). It is easier to find authors who work with Random Measures in a stricter sense. If $(M(A))_{A\in\mathcal{B}_B(\mathbb{R}^d)}$ is a process, it is required that the function $M : \Omega \times \mathcal{B}_B(\mathbb{R}^d) \to \mathbb{C}$ must be such that $M(\omega, \cdot)$ defines a measure for every $\omega \in \Omega$. A huge literature can be found for this stricter framework. See for instance the recent big treaty Kallenberg (2017) and the references therein. An older bibliographical source which is rather at the beginnings of such a theory is Morando (1969). This theory may seem more intuitive to work with since any realisation of the process is actually a measure for which all the concepts and developments of the deterministic Measure Theory can be applied, including in particular, the concepts of measure of total variation and Jordan decomposition. Nevertheless, it turns out that this theory is so *strict* that many typical models of Random Measures are excluded. For instance, Gaussian orthogonal Random Measures cannot satisfy such a restriction, unless their weight measures are sums of punctual masses (Horowitz, 1986). Hence, the Gaussian White Noise is outside this framework, and in general, every Gaussian stationary Random Field whose spectral measure has a density is also excluded. Moreover, even without the Gaussian hypothesis, it can be proven that in order to construct an orthogonal Random Measure which follows this strict definition of Random Measure, the process *must be* a Point Process (Kingman, 1967). Hence, the notion of an orthogonal Random Measure

acting "continuously" over the measurable space is lost. In order to not lose this important concept, we did not consider this framework, and we decided to restrain ourselves to the mean-square analysis where *linear operations* such as integrals and evaluations at bounded Borel sets work fine. However, we have lost in general the concept of a Random Measure of Total Variation bounding any arbitrary Random Measure (See for example Øksendal, 2003, Exercise 2.17). We do not know if a less restrictive condition such as having a *modification* which is almost-surely a measure, analogously to the idea of having an almost-surely continuous modification in the case of Random Functions, provides a more exploitable framework. We ignore if an associated theory has already been developed.

The case of Random Distributions is quite special. Although we have defined the linearity and continuity of a GeRF $(\langle Z, \varphi\rangle)_{\varphi\in\mathscr{S}(\mathbb{R}^d)}$ just in a mean-square sense, a framework with a *strict* distributional behaviour can be constructed. Let us be precise. Consider a distribution $m_Z \in \mathscr{S}'(\mathbb{R}^d)$ and a distribution $C_Z \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^d)$ defining a positive-definite Kernel. Consider the measurable space $(\mathscr{S}'(\mathbb{R}^d), \mathcal{B}(\mathscr{S}'(\mathbb{R}^d)))$, where $\mathcal{B}(\mathscr{S}'(\mathbb{R}^d))$ denotes the Borel $\sigma-$algebra of $\mathscr{S}'(\mathbb{R}^d)$ equipped with the weak-* topology. Let us set this measurable space as our *probabilisable* space $(\Omega, \mathcal{A}) = (\mathscr{S}'(\mathbb{R}^d), \mathcal{B}(\mathscr{S}'(\mathbb{R}^d)))$. It can be proven, at least in a Gaussian framework[21], that there exists a unique probability measure over $(\Omega, \mathcal{A})$, denoted by $\mathbb{P}_{m_Z,C_Z}$, such that the application $Z : (\omega, \varphi) \in \Omega \times \mathscr{S}(\mathbb{R}^d) \mapsto \langle\omega, \varphi\rangle$ defines a Gaussian process with mean m_Z and covariance C_Z. To be precise, for every $\omega \in \Omega = \mathscr{S}'(\mathbb{R}^d)$, $Z(\omega, \cdot)$ is simply the distribution ω, and for all $\varphi \in \mathscr{S}(\mathbb{R}^d)$ the random variable $Z(\cdot, \varphi)$, which we will denote by $\langle Z, \varphi\rangle$, satisfies

$$\mathbb{E}\left(e^{-i\langle Z,\varphi\rangle}\right) := \int_{\mathscr{S}'(\mathbb{R}^d)} e^{-i\langle\omega,\varphi\rangle} d\mathbb{P}_{m_Z,C_Z}(\omega) = e^{-i\langle m_Z,\varphi\rangle - \frac{1}{2}\langle C_Z,\varphi\otimes\overline{\varphi}\rangle}. \tag{3.86}$$

Hence, the random variables $(\langle Z, \varphi\rangle)_{\varphi\in\mathscr{S}(\mathbb{R}^d)}$ are Gaussian since their characteristic functions are the ones associated to the Gaussian law, following given mean and covariance structures. This result is known as the *Bochner-Minlos Theorem*, which is actually a generalization of Bochner's Theorem, since it involves a "*Fourier Transform*" of a finite measure over the space $\mathscr{S}'(\mathbb{R}^d)$, and whose result is a *positive-definite continuous functional* over $\mathscr{S}(\mathbb{R}^d)$. In conclusion, a strictly speaking *Random Distribution* can be constructed, as a process following desired mean and covariance structures and whose realisations *are* tempered distributions. Hence, every well-defined operation over tempered distributions from the deterministic world can be applied to such a process. A relatively simple proof of the Bochner-Minlos Theorem can be found in Holden et al. (2009, Appendix A), stated in the particular case where $m_Z = 0$ and C_Z is the covariance of the White Noise, although the authors mention that the right side of (3.86) can be replaced with any continuous positive-definite functional over $\mathscr{S}(\mathbb{R}^d)$ whose evaluation at 0 equals 1. In particular, C_Z can be taken to be any distribution defining a positive-definite Kernel.

[21] The Gaussianity condition is actually not needed. The only necessary mathematical tool is the definition of a continuous positive-definite functional over $\mathscr{S}(\mathbb{R}^d)$ such that its evaluation at 0 equals 1. A Gaussian functional provides easily such a functional, but there are other options. For example, we can construct functionals associated to characteristic functions of the form (B.6) presented in Appendix B.

This difference between the case of Random Distributions and the cases of Random Functions and Random Measures relies on the *lack of a Nuclear Theorem such as* Theorem 2.2.2. Indeed, the Bochner-Minlos Theorem is actually stated in a more general framework where the probability space Ω is the dual of a *Nuclear space*. A Nuclear space is, *very roughly speaking*, a space where a Nuclear Theorem such as 2.2.2 holds. The theory of Nuclear spaces was developed by Alexandre Grothendieck while searching for a general class of spaces where an analogue to the Nuclear Theorem applies (Grothendieck, 1955). See Trèves (1967, Part III) for a deep exposition of the theory of Nuclear spaces and its relation with Nuclear Theorems. Minlos then developed an extension to Bochner's Theorem considering a Fourier Transform over Nuclear spaces to show the existence of convenient probability measures over these spaces. It turns out that the *nuclearity* of a space is actually a necessary condition to obtain such a result as Bochner-Minlos Theorem (Cartier, 1963). The spaces of test-functions $\mathscr{E}(\mathbb{R}^d), \mathscr{S}(\mathbb{R}^d)$ and $\mathscr{D}(\mathbb{R}^d)$ are Nuclear spaces, for which there is a Nuclear Theorem and the Bochner-Minlos Theorem applies to define a probability measure over their dual spaces. It is known that infinite dimensional Banach spaces *are not* nuclear (Trèves, 1967, Corollary 2 to Proposition 50.2), and hence this procedure fails to determine a probability measure on the dual of such spaces. In particular, Hilbert spaces do not satisfy this property.

We conclude this section with a final argument to support our use of a mean-square analysis in this work. We *do not need* to have processes with sample-paths determining, strictly speaking, functions, measures or distributions: we only need things which *act like that* in some particular useful way and for which the operations involved in the SPDEs considered in this work are well-defined. This always holds in our mean-square framework even if the things *are not necessarily* functions, measures or distributions, since we can anyway apply the linear operators described in this chapter in complete analogy to the deterministic case. However, we must confess that we can do like this only because we restrain ourselves to SPDEs defined through linear deterministic operators: if we would like to consider non-linear or non-deterministic operators, for example taking extrema of continuous Random Functions, or working with linear SPDEs involving a multiplicative noise, then a stricter framework with more conditions on the sample-paths is needed.

3.8 Comments on stochastic integrals and non-linear SPDEs

We conclude this chapter with comments on stochastic integrals and more general theories of SPDEs than the framework used in this work. This section is quite apart from the rest of this dissertation and it can be skipped in a first reading. Here we will sometimes make some claims without giving proofs, since this subject goes beyond the scope of this dissertation.

We have defined integrals of Random Functions with respect to deterministic measures (Section 3.2.2) and of deterministic functions with respect to Random Measures (Section 3.3.3). What we have not done, which is a crucial difference between our approach and other branches of Stochastic Calculus, is the defini-

tion of a *stochastic integral of a Random Function with respect to a Random Measure*. Such a definition is actually the starting point of Itô calculus and other approaches. The key issue is that there is not a canonical way to define such an integral. We could follow, for example, an approach defining the integral through a Riemann sequence of partitions of the domain, together with associated tag points, and take a limit in some sense. However, it turns out that in general such a limit *depends on the selection of the Riemann sequence of partitions and on the selection of the tag points.* We will detail this issue. See Øksendal (2003, Chapters 2 and 3) for a complete exposition of the Itô integral over $\mathbb{R}^+$. Here we present just the main facts without deep precision nor proofs.

Let us first of all explain in broad terms the framework of Itô Calculus over $\mathbb{R}$ in its most basic form. It is *based* on Brownian Motion, a particular Gaussian Random Function over $\mathbb{R}$, but we will actually present it *in our way*. Consider a real Gaussian White Noise over $\mathbb{R}$, $(W(A))_{A\in\mathcal{B}_B(\mathbb{R})}$. If we define the Random Function $B := (B(t))_{t\in\mathbb{R}}$ as $B(t) := W([\![0,t]\!])$ if $t \geqslant 0$ and $B(t) := -W((t,0))$ if $t < 0$, it can then be proven that B is a zero-mean continuous Random Function whose derivative in distributional sense is W[22]. The covariance function of B is given by $C_B(t,s) = \min(t,s)$ if $t,s \geqslant 0$, $C_B(t,s) = \max(t,s)$ if $t,s < 0$, and $C_B(t,s) = 0$ otherwise. Hence, B is a *centred Brownian Motion*. What is referred to as an *integral with respect to a White Noise* in this dissertation (Section 3.3.3) is named *integral with respect to Brownian Motion* in Itô calculus.

Basic Itô Calculus fix W as *the* Random Measure with respect to which we make the integrations. It also describes the class of Random Functions that can be integrated with respect to W in some sense. The definition of the Itô Integral is actually quite simple. Consider a Random Function $(Z(t))_{t\in\mathbb{R}}$ which we suppose can be integrated with respect to W (Øksendal, 2003, Definition 3.14). Consider $I \subset \mathbb{R}^+$ a bounded interval. The Itô Integral of Z with respect to W over I is simply the integral obtained as a typical limit of Riemann sums using a Riemman sequence of partitions of I consistent in subintervals, and *using as tag points the left limits of the intervals in the partition.* The result is a Random Variable denoted by $\int_I Z(t)dB(t)$. The limit is defined in the sense of $L^2(\Omega,\mathcal{A},\mathbb{P})$, and Z must have some conditions for this limit to be well-defined. The selection of the left limit is crucial since the limit depends upon the tag points. Another example of stochastic integral in this aim is the *Stratonovich integral*, which defines the stochastic integral in the same analogous way but considering as tag points the *middle-points* of the intervals in the partition. Both definitions differ in general, and actually, any change in the choice of the tag points may produce a change in the resulting limit, obtaining another form of a stochastic integral. The typical example

[22] In a general deterministic framework, a *primitive in distributional sense* of a complex measure over $\mathbb{R}$ is always a function with locally bounded variation (Schwartz, 1966, Theorem II, Chapter II), which can be chosen to be right-continuous and with left-limits (a *càdlàg* function). Indeed, if $\mu \in \mathscr{M}(\mathbb{R})$, then the function $F : \mathbb{R} \mapsto \mathbb{C}$ defined as $F(t) = \mu([\![0,t]\!])\mathbf{1}_{\mathbb{R}^+}(t) - \mu((t,0))\mathbf{1}_{\mathbb{R}^-_*}(t)$ is a càdlàg function which satisfies $-\int_{\mathbb{R}} F(t)\varphi'(t)dt = \int_{\mathbb{R}} \varphi(t)d\mu(t)$ for all $\varphi \in \mathscr{D}(\mathbb{R})$. As a sketch of proof, consider for instance $\varphi \in \mathscr{D}(\mathbb{R})$ such that $\mathrm{supp}(\varphi) \subset \mathbb{R}^+$, the general case following similarly. Using Fubini's Theorem and integration by parts, one concludes that $-\int_{\mathbb{R}+} \mu([\![0,t]\!])\varphi'(t)dt = -\int_{\mathbb{R}+}\int_{\mathbb{R}+} \mathbf{1}_{[0,t]}(s)d\mu(s)\varphi'(t)dt = \int_{\mathbb{R}+} -\int_{\mathbb{R}+} \mathbf{1}_{[s,\infty)}(t)\varphi'(t)dtd\mu(s) = \int_{\mathbb{R}+}\varphi(s)d\mu(s)$. The stochastic analogue (with μ interpreted as a Random Measure) is presented in full detail in Proposition C.2.1 in Appendix C.

of this situation is obtained when trying to integrate a Brownian Motion B with respect to its own derivative which is a White Noise. Rather than exposing it right now, we will later explain *in our way* this dependence on the tag points. The non-uniqueness in this definition pushes the authors to chose between a particular kind of integral, which is often done following practical considerations. We refer to (Øksendal, 2003, end of Chapter 3) for a discussion on this selection.

Let us remark, however, a particular case of stochastic integration where this problem does not actually arise. Let $A \in \mathcal{B}_B(\mathbb{R}^d)$. Let $(V_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$ be a Riemann sequence of partitions of A, and let $(x_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$ be any collection of tag points of $(V_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$. Consider the case where $(Z(x))_{x\in\mathbb{R}^d}$ is a zero-mean continuous Random Function with covariance function C_Z and $(M(A))_{A\in\mathcal{B}_B(\mathbb{R}^d)}$ is a zero-mean Random Measure with covariance measure C_M, *independent of* Z. Hence, the cross-covariance structure between Z and M is null. In this case, we claim without giving a proof that the limits of the form

$$\lim_{N\to\infty}\sum_{j=1}^{N} M(V_j^N)Z(x_j^N) \tag{3.87}$$

exist with the limit being taken *in the sense of* $L^1(\Omega,\mathcal{A},\mathbb{P})$, and the result does neither depend on the choice of the Riemann sequence of partitions of A, nor on the choice of its tag points. The result is then a uniquely defined integrable random variable with zero-mean that we may write $\int_A Z(x)dM(x)$. It can be verified that the application $A \mapsto \int_A Z(x)dM(x)$ defines a Random Measure in a L^1-sense, with the $\sigma-$additivity satisfied in the sense of the first moment. If we suppose more conditions, for example that M and Z are Gaussian processes, the limit can also be taken in the sense of $L^2(\Omega,\mathcal{A},\mathbb{P})$, obtaining a Random Measure for which we can study the covariance structure. Following similar arguments as in Lemma A.5.2 in Appendix A.5, one can show that the covariance Kernel is of the from $(A,B) \mapsto \int_{A\times B} C_Z(x,y)dC_M(x,y)$.

The difference between the well-defined integral $\int_A Z(x)dM(x)$ and the issues exposed at the beginning of this section is the independence condition. In general, if Z and M have a dependence relationship, which could be described through a cross-covariance Kernel, then new things may arise. Consider the *measure-function cross-covariance Kernel* $K_{M,Z} : \mathcal{B}_B(\mathbb{R}^d)\times\mathbb{R}^d \to \mathbb{C}$ defined through $K_{M,Z}(A,x) = \mathbb{C}ov(M(A),Z(x))$. Then, the mean of (3.87) would be, if well-defined, the limit

$$\lim_{N\to\infty}\sum_{j=1}^{N} K_{M,Z}(V_j^N,x_j^N). \tag{3.88}$$

This would define, roughly speaking, a sort of *integral of the measure-function Kernel with respect to himself*. It is not clear if this limit is always well-defined, and it is known that in some cases where it is actually well-defined, the limit depends on the selection of the tag points. The counterexample is the already mentioned case which shows that the Itô integral and the Stratonovich integral differ. Consider $A \subset \mathbb{R}^+$ measurable and bounded with positive Lebesgue measure. Let $M = W$ be a White Noise over $\mathbb{R}$, and let $Z = B$ be

its primitive centred at 0, which is a Brownian Motion. The cross-covariance Kernel satisfies for $t \geqslant 0$, $K_{W,B}(A,t) = \mathbb{C}ov(W(A), B(t)) = \mathbb{C}ov(W(A), W([0,t])) = \mathrm{Leb}(A \cap [0,t])$. If we want to define the integral of B with respect to W over A, we could try it through a Riemman sequence of partitions of A, defining a Riemann sum as in (3.87). Consider the case where such a partition is made through intervals with tag points being their left limits, i.e. $V_j^N = \left[x_j^N, x_{j+1}^N\right)$, then

$$\lim_{N\to\infty} \sum_j K_{W,B}(V_j^N, x_j^N) = \lim_{N\to\infty} \sum_j \mathrm{Leb}(\left[x_j^N, x_{j+1}^N\right) \cap \left[0, x_j^N\right]) = 0, \tag{3.89}$$

while by taking the middle points of the intervals as tag points, writing $V_j^N = \left[a_j^N, b_j^N\right)$ and $x_j^N = \frac{b_j^N + a_j^N}{2}$, one obtains

$$\lim_{N\to\infty} \sum_j K_{W,B}(V_j^N, x_j^N) = \lim_{N\to\infty} \sum_j \mathrm{Leb}(\left[a_j^N, b_j^N\right) \cap \left[0, (b_j^N + a_j^N)/2\right)) = \lim_{N\to\infty} \sum_j \frac{b_j^N - a_j^N}{2} = \frac{\mathrm{Leb}(A)}{2} > 0. \tag{3.90}$$

This shows that the difference in the results when choosing different tag points, and hence the difference between Itô and Stratonovich Integrals is grounded on *the structure of the cross-covariance Kernel*. It is $K_{M,Z}$ and only $K_{M,Z}$ (in this mean-square analysis framework) the mathematical object which determines this distinction. An interesting question that arises is if there are other cases, besides the trivial case of non-correlation, where the cross-covariance Kernel would admit a definition of a stochastic integral without ambiguity. For instance, when the Kernel $K_{M,Z}$ is a tensor product between a measure and a continuous function, or a finite sum of such kinds of Kernels, the stochastic integrals may be uniquely defined, provided that we have a valid system of cross-covariances. In such cases, a framework where Random Functions can be integrated without problems with respect to Random Measures may be developed.

The interest of defining stochastic integrals uniquely and with enough generality does not only come from a mathematical curiosity. In fact, it can be conceived as a *subset* of the problem of defining the *product* of two different Generalized Random Fields, and thus a way to define particular classes of non-linear SPDEs or SPDEs presenting a multiplicative noise. Indeed, maintaining always the analogy with the deterministic framework, the multiplication of two distributions is not always well-defined, but there is a meaning, for example, when one distribution is a measure and the other is a continuous function (using the definition of multiplication measure such as stated in Section 2.1.2). Hence, it is natural to wonder if such a multiplication between Random Measures and Random Functions can be defined through a stochastic integral and if we can use this definition to interpret some products which appear in some SPDEs. The same idea may be applied if one of the GeRFs has a behaviour similar to a member of the class $\mathcal{O}_M(\mathbb{R}^d)$. An example of SPDE that could be included within this framework is the *Diffusion equation with random diffusivity* over the space-time $\mathbb{R}^d \times \mathbb{R}$:

$$\frac{\partial U}{\partial t} - \mathrm{div}(H \nabla U) = X, \tag{3.91}$$

where X is a GeRF over $\mathbb{R}^d \times \mathbb{R}$ and H is a *positive-definite matrix* of GeRFs over $\mathbb{R}^d$. div and ∇ are the divergence and gradient *spatial* operators respectively. This is an example of a linear equation with a *multiplicative noise*. Even in the deterministic case, the meaning of the multiplication $H\nabla U$ must be specified, requiring conditions on H and/or on the solution U. For example, if H is a matrix of measures, ∇U may be a vector of continuous functions, the resulting multiplication being a vector of measures. Even if this multiplication is well-defined, the source term X must also satisfy compatibility conditions if we want Eq. (3.91) to make sense. For example, if $H\nabla U$ is a vector of measures, X must have the behaviour of a *derivative of a measure*. All of these problems get worse when entering into the stochastic world, where even if H defines a matrix of Random Measures and ∇U defines a vector of continuous Random Functions, their product, i.e. the associated vector of stochastic integrals, is not uniquely defined and a convenient framework must be selected. Requiring in addition that X is a White Noise, which is a common practice in Stochastic Analysis, impose even more theoretical problems to this analysis. The same problem is still present if we require H to be a deterministic object but we require X to be a White Noise.

Other SPDEs presenting multiplicative noise can be found in (Holden et al., 2009). In this treaty, the authors base their work on the concept of the *Wick product*, which allows them to define a sort of multiplicative product between stochastic processes with great generality. However, this product does not necessarily follow any analogy with the product of objects in the deterministic case: this product can only be interpreted for stochastic objects, and the realisations of such objects may not possibly be interpreted as functions, measures or distributions, but rather as *averaging values over a space of random variables*. We refer to the introduction in Holden et al. (2009) for a discussion on this approach and its theoretical benefits. However, it is not obvious how to relate this approach to the practice of Geostatistics, since the covariance structure is not a basic tool of this framework, and moreover the interpretation of the realisations of the involved random objects as regionalized variables is unclear or lost.

Consider now an example of a non-linear SPDE, the Kardar-Parisi-Zhang (KPZ) Equation (Kardar et al., 1986), which is a space-time SPDE of the form

$$\frac{\partial U}{\partial t} - \nu \Delta U - \frac{\lambda}{2} |\nabla U|^2 = W, \tag{3.92}$$

where ν, λ are parameters, Δ denotes the (spatial) Laplacian operator, and $|\nabla U|^2$ denotes the squared-Euclidean norm of the (spatial) gradient of U. Hence, in this equation terms of the form $(\frac{\partial U}{\partial x_j})^2$ appear, requiring to give a definition of the *multiplication of the derivatives of U with themselves*. In order to define such a multiplication we could require, for instance, U to be a continuously differentiable Random Function, hence the multiplication of its derivatives is immediate to define. However, the presence of the White Noise at the right side requires that the behaviour of the left side must be a measure which is not determined by a continuous function, hence this regularity restriction to U may not work to define a solution to (3.92). This equation has been an inspiration for an intense work in the *SPDE community*. The analysis of

a well-defined solution to the KPZ equation in the case of spatial dimension 1 has been done by M. Hairer in (Hairer, 2013). The ideas used to define such a solution inspired the development of a more general and sophisticated framework to treat SPDEs which presents some forms of multiplicative products between Random Distributions, called the *Theory of Regularity Structures* (Hairer, 2014). The development of this theory made Hairer be one of the winners of the Fields Medal in 2014.

The problems exposed in this section are then typical, and somewhat basic[23] theoretical issues which are presented in the theory of SPDEs in the wide-sense, which engender a huge need for rather sophisticated theories. In general, it is unclear how to relate the already existing theories to the framework of Geostatistics. For these reasons, in this dissertation we deal only with linear SPDEs defined through a deterministic operator. This framework is at the same time rich and simple enough to allow us to develop and study new interesting geostatistical models without entering into the issues exposed in this section.

[23] *Basic* in the sense that they are problems at the *basis* of the theory, not in the sense that they are easy to deal with...

Chapter 4

Stationary Solutions for a class of SPDEs: existence, uniqueness and examples

SUMMARY

In this chapter we present a result concerning the existence and uniqueness of stationary solutions to a wide class of linear SPDEs. This result encompasses many of the most important cases of stationary models related to SPDEs presented in the literature. It can be considered then, as a review, as a generalization and as a simplification.

In Section 4.1 we present the motivations of these developments and the questions that are tackled.

In Section 4.2 we present the class of linear operators which will determine the class of SPDEs considered in this chapter. It consists in operators acting over the space of tempered distributions such that their Fourier Transforms are slow-growing measures. These operators are defined through the Fourier Transform and a symbol function, which is an Hermitian measurable polynomially bounded function. We describe the action of such a class of operators, which are proven to maintain the stationarity.

In Section 4.3 the main result of this chapter is presented, which is Theorem 4.3.1. We first set the class of considered SPDEs, which consists of linear SPDEs involving an operator defined through a symbol and a stationary source term. Then, we present Theorem 4.3.1, which states that the existence of a stationary solution to a SPDE in our class is equivalent to a slow-growing behaviour requirement of the multiplication between the squared-norm of the reciprocal of the symbol function and the spectral measure of the source term. The uniqueness of such a solution is equivalent to a never-null requirement for the symbol function. We make some remarks con-

cerning similar well-known results in the literature, the characterisation of some cases where existence and uniqueness is assured regardless of the source term, characterisations of stationary solutions to homogeneous cases, and some possible extensions.

In Section 4.4 we remark the case where the source term is a White Noise, which deserves to be considered as a fundamental case. We present Theorem 4.4.1 which states, under suitable conditions, that the covariance distribution of the solution of a general SPDE can be expressed as the convolution between the covariance of the source term and the covariance of the solution with a White Noise source term.

In Section 4.5 we present some examples of applications which are known in the literature and we propose some SPDEs for well-known models whose relation to some SPDEs are not broadly known. We review the popular Matérn model with its typical associated SPDE. We present the case of Matérn models without range parameter. We present the case of stationary Markov Random Fields according to Rozanov's Theory. We also show some examples where we propose some SPDEs for well-known models. We give a SPDE which describes non-exhaustively the $J-$Bessel model. Finally, we propose a SPDE whose unique stationary solution follows a Stein model in a spatio-temporal context.

We finally make a remark about the associated deterministic problem in Section 4.6. We discuss roughly some differences between the approach developed in this chapter and some typical approaches in the theory of PDEs. We also state, in the stochastic context, a result on existence and uniqueness of solutions with non-zero mean and stationary centred form.

The proof of every statement is presented in Appendix A.

4.1 Motivation

The SPDE approach in Geostatistics has been popularised in the last decade since the apparition of the seminal paper of Lindgren et al. (2011). In this work, the authors exploit an already known link between the Matérn model and a particular class of SPDEs (Whittle, 1963). The authors also notably remarked an important fact: the Matérn fields are not necessarily the *only solutions* to the related SPDEs, but rather the only *stationary solutions*. Indeed, for some cases of the involved class of SPDEs, deterministic (or even random) solutions to the associated homogeneous problem can be added, and hence the solution is not unique. We will specify those details further in Example 4.5.1.

The questions that arise are then the next ones: *when do stationary solutions to some particular classes of SPDEs exist? When does exist a unique stationary covariance model that the solutions to such SPDEs must follow? Can we fully characterise the covariance structure of these models by taking advantage from the fact that they solve a particular SPDE?*

In this section we answer these questions. For this purpose, we restrain ourselves to a particular class of linear SPDEs which are particularly adapted to treat the problem of stationary solutions, and which gives an enough rich and general framework to encompass many well-known results relating stationary covariance models to SPDEs. The results presented in this chapter will be a basis for further developments in this dissertation. It will allow us, in particular, to construct new stationary models with non-trivial properties and to verify when do stationary solutions to some important physically driven SPDEs exist.

All along this chapter we suppose that all of the random objects have null mean, except in Section 4.6. If X and Y are two GeRFs, the notation "$X = Y$" means that X is a modification of Y. We will anyway recall this particular equivalence meaning when we feel it is necessary, in order to avoid any confusion.

4.2 A class of linear operators

Let us consider the next subspace of tempered distributions:

$$\mathscr{V}'(\mathbb{R}^d) := \{T \in \mathscr{S}'(\mathbb{R}^d) \mid \mathscr{F}(T) \in \mathscr{M}_{SG}(\mathbb{R}^d)\} = \mathscr{F}^{-1}(\mathscr{M}_{SG}(\mathbb{R}^d)). \tag{4.1}$$

It is immediate that $\mathscr{V}'(\mathbb{R}^d)$ is a vector subspace of $\mathscr{S}'(\mathbb{R}^d)$. The choice of this space has been done on purpose: it was conceived considering the fact that the Fourier Transform of a stationary GeRF is a slow-growing Random Measure (Theorem 3.4.2). Hence, it is expectable that this space will be quite useful when working with stationary Random Fields. We remark from the Bochner-Schwartz Theorem 3.4.1 that every positive-definite distribution is in $\mathscr{V}'(\mathbb{R}^d)$. Actually, since a slow-growing measure can be decomposed in four positive slow-growing measures (see Section 2.1.3), $\mathscr{V}'(\mathbb{R}^d)$ is the complex span of the cone of positive-definite distributions.

We will define a class of linear operators which can be applied over distributions on $\mathscr{V}'(\mathbb{R}^d)$. Let $g : \mathbb{R}^d \to \mathbb{C}$ be a polynomially bounded measurable function. We define the operator $\mathcal{L}_g : \mathscr{V}'(\mathbb{R}^d) \to \mathscr{V}'(\mathbb{R}^d)$ as

$$\mathcal{L}_g(T) = \mathscr{F}^{-1}(g\mathscr{F}(T)), \quad \forall T \in \mathscr{V}'(\mathbb{R}^d). \tag{4.2}$$

Let us analyse this definition. First, since $\mathscr{F}(T) \in \mathscr{M}_{SG}(\mathbb{R}^d)$ and g is measurable and polynomially bounded, the multiplication $g\mathscr{F}(T)$ is a well-defined slow-growing measure (Section 2.1.3). Its Inverse Fourier Transform $\mathscr{F}^{-1}(g\mathscr{F}(T)) = \mathcal{L}_g(T)$ is well-defined in the sense of distributions and it is an element of $\mathscr{V}'(\mathbb{R}^d)$. The operator $\mathcal{L}_g$ is thus well-defined. It is also immediate, that it is a linear operator. Due to the properties of the Fourier Transform, the operator $\mathcal{L}_g$ is real if and only if g is an Hermitian function. An Hermitian polynomially bounded measurable function $g : \mathbb{R}^d \mapsto \mathbb{C}$ will be from now on said to be a **symbol function over** $\mathbb{R}^d$. In such a case, the associated operator $\mathcal{L}_g$ defined through (4.2) will be said to be an *operator defined through a symbol*, and we will say that g is *the symbol* of $\mathcal{L}_g$. We remark that every differ-

ential operator D^α, with $\alpha \in \mathbb{N}^d$, is an operator defined through a symbol. Indeed, by using the property of the Fourier Transform with respect to derivation, $\mathscr{F}(D^\alpha T) = (i\xi)^\alpha \mathscr{F}(T)$ for $T \in \mathscr{S}'(\mathbb{R}^d)$, it follows that $D^\alpha = \mathcal{L}_g$ for $g(\xi) = (i\xi)^\alpha$, $\xi \in \mathbb{R}^d$. Other examples of such operators will be worked out in Section 4.5 and further in this dissertation.

Let us now consider the application of $\mathcal{L}_g$ over a stationary GeRF Z over $\mathbb{R}^d$. We have the following property:

Proposition 4.2.1. *Let Z be a real stationary GeRF over $\mathbb{R}^d$ with covariance distribution ρ_Z and spectral measure μ_Z. Let g be a symbol function over $\mathbb{R}^d$ and let $\mathcal{L}_g$ be its associated operator. Then, $\mathcal{L}_g Z$ is a real stationary GeRF over $\mathbb{R}^d$ with covariance distribution $\rho_{\mathcal{L}_g Z} = \mathcal{L}_{|g|^2}\rho_X$ and spectral measure $\mu_{\mathcal{L}_g Z} = |g|^2\mu_Z$.*

See Appendix A.9 for a proof. Hence, operators of the form $\mathcal{L}_g$ *maintain the stationarity*, and are applicable to any real stationary GeRF without restriction. We remark that the expression of the spectral measure $\mu_{\mathcal{L}_g Z} = |g|^2\mu_Z$ is particularly simple while in general the expression of the covariance $\rho_{\mathcal{L}_g Z} = \mathcal{L}_{|g|^2}\rho_Z$ may be more complicated since $\mathcal{L}_{|g|^2}$ may not be an operator simple to deal with. It turns out that in this framework it will be easier to work with spectral measures rather than with the covariances. For simplicity, from now on every even measure in $\mathscr{M}^+_{SG}(\mathbb{R}^d)$ will be said to be a *spectral measure* over $\mathbb{R}^d$.

4.3 Associated SPDEs: an existence and uniqueness Theorem

Let g be a symbol function over $\mathbb{R}^d$. Let X be a stationary GeRF over $\mathbb{R}^d$. Consider the following SPDE which involves X as source term:

$$\mathcal{L}_g U = X. \tag{4.3}$$

We recall that with our notation, Eq. (4.3) means that the equality must be strict (see Section 3.5), meaning that $\mathcal{L}_g U$ is a modification of X. The question that arises is to establish under which conditions there exists a stationary GeRF U solution to (4.3) or not, whether it is unique and, when solutions exist, whether we can characterize their covariance structures.

In order to obtain conditions about the resolvability of Eq. (4.3), let us first of all analyse sufficient conditions to solve the weaker equation

$$\mathcal{L}_g U \overset{2nd\ o.}{=} X. \tag{4.4}$$

From Proposition 4.2.1, it follows that if U is a solution to (4.4), then the next PDE must hold:

$$\mathcal{L}_{|g|^2}\rho_U = \rho_X. \tag{4.5}$$

Equivalently, the next expression relating the spectral measures must be satisfied:

$$|g|^2\mu_U = \mu_X. \tag{4.6}$$

Eq. (4.6) is a multiplicative equation, so it is more immediate to treat. Such a problem is called in Distribution Theory a *division problem*. We can argue that there exists a stationary solution to (4.4) if there exists a *spectral measure* solution to (4.6). Intuitively, we could require μ_U to be the multiplication between $|g|^{-2}$ and μ_X. This multiplication can *always* be done, the result $|g|^{-2}\mu_X$ being in general a positive not-necessarily Radon measure. It is even since μ_X is even and g is Hermitian. Hence, a possible criterion of existence of a stationary solution to (4.3) is to require the measure $|g|^{-2}\mu_X$ to be in $\mathscr{M}^+_{SG}(\mathbb{R}^d)$. The next Theorem states that actually, this condition is necessary and sufficient for the existence of a *strict* stationary solution to (4.3), and provides in addition a criterion to determine when the solution is unique, together with the specification of its spectral measure.

Theorem 4.3.1. *Let X be a real stationary GeRF over $\mathbb{R}^d$ with spectral measure μ_X. Let g be a symbol function over $\mathbb{R}^d$ and let $\mathcal{L}_g$ be its associated operator. Then, there exist real stationary solutions to the SPDE* (4.3) *if and only if there exists $N \in \mathbb{N}$ such that*

$$\int_{\mathbb{R}^d} \frac{d\mu_X(\xi)}{|g(\xi)|^2(1+|\xi|^2)^N} < \infty. \tag{4.7}$$

In such a case, there is a unique up to a modification real stationary solution to (4.3) *if and only if $|g| > 0$. If this holds, the unique real stationary solution U to* (4.3) *has spectral measure*

$$\mu_U = |g|^{-2}\mu_X. \tag{4.8}$$

The proof of this theorem relies, roughly speaking, in the correct definition of the application of an operator of the form $\mathscr{F}^{-1}(\frac{1}{g}\mathscr{F}(\cdot))$ over X. Since $\frac{1}{g}$ is not necessarily polynomially bounded, such an operator is not necessarily of the form (4.2) and hence it cannot be in general applied to *any* stationary GeRF. The details of the proof are presented in Appendix A.10. We make the following remarks:

Remark 4.3.1. When $N = 0$ in (4.7), i.e. if $|g|^{-2}$ is integrable with respect to the measure μ_X, the measure μ_U is finite and the solution U is thus a mean-square continuous random function. This case was studied in Whittle (1963), restricted to the SPDE in the second-order sense (4.4). In his same work, Whittle mentioned that solutions corresponding to non-finite measures μ_U still make sense in some framework, the theory of which was at that time not completely available. Our work can be seen as one possible answer to this note.

Remark 4.3.2. A sufficient condition for existence and uniqueness of a strict stationary solution to (4.3), *regardless of the source term X*, is to require that $|g|$ is inferiorly bounded by the inverse of a strictly positive polynomial. Indeed, in such a case $\frac{1}{g}$ is a symbol function, and hence $\mathcal{L}_{\frac{1}{g}}$ is an operator defined through a

symbol. The operator $\mathcal{L}_g : \mathscr{V}'(\mathbb{R}^d) \to \mathscr{V}'(\mathbb{R}^d)$ is actually bijective, with inverse operator $\mathcal{L}_g^{-1} = \mathcal{L}_{\frac{1}{g}}$. This implies that Eq. (4.3) can be solved explicitly by setting simply $U = \mathcal{L}_{\frac{1}{g}} X$. U is then the unique stationary solution and its spectral measure is given by (4.8), following Proposition 4.2.1. We shall henceforth refer to this condition as the *Polynomially Bounded Reciprocal* condition on g, abbreviated as the ***PBR condition*** on g. We will also say that g *has a PBR*, in such a case.

Remark 4.3.3. When the measurable set $g^{-1}(\{0\}) = \{\xi \in \mathbb{R}^d \mid g(\xi) = 0\}$ is non-empty, the non-uniqueness is due to the existence of stationary solutions to the homogeneous problem

$$\mathcal{L}_g U_H = 0. \tag{4.9}$$

Indeed, for a spectral measure μ_{U_H} over $\mathbb{R}^d$ concentrated on $g^{-1}(\{0\})$, its associated stationary random field satisfies strictly Eq. (4.9), since $\mu_{\mathcal{L}_g U_H} = |g|^2 \mu_H = 0$. Thus, if existence is provided, the sum of any stationary solution to (4.3) with a non-trivial independent stationary solution to (4.9) is also a stationary solution to (4.3). This remark is an inspiration for describing stationary solutions to homogeneous problems, and we will use it extensively.

Remark 4.3.4. Theorem 4.3.1 has been stated under the polynomially bounded condition on g in order to freely apply $\mathcal{L}_g$ to any stationary GeRF. However, if we restrict the domain of definition of $\mathcal{L}_g$, we can include some new SPDEs, valid for more restricted classes of stationary GeRFs. For instance, let us suppose that U is a stationary Random Function following a Gaussian covariance, with the same parameters as in Eq. (3.4). Then, it is clear that U satisfies a SPDE of the form

$$\mathscr{F}^{-1}(g \mathscr{F}(U)) = W, \tag{4.10}$$

with g being the function

$$g(\xi) = (2\pi)^{\frac{d}{4}} \frac{2^{\frac{d}{4}}}{\sigma a^{\frac{d}{2}}} e^{\frac{a^2}{8}|\xi|^2}, \quad \xi \in \mathbb{R}^d. \tag{4.11}$$

This function is not polynomially bounded, hence the potential associated operator $\mathcal{L}_g$ cannot be applied to any arbitrary stationary GeRF, as well as it cannot be applied to any arbitrary distribution in $\mathscr{V}'(\mathbb{R}^d)$. However, it *can* be applied to a Random Function with this Gaussian covariance, obtaining as a result a White Noise.

4.4 A fundamental case: White Noise source term

In this section we will present a result concerning the particular case of equation (4.3) when the source term X is a real White Noise W. We will see that the covariance structure of the solution in the cases with other source terms can be related to the solution of the White Noise case in a convenient way. This can be seen

as a special and important property of the White Noise, which can be added to all the properties given in Examples 3.3.1, 3.3.4 and 3.4.7. We recall that W is a stationary GeRF with covariance distribution $\rho_W = \delta$ and with spectral measure $d\mu_W(\xi) = (2\pi)^{-\frac{d}{2}} d\xi$.

We then focus on the equation

$$\mathcal{L}_g U = W. \tag{4.12}$$

Theorem 4.3.1 allows to conclude that there exist stationary solutions to (4.12) if and only if the measure $(2\pi)^{-\frac{d}{2}}|g(\xi)|^{-2}d\xi$ is in $\mathscr{M}_{SG}^{+}(\mathbb{R}^d)$. Let us suppose this holds. From Proposition 4.2.1, every stationary solution (4.12) must have a covariance distribution satisfying

$$\mathcal{L}_{|g|^2}\rho_U = \rho_W = \delta. \tag{4.13}$$

It turns out that solutions to the deterministic equation (4.13) can be seen as *Green's Functions* of the operator $\mathcal{L}_{|g|^2}$, concept which is used in the theory of PDEs in order to obtain *fundamental solutions* to some class of PDE. These fundamental solutions are used to construct solutions to more general forms of the PDE, usually through a convolution. From this typical application of the theory of Green's Functions arises the idea that a solution to the more general case $\mathcal{L}_{|g|^2}\rho_U = \rho_X$ may be expressed as a convolution between ρ_X and the solution to (4.13). It is then expected that such a solution would be the covariance distribution of the more general SPDE (4.3). The next result presents some cases where this idea holds regardless of the source term X.

Theorem 4.4.1. *Let X be a real stationary GeRF over $\mathbb{R}^d$ with covariance distribution ρ_X. Let g be a symbol function over $\mathbb{R}^d$ satisfying at least one of the following requirements:*

1. $\frac{1}{g} \in \mathcal{O}_M(\mathbb{R}^d)$;
2. *there exists $N \in \mathbb{N}$ such that $\mathscr{F}\left((1+|x|^2)^{-N}|g|^{-2}\right) \in L^1(\mathbb{R}^d)$.*

Let $\mathcal{L}_g$ be the associated operator. Then, there exists a unique stationary solution to the SPDE (4.3)*, and its covariance distribution is given by*

$$\rho_U = \rho_U^W * \rho_X, \tag{4.14}$$

where ρ_U^W denotes the covariance distribution of the unique stationary solution to (4.12)*.*

The proof of this Theorem can be found in Appendix A.11. It is based on the idea that when there exists a stationary solution to (4.3), the measure $\mu_U := |g|^{-2}\mu_X$ is slow-growing and it is the *multiplication* of $|g|^{-2}$ and μ_X. Since $|g|^{-2}$ is, up to a multiplicative constant, the spectral measure of a solution to (4.12), the convolution relation comes from an application of an *Exchange Formula* of the Fourier Transform, if the convolvability between the Fourier Transforms of $|g|^{-2}$ and μ_X is satisfied. We make the following remarks.

Remark 4.4.1 (Remark of academic honesty). A result on Distribution Theory which we found very lately (while writing this manuscript, a few weeks before its delivery) is the one stated in Richards & Youn (1995, Example 3, Section 6, Chapter 7). Using suitable and not-quite-traditional-but-intuitive definitions of the multiplicativity and convolvability between tempered distributions, the authors state that *any* continuous and polynomially bounded function is *multiplicable* (in their sense of multiplication) with any finite measure, and hence their Fourier Transforms are *convolvable*, satisfying the Exchange Formula. Hence, with just a few more arrangements, a more general and simpler form of Theorem 4.4.1 can be stated as follows:

Let X be a real stationary GeRF over $\mathbb{R}^d$ with covariance distribution ρ_X. Let g be a symbol function over $\mathbb{R}^d$ such that $|g|^{-2}$ is continuous and polynomially bounded. Let $\mathcal{L}_g$ be its associated operator. Then, there exists a unique stationary solution to the SPDE (4.3)*, and its covariance distribution is given by $\rho_U = \rho_U^W * \rho_X$, where ρ_U^W denotes the covariance distribution of the unique stationary solution to* (4.12)*.*

Conditions 1 and 2 actually imply that $|g|^{-2}$ is continuous and polynomially bounded (see the arguments in the proof of Theorem 4.4.1 in Appendix A.11). Hence, this statement is more general than Theorem 4.4.1.

The reason why we have decided to present this result in the weaker form of Theorem 4.4.1 is founded on two motivations: first, *we do not know in detail* the theory of multiplication and convolution presented in Richards & Youn (1995), hence we cannot really justify the result; second, and more important, the result regarding the multiplicavility between polynomially bounded continuous functions and finite measures is not proven in the book Richards & Youn (1995), it is left as an exercise. Since we have not done this exercise, we do not feel with the right to state the result in its general form.

We remark however a curiosity: the continuity condition is imposed to $|g|^{-2}$ and not to g. Hence, g may have an irregular behaviour. The continuity of $|g|^{-2}$ cannot be immediately dropped out, since the Exchange Formula of the Fourier Transform does not hold in general for the multiplication (in our sense) between a measurable polynomially bounded function and a slow-growing measure. As counter-example, consider the function $f = \mathbf{1}_{\{0\}}$ and the Dirac measure δ. Then $f\delta = \delta$, so $\mathscr{F}(f\delta) = (2\pi)^{-\frac{d}{2}}$, but $\mathscr{F}(f) = 0$ in distributional sense, and hence $\mathscr{F}(f) * \mathscr{F}(\delta) = 0$.

Remark 4.4.2. The conditions on g are imposed, as already said, to obtain a condition *regardless* of the source term X, and hence ρ_X can be any positive-definite distribution. If ρ_X is in a particular class of positive-definite distributions, then other less restrictive conditions may be required on g in order to obtain an analogue result to Theorem 4.4.1. In particular, it could be argued that, as long as ρ_U^W exists and it is convolvable with ρ_X, an analogue result may hold. This can be studied, for example, in a classical framework of convolvability between functions.

It can be thus concluded that the case with a White Noise source term is of big importance in the analysis of the covariance structures of solutions to Eq. (4.3). The connection with the concept of *Green's Function* and *fundamental solutions* to PDEs justifies the use of the expression *"fundamental case"* in this context.

Because of this, many of the examples of geostatistical models related to SPDEs presented in this dissertation will be studied mainly using a White Noise source term.

4.5 Examples

The framework presented in this chapter encompasses many models already developed in the literature. Indeed, the *only* things that our analysis has provided is the formalization and generalization of the idea of defining a stationary covariance model whose spectral measure has a density with respect to other spectral measure, and a way to relate this covariance model to a particular class of SPDEs. We refer to some bibliographical sources where similar approaches have been applied to construct models, all of them considering also an associated SPDE.

- Whittle (1963) for the general case where the spectral measure of the solution μ_U is finite (Remark 4.3.1).
- Heine (1955) describing models associated to second order differential operators in dimension 2 and Vecchia (1985) for models obtained from compositions of such type of operators.
- Anh et al. (1999) and Gay & Heyde (1990) for cases associated to fractional Laplacian operators, and (Kelbert et al., 2005) for their generalization used to describe fractional forms of the Heat equation.
- Bolin & Lindgren (2011) and Lim & Teo (2009) for more general forms of the Matérn model.
- Jones & Zhang (1997) for examples on a spatio-temporal context.

In this section we will detail some of those examples and we will also present relations of some known geostatistical models with SPDEs which are not present in the literature. Other examples will be detailed further in this dissertation.

Example 4.5.1 (**Matérn Model**). As a first example, we start with the well-known and increasingly popular Matérn model (See Eq. (3.6)). The relationship between the Matérn Model and the SPDE over $\mathbb{R}^d$

$$(\kappa^2 - \Delta)^{\frac{\alpha}{2}} U = W, \tag{4.15}$$

with $\kappa > 0$, $\alpha \in \mathbb{R}$ has been established a long time ago (Whittle, 1963) and recently revisited and exploited in Lindgren et al. (2011). This relationship can be easily re-obtained from Theorem 4.3.1. Indeed, the operator $(\kappa^2 - \Delta)^{\frac{\alpha}{2}}$ is nothing but an operator of the form (4.2) with symbol function $g(\xi) = (\kappa^2 + |\xi|^2)^{\frac{\alpha}{2}}$. This function satisfies the PBR condition, and hence there exists a unique stationary solution to (4.15).

Following Eq. (4.8), the spectral measure of this solution is

$$d\mu_U^W(\xi) = \frac{d\xi}{(2\pi)^{\frac{d}{2}}(\kappa^2 + |\xi|^2)^{\alpha}}. \tag{4.16}$$

If $\alpha > \frac{d}{2}$, the measure μ_U^W is finite, and it is exactly the spectral measure of the Matérn Model (Eq. (3.6), with $a = 1$). When $\alpha \leqslant \frac{d}{2}$, we still obtain a unique stationary solution, defined as a GeRF. We refer to this model as the *generalized Matérn Model*. The associated covariance distribution $\rho_U^W = \mathscr{F}(\mu_U^W)$ is called the *generalized Matérn covariance*.

We remark that g is actually a function in $\mathcal{O}_M(\mathbb{R}^d)$, and so does its reciprocal. Hence, the operator $(\kappa^2 - \Delta)^{\frac{\alpha}{2}}$ is actually a bijective operator from $\mathscr{S}'(\mathbb{R}^d)$ to $\mathscr{S}'(\mathbb{R}^d)$, not only in $\mathscr{V}'(\mathbb{R}^d)$. In particular, g satisfies condition 1 in Theorem 4.4.1. Hence, for any real stationary GeRF X, the SPDE

$$(\kappa^2 - \Delta)^{\frac{\alpha}{2}} U = X \tag{4.17}$$

has a unique stationary solution whose covariance is the convolution between ρ_X and the generalized Matérn covariance.

We finally remark that, for example, for $\alpha = 2$, the functions of the form $f(x) = ae^{\kappa v^T x}, x \in \mathbb{R}^d$, with $a \in \mathbb{R}$ and $v \in \mathbb{R}^d$ with $|v| = 1$, are solutions to the homogeneous equation associated to Eq. (4.15). We can also make the parameters a and v be random variables, so we would obtain a Random Function solution to such homogeneous equation. However, those solutions are not stationary (they are not even *tempered*). This is the importance of the stationarity assumption in the researched solutions, as it was pointed out in Lindgren et al. (2011). ▫

Example 4.5.2 (**Matérn Model without range parameter**). The condition $\kappa > 0$ in the Matérn SPDE defined in Eq. (4.15) can be relaxed. Setting $\kappa = 0$, we obtain a fractional Laplacian operator $(-\Delta)^{\frac{\alpha}{2}}$, which is an operator of the form $\mathcal{L}_g$ with symbol function $g(\xi) = |\xi|^{\alpha}$ for $\alpha > 0$. Let us thus consider the SPDE

$$(-\Delta)^{\frac{\alpha}{2}} U = W. \tag{4.18}$$

In Theorem 4.3.1, the existence condition (4.7) requires that there exists $N \in \mathbb{N}$ such that the integral $\int_{\mathbb{R}^d}(1 + |\xi|^2)^{-N}|\xi|^{-2\alpha}d\xi$ is finite. Because of the singularity at the origin, this is only possible if $\alpha < d/2$. In this case, the spectral measure of a particular stationary solution to the SPDE (4.18) is

$$d\mu_U(\xi) = \frac{1}{(2\pi)^{\frac{d}{2}}} \frac{d\xi}{|\xi|^{2\alpha}}. \tag{4.19}$$

The associated covariance distribution is its Fourier Transform, which is the locally integrable function (see

Donoghue, 1969, Chapter 32):

$$\rho_U(h) = \frac{1}{\pi^{\frac{d}{2}}} \frac{\Gamma(\frac{d}{2} - \alpha)}{\Gamma(\alpha)} \frac{1}{|h|^{d-2\alpha}}, \quad h \in \mathbb{R}^d. \tag{4.20}$$

The function ρ_U in (4.20) is not defined at $h = 0$. It is not continuous, but it is still positive-definite in distributional sense. The associated GeRF cannot thus be interpreted as a mean-square continuous Random Function. This is an example of the kinds of covariance structures we obtain when working with non-finite spectral measures. Such models are said to have a long-range dependence behaviour. They have been studied in Anh et al. (1999) and in Gay & Heyde (1990), in which the SPDE (4.18) is specified with a slightly different definition of the operator $(-\Delta)^{\frac{\alpha}{2}}$.

We remark that the symbol function $g(\xi) = |\xi|^\alpha$ has a zero at the origin. Hence, the uniqueness condition does not hold. The stationary solution associated to the covariance (4.20) is not the unique possible solution. To describe all possible stationary solutions, we follow Remark 4.3.3 and we consider spectral measures which are supported at the origin, i.e., which are proportionals to the Dirac measure $\mu_{U_H} = a\delta$, with $a > 0$. The associated covariance distributions are then constant positive functions, and thus the associated GeRFs are *random constants*, that is, Random Functions of the form $U_H(x) = A$, for all $x \in \mathbb{R}^d$, with A being a centred random variable with variance $(2\pi)^{-\frac{d}{2}} a$. In other words, *the only stationary solutions to the homogeneous equation* $(-\Delta)^{\frac{\alpha}{2}} U_H = 0$ *are random constants*.

Another consequence of the fact that g equals zero at the origin is that Theorem 4.4.1 cannot be applied since $|g|^{-2}$ is not continuous. If it is desired to describe the covariance of a potential stationary solution to a SPDE of the form $(-\Delta)^{\frac{\alpha}{2}} U = X$ through a similar principle, a convolvability condition between ρ_X and the covariance (4.20), when it exists, must be satisfied. ▫

Example 4.5.3 (**Markov Models**). Let $p : \mathbb{R}^+ \to \mathbb{R}^+_*$ be a strictly positive polynomial over $\mathbb{R}^+$. We consider the SPDE over $\mathbb{R}^d$

$$p^{\frac{1}{2}}(-\Delta) U = W, \tag{4.21}$$

where the operator $p^{\frac{1}{2}}(-\Delta)$ is of the form (4.2) with symbol function $g(\xi) = p^{\frac{1}{2}}(|\xi|^2)$. Since p is strictly positive, the PBR condition holds. Hence, the SPDE (4.21) has a unique stationary solution with spectral measure

$$d\mu_U^W(\xi) = \frac{1}{(2\pi)^{\frac{d}{2}}} \frac{d\xi}{p(|\xi|^2)}. \tag{4.22}$$

This is a measure whose density is the reciprocal of a strictly positive and isotropic polynomial. Rozanov's Theorem (Y. A. Rozanov, 1982, Section 3.2.3) allows to conclude that this model is an *isotropic stationary Markov Random Field* (MRF). In Rozanov's Theory, a MRF is, roughly speaking, a GeRF such that for every domain of $\mathbb{R}^d$, *evaluations* of the random field on the interior of the domain are independent upon evaluations on the interior of the complement of the domain, conditionally to the behaviour of the random

field on a neighbourhood of the boundary of the domain. By *evaluations*, we mean the action of the GeRF over test-functions whose supports are included in the interior of the corresponding set. Rozanov's Theorem states that every stationary MRF has a spectral measure whose density is the inverse of a strictly positive polynomial. Thus, in the case of isotropic models, MRFs satisfy equation (4.21). An anisotropic model can be obtained by applying an anisotropy matrix to ξ in Eq. (4.22). See Y. A. Rozanov (1982) for a complete theory of MRFs which also uses the theory of GeRFs, or J. A. Rozanov (1977) for a shorter exposition which also includes the relation of MRFs with some SPDEs.

Note that g satisfies condition 1 in Theorem 4.4.1. Hence, for any real stationary GeRF X, there exists a unique stationary solution to the SPDE

$$p^{\frac{1}{2}}(-\Delta)U = X, \tag{4.23}$$

whose covariance is the convolution between ρ_X and the covariance of the MRF solution to Eq. (4.21). ▫

Example 4.5.4 (**The J-Bessel Model**)**.** Let U be a real Random Function over $\mathbb{R}^d$ following the J-Bessel model (See Eq. (3.7), we will follow the same parametrization). Since its spectral measure is proportional to the uniform measure over the $d-1$-sphere of radius $\kappa > 0$, $\partial B_\kappa^{(d)}(0)$, it follows that for any symbol function g which equals 0 over $\partial B_\kappa^{(d)}(0)$, U satisfies $\mathcal{L}_g U = 0$. This is immediate following Remark 4.3.3 since in such a case μ_U is concentrated on $g^{-1}(\{0\})$. In particular, the $J-$Bessel model satisfies the homogeneous SPDE

$$(\kappa^2 + \Delta)U = 0, \tag{4.24}$$

case for which the associated symbol function is $g(\xi) = \kappa^2 - |\xi|^2$. We remark that this is not the unique homogeneous SPDE that is satisfied by the $J-$Bessel model since there are many symbol functions which are 0 over the $d-1$-sphere. In addition, this equation is not only restricted to the $J-$Bessel model: any stationary GeRF with spectral measure concentrated on $\partial B_\kappa^{(d)}(0)$ also satisfies it. However, since the Fourier Transform of a distribution invariant under rotations is also invariant under rotations, *any stationary GeRF with isotropic covariance and satisfying* (4.24) *follows a $J-$Bessel covariance model*. We remark that SPDE (4.24) does not tell us anything about the variance of U, nor the extra isotropic condition. Hence, more restrictions on U must be required in order to fix its variance. ▫

Example 4.5.5 (**The Stein Model**)**.** Here we present a *spatio-temporal* example. We work on $\mathbb{R}^d \times \mathbb{R}$ with d being the spatial dimension. The variable $\xi \in \mathbb{R}^d$ denotes a variable of the spatial frequency domain, and $\omega \in \mathbb{R}$ denotes a variable of the temporal frequency domain. Rather than starting from a SPDE and describing its potential solution we make the procedure backwards: we *start* with a known covariance model and we propose a SPDE that a GeRF following this model must satisfy, similarly to what we have done in Example 4.5.4.

Let us consider the spectral measure over $\mathbb{R}^d \times \mathbb{R}$ proposed in Stein (2005):

$$d\mu_U(\xi,\omega) = \frac{1}{(2\pi)^{\frac{d+1}{2}}} \frac{d\xi d\omega}{(b(s^2+\omega^2)^\beta + a(\kappa^2+|\xi|^2)^\alpha)^\nu}, \tag{4.25}$$

with $a, b > 0$, $s^2 + \kappa^2 > 0$, and $\alpha, \beta, \nu \in \mathbb{R}$, satisfying that $\beta \geqslant 0$ if $s^2 = 0$ and $\alpha \geqslant 0$ if $\kappa^2 = 0$. This is always a well-defined spectral measure over $\mathbb{R}^d \times \mathbb{R}$, being finite when $\alpha, \beta, \nu > 0$ and $\frac{1}{\beta\nu} + \frac{d}{\alpha\nu} < 2$ (Stein, 2005), case in which the associated Random Function is said to follow a *Stein covariance model*. When (4.25) is not finite, we say that the associated GeRF follows a *generalized Stein covariance model*. Except for some particular values of the parameters, there is no closed-form expression for the covariance.

We then consider the spatio-temporal symbol function

$$g(\xi,\omega) = (b(s^2+\omega^2)^\beta + a(\kappa^2+|\xi|^2)^\alpha)^{\nu/2}, \quad (\xi,\omega) \in \mathbb{R}^d \times \mathbb{R}. \tag{4.26}$$

With the conditions required on the parameters, this symbol function satisfies the PBR condition, hence any SPDE involving the associated operator $\mathcal{L}_g$ has a unique stationary solution. The form of g allows us to write its associated operator in terms of fractional second-order differential operators. Using a spatio-temporal White Noise, W (i.e., with spectral measure $d\mu_W(\xi,\omega) = (2\pi)^{-(d+1)/2} d\xi d\omega$), a corresponding SPDE for the Stein model is

$$\left(b\left(s^2 - \frac{\partial^2}{\partial t^2}\right)^\beta + a\left(\kappa^2 - \Delta\right)^\alpha \right)^{\nu/2} U = W. \tag{4.27}$$

Hence, the unique stationary solution to Eq. (4.27) follows a generalized Stein covariance model.

When $\kappa, s, a, b > 0$ and α, β, ν are not null, the symbol function (4.26) satisfies condition 1 in Theorem 4.4.1. Hence, for any stationary GeRF X, the SPDE

$$\left(b\left(s^2 - \frac{\partial^2}{\partial t^2}\right)^\beta + a\left(\kappa^2 - \Delta\right)^\alpha \right)^{\nu/2} U = X \tag{4.28}$$

has a unique stationary solution whose covariance is the convolution between ρ_X and the covariance of the generalized Stein model.

We finally remark that the models proposed in Kelbert et al. (2005, Section 3) are Stein models in the particular cases where $s = 0$ and $\beta = 1$. In such cases, the SPDE (4.27) may be re-written using first-order temporal differential operators. ▫

4.6 Remark: the deterministic problem

In this section we make a simple remark about the *deterministic* problem associated to Eq. (4.3) and Theorem 4.3.1. It can be considered as a non-geostatistical analysis, but it can also be applied to describe the means of non-centred GeRFs related through a SPDE of the form (4.3), analysis that we will also do.

Considering that Theorem 4.3.1 has been obtained following the idea that $\mathscr{F}(U)$ acts as a slow-growing measure over $\mathbb{R}^d$, it is quite intuitive that we can obtain the following deterministic result. We will use lower case letters to denote deterministic distributions.

Proposition 4.6.1. *Let $f \in \mathscr{V}'(\mathbb{R}^d)$. Let g be a symbol function over $\mathbb{R}^d$ and let $\mathcal{L}_g$ be its associated operator. Then, there exist solutions in $\mathscr{V}'(\mathbb{R}^d)$ to the equation*

$$\mathcal{L}_g u = f \tag{4.29}$$

if and only if $\frac{1}{g}$ is locally integrable with respect to $\mathscr{F}(f)$ and the multiplication measure $\frac{1}{g}\mathscr{F}(f)$ is slow-growing. If this holds, there is a unique solution in $\mathscr{V}'(\mathbb{R}^d)$ to (4.29) *if and only if $|g| > 0$. In such a case, the solution is given by*

$$u = \mathscr{F}^{-1}\left(\frac{1}{g}\mathscr{F}(f)\right). \tag{4.30}$$

The proof of this result is very similar to the proof of Theorem 4.3.1 and it is presented in Appendix A.12. We can also state analogous remarks to those proposed for Theorem 4.3.1:

- There exist solutions in the space $\mathscr{F}(\mathscr{M}_F(\mathbb{R}^d))$ if and only if $|g|^{-1}$ is integrable with respect to $\mathscr{F}(f)$.
- If g satisfies the PBR condition, the solution is unique regardless of the source term f since $\mathcal{L}_g$ is bijective.
- The non-uniqueness when $g^{-1}(\{0\}) \neq \varnothing$ is explained through the existence of solutions in $\mathscr{V}'(\mathbb{R}^d)$ to the homogeneous problem when $f = 0$, which can be obtained by using a slow-growing measure concentrated on $g^{-1}(\{0\})$.
- More general results can be stated if we require some extra conditions on f.

The motivation to state this remark is to make a sort of *comparison* between the analysis we have done in this chapter and more typical analysis presented in the theory of deterministic PDEs. In the deterministic case, there are often many possible solutions to a proposed PDE, and one manner of *selecting* one of the possible solutions is by imposing a condition to the behaviour of the solution at the *boundary* of the working domain. For example, an initial condition is often used in the case of Ordinary Differential Equations or spatio-temporal PDEs, and in the case of spatial PDEs, Dirichlet or Neumann type conditions are usually a

basis of analysis for existence and uniqueness. These conditions are often inspired by our *"knowledge"* of some physical conditions in a particular phenomenon. For instance, the *initial condition* requirement is based on the idea that *we "can know"*, at least up to some precision, the present or past state of a system, while we *"cannot know"* the future state. Dirichlet conditions are used when *we "know"* the values of the interest variable at the boundary of the domain, and Neumann conditions are used when we do not necessarily know the values of the variable at the boundary but rather *we "know" some other physical condition* acting on it, like an impermeability condition. At the end of the day, all of these considerations are used in order to *fix a subspace* of possible solutions where there may be a unique solution to the PDE and hence to work with this solution. In this chapter we have done something different. The reason why we *selected* a particular solution of the PDE was not founded on physical considerations but rather on a *statistical methodology consideration*, namely, that the studied variable can be described by a stationary geostatistical model. Hence, the only assumption we do about the variable is that *it behaves in a "similar manner" all along the domain*. This was the inspiration of the use of the space $\mathscr{V}'(\mathbb{R}^d)$ as a basis. We have not imposed boundary conditions but rather the condition of belonging to the space $\mathscr{V}'(\mathbb{R}^d)$, which imposes some conditions on the increasing behaviour and on the regularity of the solution. We *do not know* if this approach is *better*, in some sense, than the classical approach of using boundary conditions and solving Cauchy problems. What we do know is that our proposition is fairly more adapted to traditional geostatistical methodologies. At the end of the story, the practice and contrast with data in some contexts and the exploitability and utility of the selected model are the *only* criteria to discriminate between one model and another, or between one methodology and another.

We finally present the following result which is an immediate consequence of Theorem 4.3.1 and Proposition 4.6.1. Here the GeRFs are not supposed to have zero mean. We omit the proof.

Theorem 4.6.1. *Let X be a real GeRF over $\mathbb{R}^d$ with mean distribution $m_X \in \mathscr{V}'(\mathbb{R}^d)$ and such that $X - m_X$ is a stationary GeRF with spectral measure μ_X. Let g be a symbol function over $\mathbb{R}^d$ and let $\mathcal{L}_g$ be its associated operator. Then, there exits a real GeRF U solution to* (4.3) *with mean distribution $m_U \in \mathscr{V}'(\mathbb{R}^d)$ and such that $U - m_U$ is stationary if, and only if, there exist $N_1, N_2 \in \mathbb{N}$ such that*

$$\int_{\mathbb{R}^d} \frac{d\,|\mathscr{F}(m_X)|\,(\xi)}{|g(\xi)|(1+|\xi|^2)^{N_1}} < \infty \quad \textit{and} \quad \int_{\mathbb{R}^d} \frac{d\mu_X(\xi)}{|g(\xi)|^2(1+|\xi|^2)^{N_2}} < \infty. \tag{4.31}$$

If this holds, up to a modification, there is a unique such a solution if and only if $|g| > 0$. In such a case, it holds that

$$m_U = \mathscr{F}^{-1}\left(\frac{1}{g}\mathscr{F}(m_X)\right) \quad \textit{and} \quad \mu_U = |g|^{-2}\mu_X, \tag{4.32}$$

where μ_U denotes the spectral measure of the stationary GeRF $U - m_U$.

Chapter 5

Spatio-temporal models driven from evolution equations

SUMMARY

In this chapter we focus on spatio-temporal geostatistical models which can be obtained through the SPDE approach.

In Section 5.1 we give a general introduction to space-time Geostatistics. We recall the most important concepts in classical space-time geostatistics and we give a brief exposition of already existing methodologies to construct space-time covariance models. We recall the concepts of separability and symmetry. We present the formalism of spatio-temporal Generalized Random Fields. We present the generalized concepts of separability and symmetry. We give simple criteria to determine if a stationary spatio-temporal GeRF is separable or symmetric through requirements on its spectral measure. We also present the analogue of spatial and temporal margins. In particular, we focus on the case where a stationary spatio-temporal GeRF can be considered as having a continuous point-wise meaning in time.

In Section 5.2 we present new stationary space-time covariance models which can be related to spatio-temporal SPDEs. The general setting consists of equations involving a temporal differential operator of arbitrary real positive order and a spatial operator defined through a symbol. We give conditions when there exists a unique stationary solution to those equations regardless of the source term and of the imaginary part of the symbol, and we specify the associated spectral measure. The separability, symmetry and time regularity is easily described through the properties of the spatial symbol function and the temporal derivative order. We remark the cases of first and second order evolution models for which the spatial covariance structure is

described. The subsections in this chapter are devoted to present particular examples with both physical and statistical interest. We present two examples of already known geostatistical models grounded on physical considerations, namely through an advection-diffusion equation with damping, and through a Langevin equation. We present the case of Evolving Matérn models, which are spatio-temporal GeRFs following a Matérn covariance model in space. We study the existence of stationary solutions to the stochastic Heat Equation, where we obtain the result that there exist stationary solutions to the Heat equation with White Noise source term only for spatial dimension higher than 2*. We study models related to the stochastic Wave equation. We show that we can construct spatio-temporal models solving the homogeneous Wave equation and following an arbitrary spatial covariance structure. These models are called Waving models. We also show that there are no stationary solutions to the Wave equation with White Noise source term.*

In Section 5.3 we present informally the resolution of a Cauchy problem involving a first order evolution equation with a particular initial condition. We start by solving the associated deterministic problem under the requirement that the spatial Fourier Transforms of the source term and of the initial condition must be slow-growing measures. The solution has a càdlàg-in-time representation. Under more restrictive conditions on the source term and on the spatial symbol function, we claim that the solution is spatio-temporally asymptotically convergent as the time flows to the unique tempered solution to the first order evolution equation whose spatio-temporal Fourier Transform is a slow-growing measure. We then present the stochastic analogue, which is done in complete similarity by using GeRFs whose spatial Fourier Transforms are slow-growing Random Measures. In the case with stationary source term and initial condition, we claim that the solution of the Cauchy problem converges spatio-temporally asymptotically to the unique stationary solution to the first order evolution SPDE as the time flows. Under suitable conditions, if the initial condition follows a suitable spatial covariance behaviour, the solution follows this same stationary space-time model. We give some examples using separable stationary source terms. In the particular case of a white in time and coloured in space source term, a Markovianity in time structure is described.

The proofs of the statements presented in Sections 5.1 and 5.2 are given in Appendix A. The complete formalisation of the claims proposed in Section 5.3 is presented in Appendix C.

In this chapter we work in a spatio-temporal context. We will always work over the space-time Euclidean domain $\mathbb{R}^d \times \mathbb{R}$, where $d \in \mathbb{N}_*$ denotes the *spatial* dimension. The variables in the initial (or physical) space-time will be denoted by $(x, t) \in \mathbb{R}^d \times \mathbb{R}$, while variables in the frequency space-time domain will be denoted by $(\xi, \omega) \in \mathbb{R}^d \times \mathbb{R}$. We suppose that all the random objects in this chapter have zero mean.

5.1 Generalities on spatio-temporal geostatistical models

5.1.1 Classical spatio-temporal models

A classical spatio-temporal geostatistical model is a real Random Function Z indexed by the space-time $\mathbb{R}^d \times \mathbb{R}$, $(Z(x,t))_{(x,t)\in\mathbb{R}^d\times\mathbb{R}} \subset L^2(\Omega, \mathcal{A}, \mathbb{P})$. Its covariance function is then a positive-definite Kernel function $C_Z : (\mathbb{R}^d \times \mathbb{R}) \times (\mathbb{R}^d \times \mathbb{R}) \to \mathbb{R}$. In principle, there is no mathematical difference in considering a covariance over $\mathbb{R}^d \times \mathbb{R}^d$ or over $(\mathbb{R}^d \times \mathbb{R}) \times (\mathbb{R}^d \times \mathbb{R})$: we have just added an extra dimension. Hence, classical spatial geostatisical models can be extended to the spatio-temporal case without technical issues. All results we have presented in the previous chapters of this dissertation can be applied to the spatio-temporal case, simply replacing $\mathbb{R}^d$ with $\mathbb{R}^d \times \mathbb{R}$; we specify the notions in the case of spatio-temporal GeRFs in Section 5.1.2. Nevertheless, in practice one needs to find models which are particularly adapted to the case of spatio-temporal phenomena. Namely, it is often expected to use models which behave differently when evolving over time than when changing the space variables. The covariance structure should then reflect these differences. It is also expected that the covariance models involve parameters which can control the statistical properties of the model when changing over time, over space, or over the whole space-time. The difficulty is then, to find valid covariance models, that is, functions which do satisfy the positive-definiteness condition, and still being manipulable enough in order to easily control the parameters of the space-time interactions. Hence, new subtleties arise in the field of spatio-temporal Geostatistics which must be taken into account.

A basic construction of a valid spatio-temporal covariance model is done through the concept of *separability*. A spatio-temporal Random Function Z is said to be *separable* or to have a *separable covariance* if there exists a spatial covariance $C_{Z_S} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ and a temporal covariance $C_{Z_T} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ such that

$$C_Z((x,t)\ ,\ (y,s)) = C_{Z_S}(x,y)C_{Z_T}(t,s), \quad \forall (x,y) \in \mathbb{R}^d \times \mathbb{R}^d, \forall (t,s) \in \mathbb{R} \times \mathbb{R}. \tag{5.1}$$

When Z is a stationary separable Random Function, the stationary covariance function can be expressed as

$$\rho_Z(h,u) = \rho_{Z_S}(h)\rho_{Z_T}(u), \tag{5.2}$$

for a spatial stationary covariance function ρ_{Z_S} and a temporal stationary covariance function ρ_{Z_T}. This kind of covariance is one of the most basic construction of valid covariance functions over the space-time. It is obtained, for example, when there exist two independent Random Functions, one over the space $(Z_S(x))_{x\in\mathbb{R}^d}$, the another over the time $(Z_T(t))_{t\in\mathbb{R}^d}$, such that $Z(x,t) = Z_S(x)Z_T(t)$.

Separability is an oversimple construction which often fails to reflect the variability of a variable which varies over the space-time. Nevertheless, it is a good starting point to construct more complicated models. An immediate extension is given by considering a so-called *product-sum* model (see for example, De Iaco et

al., 2001), which is given by a finite sum of separable models:

$$C_Z((x,t),(y,s)) = \sum_{j=1}^{N} C_{Z_S,j}(x,y) C_{Z_T,j}(t,s), \tag{5.3}$$

for finite collections of spatial and temporal covariance functions $(C_{Z_S,j})_{j\in\{1,\dots,N\}}$ and $(C_{Z_T,j})_{j\in\{1,\dots,N\}}$ respectively. The case of stationary models follows immediately.

Other more popular class of non-separable stationary covariance models is the *Gneiting class* of covariance models (Gneiting, 2002). A stationary covariance of this class is constructed through a continuous completely monotone function $f_S : \mathbb{R}^+ \to \mathbb{R}^+$ and a positive function $f_T : \mathbb{R}^+ \to \mathbb{R}^+$ with completely monotone derivative, by

$$\rho_Z(h,u) = \frac{\sigma^2}{f_T(|u|^2)^{\frac{d}{2}}} f_S\left(\frac{|h|^2}{f_T(|u|^2)}\right), \quad \forall (h,u) \in \mathbb{R}^d \times \mathbb{R}, \tag{5.4}$$

for some $\sigma^2 > 0$. The spatial and temporal behaviours of the covariance can be easily described through the specification of f_S and f_T. Hence, this construction proposes a general and flexible way of constructing non-separable models.

Other methodology for obtaining non-separable models with a practical parametrization is through the specification of a convenient spectral measure, as it is done for example in the case of the Stein model presented in Example 4.5.5. As mentioned, this controls easily the spatial and temporal regularities of the covariance structure.

Although the Gneiting and Stein classes of covariance models are rich enough to describe some statistical properties of a spatio-temporal variable, both of them have a limitation: they are *symmetric* or *fully-symmetric* models. In a symmetric model, the direction of the time evolution is ignored, obtaining equal covariance values if we look either forward or backward in time. More precisely, a spatio-temporal Random Function Z is said to be spatio-temporally *symmetric* if its covariance function satisfies

$$C_Z((x,t),(y,s)) = C_Z((x,s),(y,t)) = C_Z((y,t),(x,s)) = C_Z((y,s),(x,t)), \tag{5.5}$$

for all $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$ and for all $(t,s) \in \mathbb{R} \times \mathbb{R}$. Although the equality $C_Z((x,t),(y,s)) = C_Z((y,s),(x,t))$ comes from the definition of covariance, the stronger equality (5.5) is an extra requirement. In the stationary case, the symmetry is translated into the condition for the stationary covariance function:

$$\rho_Z(h,u) = \rho_Z(-h,u) = \rho_Z(h,-u) = \rho_Z(-h,-u), \quad \forall (h,u) \in \mathbb{R}^d \times \mathbb{R}. \tag{5.6}$$

Separable models and product-sum models are always symmetric. It is known that symmetric models fail to

describe the variability of a variable submitted to transport phenomena, as can be the case of atmospheric and environmental variables. A well-known example of a non-symmetric stationary model is the one induced by a transport phenomenon according to a constant velocity profile determined by a velocity $v \in \mathbb{R}^d$. For example, if $C_Z : (\mathbb{R}^d \times \mathbb{R}) \times (\mathbb{R}^d \times \mathbb{R}) \to \mathbb{R}$ is a symmetric covariance function, then the *transported* covariance $C_{Z_v} : (\mathbb{R}^d \times \mathbb{R}) \times (\mathbb{R}^d \times \mathbb{R}) \to \mathbb{R}$ defined through $C_{Z_v}((x,t),(y,s)) = C_Z((x-vt,t),(y-vs,s))$ is a non-symmetric covariance. If Z is stationary, the associated transported covariance is also stationary. This and other examples of non-symmetric covariance models induced by a transport phenomenon can be found in Ailliot et al. (2011).

Besides the transport approach of constructing covariance models, in the literature there are few propositions of non-symmetric covariances, often lacking in generality and/or simplicity. Stein (2005) proposes a construction based on the derivatives of a convenient particular non-symmetric covariance Kernel. Zhang & Zhang (n.d.) propose a not-so-simple construction of non-symmetric models with Matérn spatial and temporal margins based on a convenient analogy with the conditional probability density functions of a suitable random vector.

To finish this section, we recall a somewhat obvious and already implicitly introduced but important concept in the case of stationary Random Functions. If Z is a spatio-temporal real stationary Random Function with covariance function ρ_Z, then for every $t \in \mathbb{R}$, the spatial Random Function $Z(\cdot,t)$ is a stationary spatial Random Function. $Z(\cdot,t)$ is said to be a *spatial trace* of Z. All spatial traces of Z are stationary with same stationary spatial covariance function given by $\rho_{Z_S}(h) = \rho_Z(h,0)$ for all $h \in \mathbb{R}^d$. The covariance ρ_{Z_S} is called the *spatial margin* of the covariance ρ_Z. Analogously, for every $x \in \mathbb{R}^d$, the temporal Random Function $Z(x,\cdot)$ is called a *temporal trace* of Z, and all the temporal traces of Z are stationary temporal Random Functions with same covariance function given by $\rho_{Z_T}(u) = \rho_{Z_S}(0,u)$ for all $u \in \mathbb{R}$. ρ_{Z_T} is said to be the *temporal margin* of the covariance function ρ_Z.

We refer to Gneiting et al. (2006) for a more general discussion on spatio-temporal covariance models.

5.1.2 Spatio-temporal GeRFs

A generalized spatio-temporal geostatistical model is a real GeRF Z over $\mathbb{R}^d \times \mathbb{R}$. Its covariance distribution C_Z is then a real distribution belonging to $\mathscr{S}'((\mathbb{R}^d \times \mathbb{R}) \times (\mathbb{R}^d \times \mathbb{R}))$ defining a positive-definite Kernel. The covariance structure in the stationary case is described through a positive-definite even distribution $\rho_Z \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R})$ and a positive even spectral measure $\mu_Z \in \mathscr{M}^+_{SG}(\mathbb{R}^d \times \mathbb{R})$.

Let us set some notations, mostly chosen for ease of reading. We use the letters φ and ϕ to denote spatial test-functions (functions defined over $\mathbb{R}^d$), θ for temporal test-functions (defined over $\mathbb{R}$), and ψ for spatio-temporal test-functions (defined over $\mathbb{R}^d \times \mathbb{R}$). We denote by $\mathscr{F}$ the **spatio-temporal Fourier Transform**, which is applicable to distributions in $\mathscr{S}'(\mathbb{R}^d \times \mathbb{R})$. We denote by $\mathscr{F}_S$ and $\mathscr{F}_T$ the **spatial** and the **temporal**

Fourier Transforms, respectively, which applied to a spatio-temporal test-function $\psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R})$ are defined as:

$$\mathscr{F}_S(\psi)(\xi, t) := \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{-i\xi^T x} \psi(x, t) dx \quad ; \quad \mathscr{F}_T(\psi)(x, \omega) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\omega t} \psi(x, t) dx. \tag{5.7}$$

Using typical arguments which prove the continuity and bijectivity of $\mathscr{F}$ (see Donoghue, 1969, Chapter 30), one can prove that both the spatial and temporal Fourier Transforms are continuous bijective endomorphisms over $\mathscr{S}(\mathbb{R}^d \times \mathbb{R})$. The spatial and temporal Inverse Fourier Transforms, denoted by $\mathscr{F}_S^{-1}$ and $\mathscr{F}_T^{-1}$ respectively, are defined as in (5.7) without the minus sign in the exponentials. If $T \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R})$, its spatial and temporal Fourier Transforms are defined respectively through their applications to a test-function $\psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R})$ by

$$\langle \mathscr{F}_S(T), \psi \rangle := \langle T, \mathscr{F}_S(\psi) \rangle \quad ; \quad \langle \mathscr{F}_T(T), \psi \rangle := \langle T, \mathscr{F}_T(\psi) \rangle. \tag{5.8}$$

Hence, $\mathscr{F}_S$ and $\mathscr{F}_T$ are simply the adjoints of the respective spatial and temporal Fourier Transforms over $\mathscr{S}(\mathbb{R}^d \times \mathbb{R})$, and we also have that both $\mathscr{F}_S$ and $\mathscr{F}_T$ are continuous bijective endomorphisms over $\mathscr{S}'(\mathbb{R}^d \times \mathbb{R})$. The Inverse spatial and temporal Fourier Transforms, denoted by $\mathscr{F}_S^{-1}$ and $\mathscr{F}_T^{-1}$ respectively, are defined analogously.

We will use the following notation concerning tensor products. The symbol $\boxtimes$ will be reserved to denote *spatio-temporal tensor products*, that is, tensor products between two objects, one defined over the space and the other over time. Explicitly, if $S \in \mathscr{S}'(\mathbb{R}^d)$ and $T \in \mathscr{S}'(\mathbb{R})$, then $S \boxtimes T \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R})$. The same idea applies for the tensor product between spatial test-functions (resp. measures) with temporal test-functions (resp. measures): if $\varphi \in \mathscr{S}(\mathbb{R}^d)$ and $\theta \in \mathscr{S}(\mathbb{R})$, then $\varphi \boxtimes \theta \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R})$ (resp., if $\mu_S \in \mathscr{M}_{SG}(\mathbb{R}^d)$ and $\mu_T \in \mathscr{M}_{SG}(\mathbb{R})$, then $\mu_S \boxtimes \mu_T \in \mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R})$). We will always follow the *spatio-temporal writing order*: we write the spatial object on the left side of the tensor product $\boxtimes$ and the temporal on the right side. The symbol $\otimes$ will be reserved for objects acting over the same space, that is, both acting over $\mathbb{R}^d$, over $\mathbb{R}$, or over $\mathbb{R}^d \times \mathbb{R}$. For instance, if Z is a spatio-temporal GeRF, $\varphi_1, \varphi_2 \in \mathscr{S}(\mathbb{R}^d)$ and $\theta_1, \theta_2 \in \mathscr{S}(\mathbb{R})$, then we have the expression

$$\mathbb{C}ov(\langle Z, \varphi_1 \boxtimes \theta_1 \rangle, \langle Z, \varphi_2 \boxtimes \theta_2 \rangle) = \langle C_Z, (\varphi_1 \boxtimes \theta_1) \otimes \overline{(\varphi_2 \boxtimes \theta_2)} \rangle. \tag{5.9}$$

In the next sections we explain the concepts of separability, symmetry, and spatial and temporal margins in the case of a spatio-temporal GeRF.

Separability

A GeRF over $\mathbb{R}^d \times \mathbb{R}$, Z, is said to be **separable** if there exists $C_{Z_S} \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^d)$ and $C_{Z_T} \in \mathscr{S}'(\mathbb{R} \times \mathbb{R})$ such that

$$\langle C_Z, (\varphi_1 \boxtimes \theta_1) \otimes (\varphi_2 \boxtimes \theta_2) \rangle = \langle C_{Z_S}, \varphi_1 \otimes \varphi_2 \rangle \langle C_{Z_T}, \theta_1 \otimes \theta_2 \rangle, \quad \forall \varphi_1, \varphi_2 \in \mathscr{S}(\mathbb{R}^d), \theta_1, \theta_2 \in \mathscr{S}(\mathbb{R}). \tag{5.10}$$

When Z is separable, we denote it by $Z = Z_S \boxtimes Z_T$, Z_S and Z_T *representing*[1] two GeRFs over $\mathbb{R}^d$ and over $\mathbb{R}$ respectively, with covariances C_{Z_S} and C_{Z_T} respectively. In the stationary case, separability is equivalent to require that the stationary covariance distribution $\rho_Z \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R})$ is the spatio-temporal tensor product of two positive-definite even distributions, $\rho_{Z_S} \in \mathscr{S}'(\mathbb{R}^d)$, and $\rho_{Z_T} \in \mathscr{S}'(\mathbb{R})$:

$$\rho_Z = \rho_{Z_S} \boxtimes \rho_{Z_T}. \tag{5.12}$$

Consequently, the spectral measure $\mu_Z \in \mathscr{M}^+_{SG}(\mathbb{R}^d \times \mathbb{R})$ can also be expressed as the spatio-temporal tensor product of a spatial spectral measure $\mu_{Z_S} \in \mathscr{M}^+_{SG}(\mathbb{R}^d)$ and a temporal spectral measure $\mu_{Z_T} \in \mathscr{M}^+_{SG}(\mathbb{R})$,

$$\mu_Z = \mu_{Z_S} \boxtimes \mu_{Z_T}. \tag{5.13}$$

A typical separable model which can be found in the literature is the so-called *white in time and coloured in space noise*, which is a real spatio-temporal stationary stationary GeRF with an arbitrary spatial covariance distribution ρ_{Z_S} and a White Noise in time covariance $\rho_{W_T} = \delta \in \mathscr{S}'(\mathbb{R})$. Such a model is denoted by $Z_S \boxtimes W_T$. See the use of this terminology for example in Sigrist et al. (2015).

[1] We remark what do we mean with Z_S and Z_T *representing* two GeRFs rather than just *being* GeRFs. For simplicity we consider the framework of Gaussian GeRFs. If Z_S is a real Gaussian GeRF over $\mathbb{R}^d$ with covariance C_{Z_S}, and Z_T is a real Gaussian GeRF over $\mathbb{R}$ with covariance C_{Z_T} independent of Z_S, then it is possible to define a GeRF over $\mathbb{R}^d \times \mathbb{R}$ (or at least its action over test-functions in $\mathscr{S}(\mathbb{R}^d) \otimes \mathscr{S}(\mathbb{R})$), say Z, through the expression

$$\langle Z, \varphi \boxtimes \theta \rangle = \langle Z_S, \varphi \rangle \langle Z_T, \theta \rangle, \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^d), \forall \theta \in \mathscr{S}(\mathbb{R}). \tag{5.11}$$

In such a case, Z is called *the tensor product* between Z_S and Z_T, and hence can be denoted by $Z = Z_S \boxtimes Z_T$. Z has a covariance structure given by equation (5.10), but Z is not necessarily Gaussian. However, one can always construct *a Gaussian GeRF over* $\mathbb{R}^d \times \mathbb{R}$ *having covariance* C_Z, regardless of the initial spatial and temporal GeRFs Z_S and Z_T. In such a case, the Gaussian random variables of the form $\langle Z, \varphi \boxtimes \theta \rangle$ have, in principle, nothing to do with the random variables $\langle Z_S, \varphi \rangle$ and $\langle Z_T, \theta \rangle$. Moreover, if we consider Z to be a Gaussian GeRF over $\mathbb{R}^d \times \mathbb{R}$ with covariance given by (5.10), it is not clear at all if we can construct two GeRFs Z_S and Z_T, Gaussian or not, such that (5.11) holds. Hence, in this dissertation the notation $Z = Z_S \boxtimes Z_T$, when applied to GeRFs, is merely symbolic and it does not mean that Z *is* the tensor product between two GeRFs: it rather symbolizes a separability condition on the *covariance structure* of Z, and not on Z itself. The same issue is present in the case of Random Functions and Random Measures.

Symmetry

A real GeRF Z over $\mathbb{R}^d \times \mathbb{R}$ is said to be **symmetric** if

$$\begin{aligned}\langle C_Z, (\varphi_1 \boxtimes \theta_1) \otimes (\varphi_2 \boxtimes \theta_2)\rangle &= \langle C_Z, (\varphi_1 \boxtimes \theta_2) \otimes (\varphi_2 \boxtimes \theta_1)\rangle \\ &= \langle C_Z, (\varphi_2 \boxtimes \theta_1) \otimes (\varphi_1 \boxtimes \theta_2)\rangle \\ &= \langle C_Z, (\varphi_2 \boxtimes \theta_2) \otimes (\varphi_1 \boxtimes \theta_1)\rangle,\end{aligned} \tag{5.14}$$

for all test-functions $\varphi_1, \varphi_2 \in \mathscr{S}(\mathbb{R}^d)$ and $\theta_1, \theta_2 \in \mathscr{S}(\mathbb{R})$. In the stationary case, the symmetry is equivalent to have for the covariance $\rho_Z \in \mathscr{S}'(\mathbb{R}^d)$:

$$\langle \rho_Z, \varphi \boxtimes \theta\rangle = \langle \rho_Z, \varphi \boxtimes \check{\theta}\rangle = \langle \rho_Z, \check{\varphi} \boxtimes \theta\rangle = \langle \rho_Z, \check{\varphi} \boxtimes \check{\theta}\rangle, \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^d), \theta \in \mathscr{S}(\mathbb{R}). \tag{5.15}$$

Because of the properties of the Fourier Transform with respect to reflections, we obtain that a real stationary GeRF over $\mathbb{R}^d \times \mathbb{R}$ is symmetric if and only if its spectral measure $\mu_Z \in \mathscr{M}^+_{SG}(\mathbb{R}^d \times \mathbb{R})$ satisfies

$$d\mu_Z(\xi, \omega) = d\mu_Z(\xi, -\omega) = d\mu_Z(-\xi, \omega) = d\mu_Z(-\xi, -\omega), \tag{5.16}$$

condition which is more explicitly expressed through

$$\mu_Z(A \times B) = \mu_Z(A \times (-B)) = \mu_Z((-A) \times B) = \mu_Z((-A) \times (-B)), \quad \forall A, B \in \mathcal{B}_B(\mathbb{R}^d), \tag{5.17}$$

or through,

$$\int_{\mathbb{R}^d \times \mathbb{R}} \psi(\xi, \omega) d\mu_Z(\xi, \omega) = \int_{\mathbb{R}^d \times \mathbb{R}} \psi(\xi, -\omega) d\mu_Z(\xi, \omega) = \int_{\mathbb{R}^d \times \mathbb{R}} \psi(-\xi, \omega) d\mu_Z(\xi, \omega) = \int_{\mathbb{R}^d \times \mathbb{R}} \psi(-\xi, -\omega) d\mu_Z(\xi, \omega), \tag{5.18}$$

for all $\psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R})$. Hence, when Z is stationary and symmetric, its spectral measure is not only even in the sense of a measure over $\mathbb{R}^d \times \mathbb{R}$, having $\check{\mu_Z} = \mu_Z$, but it is also invariant under *partial reflections* of the space and time components. We say that such a measure over $\mathbb{R}^d \times \mathbb{R}$ *depends on the temporal frequency variable only through its absolute value*. The usage of this characterization of μ_Z is inspired by the case where μ_Z has a density, say $d\mu_Z(\xi, \omega) = f_{\mu_Z}(\xi, \omega) d\xi d\omega$. Indeed, in such a case, if μ_Z satisfies (5.18), then f_{μ_Z} satisfies $f_{\mu_Z}(\xi, \omega) = f_{\mu_Z}(\xi, -\omega) = f_{\mu_Z}(\xi, |\omega|)$ almost everywhere. The usage of the expression "*depending on the temporal variable only through its absolute value*" can be justified in more general cases using disintegration expressions of spatio-temporal measures. We will not enter into these details in this dissertation.

It turns out that an easy manner to obtain non-symmetric models is through the specification of a spectral measure over $\mathbb{R}^d \times \mathbb{R}$ not depending on its temporal frequency variable only through its absolute value. For instance, if we define a positive even and integrable function $f_{\mu_Z} : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^+$ such that $f_{\mu_Z}(\xi, \omega) \neq f_{\mu_Z}(\xi, -\omega)$ for some values (ξ, ω) in a set of non-null Lebesgue measure, then its spatio-

temporal Fourier Transform provides a non-symmetric stationary covariance function. In Section (5.2) we provide non-symmetric stationary covariance models, whose non-symmetry can be verified immediately through this criterion.

Spatial and Temporal Margins

In the case of a spatio-temporal GeRF Z, the spatial and temporal traces or margins are not immediate to describe since we cannot always evaluate at a fixed point $x \in \mathbb{R}^d$ or $t \in \mathbb{R}$. An approach fixing a spatial test-function $\varphi \in \mathscr{S}(\mathbb{R}^d)$ and then analysing the structure of the associated temporal GeRF $\langle Z, \varphi \boxtimes \cdot \rangle$ can always be done, as well as in the other sense fixing a temporal test-function $\theta \in \mathscr{S}(\mathbb{R})$. We will not enter into these details. We will rather present a case which is a kind of *middle ground* between the case of Random Functions and GeRFs, in the context of spatio-temporal stationary GeRFs.

Let Z be a real stationary GeRF over $\mathbb{R}^d \times \mathbb{R}$, and consider its spectral measure $\mu_Z \in \mathscr{M}^+_{SG}(\mathbb{R}^d \times \mathbb{R})$. In the case of a continuous stationary Random Function, μ_Z is finite. In the case of a general GeRF, μ_Z is slow-growing, not necessarily finite. We say that μ_Z is **temporally integrable** if it satisfies

$$\mu_Z(A \times \mathbb{R}) < \infty, \quad \forall A \in \mathcal{B}_B(\mathbb{R}^d). \tag{5.19}$$

When μ_Z is temporally integrable, the covariance distribution ρ_Z has a *continuous meaning in time*. Let us explain this notion. Since μ_Z is temporally integrable, its temporal Fourier Transform $\mathscr{F}_T(\mu_Z)$, which is a tempered distribution over $\mathbb{R}^d \times \mathbb{R}$, can be identified with a *measure-function Kernel*, $\mathscr{F}_T(\mu_Z) : \mathcal{B}_B(\mathbb{R}^d) \times \mathbb{R} \to \mathbb{C}$, defined through

$$\mathscr{F}_T(\mu_Z)(A, u) = \frac{1}{\sqrt{2\pi}} \int_{A \times \mathbb{R}} e^{-iu\omega} d\mu_Z(\xi, \omega), \quad A \in \mathcal{B}_B(\mathbb{R}^d), u \in \mathbb{R}. \tag{5.20}$$

Equivalently, $\mathscr{F}_T(\mu_Z)$ can be identified with a *distribution-function* Kernel, $\mathscr{F}_T(\mu_Z) : \mathscr{S}(\mathbb{R}^d) \times \mathbb{R} \to \mathbb{C}$ through

$$\mathscr{F}_T(\mu_Z)(\varphi, u) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^d \times \mathbb{R}} e^{-iu\omega} \varphi(\xi) d\mu_Z(\xi, \omega), \quad \varphi \in \mathscr{S}(\mathbb{R}^d), u \in \mathbb{R}. \tag{5.21}$$

A typical application of Dominated Convergence Theorem allows to conclude that for every $\varphi \in \mathscr{S}(\mathbb{R}^d)$, the function $u \in \mathbb{R} \mapsto \mathscr{F}_T(\mu_Z)(\varphi, u)$ is continuous. Since $\rho_Z = \mathscr{F}(\mu_Z) = \mathscr{F}_S(\mathscr{F}_T(\mu_Z))$, one may define for every $u \in \mathbb{R}$ the spatial distribution $\rho_Z^u \in \mathscr{S}'(\mathbb{R}^d)$:

$$\rho_Z^u = \mathscr{F}_S(\mathscr{F}_T(\mu_Z)(\cdot, u)), \tag{5.22}$$

whose explicit expression is

$$\langle \rho_Z^u, \varphi \rangle = \mathscr{F}_T(\mu_Z)(\mathscr{F}_S(\varphi), u) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^d \times \mathbb{R}} e^{-iu\omega} \mathscr{F}_S(\varphi)(\xi) d\mu_Z(\xi, \omega), \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^d). \tag{5.23}$$

The following fact follows from Fubini's Theorem: there exists a family of spatial tempered distributions $(\rho_Z^u)_{u \in \mathbb{R}} \subset \mathscr{S}'(\mathbb{R}^d)$ such that:

- for all $\varphi \in \mathscr{S}(\mathbb{R}^d)$, the function $u \in \mathbb{R} \mapsto \langle \rho_Z^u, \varphi \rangle$ is continuous.
- for all $\varphi \in \mathscr{S}(\mathbb{R}^d)$ and for all $\theta \in \mathscr{S}(\mathbb{R})$, it holds that

$$\langle \rho_Z, \varphi \boxtimes \theta \rangle = \int_{\mathbb{R}} \langle \rho_Z^u, \varphi \rangle \theta(u) du. \tag{5.24}$$

In general, any spatio-temporal tempered distribution $\rho_Z \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R})$ satisfying these conditions is known as a **continuous-in-time distribution**, and the family of spatial distributions $(\rho_Z^u)_{u \in \mathbb{R}} \subset \mathscr{S}'(\mathbb{R}^d)$ is known as the **continuous-in-time representation of** ρ_Z.

When a spatio-temporal real stationary GeRF Z has a temporally integrable spectral measure, it can be proven that Z *itself* has a *continuous meaning in time*. This can be seen intuitively from the fact that its covariance distribution ρ_Z is continuous in the time component. Let us explain this formally. We say that Z is **continuous in time** or that it has a **continuous-in-time representation** if there exists a family of spatial GeRFs $(Z_t)_{t \in \mathbb{R}}$ such that

- for all $\varphi \in \mathscr{S}(\mathbb{R}^d)$, the Random Function $t \in \mathbb{R} \mapsto \langle Z_t, \varphi \rangle$ is continuous in mean-square.
- for all $\varphi \in \mathscr{S}(\mathbb{R}^d)$ and for all $\theta \in \mathscr{S}(\mathbb{R})$, it holds that

$$\langle Z, \varphi \boxtimes \theta \rangle = \int_{\mathbb{R}} \langle Z_t, \varphi \rangle \theta(t) dt. \tag{5.25}$$

We obtain the following result.

Proposition 5.1.1. *Let Z be a real stationary GeRF over $\mathbb{R}^d \times \mathbb{R}$ such that its spectral measure μ_Z is temporally integrable. Then, Z has a continuous-in-time representation, $(Z_t)_{t \in \mathbb{R}}$. Moreover, if ρ_Z is the stationary covariance distribution of Z and $(\rho_Z^u)_{u \in \mathbb{R}}$ is its continuous-in-time representation, then it holds that*

$$\mathbb{C}ov(\langle Z_t, \varphi \rangle, \langle Z_s, \phi \rangle) = \langle \rho_Z^{t-s}, \varphi * \check{\bar{\phi}} \rangle, \quad \forall \varphi, \phi \in \mathscr{S}(\mathbb{R}^d), \forall t, s \in \mathbb{R}. \tag{5.26}$$

The proof of this result is given in Appendix A.13. Hence, we are able to *evaluate* the spatio-temporal GeRF Z at time locations, as it would be a temporal function, without technical issues. Every member of the

family $(Z_t)_{t\in\mathbb{R}}$ is called a **spatial trace** of Z. It is not hard to conclude from Proposition 5.1.1 that for every $t \in \mathbb{R}$, the spatial GeRF Z_t is a stationary real GeRF over $\mathbb{R}^d$, with covariance distribution $\rho_Z^{t-t} = \rho_Z^0$ and hence, as expected, all spatial traces have the same spatial covariance structure, given by ρ_Z^0. We denote by $\rho_{Z_S} := \rho_Z^0$ and we call it the **spatial margin** of the covariance distribution ρ_Z. The spectral measure of the spatial traces can be obtained by evaluating (5.20) at 0, obtaining the measure defined through

$$\mu_{Z_S}(A) := \frac{1}{\sqrt{2\pi}}\mu_Z(A \times \mathbb{R}), \quad \forall A \in \mathcal{B}_B(\mathbb{R}^d). \tag{5.27}$$

Hence, we obtain the spectral measure of the spatial traces of Z through the temporal integration of its spectral measure. It follows that $\rho_{Z_S} = \mathscr{F}_S(\mu_{Z_S})$. We write generically Z_S representing any spatial trace of Z.

It is also concluded from Proposition 5.1.1 that for every $\varphi \in \mathscr{S}(\mathbb{R}^d)$, the continuous Random Function $t \mapsto \langle Z_t, \varphi\rangle$ is stationary. Its covariance function, denoted by $\rho_{Z_T}^{\varphi}$, is given by $\rho_{Z_T}^{\varphi}(u) = \langle \rho_Z^u, \varphi * \check{\bar{\varphi}}\rangle$ for all $u \in \mathbb{R}$. We call $\rho_{Z_T}^{\varphi}$ the $\varphi-$**temporal margin** of the covariance distribution ρ_Z. If we change φ with another test-function $\phi \in \mathscr{S}(\mathbb{R}^d)$, the covariance $\rho_{Z_T}^{\phi}$ is in general a different distribution. An equality is anyway present when ϕ is a translation of φ. We do not enter in details about the spectral measures associated to these temporal margins.

When μ_Z is temporally integrable, it is possible to prove that the Fourier Transform of Z, say $M_Z = \mathscr{F}(Z)$, which is a slow-growing orthogonal Random Measure, is also *temporally integrable* in the sense that the random variables of the form $M_Z(A \times \mathbb{R})$, with $A \in \mathcal{B}_B(\mathbb{R}^d)$, are well-defined square-integrable random variables. This notion can be extended immediately to every spatio-temporal Random Measure. We do not enter in details.

We finally remark that this procedure can be done analogously in *the other sense*, by using spatio-temporal spectral measures *spatially integrable*, obtaining a stationary GeRF which is *continuous in space*.

5.2 Evolution equations: new stationary spatio-temporal models

In the most general sense, any spatio-temporal PDE or SPDE deserves to be called an *evolution equation*. However, in this section we will restrict this name to a particular class of SPDEs which involves an operator which is the sum of a temporal differential operator of arbitrary order (including fractional operators), and a purely-spatial operator defined through a symbol. We will also restrict our analysis to stationary models, in order to apply the results of Chapter 4.

We consider thus SPDEs over $\mathbb{R}^d \times \mathbb{R}$ of the form:

$$\frac{\partial^\beta U}{\partial t^\beta} + \mathcal{L}_g U = X, \tag{5.28}$$

where X is a real stationary spatio-temporal GeRF, $\beta > 0$ and $g : \mathbb{R}^d \to \mathbb{C}$ is a *spatial* symbol function. We denote by g_R the real part of g and g_I the imaginary part of g. For this class of SPDEs, we study in detail several examples of physical and statistical interest. They involve for example Langevin-type equations, advection, diffusion and wave propagation phenomena.

First of all, for $\beta > 0$, we specify the definition of the fractional differential operator $\frac{\partial^\beta}{\partial t^\beta}$:

$$\frac{\partial^\beta}{\partial t^\beta} := \mathscr{F}_T^{-1}((i\omega)^\beta \mathscr{F}_T(\cdot)). \tag{5.29}$$

Hence, $\frac{\partial^\beta}{\partial t^\beta}$ is nothing but an operator defined through a symbol (Eq. (4.2)), specifically through the symbol function over $\mathbb{R}$:

$$\omega \mapsto (i\omega)^\beta := |\omega|^\beta e^{i\,\mathrm{sgn}(\omega)\beta\frac{\pi}{2}}. \tag{5.30}$$

The function (5.30) is Hermitian, continuous and bounded by a polynomial for every $\beta > 0$, so it is indeed a well-defined symbol function. Similar definitions of a fractional differential operator can be found in Mainardi et al. (2007). We call a **fractional order evolution model** every real spatio-temporal stationary solution of the SPDE (5.28) with $\beta \notin \mathbb{N}$. For $\beta \in \mathbb{N}$, (5.29) coincides with a classical differential operator. The corresponding stationary solutions are called **β-th order evolution model**.

The spatio-temporal symbol function of the operator involved in (5.28) is the function

$$(\xi,\omega) \in \mathbb{R}^d \times \mathbb{R} \mapsto (i\omega)^\beta + g(\xi) = |\omega|^\beta \cos\left(\frac{\beta\pi}{2}\right) + g_R(\xi) + i\left(\mathrm{sgn}(\omega)|\omega|^\beta \sin\left(\frac{\beta\pi}{2}\right) + g_I(\xi)\right). \tag{5.31}$$

Theorem 4.3.1 allows us to conclude that there exists stationary solutions to (5.28) if and only if the measure

$$d\mu_U(\xi,\omega) = \frac{d\mu_X(\xi,\omega)}{|(i\omega)^\beta + g(\xi)|^2}, \tag{5.32}$$

is slow-growing. We will focus on the result stated on Remark 4.3.2 and look at for conditions on g such that (5.31) satisfies the PBR condition and thus to have a unique stationary solution regardless of the source term X. The next proposition, proven in Appendix A.14, allows us to identify the cases where the PBR condition holds regardless of the imaginary part g_I.

Proposition 5.2.1. *Let $g_R : \mathbb{R}^d \to \mathbb{R}$ be an even and polynomially bounded measurable function. Then, the spatio-temporal function defined through* (5.31) *satisfies the PBR condition for every odd polynomially bounded measurable function $g_I : \mathbb{R}^d \to \mathbb{R}$ if and only if g_R satisfies the PBR condition and $g_R \cos(\frac{\beta\pi}{2}) \geqslant 0$.*

We suppose that the conditions on g_R in Proposition 5.2.1 hold. Let us study the properties of this kind of model. For simplicity, we restrict ourselves to the cases where X is a separable model $X = X_S \boxtimes X_T$.

The spectral measure of the unique stationary solution to (5.28) is then

$$d\mu_U(\xi,\omega) = \frac{d\mu_{X_S}(\xi)d\mu_{X_T}(\omega)}{|\omega|^{2\beta} + 2|\omega|^{\beta}\left(g_R(\xi)\cos\left(\frac{\beta\pi}{2}\right) + \operatorname{sgn}(\omega)g_I(\xi)\sin\left(\frac{\beta\pi}{2}\right)\right) + |g(\xi)|^2}. \tag{5.33}$$

A separable model is obtained when $g_I = 0$ and g_R is a constant function. Otherwise, the model is not separable. The function sgn in (5.33) allows to identify the cases where the spectral measure does not depend on the argument ω only through $|\omega|$ and thus the symmetry of the model can be controlled. A symmetric model is then obtained when β is an even integer or when the function g_I is null. A non-symmetric model is obtained otherwise. In this case the non-symmetry can be parametrized by controlling the function g_I. This fact is what gives importance to Proposition 5.2.1 since in the cases considered in this Proposition the function g_I can be controlled freely. The mean-square temporal regularity of the associated random field depends on the parameter β, as it can be seen by analysing the temporal-integrability of the measure μ_U. Thus, this model allows a practical control of the separability, symmetry and regularity conditions.

The covariance structure of a spatial trace of this model can be described if the measure μ_U is temporally-integrable, that is, if $\int_{\mathbb{R}} |(i\omega)^{\beta} + g(\xi)|^{-2} d\mu_{X_T}(\omega) < \infty$. Let us restrict ourselves to the case where X is a White Noise in time, $X = X_S \boxtimes W_T$, i.e. $d\mu_X(\xi,\omega) = d\mu_{X_S}(\xi)d\mu_{W_T}(\omega) = d\mu_{X_S}(\xi)(2\pi)^{-\frac{1}{2}}d\omega$. In that case, the measure μ_U is temporally-integrable when $\beta > \frac{1}{2}$. The spectral measure of the spatial traces can be obtained by calculating the corresponding integral. We study the spatial structure in the case $g_I = 0$. The general case with $g_I \neq 0$ is much more technical and we have not found simple and enlightening expressions for the spatial covariance behaviour, so it is not presented in this dissertation. The spectral measure of a spatial trace U_S is then

$$d\mu_{U_S}(\xi) = \frac{1}{2\pi}\int_{\mathbb{R}} \frac{d\omega}{|\omega|^{2\beta} + 2|\omega|^{\beta} g_R(\xi)\cos\left(\frac{\pi}{2}\beta\right) + g_R^2(\xi)} d\mu_{X_S}(\xi) \tag{5.34}$$

$$= \frac{|g_R(\xi)|^{\frac{1}{\beta}-2}}{\pi\beta} \underbrace{\int_0^{\infty} \frac{\theta^{\frac{1}{\beta}-1}}{\theta^2 + 2\theta\operatorname{sgn}(g_R)\cos\left(\frac{\pi}{2}\beta\right) + 1} d\theta}_{=I_\beta} d\mu_{X_S}(\xi), \tag{5.35}$$

where we have used the parity of the function with respect to ω and then used the change of variable $\omega = (|g_R(\xi)|\theta)^{\frac{1}{\beta}}$. The integral I_β does not depend on ξ since g_R does not change in sign. This integral can be computed (see for instance Gradshteyn & Ryzhik, 2014, 3.252.12). In particular, $I_1 = I_2 = \pi/2$. Then, the spatial traces of the solution satisfy the spatial SPDE

$$\sqrt{\frac{\pi\beta}{I_\beta}}\mathcal{L}_{|g_R|^{1-\frac{1}{2\beta}}} U_S \overset{2nd\ o.}{=} X_S. \tag{5.36}$$

This model has a continuous point-wise meaning when the function $|g_R|^{\frac{1}{\beta}-2}$ is integrable with respect to the measure μ_{X_S}, case in which the measure μ_U is a finite measure.

Condition 1 in Theorem 4.4.1 does not hold when $\beta \notin \mathbb{N}$ since the symbol function (5.31) is not smooth. The case $\beta \in \mathbb{N}$ can be worked out supposing some regularity conditions on g. We present the corresponding analysis for the cases $\beta \in \{1, 2\}$. We are not going to focus on condition 2 in Theorem 4.4.1. We remark that if the statement that we have considered in Remark 4.4.1 holds, then it is sufficient to consider g continuous satisfying conditions in Proposition 5.2.1 to obtain a convolution result such as in Theorem 4.4.1. Indeed, in such a case, the function $(\xi, \omega) \mapsto |(i\omega)^\beta + g(\xi)|^{-2}$ is continuous and polynomially bounded.

A **first order evolution model** is a stationary solution of Eq. (5.28) when $\beta = 1$. Let us set $X = W$, the spatio-temporal White Noise. The spectral measure is then

$$d\mu_U^W(\xi, \omega) = \frac{1}{(2\pi)^{\frac{d+1}{2}}} \frac{d\xi d\omega}{(\omega + g_I(\xi))^2 + g_R^2(\xi)}. \tag{5.37}$$

From this we obtain that its covariance is of the form

$$\rho_U^W(h, u) = \mathscr{F}_S \left(\xi \mapsto \frac{1}{(2\pi)^{\frac{d}{2}}} \frac{e^{iug_I(\xi) - |u||g_R(\xi)|}}{2|g_R(\xi)|} \right)(h). \tag{5.38}$$

This model can then be seen as a *mixture of (complex) exponentials*. For ease of reading, we have used a functional notation for the variables (h, u) in (5.38), but ρ_U^W is not necessarily a function. Generally, it is a tempered distribution, and it depends on g_R if this distribution can be identified with a continuous function or not. A continuous function is obtained when $|g_R|^{-1}$ is an integrable function. The spatial margin of ρ_U^W is obtained by setting $u = 0$ in (5.38). We see that it does not depend on g_I. Thus, Eq. (5.36) can be used to describe the spatial behaviour of the model for the case $X_S = W_S$, including the cases where $g_I \neq 0$.

A similar analysis can also be done easily for the case $X = X_S \boxtimes W_T$, a coloured in space and White in time noise. For that, it is enough to replace the term $d\xi$ in (5.37) with $(2\pi)^{d/2} d\mu_{X_S}(\xi)$. The covariance $\rho_U^{X_S \boxtimes W_T}$ in such a case is given by

$$\rho_U^{X_S \boxtimes W_T}(h, u) = \mathscr{F}_S \left(\xi \mapsto \frac{e^{iug_I(\xi) - |u||g_R(\xi)|}}{2|g_R(\xi)|} d\mu_{X_S}(\xi) \right)(h). \tag{5.39}$$

Setting $u = 0$, the covariance of a spatial trace is simply

$$\rho_{U_S}^{X_S \boxtimes W_T}(h) = \mathscr{F}_S \left(\xi \mapsto \frac{1}{2} |g_R(\xi)|^{-1} d\mu_{X_S}(\xi) \right)(h), \tag{5.40}$$

from where we obtain immediately that the spatial traces follow the SPDE in second-order sense (c.f. Eq.

(5.36))

$$\sqrt{2}\mathcal{L}_{|g_R|^{-\frac{1}{2}}} \overset{2nd\ o.}{=} X_S \tag{5.41}$$

We thus obtain a particular description which holds for first order evolution models: *the spatial behaviour is completely described by g_R, while the spatio-temporal non-symmetry is described by g_I*. This does not necessarily hold for other values of β, as it will be shown for the case $\beta = 2$.

Condition 1 in Theorem 4.4.1 can be applied if g_R, g_I and $1/g_R$ are in $\mathcal{O}_M(\mathbb{R}^d)$, since in this case the reciprocal of the spatio-temporal symbol function $(\xi, \omega) \mapsto i\omega + g(\xi)$ is in $\mathcal{O}_M(\mathbb{R}^d \times \mathbb{R})$. We obtain in that case that the covariance of the solution with an arbitrary source term X is the convolution $\rho_U^W * \rho_X$, with ρ_U^W given by (5.38).

A **second order evolution model** is a stationary solution of Eq. (5.28) when $\beta = 2$. Consider again $X = W$. Since g_R satisfies conditions in Proposition 5.2.1, in particular $g_R < 0$. The spectral measure is then

$$d\mu_U^W(\xi, \omega) = \frac{1}{(2\pi)^{\frac{d+1}{2}}} \frac{d\xi d\omega}{(\omega^2 - g_R(\xi))^2 + g_I^2(\xi)}, \tag{5.42}$$

and the covariance distribution ρ_U^W is the Fourier Transform of μ_U^W. To simplify the notation, consider the complex spatial function $\gamma : \xi \mapsto \mathbb{C}$ defined through

$$\gamma(\xi) = \sqrt{\frac{|g(\xi)| + g_R(\xi)}{2}} + i\sqrt{\frac{|g(\xi)| - g_R(\xi)}{2}}, \quad \forall \xi \in \mathbb{R}^d. \tag{5.43}$$

The function γ is never null since $g_R < 0$. Let us denote by γ_R and γ_I the real and imaginary parts of γ respectively. The covariance ρ_U^W is then

$$\rho_U^W(h, u) = \mathscr{F}_S\left(\xi \mapsto \frac{e^{-(|\gamma_I(\xi)| + i\gamma_R(\xi))|u|}}{(2\pi)^{\frac{d}{2}} 8|\gamma_I(\xi)|^2}\left[\frac{1}{|\gamma_I(\xi)| + i\gamma_R(\xi)} + \frac{e^{i2\gamma_R(\xi)|u|}}{|\gamma_I(\xi)| - i\gamma_R(\xi)} + \frac{e^{i2\gamma_R(\xi)|u|} - 1}{i\gamma_R(\xi)}\right]\right)(h). \tag{5.44}$$

The term $(e^{i2\gamma_R(\xi)|u|} - 1)/i\gamma_R(\xi)$ is interpreted to be equal to $2|u|$ when $\gamma_R(\xi) = 0$, which corresponds to $g_I(\xi) = 0$. This covariance distribution is a continuous function if the function $|\gamma_I|^{-1}|\gamma|^{-2}$ is integrable over $\mathbb{R}^d$, which is equivalent to require that the function $|g|^{-1}(|g| - g_R)^{-\frac{1}{2}}$ is integrable over $\mathbb{R}^d$. Contrarily to the case of first order evolution models, this model is always symmetric and the covariance structure of the spatial traces depends on both g_R and g_I, as it can be seen by evaluating (5.44)) at $u = 0$. Thus, Eq. (5.36) does not hold for $g_I \neq 0$. The spectral measure of a spatial trace is

$$d\mu_{U_S}^W(\xi) = \frac{d\xi}{(2\pi)^{\frac{d}{2}} 2\sqrt{2}|g(\xi)|\sqrt{|g(\xi)| - g_R(\xi))}}, \tag{5.45}$$

from which we obtain that a spatial trace U_S satisfies the spatial SPDE

$$\sqrt{2\sqrt{2}}\mathcal{L}_{\sqrt{|g|\sqrt{|g|-g_R}}}U_S \overset{2nd\ o.}{=} W_S, \tag{5.46}$$

where W_S is a spatial White Noise. An analogue expression is obtained in the case $X = X_S \boxtimes W_T$, by replacing W_S by X_S in (5.46)) and $d\xi$ by $(2\pi)^{d/2}d\mu_{X_S}(\xi)$ in (5.42)) and (5.45)). When X is a general spatio-temporal stationary GeRF, a sufficient condition to apply Theorem 4.4.1 is that g_R, $1/g_R$ and g_I are in the space $\mathcal{O}_M(\mathbb{R}^d)$. In this case, the only stationary solution to the SPDE (5.28) with $\beta = 2$ has a covariance of the form $\rho_U = \rho_U^W * \rho_X$, where ρ_U^W is given by (5.44)).

We now present some particular models inspired by physical and statistical literature. In some cases Proposition 5.2.1 can be applied. In other cases, there is no uniqueness and sometimes not even existence of stationary solutions.

5.2.1 Some examples from the literature

We present two briefs examples of models which are inspired by physical consideration and have been used to define geostatistical models.

Example 5.2.1 (**Advection-diffusion equation.**). Sigrist et al. (2015) propose estimation methods and simulation algorithms for the unique stationary solution of the SPDE over $\mathbb{R}^d \times \mathbb{R}$:

$$\frac{\partial U}{\partial t} + \kappa^2 U + v^T \nabla U - \operatorname{div}(\Sigma \nabla U) = X_S \boxtimes W_T, \tag{5.47}$$

where $\kappa > 0$ is a damping parameter, $v \in \mathbb{R}^d$ is a velocity vector and Σ is a symmetric positive-definite matrix controlling non-isotropic diffusion. W_T is a temporal White Noise and X_S represents a stationary spatial random field. This equation, known as the *advection-diffusion equation*, is a particular first order evolution model. Its spatial symbol function is

$$g(\xi) = \kappa^2 + \xi^T \Sigma \xi + i v^T \xi,$$

for which conditions in Proposition 5.2.1 are satisfied. Without advection ($v = 0$), this equation was studied in Whittle (1963) in a non-generalized framework. Sigrist et al. (2015) consider a Matérn Model for X_S, with smoothness parameter equals to 1, corresponding to $\alpha = 2$ in (3.6) when $d = 2$. The spatial behaviour of this model is described by the SPDE (5.36) for $\beta = 1$.

Example 5.2.2 (**A Langevin Equation**). Using linear response theory, Hristopulos & Tsantili (2016) pro-

pose stationary random fields which are solutions to the following Langevin equation

$$\frac{\partial U}{\partial t} + \frac{D}{2k^d\eta_0}\left(1 - \eta_1 k^2\Delta + \nu k^4\Delta^2\right)U = W, \tag{5.48}$$

with $D, k, \eta_0 > 0$, $\eta_1, \nu \geqslant 0$. The parameter ν is called the *curvature coefficient*. For simplicity, let $C = D/(2k^d\eta_0)$. For this first order evolution model, the spatial symbol function is

$$g(\xi) = C\left(1 + \eta_1 k^2|\xi|^2 + \nu k^2|\xi|^4\right),$$

which satisfies conditions of Proposition 5.2.1. Hence, (5.48) has a unique stationary solution, whose spectral measure can be obtained using the general expression of first order evolution model in (5.37). Hristopulos & Tsantili (2016) provide expressions of the related covariance structures, which are functions for $d \leqslant 3$, and which can be obtained through formulas similar to (5.38) in combination with the Fourier Transform of radial functions. The spatial behavior of this model can be described following equation (5.36), with spatial White Noise source term, $X_S = W_S$.

5.2.2 Evolving Matérn model

In the most general term, we call *Evolving Matérn model* every spatio-temporal GeRF such that its spatial traces follow Matérn covariance models. In the case of stationary solutions to Eq. (5.28), evolving Matérn models can be obtained by adequately controlling g, X or both. In this section we focus on stationary solutions to equations of the form

$$\frac{\partial^\beta U}{\partial t^\beta} + s_\beta a(\kappa^2 - \Delta)^{\frac{\alpha}{2}}U = W, \tag{5.49}$$

where W is as usual a spatio-temporal White Noise, $\kappa^2, a > 0, \alpha \in \mathbb{R}$, and s_β is a parameter that takes the value 1 or -1 depending conveniently on β in order to obtain conditions in Proposition 5.2.1 for $g(\xi) = s_\beta a(\kappa^2 + |\xi|^2)^{\frac{\alpha}{2}}$. There is then a unique stationary solution to (5.49). Its spectral measure is

$$d\mu_U(\xi,\omega) = \frac{1}{(2\pi)^{\frac{d+1}{2}}}\frac{d\xi d\omega}{|\omega|^{2\beta} + 2|\omega|^\beta a(\kappa^2 + |\xi|^2)^{\frac{\alpha}{2}}|\cos\left(\frac{\beta\pi}{2}\right)| + a^2(\kappa^2 + |\xi|^2)^\alpha}. \tag{5.50}$$

Following Eq. (5.36), when $\beta > \frac{1}{2}$ the spatial traces of this model follow the spatial SPDE

$$\sqrt{\frac{\pi\beta}{I_\beta}}a^{1-\frac{1}{2\beta}}(\kappa^2 - \Delta)^{\frac{\alpha}{2}\left(1-\frac{1}{2\beta}\right)}U_S \overset{2nd\ o.}{=} W_S, \tag{5.51}$$

where W_S is a spatial White Noise. Direct identification between (5.51) and the SPDE (4.15) in Example 4.5.1 indicates that the spatial covariance is a Matérn covariance. It has a functional meaning when $\alpha\,(1-1/(2\beta)) > d/2$. Explicit expressions of the covariance can be obtained using expressions of the Fourier Transform of radial functions (Donoghue, 1969, chapter 41).

In particular, for $\beta = 1$, we get

$$\rho_U(h,u) = \frac{1}{(2\pi)^{\frac{d}{2}}|h|^{\frac{d-2}{2}}}\int_0^\infty J_{\frac{d-2}{2}}(|h|r)\frac{e^{-a(\kappa^2+r^2)^{\frac{\alpha}{2}}|u|}}{2a(\kappa^2+r^2)^{\frac{\alpha}{2}}}r^{\frac{d}{2}}dr, \tag{5.52}$$

where J_b denotes the Bessel function of the first kind of order b. This model has also been proposed in Jones & Zhang (1997), in which an approach similar to our framework was followed for first order evolution equations. This is a symmetric non-separable model which can be identified as a mixture of a $J-$Bessel model in space with an exponential model in time.

Notice that in this case $\beta = 1$ we can add a non-null imaginary part g_I to the symbol function without changing the spatial behaviour, thereby generating non-symmetric evolving Matérn models. This can be concluded from our development explained above concerning first order evolution models. However, in such a case the expression (5.52) no longer applies.

For $\beta = 2$, one gets

$$\rho_U(h,u) = \frac{1}{(2\pi)^{\frac{d}{2}}|h|^{\frac{d-2}{2}}}\int_0^\infty J_{\frac{d-2}{2}}(|h|r)\frac{e^{-\sqrt{a}(\kappa^2+r^2)^{\frac{\alpha}{4}}|u|}(1+\sqrt{a}(\kappa^2+r^2)^{\frac{\alpha}{4}}|u|)}{4a\sqrt{a}(\kappa^2+r^2)^{\frac{3\alpha}{4}}}r^{\frac{d}{2}}dr. \tag{5.53}$$

This covariance is a mixture of $J-$Bessel model in space and a Matérn model in time since the spectral measure (5.42) has the form of a Matérn spectral measure in ω (we recall that $g_R < 0$ for $\beta = 2$ and $g_I = 0$). This covariance has a functional meaning for $\alpha > \frac{2d}{3}$. Notice that this mixture property between a $J-$Bessel model in space and a Matérn model in time does not hold for $\beta \notin \{1,2\}$, the spectral measure (5.50) having not the form of a Matérn spectral measure in the variable ω.

Notice that both g_R and $1/g_R$ are in $\mathcal{O}_M(\mathbb{R}^d)$. Thus, for $\beta \in \{1,2\}$ Theorem 4.4.1 can be applied. In these cases, the covariance of the solution to an equation of the form (5.49) with an arbitrary source term X is the convolution between (5.52) for $\beta = 1$ (respectively (5.53) for $\beta = 2$) and ρ_X.

Some classes of models which are evolving Matérn models can be found in the literature. For instance, for the cases $\beta \in \mathbb{N}$ these models are Stein models. See the correspondences between Eq. (4.25) and Eq. (5.50) in those cases, considering the temporal scale parameter $s = 0$. The advection-diffusion equation exposed in Example 5.2.1 also provides evolving Matérn models. The Langevin equation presented in Example 5.2.2 provides evolving Matérn models when the curvature coefficient ν equals 0 and $\eta_1 > 0$.

We finally remark that we also obtain evolving Matérn models when replacing W in Eq. (5.49) with a coloured in space and white in time noise, $X_S \boxtimes W_T$, with X_S following a Matérn model with scale parameter κ. Another example of evolving Matérn model is the Waving Matérn model, which will be exposed in Section 5.2.4.

5.2.3 Heat equation

We now consider the stochastic Heat (or Diffusion) Equation over $\mathbb{R}^d \times \mathbb{R}$

$$\frac{\partial U}{\partial t} - a\Delta U = X, \tag{5.54}$$

where $a > 0$ is the *diffusivity* parameter. It is a first order evolution model with spatial symbol function $g(\xi) = a|\xi|^2$. In this case, the spatio-temporal symbol function $(\xi, \omega) \mapsto i\omega + a|\xi|^2$ is not strictly positive, the origin being the only zero of g. From Theorem 4.3.1, there is no uniqueness of stationary solutions, if they exist. Following Remark 4.3.3, since the only zero point of the symbol function is the origin, the stationary solutions to the homogeneous problem

$$\frac{\partial U_H}{\partial t} - a\Delta U_H = 0 \tag{5.55}$$

must have spectral measures supported on the origin, hence proportional to the Dirac measure. We conclude that the only stationary solutions to the homogeneous Heat Equation are random constants (c.f. Example 4.5.2).

Because of the singularity at the origin of the function $|g|^{-2}$, the existence condition (4.7) does not always hold. Existence needs to be checked for each source term X. Let us first consider the case where the source term is a spatio-temporal White Noise. Equation (5.54) becomes

$$\frac{\partial U}{\partial t} - a\Delta U = W. \tag{5.56}$$

Using Theorem 4.3.1, one concludes (see Appendix A.15.1) that *there exist stationary solutions to the stochastic Heat equation* (5.56) *only for spatial dimensions* $d \geqslant 3$. In addition, in these cases, the solutions can only be conceived as GeRFs and never as continuous Random Functions. When $d = 3$, computations reported in Appendix A.15.2 show that the covariance structure of a particular solution is described by

$$\rho_U^W(h, u) = \frac{1}{(2\pi)^{\frac{d+1}{2}}} \frac{\pi}{2a|h|} \operatorname{erf}\left(\frac{|h|}{2\sqrt{a|u|}}\right). \tag{5.57}$$

Since the covariance (5.38) is not defined at $|h| = |u| = 0$, it must be interpreted in a suitable distributional

sense which is also explained in Appendix A.15.2. A spatial trace of the stationary GeRF associated to (5.57), U_S, can be described evaluating this covariance in $u = 0$ with $h \neq 0$. We obtain that U_S satisfies the spatial SPDE

$$\sqrt{2}(-\Delta)^{\frac{1}{2}} U_S \overset{2nd\ o.}{=} W_S, \tag{5.58}$$

where W_S is a spatial White Noise. In other words, U_S is a Matérn model without range parameter as presented in Example 4.5.2 (see Eq. (4.18)).

When X is an arbitrary source term, Theorem 4.4.1 cannot be applied for spatial dimensions smaller than 3. For $d = 3$, a convolvability condition between ρ_X and (5.57) must be satisfied. Nevertheless, the existence of a solution can be ensured independently of the existence of solutions with White Noise source term by imposing some sufficient conditions on μ_X such that the existence criterion (4.7) in Theorem 4.3.1 holds. For example, one could require μ_X to be concentrated on the complementary of some neighbourhood of the origin.

5.2.4 Wave equation and Waving models

As a final example we consider the stochastic wave equation

$$\frac{\partial^2 U}{\partial t^2} - c^2 \Delta U = X, \tag{5.59}$$

where X is a real stationary random field and $c > 0$ is the *wave propagation velocity*. This is a second order evolution model with spatial symbol function $g(\xi) = c^2|\xi|^2$. The null-set of the associated spatio-temporal symbol function $(\xi, \omega) \mapsto -\omega^2 + c^2|\xi|^2$ is the spatio-temporal cone $\mathcal{C}^c = \{(\xi, \omega) \in \mathbb{R}^d \times \mathbb{R} \mid |\omega| = c|\xi|\}$ (see Example 2.1.3). As a consequence, uniqueness of a potential stationary solution does not hold.

We call a **Waving model** any spatio-temporal real stationary GeRF solution to the homogeneous Wave equation

$$\frac{\partial^2 U_H}{\partial t^2} - c^2 \Delta U_H = 0. \tag{5.60}$$

Following Remark 4.3.3, the spectral measure of such a model must be concentrated on $\mathcal{C}^c$. Following Eq. (2.41), if μ_{U_H} is the spectral measure of a stationary solution to (5.60), then μ_{U_H} is of the form

$$\langle \mu_{U_H}, \psi \rangle = \int_{\mathbb{R}^d} \psi(\xi, c|\xi|) d\mu_1(\xi) + \int_{\mathbb{R}^d \setminus \{0\}} \psi(\xi, -c|\xi|) d\mu_2(\xi), \quad \forall \psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R}), \tag{5.61}$$

for some measures $\mu_1, \mu_2 \in \mathscr{M}^+(\mathbb{R}^d)$. Since μ_{U_H} must be slow-growing, then $\mu_1, \mu_2 \in \mathscr{M}^+_{SG}(\mathbb{R}^d)$ (Proposition 2.1.3), and since μ_{U_H} must be even, it turns out that μ_1 and μ_2 are equal over $\mathbb{R}^d \setminus \{0\}$ and both are

even. Hence, there exists an even measure $\mu_{U_H^S} \in \mathscr{M}_{SG}^+(\mathbb{R}^d)$, i.e. a spatial spectral measure, such that

$$\langle \mu_{U_H}, \psi \rangle = \sqrt{2\pi} \int_{\mathbb{R}^d} \frac{\psi(\xi, c|\xi|) + \psi(\xi, -c|\xi|)}{2} d\mu_{U_H^S}(\xi), \quad \forall \psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R}). \tag{5.62}$$

Hence, all stationary solutions of (5.60) have a spectral measure of the form (5.62), and conversely, every measure of the form (5.62), with $\mu_{U_H^S}$ being a spatial spectral measure, is a spectral measure over $\mathbb{R}^d \times \mathbb{R}$ whose associated stationary GeRFs are solutions to (5.60). The factors $\sqrt{2\pi}$ and $\frac{1}{2}$ are included for convenience. Another way of expressing measures of this form is through the *disintegration language:*

$$d\mu_{U_H}(\xi, \omega) = \sqrt{2\pi} d \left(\frac{\delta_{-c|\xi|} + \delta_{c|\xi|}}{2} \right) (\omega) d\mu_{U_H^S}(\xi). \tag{5.63}$$

The associated covariance distribution over $\mathbb{R}^d \times \mathbb{R}$ is its Fourier Transform, which is

$$\rho_{U_H}(h, u) = \mathscr{F}_S \left(\xi \mapsto \cos(c|\xi||u|) d\mu_{U_H^S}(\xi) \right) (h). \tag{5.64}$$

By setting $u = 0$, it follows that the covariance of a spatial trace is the spatial Fourier Transform of $\mu_{U_H^S}$, and hence $\mu_{U_H^S}$ is the spectral measure of the spatial traces, describing then the spatial behaviour of the solution U_H. We conclude that a Waving model can follow any arbitrary spatial covariance model, which can be chosen freely by fixing the spatial spectral measure $\mu_{U_H^S}$. In addition, if this spatial spectral measure is finite, the associated Waving model is a continuous Random Function over $\mathbb{R}^d \times \mathbb{R}$, since in such a case the measure (5.63) is finite.

We consider as example the case of **Waving Matérn models**, which are Waving models which follow a Matérn model in space. Hence, their spectral measures must be of the form (5.62), with $d\mu_{U_H^S}$ being of the form

$$d\mu_{U_H^S}(\xi) = \frac{d\xi}{(2\pi)^{\frac{d}{2}} a(\kappa^2 + |\xi|^2)^\alpha}, \tag{5.65}$$

with $a, \kappa > 0$ and $\alpha \in \mathbb{R}$. The associated covariance is

$$\rho(h, u) = \mathscr{F}_S \left(\xi \mapsto \frac{\cos(c|\xi||u|)}{(2\pi)^{\frac{d}{2}} a(\kappa^2 + |\xi|^2)^\alpha} \right) (h). \tag{5.66}$$

Let us now go back to the existence of stationary solutions to (5.59) in a non-homogeneous form. Consider the case $X = W$, i.e.

$$\frac{\partial^2 U}{\partial t^2} - c^2 \Delta U = W. \tag{5.67}$$

Since the function $(\xi, \omega) \mapsto (-\omega^2 + c^2|\xi|^2)^{-2}$ is not locally integrable, by applying Theorem 4.3.1 we conclude that *there are no stationary solutions to the stochastic wave equation* (5.67). Hence, we cannot

apply Theorem 4.4.1 to relate the covariance of a possible stationary solution of (5.59) to the covariance of the solution with White Noise source term. The existence of a stationary solution to (5.59) must be then studied for every particular case of X. Notice however that the existence is guaranteed when the support of the spectral measure of the source term μ_X and the spatio-temporal cone $\mathcal{C}^c$ are separated by neighbourhoods.

5.3 Further developments on first order evolution models

In this section we discuss informally the particular case of first order evolution models satisfying a particular initial condition. The resolution of an associated Cauchy problem in a general case provides spatio-temporal geostatistical models which are in general non-separable, non-symmetric, and non-stationary. Under suitable conditions, we are able to prove an asymptotic convergence to a spatio-temporal stationary solution when the time flows enough.

The formal definitions and proofs of the statements proposed in this section are presented in Appendix C.

5.3.1 The deterministic problem

We will follow a more typical approach in analysis of PDEs and SPDEs. We will first of all consider the deterministic problem of finding a solution to the Cauchy problem over $\mathbb{R}^d \times \mathbb{R}^+$:

$$\begin{cases} \dfrac{\partial U}{\partial t} + \mathcal{L}_g U = X \\ U\big|_{t=0} = U_0 \end{cases} . \tag{5.68}$$

Here X is a distribution over $\mathbb{R}^d \times \mathbb{R}$ and U_0 is a distribution over $\mathbb{R}^d$, both of them belonging to suitable subspaces of tempered distributions. $g : \mathbb{R}^d \to \mathbb{C}$ is a continuous spatial symbol function with real part g_R and imaginary part g_I for which we suppose in addition that $g_R \geqslant 0$. The fact that we require an *initial* condition to be satisfied implicitly requires that the solution must have a *functional meaning in time*, or at least at a neighbourhood of $t = 0$. This condition is obtained by requiring suitable conditions on X.

We require X to be in the subspace of tempered distributions such that their spatial Fourier Transforms are slow-growing measures over $\mathbb{R}^d \times \mathbb{R}^+$. U_0 is also required to be such that its (spatial) Fourier Transform is a slow-growing measure over $\mathbb{R}^d$. We apply then a spatial Fourier Transform to (5.68), and we obtain the transformed Cauchy problem

$$\begin{cases} \dfrac{\partial V}{\partial t} + gV = Y \\ V\big|_{t=0} = V_0 \end{cases} , \tag{5.69}$$

where $Y = \mathscr{F}_S(X)$ is a slow-growing measure over $\mathbb{R}^d \times \mathbb{R}^+$ and $V_0 = \mathscr{F}_S(U_0)$ is a slow-growing measure

over $\mathbb{R}^d$. In such a case we can prove the existence of a unique solution V to (5.69), which is a slow-growing measure over $\mathbb{R}^d \times \mathbb{R}^+$ and which can be described in a particular convenient way which will be specified further. For now, we remark that in a classical case where Y and V_0 are sufficiently regular functions (for example in $\mathscr{S}(\mathbb{R}^d \times \mathbb{R})$ and $\mathscr{S}(\mathbb{R}^d)$ respectively), the solution to (5.69) is given by

$$V(\xi, t) = e^{-tg(\xi)} V_0(\xi) + \int_0^t e^{-(t-s)g(\xi)} Y(\xi, s) ds, \tag{5.70}$$

and the solution to (5.69) is the inverse spatial Fourier Transform of V. Supposing that the function $\xi \mapsto e^{-tg(\xi)}$ is integrable for every $t > 0$, the solution can be expressed through the *Duhamel's Formula*:

$$U(x, t) = (\Phi_t * V_0)(x) + \int_0^t \left(\Phi_{t-s} \overset{(\mathbb{R}^d)}{*} Y(\cdot, s) \right)(x) ds, \tag{5.71}$$

where

$$\Phi_u(x) = (2\pi)^{\frac{d}{2}} \mathscr{F}_S(\xi \mapsto e^{-ug(\xi)})(x), \quad \forall u > 0. \tag{5.72}$$

In (5.71), the symbol $\overset{(\mathbb{R}^d)}{*}$ denotes a *spatial* convolution between the spatial function Φ_{t-s} and the spatial function $Y(\cdot, s)$. We remark that Φ_0 can be interpreted as the Dirac measure at 0. The regular case of formula (5.71) is actually restrictive and it does not include very interesting cases. For instance, in the *stochastized* version of the PDE, we are interested in the cases where Y behaves as a measure, for example when using a White Noise. Hence, our requirement that Y must be a slow-growing measure over $\mathbb{R}^d \times \mathbb{R}^+$ is more adapted. We will nevertheless be inspired by expression (5.71) in order to find a solution in our more general approach.

We recall that a function $f : \mathbb{R} \to \mathbb{C}$ is said to be *càdlàg* if it is right-continuous with left-limits. We recall that for every measure μ over $\mathbb{R}$ there exists a unique càdlàg function f such that $f(0) = \mu(\{0\})$ and which is a distributional primitive of μ (see footnote 22 in Chapter 3). Since in problem (5.69) we expect a solution whose derivative has the behaviour of a slow-growing measure, it is not surprising that the solution V may have a functional càdlàg meaning in time. Indeed, consider V the slow-growing measure over $\mathbb{R}^d \times \mathbb{R}^+$ solution to (5.69). It can be proven that it has a *càdlàg-in-time representation*. By this we mean the following: there exists a family of slow-growing measures over $\mathbb{R}^d$, $(V_t)_{t \in \mathbb{R}^+} \subset \mathscr{M}_{SG}(\mathbb{R}^d)$ such that

- for all $\varphi \in C_{FD}(\mathbb{R}^d)$, the function $t \in \mathbb{R}^+ \mapsto \langle V_t, \varphi \rangle$ is càdlàg.

- for all $\varphi \in C_{FD}(\mathbb{R}^d)$ and for all $\theta \in C_{FD}(\mathbb{R}^+)$,[2] we have

$$\langle V, \varphi \boxtimes \theta \rangle = \int_{\mathbb{R}^+} \langle V_t, \varphi \rangle \theta(t) dt. \tag{5.73}$$

It can be shown that the members in the family $(V_t)_{t \in \mathbb{R}^+}$ can be expressed as

$$V_t = e^{-tg}(V_0 - Y(\cdot \times \{0\})) + \int_{[0,t]} e^{-(t-s)g} Y(\cdot \times ds). \tag{5.74}$$

Considering that Y is a slow-growing measure over $\mathbb{R}^d \times \mathbb{R}^+$, expression $Y(\cdot \times \{0\})$ denotes a spatial slow-growing measure. We can express (5.74) in a more explicit way through the action of V_t over every bounded Borel set $A \in \mathcal{B}_B(\mathbb{R}^d)$ through

$$V_t(A) = \int_A e^{-tg(\xi)} d(V_0 - Y(\cdot \times \{0\}))(\xi) + \int_{A \times [0,t]} e^{-(t-s)g(\xi)} dY(\xi, s), \tag{5.75}$$

or equivalently through its action on a spatial test-function $\varphi \in C_{FD}(\mathbb{R}^d)$ through

$$\langle V_t, \varphi \rangle = \int_{\mathbb{R}^d} e^{-tg(\xi)} \varphi(\xi) d(V_0 - Y(\cdot \times \{0\}))(\xi) + \int_{\mathbb{R}^d \times [0,t]} e^{-(t-s)g(\xi)} \varphi(\xi) dY(\xi, s). \tag{5.76}$$

Finally, the solution of (5.69) is simply the inverse spatial Fourier Transform of V, $U = \mathscr{F}_S^{-1}(V)$. It can also be shown that there exists a càdlàg-in-time representation of U, which consists in a family of spatial tempered distributions in the space $\mathscr{V}'(\mathbb{R}^d)$ (See the definition in Chapter 4), $(U_t)_{t \in \mathbb{R}^+}$, such that

- for all $\varphi \in \mathscr{S}(\mathbb{R}^d)$, the function $t \in \mathbb{R}^+ \mapsto \langle U_t, \varphi \rangle$ is càdlàg.
- for all $\varphi \in \mathscr{S}(\mathbb{R}^d)$ and for all $\theta \in C_{FD}(\mathbb{R}^+)$, we have

$$\langle U, \varphi \boxtimes \theta \rangle = \int_{\mathbb{R}^+} \langle U_t, \varphi \rangle \theta(t) dt. \tag{5.77}$$

This family is simply obtained through $U_t = \mathscr{F}_S^{-1}(V_t)$, for all $t \in \mathbb{R}^+$. Hence, we can write

$$U_t = \mathscr{F}_S^{-1}\left(e^{-tg}(V_0 - Y(\cdot \times \{0\})) + \int_{[0,t]} e^{-(t-s)g} Y(\cdot \times ds) \right), \quad \forall t \in \mathbb{R}^+. \tag{5.78}$$

We remark that, differently to the regular case of functions (5.71), we cannot simply express U_t as a spatial convolution of distributions, since we are not aware if an Exchange Formula holds for the multiplication

[2] We denote by $C_{FD}(\mathbb{R}^+)$ the space of continuous functions defined over $\mathbb{R}^+$ with fast decreasing behaviour. That is, the continuous functions $\theta : \mathbb{R}^+ \to \mathbb{C}$ such that for every $N \in \mathbb{N}$ we have $\sup_{t \in \mathbb{R}^+} \left| (1+t^2)^N \theta(t) \right| < \infty$.

between the spatial function $\xi \mapsto e^{-tg(\xi)}$ and any arbitrary slow-growing measure. This holds, for example, when g is such that $\xi \mapsto e^{-tg(\xi)} \in \mathcal{O}_M(\mathbb{R}^d)$ for all $t \in \mathbb{R}^+$, case in which we can use an Exchange Formula at least on expression $e^{-tg}(V_0 - Y(\cdot \times \{0\}))$. It is not clear if we can do so in the case of expression $\int_{[0,t]} e^{-(t-s)g(\cdot)} Y(\cdot \times ds)$, even with this extra supposition on g.

5.3.2 Asymptotic behaviour

We consider now the slightly more restrictive case where X in (5.68) is required that both its spatial Fourier Transform and its spatio-temporal Fourier Transform are slow-growing measures over the whole space $\mathbb{R}^d \times \mathbb{R}$. Hence, $Y = \mathscr{F}_S(X) \in \mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R})$ and in addition $X \in \mathscr{V}'(\mathbb{R}^d \times \mathbb{R})$, the space of spatio-temporal tempered distributions whose spatio-temporal Fourier Transform is in $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R})$. In such a case, we can analyse the equation

$$\frac{\partial U}{\partial t} + \mathcal{L}_g U = X \tag{5.79}$$

simply through the analysis of the spatio-temporally Fourier Transformed problem

$$(i\omega + g(\xi))M_U = M_X, \tag{5.80}$$

where we have denoted by $M_X = \mathscr{F}(X)$ and $M_U = \mathscr{F}(U)$. Let us suppose that there exists a positive constant $\kappa > 0$ such that $g_R \geqslant \kappa$. If we consider solutions to (5.79) which are in $\mathscr{V}'(\mathbb{R}^d \times \mathbb{R})$ we can simply follow the approach of Proposition 4.6.1, and argue that there exists a solution $U \in \mathscr{V}'(\mathbb{R}^d \times \mathbb{R})$ if and only if $\frac{1}{i\omega + g(\xi)} M_X \in \mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R})$. We can apply Proposition 5.2.1 for $\beta = 1$ to argue that this holds in this case since $g_R \geqslant \kappa$. The solution is also unique since the spatio-temporal symbol function $i\omega + g(\xi)$ is never null. We will denote this solution by U^∞, which we know it is given by

$$U^\infty = \mathscr{F}^{-1}\left(\frac{1}{i\omega + g(\xi)} M_X\right) \tag{5.81}$$

The question that arises is if there is any relation between this solution $U^\infty \in \mathscr{V}'(\mathbb{R}^d \times \mathbb{R})$ and the solution to the Cauchy problem (5.68). The answer is yes, and it is described through a *spatio-temporal asymptotic convergence*. Under the already proposed extra requirements for X, if U is the solution to (5.68), it can be proven that for every $\epsilon > 0$ and for every $\varphi \in \mathscr{S}(\mathbb{R}^d)$, there exists $t_{\epsilon,\varphi} \in \mathbb{R}^+$ such that

$$|\langle U^\infty - U, \varphi \boxtimes \theta \rangle| < \epsilon, \quad \forall \theta \in \mathscr{S}(\mathbb{R}) \text{ such that } \operatorname{supp}(\theta) \subset \mathbb{R}^d \times [t_{\epsilon,\varphi}, \infty) \text{ and } \int_{\mathbb{R}^+} |\theta(t)| dt = 1. \tag{5.82}$$

Hence, the solution U^∞ describes how the solution U behaves spatio-temporally after enough time. This convergence does not depend on the initial condition U_0 used in the Cauchy problem (5.68).

In the case where M_U is *temporally integrable*, which is the case when

$$\int_{A\times\mathbb{R}} \frac{1}{i\omega + g(\xi)} dM_X(\xi,\omega) < \infty \quad , \forall A \in \mathcal{B}_B(\mathbb{R}^d), \tag{5.83}$$

it can be proven that U^∞ has a continuous-in-time representation $(U_t^\infty)_{t\in\mathbb{R}^+} \subset \mathscr{V}'(\mathbb{R}^d)$. If we consider the case where the initial condition to the Cauchy problem (5.68) is set to be $U_0 = U_0^\infty$, then it can be proven that the solution U equals U^∞ over $\mathbb{R}^d \times \mathbb{R}^+$.

5.3.3 *Stochastized* version

We consider now a *stochastized* version of problem (5.68), with analogue conditions to the deterministic case. Namely, X is now a real GeRF over $\mathbb{R}^d \times \mathbb{R}$ such that its spatial Fourier Transform $Y = \mathscr{F}_S(X)$ is a slow-growing Random Measure over $\mathbb{R}^d \times \mathbb{R}^+$. We suppose also that U_0 is a real GeRF over $\mathbb{R}^d$ such that its (spatial) Fourier Transform $V_0 = \mathscr{F}_S(U_0)$ is a slow-growing Random Measure over $\mathbb{R}^d$. The strict solution to this problem is constructed with complete analogy to the deterministic case. Let us explain it roughly. It can be shown that the problem has a unique solution U, which is a GeRF over $\mathbb{R}^d \times \mathbb{R}$ such that its spatial Fourier Transform is a slow-growing Random Measure over $\mathbb{R}^d \times \mathbb{R}^+$. It can also be proven that U has a càdlàg-in-time representation, that is, there exists a family of GeRFs over $\mathbb{R}^d$, $(U_t)_{t\in\mathbb{R}^+}$ with analogue properties to the continuous-in-time representation of a GeRF presented in Section 5.1.2, but requiring the associated random functions $t \mapsto \langle U_t, \varphi\rangle$ to be càdlàg in mean-square rather than continuous, for every $\varphi \in \mathscr{S}(\mathbb{R}^d)$. All members of the family $(U_t)_{t\in\mathbb{R}^+}$ satisfy that their (spatial) Fourier Transforms are slow-growing Random Measures over $\mathbb{R}^d$. For $V = \mathscr{F}_S(U)$, which is a slow-growing Random Measure over $\mathbb{R}^d \times \mathbb{R}^+$, we also obtain a càdlàg-in-time representation through $V_t = \mathscr{F}_S(U_t)$. The stochastic interpretation of expressions (5.75) and (5.76) are well-defined as stochastic integrals, and we can obtain the associated covariances expressions.

Since we have supposed that $Y = \mathscr{F}_S(X)$ is a slow-growing Random Measure over $\mathbb{R}^d \times \mathbb{R}^+$, its covariance measure C_Y is in $\mathscr{M}_{SG}((\mathbb{R}^d \times \mathbb{R}^+) \times (\mathbb{R}^d \times \mathbb{R}^+))$. Analogously, since $V_0 = \mathscr{F}_S(U_0)$ is also a slow-growing Random Measure, we have $C_{V_0} \in \mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^d)$. For simplicity, let us suppose that U_0 and X are independent. Let V be the solution of the stochastic transformed problem (5.69), for which it can be proven it is a slow-growing Random Measure. Then, the covariance of V, C_V, which is in $\mathscr{M}_{SG}((\mathbb{R}^d \times \mathbb{R}^+) \times (\mathbb{R}^d \times \mathbb{R}^+))$, can be described, for example, by analysing the random variables of the form (5.76), obtaining for two spatial test-functions $\varphi, \phi \in C_{FD}(\mathbb{R}^d)$ and for two time locations $t, s \in \mathbb{R}^+$,

$$\mathbb{C}ov(\langle V_t, \varphi\rangle, \langle V_s, \phi\rangle) = \int_{\mathbb{R}^d\times\mathbb{R}^d} e^{-tg(\xi)-s\overline{g(\eta)}}\varphi(\xi)\overline{\phi}(\eta) dC_{V_0 - Y(\cdot\times\{0\})}(\xi,\eta) \quad +$$
$$\int_{(\mathbb{R}^d\times[0,t])\times(\mathbb{R}^d\times[0,s])} \left(e^{-(t-u)g(\xi)} - e^{-tg(\xi)}\mathbf{1}_{\{0\}}(u)\right)\left(e^{-(s-v)\overline{g}(\eta)} - e^{-s\overline{g}(\eta)}\mathbf{1}_{\{0\}}(v)\right)\varphi(\xi)\overline{\phi}(\eta) dC_Y((\xi,u),(\eta,v)). \tag{5.84}$$

It turns out that the covariance measure C_V can be also expressed through a family of spatial covariance measures $(C_V^{(t,s)})_{(t,s)\in\mathbb{R}^+\times\mathbb{R}^+} \subset \mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^d)$, which satisfies:

- for all $\varphi, \phi \in C_{FD}(\mathbb{R}^d)$ and for all $t \in \mathbb{R}^+$, the function $s \in \mathbb{R}^+ \mapsto \langle C_V^{(t,s)}, \varphi \otimes \overline{\phi}\rangle$ is càdlàg.

- for all $\varphi, \phi \in C_{FD}(\mathbb{R}^d)$ and for all $s \in \mathbb{R}^+$, the function $t \in \mathbb{R}^+ \mapsto \langle C_V^{(t,s)}, \varphi \otimes \overline{\phi}\rangle$ is càdlàg.

- for all $\varphi, \phi \in C_{FD}(\mathbb{R}^d)$ and for all $\theta_1, \theta_2 \in C_{FD}(\mathbb{R}^+)$, it holds that

$$\langle C_V, (\varphi \boxtimes \theta_1) \otimes \overline{(\phi \boxtimes \theta_2)}\rangle = \int_{\mathbb{R}^+\times\mathbb{R}^+} \langle C_V^{(t,s)}, \varphi \otimes \overline{\phi}\rangle \theta_1(t)\overline{\theta_2}(s) d(t,s). \tag{5.85}$$

The family $(C_V^{(t,s)})_{(t,s)\in\mathbb{R}^+\times\mathbb{R}^+}$ is of course determined by expression (5.84) through

$$\langle C_V^{(t,s)}, \varphi \boxtimes \overline{\phi}\rangle = \mathbb{C}ov(\langle V_t, \varphi\rangle, \langle V_s, \phi\rangle), \quad \forall \varphi, \phi \in C_{FD}(\mathbb{R}^d), \tag{5.86}$$

which we may write similarly as expression (5.74):

$$\begin{aligned} C_V^{(t,s)} &= (e^{-tg} \otimes e^{-s\overline{g}}) C_{V_0 - Y(\cdot\times\{0\})} \\ &\quad + \int_{[0,t]\times[0,s]} \left(e^{-(t-u)g} - e^{-tg}\mathbf{1}_{\{0\}}(u)\right) \otimes \left(e^{-(s-v)\overline{g}} - e^{-s\overline{g}}\mathbf{1}_{\{0\}}(v)\right) dC_Y((\cdot, du)\times(\cdot, dv)). \end{aligned} \tag{5.87}$$

The solution U can also be expressed through a càdlàg-in-time representation, defining the family $(U_t)_{t\in\mathbb{R}^+}$ through $U_t = \mathscr{F}_S^{-1}(V_t)$ for every $t \in \mathbb{R}^+$. The covariance structure of U is described by a distribution C_U over $(\mathbb{R}^d \times \mathbb{R}^+) \times (\mathbb{R}^d \times \mathbb{R}^+)$, which can also be described by a family $(C_U^{(t,s)})_{(t,s)\in\mathbb{R}^+\times\mathbb{R}^+} \subset \mathscr{V}'(\mathbb{R}^d \times \mathbb{R}^d)$. This family is defined by

$$\langle C_U^{(t,s)}, \varphi \otimes \overline{\phi}\rangle = \langle C_V^{(t,s)}, \mathscr{F}_S^{-1}(\varphi) \otimes \overline{\mathscr{F}_S^{-1}(\phi)}\rangle, \tag{5.88}$$

hence, it holds that

$$\begin{aligned} \langle C_U, (\varphi \boxtimes \theta_1) \otimes \overline{(\phi \boxtimes \theta_2)}\rangle &= \int_{\mathbb{R}^+\times\mathbb{R}^+} \langle C_U^{(t,s)}, \varphi \otimes \overline{\phi}\rangle \theta_1(t)\overline{\theta_2}(s) d(t,s) \\ &= \int_{\mathbb{R}^+\times\mathbb{R}^+} \langle C_V^{(t,s)}, \mathscr{F}_S^{-1}(\varphi) \otimes \overline{\mathscr{F}_S^{-1}(\phi)}\rangle \theta_1(t)\overline{\theta_2}(s) d(t,s), \end{aligned} \tag{5.89}$$

for every $\varphi, \phi \in \mathscr{S}(\mathbb{R}^d)$ and $\theta_1, \theta_2 \in C_{FD}(\mathbb{R}^+)$.

5.3.4 Asymptotic convergence to the stationary solution

We consider now a particular case for X and U_0 for which we obtain a stochastic analogue result to the convergence when the time flows long enough to a convenient solution of the equation (5.79).

Let X be a real stationary GeRF over $\mathbb{R}^d \times \mathbb{R}$ such that its *spatial* Fourier Transform is a Random Measure over $\mathbb{R}^d \times \mathbb{R}$. Of course, the spatio-temporal Fourier Transform of X, say $M_X = \mathscr{F}(X)$, is also a Random Measure (Theorem 3.4.2), but the requirement that its spatial Fourier Transform is also a Random Measure is an extra supposition. Examples of such GeRFs are separable models between any spatial stationary GeRF and a temporal continuous stationary Random Function, or a coloured in space and white in time noise.

From Proposition 5.2.1 it can be concluded that if we suppose that $g_R \geqslant \kappa$ for some $\kappa > 0$, there exists a unique stationary solution to the SPDE (5.79). Let us call U^{stat} this stationary solution, which is given by

$$U^{stat} = \mathscr{F}^{-1}\left(\frac{1}{i\omega + g(\xi)}\mathscr{F}(X)\right) = \mathscr{F}^{-1}\left(\frac{1}{i\omega + g(\xi)}M_X\right). \tag{5.90}$$

Consider now U to be the solution to the stochastic Cauchy problem (5.68), where U_0 is a real stationary GeRF over $\mathbb{R}^d$ independent of X. We remark that in such a problem only the values that X takes over $\mathbb{R}^d \times \mathbb{R}^+$ intervene, and not the values over the negative time. Then, the following fact about the asymptotic convergence for large enough t is obtained: for all $\epsilon > 0$ and for all $\varphi \in \mathscr{S}(\mathbb{R}^d)$, there exists $t_{\epsilon,\varphi} \in \mathbb{R}^+$ such that

$$\mathbb{E}\left(\left|\langle U - U^{stat}, \varphi \boxtimes \theta\rangle\right|^2\right) < \epsilon, \quad \forall \theta \in \mathscr{S}(\mathbb{R}^d) \text{ such that } \operatorname{supp}(\theta) \subset [t_{\epsilon,\varphi}, \infty) \text{ and } \int_{\mathbb{R}^+} |\theta(t)|dt = 1. \tag{5.91}$$

Hence, the solution U is arbitrarily close in a mean-square sense to the stationary solution U^{stat} for large enough times.

5.3.5 Some examples. Time Markovianity.

In order to show some particular examples, we will consider the case with the requirement presented in Section 5.3.4. We will always suppose that there exists $\kappa > 0$ such that $g_R \geqslant \kappa$.

Consider U_0 a spatial real stationary GeRF with spectral measure μ_{U_0}. Let X be a spatio-temporal real stationary GeRF with separable form $X = X_S \boxtimes X_T$ (we recall that this is not a tensor product *stricto sensus*), where X_S represents any spatial real stationary GeRF with spectral measure μ_{X_S} and X_T represents a temporal real stationary GeRF which can be either a continuous Random Function over time or a White Noise in time, hence with spectral measure μ_{X_T} either finite or either proportional to the Lebesgue measure. We suppose U_0 and X independent. If U^{stat} is the unique stationary solution to (5.79), its spectral measure

is

$$d\mu_{U^{stat}}(\xi,\omega) = \frac{d\mu_{X_S}(\xi)d\mu_{X_T}(\omega)}{|i\omega + g(\xi)|^2}, \tag{5.92}$$

and in this case, this spectral measure is always temporally integrable. The spectral measure of the spatial traces of U^{stat} is given by

$$\mu_{U^{stat}_S}(A) = \frac{1}{\sqrt{2\pi}} \int_A \int_{\mathbb{R}} \frac{d\mu_{X_T}(\omega)}{|i\omega + g(\xi)|^2} d\mu_{X_S}(\xi), \quad \forall A \in \mathcal{B}_B(\mathbb{R}^d). \tag{5.93}$$

Consider the solution U to the stochastic Cauchy problem (5.68). The analysis of its covariance structure is easier through the analysis of the covariance structure of $V = \mathscr{F}_S(U)$ which is the solution to the stochastic transformed problem (5.69). V has a continuous-in-time representation $(V_t)_{t\in\mathbb{R}^+}$ with the covariance structure given by (5.84). In our case $V_0 = \mathscr{F}_S(U_0)$ is orthogonal, as well as the spatial behaviour of $Y = \mathscr{F}_S(X) \overset{2nd\ o.}{=} \mathscr{F}_S(X_S)\boxtimes X_T$. In addition, with our suppositions $Y(\cdot \times \{0\})$ is null almost-surely, since its acts as a continuous Random Function or as a White Noise in time. We obtain thus the next expression for the covariance structure of V for every pair $(t,s) \in \mathbb{R}^+ \times \mathbb{R}^+$ and every $\varphi, \phi \in \mathscr{S}(\mathbb{R}^d)$:

$$\begin{aligned} \mathbb{C}ov(\langle V_t, \varphi\rangle, \langle V_s, \phi\rangle) &= (2\pi)^{\frac{d}{2}} \int_{\mathbb{R}^d} e^{-tg(\xi) - s\overline{g}(\xi)} \varphi(\xi)\overline{\phi}(\xi) d\mu_{U_0}(\xi) \\ &+ (2\pi)^{\frac{d}{2}} \int_{\mathbb{R}^d} \int_{[0,t]\times[0,s]} e^{-(t-u)g(\xi) - (s-v)\overline{g}(\xi)} \varphi(\xi)\overline{\phi}(\xi) dC_{X_T}(u,v) d\mu_{X_S}(\xi). \end{aligned} \tag{5.94}$$

Here C_{X_T} denotes de covariance distribution of X_T, which under our assumptions, is either a continuous temporal stationary covariance function, with $dC_{X_T}(u,v) = \rho_{X_T}(u-v)dudv$, or the covariance distribution of a temporal White Noise, $C_{X_T} = \delta(u-v) = \mathrm{Leb}\,\delta^{\{u=v\}}$. The solution U has a continuous-in-time representation $(U_t)_{t\in\mathbb{R}^+}$ given by $U_t = \mathscr{F}_S^{-1}(V_t)$, and hence

$$\begin{aligned} \mathbb{C}ov(\langle U_t, \varphi\rangle, \langle U_s, \phi\rangle) &= (2\pi)^{\frac{d}{2}} \int_{\mathbb{R}^d} e^{-tg(\xi) - s\overline{g}(\xi)} \mathscr{F}_S^{-1}(\varphi)(\xi)\overline{\mathscr{F}_S^{-1}(\phi)}(\xi) d\mu_{U_0}(\xi) \\ &+ (2\pi)^{\frac{d}{2}} \int_{\mathbb{R}^d} \int_{[0,t]\times[0,s]} e^{-(t-u)g(\xi) - (s-v)\overline{g}(\xi)} \mathscr{F}_S^{-1}(\varphi)(\xi)\overline{\mathscr{F}_S^{-1}(\phi)}(\xi) dC_{X_T}(u,v) d\mu_{X_S}(\xi). \end{aligned} \tag{5.95}$$

As stated in the previous section, the solution U converges to the stationary solution U^{stat} as t becomes large enough.

In some particular cases, it holds that if we consider a stationary initial condition U_0 following the spatial spectral measure (5.93), the covariance structure of the solution U is stationary with the same spectral

measure as U^{stat}, given by (5.92). This is the example where $X_T = W_T$. In such a case, equation (5.94) is

$$\begin{aligned}
&\mathbb{C}ov(\langle V_t, \varphi\rangle, \langle V_s, \phi\rangle) \\
&\quad = (2\pi)^{\frac{d}{2}} \int_{\mathbb{R}^d} e^{-tg(\xi)-s\overline{g}(\xi)} \varphi(\xi)\overline{\phi}(\xi) d\mu_{U_0}(\xi) \\
&\qquad + (2\pi)^{\frac{d}{2}} \int_{\mathbb{R}^d} \int_{[0,t\wedge s]} e^{-(tg(\xi)-s\overline{g}(\xi))+2ug_R(\xi)} \varphi(\xi)\overline{\phi}(\xi) du d\mu_{X_S}(\xi). \\
&\quad = (2\pi)^{\frac{d}{2}} \int_{\mathbb{R}^d} e^{-tg(\xi)-s\overline{g}(\xi)} \varphi(\xi)\overline{\phi}(\xi) d\mu_{U_0}(\xi) \\
&\qquad + (2\pi)^{\frac{d}{2}} \left[\int_{\mathbb{R}^d} \frac{e^{-i(t-s)g_I(\xi)-|t-s|g_R(\xi)}}{2g_R(\xi)} \varphi(\xi)\overline{\phi}(\xi) d\mu_{X_S}(\xi) - \int_{\mathbb{R}^d} \frac{e^{-(tg(\xi)+s\overline{g}(\xi))}}{2g_R(\xi)} \varphi(\xi)\overline{\phi}(\xi) d\mu_{X_S}(\xi) \right],
\end{aligned} \tag{5.96}$$

where we have solved the temporal integral and used the expression $t \wedge s = \frac{t+s-|t-s|}{2}$. We remark the similarities between the "time stationary term"

$$(2\pi)^{\frac{d}{2}} \int_{\mathbb{R}^d} \frac{e^{-i(t-s)g_I(\xi)-|t-s|g_R(\xi)}}{2g_R(\xi)} \varphi(\xi)\overline{\phi}(\xi) d\mu_{X_S}(\xi), \tag{5.97}$$

which depends only on the gap $t-s$, and the expression of the stationary covariance (5.39). The asymptotic convergence to the stationary covariance model determined by (5.92) can be simply obtained in this case by analysing expression (5.96) and remarking that the first and third terms tend to 0 as $t, s \to \infty$. It is also more obvious to see that if we chose the right model for U_0 through the identification $\mu_{U_0} = \frac{\mu_{X_S}}{2g_R}$, the solution follows a stationary model whose covariance is the same as in (5.39).

The case $X = X_S \boxtimes W_T$ presents also another particularity. Let us consider, for simplicity and rigorousness, that all the real GeRFs involved are Gaussian. Let us write $M_{X_S} = \mathscr{F}_S(X)$. The orthogonality in time of $X_T = W_T$ induces a *Markovianity in time*, in the classical sense of Markovianity. Precisely, if U is the solution to the stochastic Cauchy problem (5.68), and $(U_t)_{t\in\mathbb{R}^+}$ is its continuous-in-time representation, then for any $s > 0$, the GeRFs $(U_t)_{t\geqslant s}$ are independent to the GeRFs $(U_t)_{t\in[0,s)}$, conditionally to the GeRF U_s. To see this, we can consider the solution to the transformed problem $V = \mathscr{F}_S(U)$, represented through its continuous-in-time representation $(V_t)_{t\in\mathbb{R}^+}$, for a $\Delta t > 0$, and for any $\varphi \in \mathscr{S}(\mathbb{R}^d)$ through

$$\begin{aligned}
\langle V_{t+\Delta t}, \varphi\rangle &= \int_{\mathbb{R}^d} e^{-(t+\Delta t)g(\xi)} \varphi(\xi) dV_0(\xi) + \int_{\mathbb{R}^d\times[0,t+\Delta t]} e^{-(t+\Delta t-s)} \varphi(\xi) d(M_{X_S} \boxtimes W_T)(\xi, s) \\
&= \langle V_t, e^{-\Delta t g}\varphi\rangle + \int_{\mathbb{R}^d\times(t,t+\Delta t]} e^{-(t+\Delta t-s)} \varphi(\xi) d(M_{X_S} \boxtimes W_T)(\xi, s).
\end{aligned} \tag{5.98}$$

Since $Y = M_{X_S} \boxtimes W_T$ is an orthogonal Random Measure, and the expression $\langle V_t, e^{-\Delta t g}\varphi\rangle$ only considers

integrations with respect to Y up to the time t, then the expression

$$\int_{\mathbb{R}^d \times (t,t+\Delta t]} e^{-(t+\Delta t-s)} \varphi(\xi) d(M_{X_S} \boxtimes W_T)(\xi, s)$$

is uncorrelated with $\langle V_t, e^{-\Delta t g} \varphi \rangle$, and hence they are independent in this Gaussian framework. The temporal Markovianity of the solution U is obtained immediately through the relation $U_t = \mathscr{F}_S^{-1}(V_t)$ for every $t \in \mathbb{R}^+$. We remark however, that we have been able to prove this Markovian structure only when X has a White Noise in time covariance structure. Indeed, since we have required X to be stationary, and that $\mathscr{F}_S(X)$ is an orthogonal Random Measure, its temporal covariance structure must be such that it is a stationary orthogonal Random Measure, and the unique covariance structure which satisfies this is the one of the White Noise (see Example 3.4.7), up to a multiplicative constant.

Chapter 6

Simulations

SUMMARY

In this chapter we present a method of non-conditional simulation of general stationary GeRFs. This method is based on a convenient approximation of the Fourier Transform of the field. We show that this method is immediately adapted to the models already presented in this dissertation. Taking advantage of numerical algorithms such as the Fast Fourier Transform, efficient simulation methods can be achieved. This method is not new, and it has been already applied to obtain efficient simulations of stationary Random Fields related to SPDEs, although a theoretical proof of its performance lacks in the literature.

In the introductory Section 6.1 we describe roughly and without technical details some methodologies of simulation of Random Fields within the SPDE approach. They consist in using numerical solvers for PDEs applied in a stochastic framework. We describe roughly the Finite Element Method, together with its advantages and issues. We also mention spectral methods, that is, methods based on the development of the Random Field in a suitable basis of orthonormal functions. We then present the motivations which lead us to select the method presented in this chapter. They rely mainly on its adaptability to the models presented in this dissertation and to the capacity of simulating general classes of models over large grids with an efficient computing time.

In Section 6.2 we present the theoretical foundations of the method. In the case of a stationary GeRF, we show that this method provides approximations which converge in a weak sense to the desired solution. We are also able to prove a mean-square-uniformly-on-compacts convergence in the case of continous Random Functions. We show how to apply this method in the cases of SPDEs presented in this dissertation. We are able to prove convergence of the approximation

under suitable conditions. We also show how to apply this method to obtain simulations of first order evolution models and Waving models.

In Section 6.3 we present the implementation of this method together with the technical details that must be considered. We propose a qualitative error analyse in the case of the Matérn model, for which the mean-square-uniformly-on-compacts convergence of the covariances is theoretically proven. We propose illustrations of different types of models that can be easily simulated with this method. In particular, cases with different forms of advections, asymmetries, and separated regularities along different directions are illustrated. We show illustrations of first order evolution models, together with the theoretical asymptotic convergence exposed in Section (5.3)*. We also present illustrations of a Waving model following a Matérn covariance in space.*

We finish in Section 6.4 with some final words. We discuss the advantages and disadvantages of this method, together with propositions to improve it.

The proofs of the theoretical results are presented in Appendix A.

6.1 Introduction

Within the SPDE approach, the simulation of a Random Field related to a SPDE can be performed through the use of numerical solvers to PDEs. Such an approach is performed by considering a suitably discretized or approximated version of the PDE and its solution, restricting the space of possible solutions to spaces of finite dimension. The application to the stochastic framework is done simply by replacing the deterministic functions with Random Functions of GeRFs, which finally consists in simulating a suitable random vector.

The most popular method used within the SPDE approach in Geostatistics is the FEM. This method considers an approximation of the solution of the PDE expressed as a finite linear combination of suitable functions defined over the space. These functions are determined by a triangulation of the working domain, usually consisting in a mesh of triangular elements with associated nodes an edges. There are many bibliographical sources on this method, both for the deterministic framework of PDEs and the stochastic one. We suggest Zienkiewicz et al. (2013) and Braess (2007) as treaties on this practice in the deterministic case. In the probabilist community, this method is widely used to analyse approximations of solutions to SPDEs. See Stefanou (2009) for a general review and Barth & Lang (2012) for application examples. We will not enter into the technical details of this method. However, we remark its main advantages within the needs of the geostatistical community, together with its limitations when trying to apply it to more general cases.

In the geostatistical community, the use of this method was popularised by Lindgren et al. (2011), in which the main interesting properties while solving the equation associated to the Matérn model are pre-

sented. As mentioned in the introductory chapter, when the Matérn model has a Markovian behaviour ($\alpha \in \mathbb{N}$ in Eq. (4.15), applying Rozanov's Theorem presented in Example 4.5.3), the precision matrices obtained when applying the FEM are sparse, condition which allows a fast computational treatment of the model. Another particularity of the FEM, which holds regardless of the Markovianity of the Random Field, is that the values obtained by the approximation at the nodes of the triangulation mesh are given immediately by the method. Hence, when facing a particular data base with values located at arbitrary points in the space, the FEM can be easily adapted for inference and conditional simulation methods by identifying the sampling location points with nodes in the triangulation mesh. Inference methods and conditional simulations can be then performed, the required precision matrices being already obtained once the SPDE is discretized.

Although the FEM presents many advantages, it is not immediately adapted to more general models related to SPDEs which do not involve classical differential operators. For instance, Lindgren et al. (2011) apply the method for the Matérn model only for integer values of α. Other values with fractional regularities must be treated differently through suitable adaptations. In the commentary section of Lindgren et al. (2011), the authors propose to approximate the target Random Field in the case of fractional α by a suitable Markov Random Field, with a spectral density defined by the inverse of a suitable strictly positive polynomial determined in order to obtain an appropiate approximation. In Bolin & Kirchner (2017) a method of adaptation of the FEM to cases of the Matérn model with fractional regularity parameter is proposed, performed through a rational approximation of the SPDE. The methodology allows to obtain simulation and inference methods maintaining the computational benefits of the Markovian case. In general terms, this method needs ad-hoc adaptations when facing different types of SPDEs. Hence, the generality of this method is limited. For non-Markovian models, the sparse condition on the precision matrices is also lost in general, needing an extra special treatment. The necessity of suitable adaptations are more intricate when considering spatio-temporal PDEs, specially if they involve fractional operators of different orders in time and space, as the models presented in Chapter 5.

Another approach of numerical resolution of PDEs, and hence to SPDEs in a stochastic framework is done through **spectral methods**. The term *spectral* has different meanings depending on the community. In the geostatistical community, the term *spectral* is often used for methods of simulations or inference based on a suitable utilisation of the spectral measure of a stationary Random Field. See the usage, for instance, in Chilès & Delfiner (1999, Section 7.5.3), Lantuéjoul (2013, Section 15.2.3), and Emery et al. (2016). In the *PDE community*, the term is used for methods of numerical resolution of PDEs based on the development of the functions in a basis of linearly independent functions generating the space of possible solutions, often taken to be a complete orthonormal basis with respect to the interior product of a suitable Hilbert space to which the theoretical solution belongs. The solution, the source term and other functions involved in the PDE can be formally developed in this basis through an infinite (countable) linear combination. The typical approach is to *truncate* the infinite development at a large enough finite order, obtaining a development on a subspace of finite dimension of the original vector space, generated by a finite sub-basis of the original

infinite basis. The coefficients associated to each orthonormal function are then determined by the PDE, and can be obtained solving linear systems in the case of linear PDEs, analogously to the FEM. The convergence of these approximations to the theoretical solution, when it exists, are considered in the sense of the norm of the associated Hilbert space. The initial basis of orthogonal functions can be selected in many ways. One possible approach is to select them considering the geometry of the domain where the PDE is analysed. Typical approaches are, for example, to chose an orthonormal basis of the space $L^2(D)$ for a domain $D \subset \mathbb{R}^d$. Spaces of the form $L^2(D, \lambda)$, with λ being a suitable positive measure over the domain D are also considered, the resulting functions in the basis being usually identified with convenient polynomials such as Chebyshev polynomials or Hermite polynomials. Sobolev spaces are also used when facing fractional differential operators or in contexts where the solution has a fractional differentiability order. When working in particular geometrical settings, ad-hoc basis of functions can be selected. For example when considering PDEs defined over the unitary sphere in $\mathbb{R}^d$, $\partial B_1^{(d)}(0)$, the spherical harmonic functions are often used (Dai & Xu, 2013). Another interesting approach is to select a basis of functions which is not independent of the operator involved on the PDE. For instance, when facing an equation involving different forms of the Laplace operator over bounded domains, a typical approach consists in considering the basis of eigenfunctions of minus the Laplacian $-\Delta$; see for instance the developments in M. D. Ruiz-Medina et al. (2016). We refer to Gottlieb & Orszag (1977) for a simple introduction on spectral methods in a deterministic framework, presenting applications to typical PDEs and showing the advantages and disadvantages that such a method may present in particular contexts. We refer to Canuto et al. (2006) for a deeper exposition.

In a stochastic framework, spectral methods (in the PDE sense) are widely used in the probabilist community to analyse approximations of solutions to specific SPDEs. See for instance Kærgaard (2013) for a source with an explicit theoretical background plus applications, and Bréhier et al. (2016) for the study of the resolution of specific space-time SPDEs with this approach. In Lang et al. (2015) applications for the case of differential equations defined over the sphere can be found. In the geostatistical community this approach has not been widely exploited, at the best of our knowledge. Some examples using a wavelet basis can be found in M. D. Ruiz-Medina et al. (2016). An important example of spectral method is obtained when using the Karhunen-Loève expansion of a stochastic process, where the basis of orthonormal functions is taken to be adapted to the covariance model of the process; see Loève (1978, Chapter XI) and Yaglom (1987, Chapter 4, Section 26.1).

The question that arises is how to discriminate between the already existent simulation methods based on PDE solvers and identify which ones are more easily adaptable to the framework presented in this dissertation. The FEM lacks in generality for cases with complicated operators involved, needing ad-hoc adaptations. A spectral method (in the PDE sense) requires to fix a particular basis of functions which may be useful in some particular settings, but that could be less adapted to different kinds of equations.

In this chapter we have decided to apply an already existent method which may be catalogued as a

spectral method both in the geostatistical and PDE sense. It is intimately related to the oldest spectral method: the one consisting in the development of a periodic and square-integrable over a rectangle function in the *Fourier basis*. It is not, however, *exactly* this method. As mentioned, in a usual spectral method the basis of orthonormal functions is fixed, and the approximation is done by truncating the development of the target function up to some large enough order. In contrast, the method presented here *changes the basis of functions when changing the approximation order*. We remark that this is also the case when doing an approximation through the FEM. This method is based on the approximation of the Fourier Transform of a stationary Random Field, which is an orthogonal Random Measure. Under suitable arrangements which will be specified in this chapter, we can obtain an easily computable form of a Discrete Fourier Transform, and hence apply the Fast Fourier Transform algorithm (FFT) (Cooley & Tukey, 1965) to obtain a fast simulation even for large simulation grids.

This method is not new at all. We refer to Pardo-Iguzquiza & Chica-Olmo (1993) for a detailed exposition of the method considering a geostatistical approach, that is, done under the context and needs of geostatistical simulations and applications. A general description of this method is also presented in Chilès & Delfiner (1999, Section 7.5.3), where adequate bibliographical sources concerning the details of this method are presented. In Lang & Potthoff (2011) this method is also presented in the context of the numerical resolution of SPDEs. The equations considered therein are almost of the same form as the ones we presented in Chapter 4. The differences rely on our use of complex symbol functions and arbitrary stationary source terms, while in Lang & Potthoff (2011) the exposition of the method is restricted to positive symbol functions and White Noise source terms. In practice, these restrictions do not really pose a real problem: it can be proven that by using positive symbol functions we obtain the same desired covariance structures for the solutions (the final spectral measure is determined by $|g|$), and for many applications and interesting models, such as those presented in Chapter 5, the restriction to the case of a White Noise source term is not an issue. Nevertheless, even if this method is old, in the literature there is a lack of theoretical justifications of the convergence of the approximations to the target model to be simulated. Pardo-Iguzquiza & Chica-Olmo (1993) compare experimental variograms obtained from the simulations with theoretical models with satisfactory results. Lang & Potthoff (2011) illustrate the convergence to theoretical covariance expressions when increasing the approximation order in the case of a Matérn covariance model. However, none of these sources presents a rigorous mathematical proof of some form of convergence of the method when increasing the approximation order.

The framework of GeRFs exposed in this dissertation has allowed us to prove the theoretical convergence of this method in quite general cases. When the Random Fields are interpreted as GeRFs, we are able to prove the convergence to the target solution in a mean-square-$\mathscr{S}'(\mathbb{R}^d)$-weak-$*$ sense, which is probably the weakest form of convergence which may be achieved without using more general theories. Under suitable conditions, we have been able to prove a mean-square-uniformly-on-compacts convergence in the case of stationary Random Functions. This implies that the covariance functions of the approximations converge

uniformly on compacts to the covariance function of the target Random Function. This convergence is more adapted to geostatistical needs than other kinds of convergences in the sense of Hilbert spaces obtained when applying typical spectral methods, since a point-wise convergence guarantees, for instance, that the variance of the Random Field will be well-approximated when considering large enough approximation orders. It also allows us to prove the convergence of the approximation when considering continuous non-differentiable Random Functions, case in which, for instance, typical spectral methods based on the Fourier basis may fail to approximate the solution (See Deitmar, 2005, Chapter 1).

The attractiveness of this method relies mainly on two aspects: first, it is computationally fast thanks to the orthogonal structure of the Fourier Transform of a stationary Random Field and the application of the FFT algorithm; second, it is general and immediately adaptable to the context of SPDEs presented in this dissertation. Indeed, this method, as it will be seen, is ad-hoc for cases of equations of the form (4.3), and hence it allows us to simulate approximations of quite general and interesting Random Fields, whether if its associated SPDE involves a classical differential operator or not, contrarily to the case of the FEM. We are thus able to illustrate $2D$-versions of the models developed in Chapter 5 without technical issues or adaptability needs.

We could refer to Pardo-Iguzquiza & Chica-Olmo (1993) and to Lang & Potthoff (2011) for the implementation details. However, it seems more convenient to present them *in our way* and within the context of this dissertation, so the theoretical proofs and practical implementation issues will be exposed with more clarity. We will thus, present all the technical details of this method and its implementation. For spatio-temporal models, we adapt this method to obtain simulation techniques of first order evolution models inspired by the developments in Section 5.3 in Chapter 5. The method is a generalization of the methodology proposed in Sigrist et al. (2015) in the case of the advection-diffusion equation (Example 5.2.1), which consists in a Fourier Analysis-based spectral method in space using FFT, with an explicit resolution of the equation over time. We also present the adaptation of this method to simulate Waving models.

6.2 A Spectral Method based on the Fourier Transform

The method we present here is not exactly based on the development of the stochastic process on the Fourier basis but rather on an approximation of its Fourier Transform. We will see that both approaches are intimately related but they are not exactly the same. We restrict ourselves to stationary Random Fields. In such a case, the Fourier Transform of the process is an orthogonal Random Measure, finite if the process is a Random Function, and slow-growing if the process is a GeRF. The approach is then to approximate *this* Random Measure and then apply a Fourier Transform which can be expressed in a convenient discretized manner.

6.2.1 Theoretical foundations of the method

Let us consider our definition of a Riemann sequence of partitions of a bounded Borel set of $\mathbb{R}^d$ which we introduced in Section 3.2.2. Such a sequence consists of a collection of bounded subsets of $\mathbb{R}^d$ satisfying suitable properties. Rather than recalling the details, we will give now a definition of a **Riemannn sequence of partitions growing to the whole space** $\mathbb{R}^d$. A sequence of finite collections of bounded Borel subsets of $\mathbb{R}^d$, $(V_j^N)_{j\in\{1,...,N\},N\in\mathbb{N}_*}$, is said to be a Riemann sequence of partitions growing to $\mathbb{R}^d$ if

- $V_j^N \cap V_k^N = \varnothing$, for all $j,k \in \{1,...,N\}$ such that $j \neq k$, for all $N \in \mathbb{N}_*$,
- $\max\limits_{j\in\{1,...,N\}} \operatorname{diam}(V_j^N) \to 0$ as $N \to \infty$,
- for all $K \subset \mathbb{R}^d$ compact, there exists $N_0 \in \mathbb{N}$ such that for all $N \geqslant N_0$, $K \subset \bigcup_{j=1}^N V_j^N$.

Hence, this sequence forms partitions of bounded subsets of $\mathbb{R}^d$ whose union *grows* to the space $\mathbb{R}^d$ as N grows, and such that the *size* of each set in the partition decreases to 0 as N grows. For every $N \in \mathbb{N}_*$, we denote by $D_N := \mathbb{R}^d \backslash \bigcup_{j=1}^N V_j^N$. The sequence of sets $(D_N)_{N\in\mathbb{N}_*}$ *decreases to* $\varnothing$ as N grows, in the sense that $\bigcap_{N\in\mathbb{N}_*} D_N = \varnothing$. To the sequence of partitions $(V_j^N)_{j\in\{1,...,N\},N\in\mathbb{N}_*}$, we associate an arbitrary sequence of finite collections of points in $\mathbb{R}^d$, $(\xi_j^N)_{j\in\{1,...,N\},N\in\mathbb{N}_*}$[1], satisfying $\xi_j^N \in V_j^N$ and called the *tag points* of $(V_j^N)_{j\in\{1,...,N\},N\in\mathbb{N}_*}$. We denote by $\ell_N := \max_{j\in\{1,...,N\}} \operatorname{diam}(V_j^N)$ for every $N \in \mathbb{N}_*$. Hence $(\ell_N)_{N\in\mathbb{N}_*}$ is a sequence of positive real numbers which converges to 0. For every $N \in \mathbb{N}_*$, we will consider an additional tag point $d_N \in D_N$, which will play an auxiliary role. The definition of a Riemann sequence of partitions *growing to any other unbounded Borel set of* $\mathbb{R}^d$ is completely analogous.

Consider now a (deterministic) measure over $\mathbb{R}^d$, $\mu \in \mathscr{M}(\mathbb{R}^d)$. We can consider an *approximation* of μ by defining:

$$\mu_N = \sum_{j=1}^N \mu(V_j^N)\delta_{\xi_j^N}. \tag{6.1}$$

For every $N \in \mathbb{N}_*$, μ_N is a finite measure (it is actually compactly supported), and it is easy to prove (see Lemma A.5.1, using $\operatorname{supp}(\varphi)$ as A) that for any $\varphi \in C_c(\mathbb{R}^d)$, $\langle \mu_N, \varphi\rangle \to \langle \mu, \varphi\rangle$ as N grows. Hence, the sequence of measures $(\mu_N)_{N\in\mathbb{N}_*}$ converges to μ in the sense of the weak-$*$ topology on the space $\mathscr{M}(\mathbb{R}^d) = C_c'(\mathbb{R}^d)$.

Let us consider now the stochastic case. We consider the case of a real stationary GeRF over $\mathbb{R}^d$, Z with stationary covariance distribution ρ_Z and spectral measure μ_Z. Following Theorem 3.4.2, its Fourier Transform, which will be denoted by $M_Z = \mathscr{F}(Z)$, is a complex Hermitian slow-growing orthogonal Random Measure. We are going to interpret M_Z both as a Random set-function (considering the random

[1] In this context, we will use the notation ξ for the variables in $\mathbb{R}^d$ since the Riemann sequence of partitions will be actually constructed over the frequency space.

variables of the form $M_Z(A)$ for $A \in \mathcal{B}_B(\mathbb{R}^d)$) and as a linear functional (considering the random variables $\langle M_Z, \varphi \rangle$ for every φ integrable with respect to M_Z).

For every $N \in \mathbb{N}_*$ let us consider the complex Random Measure defined as

$$M_{Z_N} = \sum_{j=1}^{N} M_Z(V_j^N)\delta_{\xi_j^N}. \tag{6.2}$$

M_{Z_N} is compactly supported, hence it is a finite complex Random Measure. In addition, M_{Z_N} is orthogonal. Indeed, let us consider $A, B \in \mathcal{B}_B(\mathbb{R}^d)$. Since M_Z is an orthogonal Random Measure with weight $(2\pi)^{\frac{d}{2}}\mu_Z$, and since the class of sets $(V_j^N)_{j\in\{1,\dots,N\}}$ forms a partition, we conclude that

$$\begin{aligned}
\mathbb{C}ov(M_{Z_N}(A), M_{Z_N}(B)) &= \sum_{j=1}^{N}\sum_{k=1}^{N} \delta_{\xi_j^N}(A)\delta_{\xi_k^N}(B)\mathbb{C}ov(M_Z(V_j^N), M_Z(V_k^N)) \\
&= \sum_{j=1}^{N}\sum_{k=1}^{N} \delta_{\xi_j^N}(A)\delta_{\xi_k^N}(B)(2\pi)^{\frac{d}{2}}\mu_Z(V_j^N \cap V_k^N) \\
&= (2\pi)^{\frac{d}{2}} \sum_{j=1}^{N} \delta_{\xi_j^N}(A)\delta_{\xi_j^N}(B)\mu_Z(V_j^N) \\
&= (2\pi)^{\frac{d}{2}} \sum_{j=1}^{N} \delta_{\xi_j^N}(A \cap B)\mu_Z(V_j^N) = (2\pi)^{\frac{d}{2}} \left(\sum_{j=1}^{N} \mu_Z(V_j^N)\delta_{\xi_j^N}\right)(A \cap B).
\end{aligned} \tag{6.3}$$

Hence M_{Z_N} is an orthogonal Random Measure with weight $\nu_{Z_N} = (2\pi)^{\frac{d}{2}} \sum_{j=1}^{N} \mu_Z(V_j^N)\delta_{\xi_j^N}$.

Since both M_Z and M_{Z_N} are slow-growing, we can analyse the random variables of the form $\langle M_Z, \varphi \rangle - \langle M_{Z_N}, \varphi \rangle$ for any $\varphi \in \mathscr{S}(\mathbb{R}^d)$. Hence, we can compare their respective Inverse Fourier Transforms. We define

$$Z_N(x) = \mathscr{F}^{-1}(M_{Z_N})(x) = \frac{1}{(2\pi)^{\frac{d}{2}}} \sum_{j=1}^{N} M_Z(V_j^N)e^{ix^T\xi_j^N}, \quad x \in \mathbb{R}^d. \tag{6.4}$$

Z_N is a complex Random Function, and since it is the Fourier Transform of a finite Random Measure, it is continuous. If we want it to be a real Random Function, we need to choose conveniently the collection $(V_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$ and the tag points $(\xi_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$ in order to make M_{Z_N} be an Hermitian Random Measure. In such a case, Z_N is a real stationary Random Function. We will require the Hermitian condition on M_{Z_N} in the implementation Section 6.3, but here we will simply work with the *complex*[2] stationary

[2]Without much detail, we can define a complex continuous stationary Random Function over $\mathbb{R}^d$ as a Random Function which is a Fourier Transform of a complex finite orthogonal Random Measure. A complex continuous stationary Random Function has a spectral measure which is positive and finite but not necessarily even, and a continuous stationary covariance function which is positive-definite, neither necessarily real nor even but always Hermitian.

Random Function Z_N. From (6.3) we conclude that the spectral measure of Z_N is

$$\mu_{Z_N} = \sum_{j=1}^{N} \mu_Z(V_j^N)\delta_{\xi_j^N}. \tag{6.5}$$

The covariance function of Z_N is then

$$\rho_{Z_N}(h) = \frac{1}{(2\pi)^{\frac{d}{2}}} \sum_{j=1}^{N} \mu_Z(V_j^N) e^{ih^T\xi_j^N}. \tag{6.6}$$

The next result states that the Random Functions $(Z_N)_{N\in\mathbb{N}_*}$ approach Z in some sense.

Proposition 6.2.1. *Let Z be a real stationary GeRF over $\mathbb{R}^d$ with spectral measure μ_Z. Let $M_Z = \mathscr{F}(Z)$. Let $(Z_N)_{N\in\mathbb{N}_*}$ be the sequence of Random Functions over $\mathbb{R}^d$ defined through* (6.4) *for an arbitrary Riemann sequence of partitions growing to $\mathbb{R}^d$, $(V_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$ and for arbitrary tag points $(\xi_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$. Then, Z_N converges to Z in a mean-square-$\mathscr{S}'(\mathbb{R}^d)$-weak-$*$ sense, that is,*

$$\mathbb{E}\left(|\langle Z,\varphi\rangle - \langle Z_N,\varphi\rangle|^2\right) \to 0, \quad \text{as } N\to\infty, \forall\varphi\in\mathscr{S}(\mathbb{R}^d). \tag{6.7}$$

Proposition 6.2.1 is proven in Appendix A.16. The result stated in this Proposition gives us an idea of how to construct Random Fields which converge to a desired GeRF in a weak sense. It is then expected that if we require more conditions on Z, stronger forms of convergence may arise, which may be useful to describe. In this aim, let us suppose now that Z is a real continuous stationary Random Function, and let us follow the same procedure as in the generalized case. The Fourier Transform of Z, M_Z, is now a finite Random Measure, and the random variables $M_Z(\mathbb{R}^d)$ and $M_Z(D_N)$ have finite variance.

Theorem 6.2.1. *Let Z be a real stationary continuous Random Function over $\mathbb{R}^d$ with Fourier Transform $M_Z = \mathscr{F}(Z)$. Let $(Z_N)_{N\in\mathbb{N}_*}$ be the sequence of Random Functions over $\mathbb{R}^d$ defined through* (6.4) *for an arbitrary Riemann sequence of partitions growing to $\mathbb{R}^d$, $(V_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$ and for arbitrary tag points $(\xi_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$. Then, Z_N converges to Z in a mean-square-uniformly on compacts sense, that is,*

$$\sup_{x\in K}\mathbb{E}\left(|Z(x)-Z_N(x)|^2\right) \to 0, \quad \text{as } N\to\infty, \forall K\subset\mathbb{R}^d \text{ compact.} \tag{6.8}$$

It is not hard to conclude that Theorem 6.2.1 implies that the sequence of covariance functions $(\rho_{Z_N})_{N\in\mathbb{N}_*}$ converges to ρ_Z uniformly on compact sets. We give a proof of Theorem 6.2.1 in Appendix A.17. In such proof, the following vanishing bound for the mean-square-uniformly on compacts convergence (6.8) is

proven:

$$\sup_{x\in K}\mathbb{E}\left(|Z(x)-Z_N(x)|^2\right)\leqslant\frac{1}{(2\pi)^{\frac{d}{2}}}\left[4\ell_N^2\mu_Z(\mathbb{R}^d)\sup_{x\in K}|x|^2+\mu_Z(D_N)\right],\quad\forall K\subset\mathbb{R}^d\text{ compact.}\tag{6.9}$$

Hence, the rate of convergence of Z_N to Z is determined by the selection of the Riemann sequence of partitions $(V_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$ and by "how fast μ_Z decays at infinity". Indeed, the two elements which determine the rate of convergence are ℓ_N and $\mu_Z(D_N)$. The term $\mu_Z(D_N)$ depends on the decreasing behaviour of μ_Z at infinity, and as we have seen in Section 3.2.2, this is closely related with the regularity of the Random Function Z. The less regular Z is, the slower the term $\mu_Z(D_N)$ goes to 0. We remark also that all Random Functions in $(Z_N)_{N\in\mathbb{N}_*}$ are smooth in mean-square, since their Fourier Transforms are compactly supported Random Measures. Hence, it is expected that approximation methods based on this Theorem work better for regular Random Functions Z.

Remark 6.2.1. The Random Function Z_N is periodic. This follows immediately from the fact that the functions of the form $e^{ix^T\xi_j^N}$ are all periodic. Hence, when doing computational implementations of this method, one must be aware of simulating over a domain of $\mathbb{R}^d$ where it is assured that an undesired periodic behaviour will not be present.

Remark 6.2.2. Z_N does not have the same variance as Z. Its variance is always smaller since

$$\begin{aligned}\mathbb{V}ar(Z(x))-\mathbb{V}ar(Z_N(x))&=\frac{1}{(2\pi)^{\frac{d}{2}}}\left(\mu_Z(\mathbb{R}^d)-\mu_{Z_N}(\mathbb{R}^d)\right)\\&=\frac{1}{(2\pi)^{\frac{d}{2}}}\left(\mu_Z(\mathbb{R}^d)-\mu_Z(\bigcup_{j=1}^N V_j^N)\right)\\&=\frac{1}{(2\pi)^{\frac{d}{2}}}\mu_Z(D_N)\geqslant 0.\end{aligned}\tag{6.10}$$

The rate of convergence of the difference of the variances is then determined by $\mu_Z(D_N)$. In cases where Z is not sufficiently regular, the differences between the variances may be considerably high, generating issues for some statistical purposes. We propose two options to construct an approximation of Z which has the same variance as Z:

- through the addition of the stationary Random Function

$$R_N(x)=M_Z(D_N)\frac{e^{ix^Td_N}}{(2\pi)^{\frac{d}{2}}}=\mathscr{F}^{-1}\left(M_Z\left(D_N\right)\delta_{d_N}\right),\tag{6.11}$$

  the stationarity of this function being guaranteed since $M_Z(D_N)\delta_{d_N}$ is an orthogonal Random Measure, which can be proved using the same arguments as in (6.3);

- through a convenient normalization. Precisely, if we set $\sigma_Z^2 := \mathbb{V}ar(Z(x)) = \rho_Z(0)$ and

$$\sigma_{Z_N}^2 := \mathbb{V}ar(Z_N(x)) = \rho_{Z_N}^2(0) = \frac{1}{(2\pi)^{\frac{d}{2}}} \sum_{j=1}^{N} \mu_Z(V_j^N) = \frac{1}{(2\pi)^{\frac{d}{2}}} \mu_Z\left(\bigcup_{j=1}^{N} V_j^N\right), \tag{6.12}$$

  we consider the sequence of stationary Random Functions $(\frac{\sigma_Z}{\sigma_{Z_N}} Z_N)_{N\in\mathbb{N}_*}$.

Proposition 6.2.2. *Both sequences of Random Functions $(Z_N + R_N)_{N\in\mathbb{N}_*}$ and $(\frac{\sigma_Z}{\sigma_{Z_N}} Z_N)_{N\in\mathbb{N}_*}$ converge to Z in a mean-square-uniformly on compacts sense, and all of the Random Functions in the sequences have the same variance as Z.*

A proof for Proposition is given in Appendix A.18. In such a proof we obtain also the following bounds for the mean-square-uniform on compacts convergence:

$$\sup_{x\in K} \mathbb{E}\left(|Z(x) - (Z_N(x) + R_N(x))|^2\right) \leqslant \frac{2}{(2\pi)^{\frac{d}{2}}} \left[2\ell_N^2 \mu_Z(\mathbb{R}^d) \sup_{x\in K} |x|^2 + \mu_Z(D_N)\right], \quad \forall K \subset \mathbb{R}^d \text{ compact} \tag{6.13}$$

and

$$\begin{aligned} &\sup_{x\in K} \mathbb{E}\left(\left|Z(x) - \frac{\sigma_Z}{\sigma_N} Z_N(x)\right|^2\right) \\ &\leqslant \frac{2}{(2\pi)^{\frac{d}{2}}} \left[\left|\sqrt{\mu_Z(\mathbb{R}^d)} - \sqrt{\mu_Z(\bigcup_{j=1}^{N} V_j^N)}\right|^2 + \left(4\ell_N^2 \mu_Z(\mathbb{R}^d) \sup_{x\in K} |x|^2 + \mu_Z(D_N)\right)\right], \quad \forall K \subset \mathbb{R}^d \text{ compact.} \end{aligned} \tag{6.14}$$

The bound (6.13) is larger than the bound (6.9), with the addition of $(2\pi)^{-\frac{d}{2}} \mu_Z(D_N)$. The bound (6.14) is the largest of the three. This implies that even if we have corrected the variance, the approximation is worse than originally in the sense of the mean-square-uniform convergence of compact sets, and hence other properties of the target Random Function such as its regularity or the practical range may be worse reproduced.

Remark 6.2.3. Consider the classical approach of developing a square-integrable function over the interval $[0, 2\pi]$ in its Fourier basis. Such an approach consists in approximating a function $f \in L^2([0, 2\pi])$ by truncating the development of f on the basis of functions of the form $(\frac{e^{ixn}}{\sqrt{2\pi}})_{\mathbf{n}\in\mathbb{Z}}$. The expression of a truncated expansion is given by

$$f_N = \sum_{n=-N}^{N} \frac{1}{2\pi} (f, e^{ixn})_{L^2([0,2\pi])} e^{ixn}, \tag{6.15}$$

for some $N \in \mathbb{N}$. We remark some similarities with respect to Eq. (6.4) for $d = 1$. The big difference is that in the case of the Fourier basis the distances between the associated tag points are constant and not depending of N. Hence, the distances are not bounded by a term such as ℓ_N which goes to 0. Such an approach would not necessarily provide a convergence in a mean-square-uniformly on compacts sense, not even a point-wise

convergence in general.

Remark 6.2.4. We mention the most important fact of this approximation method. Since the Random Measure M_Z is orthogonal, the random variables of the form $M_Z(V_j^N)$ and $M_Z(D_N)$ are mutually non-correlated. Hence, they are easy to simulate. Two difficulties are anyway still present. The first is calculating their variances. Since M_Z is an orthogonal Random Measure with weight $(2\pi)^{\frac{d}{2}}\mu_Z$, the variances are given by

$$\mathbb{V}ar(M_Z(V_j^N)) = (2\pi)^{\frac{d}{2}}\mu_Z(V_j^N) \tag{6.16}$$

and

$$\mathbb{V}ar(M_Z(D_N)) = (2\pi)^{\frac{d}{2}}\mu_Z(D_N). \tag{6.17}$$

If we can rely on a closed and easily computable form for $\mu_Z(V_j^N)$ and $\mu_Z(D_N)$, we are then able to obtain the variances without practical problems. If not, other approaches must be followed to obtain the variances. For example, one may use a computational method to approximate the integrals. This produces an extra error in the approximation and increases the computational cost of the method. Another option is giving an easily computable approximation of the integral, but this may produce extra errors in the approximation. We discuss an example of this option in Section 6.2.2.

The second difficulty comes from the Hermitian condition on M_Z. As already mentioned, the random variables of the form $M_Z(V_j^N)$ are mutually uncorrelated but, since they are complex, this does not imply independence, even in a Gaussian framework. Hence, the procedure is not as simple as simply simulating a vector of independent random variables without any special regard. For instance, if for two different indices $j, k \in \{1, ..., N\}$, the associated sets V_j^N and V_k^N satisfy $V_j^N = -V_k^N$, then the Hermitianity of M_Z implies that $M_Z(V_j^N) = \overline{M_Z(V_k^N)}$, and hence $M_Z(V_j^N)$ and $M_Z(V_k^N)$ cannot be independent. This *detail* can be tackled in many manners. We will show an example on how to do this in the implementation section 6.3.

Once these difficulties are tackled, expression (6.4) can be computed by interpreting it as a discrete Fourier Transform and applying convenient numerical algorithms for its computation. This will be detailed in the implementation section 6.3.

6.2.2 Application to SPDEs

The results presented in Section 6.2.1 can be applied to develop simulation methods of approximations of stationary GeRFs or Random Functions with a big generality, whether the GeRF being concerned by a SPDE or not. However, it also gives us an inspiration to develop numerical methods to solve some classes of SPDEs. In this section we explain how to do this.

Let us consider the case of the equation studied in Chapter 4:

$$\mathcal{L}_g U = X, \tag{6.18}$$

where X is an arbitrary stationary GeRF over $\mathbb{R}^d$ with spectral measure μ_X. We will suppose that the symbol function g satisfies the PBR condition and that it is continuous. We recall that this implies that $|g| > 0$. Hence, there exists a unique stationary solution to (6.18) and it is simply given by

$$U = \mathcal{L}_{\frac{1}{g}} X. \tag{6.19}$$

Denoting $M_X = \mathscr{F}(X)$, we construct an approximation of the source term X following the principles exposed in the previous section. Let $(V_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$ be a Riemann sequence of partitions growing to $\mathbb{R}^d$, and let $(\xi_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$ be a collection of associated tag points. We set

$$X_N(x) = \frac{1}{(2\pi)^{\frac{d}{2}}} \sum_{j=1}^{N} M_X(V_j^N) e^{ix^T\xi_j^N} = \mathscr{F}^{-1}\left(\sum_{j=1}^{N} M_X(V_j^N)\delta_{\xi_j^N}\right). \tag{6.20}$$

Then, we propose an approximation of U through the Random Function:

$$U_N(x) = \mathcal{L}_{\frac{1}{g}}(X_N)(x). \tag{6.21}$$

Since for every $\xi \in \mathbb{R}^d$, one has $\frac{1}{g}\delta_\xi = \frac{1}{g(\xi)}\delta_\xi$, we obtain

$$U_N(x) = \mathscr{F}^{-1}\left(\sum_{j=1}^{N} \frac{M_X(V_j^N)}{g(\xi_j^N)}\delta_{\xi_j^N}\right) = \frac{1}{(2\pi)^{\frac{d}{2}}} \sum_{j=1}^{N} \frac{M_X(V_j^N)}{g(\xi_j^N)} e^{ix^T\xi_j^N}. \tag{6.22}$$

U_N is a stationary Random Function with spectral measure

$$\mu_{U_N} = \sum_{j=1}^{N} \frac{\mu_X(V_j^N)}{g(\xi_j^N)}\delta_{\xi_j^N}, \tag{6.23}$$

and with covariance function

$$\rho_{U_N}(h) = \frac{1}{(2\pi)^{\frac{d}{2}}} \sum_{j=1}^{N} \frac{\mu_X(V_j^N)}{g(\xi_j^N)} e^{ih^T\xi_j^N}. \tag{6.24}$$

By definition, it is immediate that U_N satisfies the SPDE:

$$\mathcal{L}_g U_N = X_N, \tag{6.25}$$

which may be seen as an *approximation* of the original SPDE (6.18). The next result shows that the sequence

of Random Functions $(U_N)_{N\in\mathbb{N}_*}$ converges in a mean-square-$\mathscr{S}'(\mathbb{R}^d)$-weak-$*$ sense to the solution of (6.18).

Proposition 6.2.3. *Let X be a real stationary GeRF over $\mathbb{R}^d$ with spectral measure μ_X. Let $M_X = \mathscr{F}(X)$. Let g be a continuous symbol function satisfying the PBR condition. Let $(U_N)_{N\in\mathbb{N}_*}$ be the sequence of Random Functions over $\mathbb{R}^d$ defined through* (6.22) *for an arbitrary Riemann sequence of partitions growing to $\mathbb{R}^d$, $(V_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$ and for arbitrary tag points $(\xi_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$. Then, $(U_N)_{N\in\mathbb{N}_*}$ converges to the unique stationary solution to* (6.18) *in a mean-square-$\mathscr{S}'(\mathbb{R}^d)$-weak-$*$ sense. That is, if U is such a solution, then*

$$\mathbb{E}\left(|\langle U,\varphi\rangle - \langle U_N,\varphi\rangle|^2\right) \to 0, \quad \text{as } N\to\infty, \forall\varphi\in\mathscr{S}(\mathbb{R}^d). \tag{6.26}$$

Similarly to what has been done in Section 6.2.1, we can also prove a stronger convergence when the objective GeRF U is a stationary continuous Random Function. However, we need more conditions on the symbol function g. We recall that the solution U to (6.18) is a stationary continuous Random Function if and only if $|g|^{-2}$ is integrable with respect to the spectral measure of X (Remark 4.3.1). The next result is proven in Appendix A.20.

Theorem 6.2.2. *Let X be a real stationary GeRF over $\mathbb{R}^d$ with spectral measure μ_X. Let $M_X = \mathscr{F}(X)$. Let g be a continuous symbol function such that $|g|^{-2}$ is integrable with respect to μ_X and such that there exist $\alpha\in\mathbb{R}$ and two constants $C_1, C_2 > 0$ satisfying*

$$C_1(1+|\xi|^2)^\alpha \leqslant |g(\xi)| \leqslant C_2(1+|\xi|^2)^\alpha, \quad \forall\xi\in\mathbb{R}^d. \tag{6.27}$$

Let $(U_N)_{N\in\mathbb{N}_*}$ be the sequence of Random Functions over $\mathbb{R}^d$ defined through (6.22) *for an arbitrary Riemann sequence of partitions growing to $\mathbb{R}^d$, $(V_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$ and for an arbitrary collection of tag points $(\xi_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$. Then, the sequence $(U_N)_{N\in\mathbb{N}_*}$ converges in a mean-square-uniformly on compacts sense to the unique real stationary solution to* (6.18)*, which is a continuous Random Function U. Explicitly,*

$$\sup_{x\in K}\mathbb{E}\left(|U(x)-U_N(x)|^2\right) \to 0, \quad \text{as } N\to\infty, \forall K\subset\mathbb{R}^d \text{ compact.} \tag{6.28}$$

Remark 6.2.5. The proof of Theorem 6.2.2 relies on the Dominated Convergence Theorem, and it does not provide a vanishing bound to measure the error of the approximation. It can be verified that the error is higher than the error when approximating a GeRF through the approach of Theorem 6.2.1.

Remark 6.2.6. We can highlight some cases where the condition (6.27) is not necessary and still having a mean-square-uniform convergence on compact sets. Always supposing that g is a continuous symbol function satisfying the PBR condition and such that $|g|^{-2}$ is integrable with respect to μ_X, under the following cases the convergence is also guaranteed:

- $\frac{1}{g}$ is uniformly continuous and μ_X is finite.

- μ_X has compact support.

- The tag points are not arbitrary and they are chosen such that $|g(\xi)| \leqslant |g(\xi_j^N)|$ for all $\xi \in V_j^N$.

Under these conditions the convergence is easily verified by bounding expressions of the form

$$\sum_{j=1}^{N} \int_{V_j^N} \left| \frac{1}{g(\xi)} - \frac{1}{g(\xi_j^N)} \right|^2 d\mu_X(\xi) \tag{6.29}$$

which appear in the proof of Theorem 6.2.2. On the first two cases, a vanishing bound for the error can be obtained using the uniform continuity of $\frac{1}{g}$.

Remark 6.2.7. Rather than *solving the SPDE* (6.18) one could follow the approach in Theorem 6.2.1 to construct an approximation of a Random Function with the same covariance structure as U. This is highly recommended if the integrals of the form $\int_{V_j^N} \frac{d\mu_X(\xi)}{|g(\xi)|^2}$ are well-known or can be exactly computed. The approach using the approximation (6.22) is actually nothing but following the approach in Eq. (6.4) with an approximation of the variances of the random variables $M_U(V_j^N)$. Indeed, one could argue that a good approximation for such variances is given by

$$\mathbb{V}ar(M_U(V_j^N)) = (2\pi)^{\frac{d}{2}} \mu_U(V_j^N) = (2\pi)^{\frac{d}{2}} \int_{V_j^N} \frac{d\mu_X(\xi)}{|g(\xi)|^2} \approx (2\pi)^{\frac{d}{2}} \frac{\mu_X(V_j^N)}{|g(\xi_j^N)|^2}. \tag{6.30}$$

The last expression coincides with the variance of a random variable of the form $\frac{M_X(V_j^N)}{g(\xi_j^N)}$ in Eq. (6.22), hence the approach of solving the approximative SPDE (6.25) is equivalent in law to use an approximation of the form (6.4) with an approximative computation for the variances. We have noticed in Remark 6.2.5 that such an approximation produces difficulties. Theorem 6.2.2 proposes a convergence with a restrictive condition on g, for which we do not have a bound to measure the error. It is then preferable, if possible, to use the approach of Theorem 6.2.1 with an exact computation of the variances. However, a subtlety still remains: we have not solved the SPDE (6.18), and hence we have not simulated approximations of *both Random Fields* U and X. If U is the only Random Function of interest, this is not so much of an issue. However, if we want to simulate the couple (U, X) in a bivariate modelling approach (Section 3.6), the only simulation of U using Theorem 6.2.1 is insufficient. Given a good approximation U, say U_N, we still can simulate an approximation of X through $X_N = \mathcal{L}_g U_N$, the approach being, at the end of the story, the same as in Theorem 6.2.2. Hence, a non-controlled error will be present in some of the two simulated Random Fields. We suggest then, if possible, to use the approximation (6.4) without approximative variances to approach the one we are more interested in, if we are not in a bivariate context.

6.2.3 Adaptation to first order evolution models

In this section we show how to apply the results presented in Section 6.2.1 to solve numerically the Cauchy problem associated to first order evolution models studied in Section 5.3. The approach is a generalization of the method used in Sigrist et al. (2015) to simulate solutions to the stochastic advection-diffusion equation, where the approach is to do a spatial FFT combined with an exact expression of the solution in time. Here we give the details of such an approach. We show that it can be generalized to the cases of other equations (just by changing the function g), and we prove the convergence of the approximations to the theoretical solution in a weak sense. In this section we work in a spatio-temporal framework, hence we use the notational conventions pointed out at the beginning of Chapter 5.

We consider thus the Cauchy problem as presented in Section 5.3:

$$\begin{cases} \dfrac{\partial U}{\partial t} + \mathcal{L}_g U = X \\ U\big|_{t=0} = U_0 \end{cases} \tag{6.31}$$

We recall that here g is a spatial symbol function for which we suppose in addition that $g_R \geqslant 0$, being g_R its real part. We will focus on the resolution of the spatial-Fourier transformed problem

$$\begin{cases} \dfrac{\partial V}{\partial t} + gV = Y \\ V\big|_{t=0} = V_0 \end{cases} , \tag{6.32}$$

where $Y = \mathscr{F}_S(X)$ and $V_0 = \mathscr{F}_S(U_0)$. We suppose that we are in the case presented in Section 5.3.3, where X is a real GeRF such that $Y = \mathscr{F}_S(X)$ is a slow-growing Random Measure over $\mathbb{R}^d \times \mathbb{R}^+$, and U_0 is a real GeRF such that $V_0 = \mathscr{F}_S(U_0)$ is a slow-growing Random Measure over $\mathbb{R}^d$. The principle is to use an approximation of Y and V_0 constructed using a Riemann sequence of partitions growing to the space $\mathbb{R}^d$. Then, we propose a *spatial* approximation of the solution to (6.32), and we can thus give an explicit expression for its time evolution.

Let $(V_j^N)_{j \in \{1,\dots,N\}, N \in \mathbb{N}_*}$ be a Riemann sequence of partitions growing to $\mathbb{R}^d$, and let $(\xi_j^N)_{j \in \{1,\dots,N\}, N \in \mathbb{N}_*}$ be a collection of associated tag points. We introduce the following slow-growing Random Measure, defined as a set-function:

$$Y_N(A \times B) := \sum_{j=1}^{N} Y(V_j^N \times B)\delta_{\xi_j^N}(A), \quad \forall A \in \mathcal{B}_B(\mathbb{R}^d), B \in \mathcal{B}_B(\mathbb{R}^+). \tag{6.33}$$

Equivalently, Y_N can be described through its action to test-functions $\psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R})$ by

$$\langle Y_N, \psi \rangle = \sum_{j=1}^{N} \int_{\mathbb{R}^d \times \mathbb{R}^+} \mathbf{1}_{V_j^N}(\xi) \psi(\xi_j^N, t) dY(\xi, t). \tag{6.34}$$

The initial condition V_0 is also approximated through our method. We consider thus the Random Measure

$$V_{0,N} := \sum_{j=1}^{N} V_0(V_j^N) \delta_{\xi_j^N}. \tag{6.35}$$

We consider then the solution to the approximated transformed problem

$$\begin{cases} \dfrac{\partial V_N}{\partial t} + g V_N = Y_N \\ V_N\big|_{t=0} = V_{0,N} \end{cases} . \tag{6.36}$$

As we have seen in Section 5.3.3, there is a unique solution V_N to (6.36) which is a slow-growing Random Measure with a càdlàg-in-time representation. We can hence express this solution as a collection of Random Measures over $\mathbb{R}^d$, $(V_{N,t})_{t \in \mathbb{R}^+}$, having a càdlàg behaviour in time. The expression is given by

$$V_{N,t}(A) = \int_A e^{-tg(\xi)} d(V_{0,N} - Y_N(\cdot \times \{0\}))(\xi) + \int_{A \times [0,t]} e^{-(t-s)g(\xi)} dY_N(\xi, s), \tag{6.37}$$

for all $A \in \mathcal{B}_B(\mathbb{R}^d)$. Such expression, which may be quite complicated to compute for a general Y, can now be expressed in a simpler manner due to the definition of Y_N using Dirac delta measures:

$$V_{N,t} = \sum_{j=1}^{N} \left(e^{-tg(\xi_j^N)} \left(V_0(V_j^N) - Y(V_j^N \times \{0\}) \right) + \int_{[0,t]} e^{-(t-s)g(\xi_j^N)} dY(V_j^N \times \cdot)(s) \right) \delta_{\xi_j^N}, \tag{6.38}$$

or more explicitly,

$$V_{N,t}(A) = \sum_{j=1}^{N} \left(e^{-tg(\xi_j^N)} \left(V_0(V_j^N) - Y(V_j^N \times \{0\}) \right) + \int_{[0,t]} e^{-(t-s)g(\xi_j^N)} dY(V_j^N \times \cdot)(s) \right) \delta_{\xi_j^N}(A) \tag{6.39}$$

for all $A \in \mathcal{B}_B(\mathbb{R}^d)$. Indeed, in this case we have used the well-defined Random Measures over $\mathbb{R}^+$ determined by the collection of random variables $(Y(V_j^N \times B))_{B \in \mathcal{B}_B(\mathbb{R}^+)}$. Since Y is a Random Measure over $\mathbb{R}^d \times \mathbb{R}^+$, the covariance Kernel of $Y(V_j^N \times \cdot\,)$ defines a measure over $\mathbb{R}^+ \times \mathbb{R}^+$ for every $N \in \mathbb{N}_*$ and for every $j \in \{1, ..., N\}$. Hence, the expressions of the form

$$\int_{[0,t]} e^{-(t-s)g(\xi_j^N)} dY(V_j^N \times \cdot)(s) \tag{6.40}$$

are simply the stochastic integrals of the temporal deterministic functions of the form $s \mapsto e^{-(t-s)g(\xi_j^N)}$ with respect to the temporal Random Measure $Y(V_j^N \times \cdot\,)$ over the interval $[0,t]$. In addition, since $g_R \geqslant 0$, the functions $s \mapsto e^{-(t-s)g(\xi_j^N)}$ are bounded over $[0,t]$, which guarantees the good definition of this stochastic integral.

Finally, we consider the non-transformed approximated problem

$$\begin{cases} \dfrac{\partial U_N}{\partial t} + \mathcal{L}_g U_N = X_N \\ U_N\big|_{t=0} = U_{0,N} \end{cases} \tag{6.41}$$

where $X_N = \mathscr{F}_S^{-1}(Y_N)$ and $U_{0,N} = \mathscr{F}_S^{-1}(V_{0,N})$. The solution is given by the spatial Fourier Transform of (6.38), which is a Random Function over $\mathbb{R}^d \times \mathbb{R}^+$, determined by

$$U_N(x,t) = \frac{1}{(2\pi)^{\frac{d}{2}}} \sum_{j=1}^{N} \left(e^{-tg(\xi_j^N)} \left(V_0(V_j^N) - Y(V_j^N \times \{0\}) \right) + \int_{[0,t]} e^{-(t-s)g(\xi_j^N)} dY(V_j^N \times \cdot)(s) \right) e^{ix^T \xi_j^N}. \tag{6.42}$$

U_N is a Random Function smooth in mean-square in space and càdlàg in time.

In expression (6.38) we have used the random variables of the form $V_{0,N}(V_j^N)$, $Y(V_j^N \times \{0\})$ and the stochastic integral (6.40). For general slow-growing Random Measures Y and V_0, these random variables are not necessarily non-correlated, hence we have not win that much in simplicity when looking for a method to simulate the solution (6.38). This situation is avoided when requiring the extra conditions on X and U_0 which were presented in Section 5.3.4. Namely, that X and U_0 are stationary, and that $Y = \mathscr{F}_S(X)$ is a slow-growing Random Measure. For simplicity we will suppose that $Y(V_j^N \times \{0\}) = 0$. In such a case, expression (6.38) gives

$$V_{N,t} = \sum_{j=1}^{N} \left(e^{-tg(\xi_j^N)} V_0(V_j^N) + \int_{[0,t]} e^{-(t-s)g(\xi_j^N)} dY(V_j^N \times \cdot)(s) \right). \tag{6.43}$$

We suppose, in addition, that X and U_0 are independent. Consequently, the random variables of the form $V_0(V_j^N)$ are independent of the random variables of the stochastic integrals of the form (6.40). Since X is stationary and Y is the spatial Fourier Transform of X, Y must have an orthogonal behaviour in space, and hence for a fixed $t \in \mathbb{R}^+$ and for every $N \in \mathbb{N}_*$, the collection of random variables

$$\left(\int_{[0,t]} e^{-(t-s)g(\xi_j^N)} dY(V_j^N \times \cdot)(s) \right)_{j \in \{1,\ldots,N\}} \tag{6.44}$$

are mutually uncorrelated. With these extra suppositions, the simulation of $V_{N,t}$ *for a fixed* $t \in \mathbb{R}^+$ is easy to compute, maintaining anyway the subtleties presented in Remark 6.2.4.

The next proposition gives sufficient conditions when this approximation procedure converges to the desired solution U in a weak sense. The proof is presented in Appendix A.21.

Proposition 6.2.4. *Let X be a real stationary GeRF over $\mathbb{R}^d \times \mathbb{R}$ such that X is separable in the form $X = X_S \boxtimes X_T$ (symbolically), where X_T represents a stationary Random Measure over $\mathbb{R}$. Let $g : \mathbb{R}^d \to \mathbb{C}$ be a continuous spatial symbol function such that $g_R \geqslant 0$. Let U_0 be a real stationary GeRF over $\mathbb{R}^d$. Let $(U_N)_{N\in\mathbb{N}*}$ be a sequence of Random Functions defined as in* (6.42) *for an arbitrary Riemann sequence of partitions growing to $\mathbb{R}^d$, $(V_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$ and for arbitrary tag points $(\xi_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$. Let U be the solution to the Cauchy problem* (6.31)*, and let $(U_t)_{t\in\mathbb{R}^+}$ be its càdlàg-in-time representation. Then, $U_N \to U$ in a mean-square-$\mathscr{S}'(\mathbb{R}^d)$-weak-$*$ sense in space and in point-wise sense in time. Explicitly,*

$$\mathbb{E}\left(\left|\langle U_t, \varphi\rangle - \int_{\mathbb{R}^d} U_N(x,t)\varphi(x)dx\right|^2\right) \to 0, \quad \text{as } N \to \infty, \forall\varphi \in \mathscr{S}(\mathbb{R}^d), \forall t \in \mathbb{R}^+. \tag{6.45}$$

Remark 6.2.8. Proposition (6.2.4) also holds when X has a product-sum form, say

$$X = \sum_{k=1}^{M} X_S^k \boxtimes X_T^k, \quad M \in \mathbb{N}_*, \tag{6.46}$$

provided that all of the temporal parts $(X_T^k)_{k\in\{1,\dots,M\}}$ are stationary Random Measures. We recall that in a product-sum model, the representation (6.46) means that the covariance of X can be expressed as the covariance of sums of tensor products as in (6.46), being the families of GeRFs $(X_S^k)_{k\in\{1,\dots,M\}}$ and $(X_T^k)_{k\in\{1,\dots,M\}}$ all mutually independent.

Remark 6.2.9. Under the conditions of Proposition 6.2.4, and if we require in addition that $g_R \geqslant \kappa > 0$ for some $\kappa > 0$, as we have pointed out in Section 5.3.4, the covariance of the solution to the problem (6.31) converges spatio-temporally to the covariance of the unique stationary solution of the associated equation (5.79) as the time flows. In the approximated case, the solution to the approximated Cauchy problem (6.41) converges spatio-temporally to the unique stationary solution of the *approximated* SPDE

$$\frac{\partial U_N}{\partial t} + \mathcal{L}_g(U_N) = X_S^N \boxtimes X_T, \tag{6.47}$$

where X_S^N represents a *spatial* approximation of the spatial trace X_S, through the method exposed in Section 6.2.1. Hence, we obtain a convergence to an approximation of the stationary solution U^{stat}.

Let us now consider the problem of simulating U_N at different time locations. In this case we can follow a recursive approach to obtain a practical expression for V_N. Let $t \in \mathbb{R}^+$ and $\Delta t > 0$. With some simple

algebraic calculations, one obtains the next formula for $V_{N,t+\Delta t}$ as a function of $V_{N,t}$:

$$V_{N,t+\Delta t}(A) = \int_A e^{-\Delta t g(\xi)} dV_{N,t}(\xi) + \int_{A\times(t,t+\Delta t]} e^{-(t+\Delta t-s)g(\xi)} dY_N(\xi,s), \quad \forall A \in \mathcal{B}_B(\mathbb{R}^d). \tag{6.48}$$

Equivalently, we can express $V_{N,t+\Delta t}$ as

$$\begin{aligned} V_{N,t+\Delta t} = &\sum_{j=1}^{N} \left(e^{-\Delta t g(\xi_j^N)} \left[e^{-t g(\xi_j^N)} V_0(V_j^N) + \int_{[0,t]} e^{-(t-s)g(\xi_j^N)} dY(V_j^N \times \cdot)(s) \right] \right) \delta_{\xi_j^N} \\ &+ \sum_{j=1}^{N} \left(\int_{(t,t+\Delta t]} e^{-(t+\Delta t-s)g(\xi_j^N)} dY(V_j^N \times \cdot)(s) \right) \delta_{\xi_j^N}. \end{aligned} \tag{6.49}$$

If we have already a simulation of $V_{N,t}$ through the specification of the collection of random variables

$$\left(e^{-t g(\xi_j^N)} V_0(V_j^N) + \int_{[0,t]} e^{-(t-s)g(\xi_j^N)} dY(V_j^N \times \cdot)(s) \right)_{j\in\{1,\dots,N\}}, \tag{6.50}$$

we can use these random variables to calculate the first sum in (6.49). The second sum contains the random variables of the form

$$\int_{(t,t+\Delta t]} e^{-(t+\Delta t-s)g(\xi_j^N)} dY(V_j^N \times \cdot)(s), \tag{6.51}$$

which may be seen as *innovation terms*. Such terms are not necessarily uncorrelated with the random variables of the form (6.50). The dependence structure between them is determined by the covariance structure of Y, which is not necessarily orthogonal in time. A case where there is no correlation between the innovation terms (6.51) and the random variables (6.50) is when Y has the structure of a orthogonal Random Measure in time, case in which, as we have seen in Section 5.3.5, the solution V has a Markovian behaviour in time when working in a Gaussian framework. Consider then the case where Y is an orthogonal Random Measure over $\mathbb{R}^d \times \mathbb{R}$. If $\nu_Y \in \mathscr{M}_{SG}^+(\mathbb{R}^d \times \mathbb{R})$ is its weight, the variances of the involved random variables can be expressed as

$$\begin{aligned} &\mathbb{V}ar\left(e^{-t g(\xi_j^N)} V_0(V_j^N) + \int_{[0,t]} e^{-(t-s)g(\xi_j^N)} dY(V_j^N \times \cdot)(s) \right) \\ &= e^{-2t g_R(\xi_j^N)} (2\pi)^{\frac{d}{2}} \mu_{U_0}(V_j^N) + \int_{[0,t]} e^{-2(t-s)g_R(\xi_j^N)} d\nu_Y(V_j^N \times \cdot)(s), \end{aligned} \tag{6.52}$$

and

$$\mathbb{V}ar\left(\int_{(t,t+\Delta t]} e^{-(t+\Delta t-s)g(\xi_j^N)} dY(V_j^N \times \cdot)(s) \right) = \int_{(t,t+\Delta t]} e^{-2(t+\Delta t-s)g_R(\xi_j^N)} d\nu_Y(V_j^N \times \cdot)(s), \tag{6.53}$$

where μ_{U_0} is the spectral measure of U_0. Hence, when Y is orthogonal we have an easy way to simulate the involved random variables and hence to obtain a non-limited in time simulation of the spatio-temporal model. We remark that if the variances (6.52) and (6.53) are known exactly, the simulation does not accumulate time errors when evolving in time: the solution is the strict solution to the approximated problem (6.32). The errors in the approximation are completely determined by the spatial approximation.

We remark that, as we have mentioned in 5.3.5, the conditions that X is stationary and that $Y = \mathscr{F}_S(X)$ is orthogonal require that X must be both orthogonal and stationary in time, and hence, possibly, the only kind of models which satisfy this property are those who have the behaviour of a White Noise in time. In the implementation section 6.3 we consider the case where X is a coloured in space and white in time noise, $X = X_S \boxtimes W_T$.

6.2.4 Adaptation to Waving models

Let us consider the application of these principles to obtain approximations of real stationary solutions to some homogeneous SPDEs. We consider the homogeneous SPDE

$$\mathcal{L}_g U = 0, \tag{6.54}$$

for which we look to approximate a particular stationary solution. As stated in Remark 4.3.3, this is only possible in the case when g has null values, and in such a case the stationary solutions have spectral measures concentrated on the subset $g^{-1}(\{0\})$. Hence, a general expression for U is given by

$$U = \mathscr{F}^{-1}(M_U), \tag{6.55}$$

where M_U is an Hermitian slow-growing orthogonal Random Measure concentrated on $g^{-1}(\{0\})$. The general principle is to use a Riemann sequence of partitions of $g^{-1}(\{0\})$ (or growing to $g^{-1}(\{0\})$ if it is not bounded) together with tag points, and then defining a Random Function U_N following the same principle as in Eq. (6.4), using an orthogonal Random Measure M_U concentrated on $g^{-1}(\{0\})$. Hence, $\mathscr{F}(U_N)$ is an orthogonal Random Measure concentrated on $g^{-1}(\{0\})$, and hence it is immediate that U_N satisfies Eq. (6.54). However, we still need a criterion to *select* a particular solution. Since there are many possible spectral measures concentrated on the set $g^{-1}(\{0\})$, there are also many possible orthogonal Random Measures M_U which can be chosen. Such a selection is arbitrary and it is done in order to obtain a model with desired extra properties, besides the fact of solving Eq. (6.54).

In this section we will focus on the case of Waving models (Section 5.2.4). The selection of a particular stationary solution for the associated homogeneous Wave equation is done in order to make the solution follow a desired spatial covariance model. We maintain the notational conventions of the spatio-temporal

framework as considered in Chapter 5.

Let U be a real stationary GeRF over $\mathbb{R}^d \times \mathbb{R}$ solution to the homogeneous Wave equation

$$\frac{\partial^2 U}{\partial t^2} - c^2 \Delta U = 0. \tag{6.56}$$

As we have seen in Section 5.2.4, U must have a spectral measure of the form

$$d\mu_U(\xi, \omega) = \sqrt{2\pi} d \left(\frac{\delta_{-c|\xi|} + \delta_{c|\xi|}}{2} \right) (\omega) d\mu_{U_S}(\xi), \tag{6.57}$$

where μ_{U_S} is a spectral measure over $\mathbb{R}^d$. μ_{U_S} describes the spatial behaviour of U. Let $M_U = \mathscr{F}(U)$. M_U is an orthogonal Random Measure concentrated on the spatio-temporal cone $\mathcal{C}^c = \{(\xi, \omega) \in \mathbb{R}^d \times \mathbb{R} \mid |\omega| = c|\xi|\}$. We consider a *spatial* Riemann sequence of partitions growing to $\mathbb{R}^d$, $(V_j^N)_{j \in \{1,...,N\}, N \in \mathbb{N}_*}$, with associated tag points $(\xi_j^N)_{j \in \{1,...,N\}, N \in \mathbb{N}_*}$. Starting from this spatial sequence of partitions, we construct a *spatio-temporal* Riemann sequence of partitions growing to $\mathcal{C}^c$ as follows: for every $N \in \mathbb{N}_*$ and for every $j \in \{1, ..., N\}$, we define the sets

$$B_{j,+}^N := \left(V_j^N \times \mathbb{R}^+\right) \cap \mathcal{C}^c \quad ; \quad B_{j,-}^N := \left(V_j^N \times \mathbb{R}_*^-\right) \cap \mathcal{C}^c. \tag{6.58}$$

The sequence of collections of sets $(B_{j,+}^N)_{j \in \{1,...,N\}, N \in \mathbb{N}_*}$ forms a Riemann sequence of partitions growing to the "*positive-temporal-frequency part*" of the spatio-temporal cone, $\mathcal{C}^c \cap (\mathbb{R}^d \times \mathbb{R}^+)$. Analogously, the collection of sets $(B_{j,-}^N)_{j \in \{1,...,N\}, N \in \mathbb{N}_*}$ forms a Riemann sequence of partitions growing to $\mathcal{C}^c \cap (\mathbb{R}^d \times \mathbb{R}_*^-)$. To every set of the form $B_{j,+}^N$ we associate the tag point $(\xi_j^N, c|\xi_j^N|) \in \mathcal{C}^c \cap (\mathbb{R}^d \times \mathbb{R}^+)$, while to a set of the form $B_{j,-}^N$ we associate the tag point $(\xi_j^N, -c|\xi_j^N|) \in \mathcal{C}^c \cap (\mathbb{R}^d \times \mathbb{R}_*^-)$. Since M_U is concentrated on $\mathcal{C}^c$, we have that $M_U(V_j^N \times \mathbb{R}^+) = M_U(B_{j,+}^N)$ and $M_U(V_j^N \times \mathbb{R}_*^-) = M_U(B_{j,-}^N)$. We propose then the next approximation for M_U:

$$M_{U_N} := \sum_{j=1}^{N} M_U(V_j^N \times \mathbb{R}^+) \delta_{(\xi_j^N, c|\xi_j^N|)} + M_U(V_j^N \times \mathbb{R}_*^-) \delta_{(\xi_j^N, -c|\xi_j^N|)}. \tag{6.59}$$

We recall that the Dirac measures $\delta_{(\xi_j^N, c|\xi_j^N|)}$ denotes a measure over $\mathbb{R}^d \times \mathbb{R}$. M_{U_N} is then a compactly supported orthogonal Random Measure concentrated on $\mathcal{C}^c$ and hence its spatio-temporal Inverse Fourier Transform is a stationary complex Random Function which satisfies Eq. (6.56). Such Random Function is determined by

$$U_N(x,t) = \frac{1}{(2\pi)^{\frac{d+1}{2}}} \sum_{j=1}^{N} \left(e^{itc|\xi_j^N|} M_U(V_j^N \times \mathbb{R}^+) + e^{-itc|\xi_j^N|} M_U(V_j^N \times \mathbb{R}_*^-) \right) e^{ix^T \xi_j^N}. \tag{6.60}$$

Let us make explicit the variances of the random variables of the form $M_U(V_j^N \times \mathbb{R}^+)$ and $M_U(V_j^N \times \mathbb{R}_*^-)$. M_U is an orthogonal Random Measure over $\mathbb{R}^d \times \mathbb{R}$ with weight $(2\pi)^{\frac{d+1}{2}}\mu_U$, with μ_U given by (6.57). Hence, for every $B \in \mathcal{B}_B(\mathbb{R}^d \times \mathbb{R})$,

$$\mathbb{V}ar(M_U(B)) = (2\pi)^{\frac{d+1}{2}}\mu_U(B). \tag{6.61}$$

Following the disintegration expression (6.57), one has

$$\begin{aligned}
\mathbb{V}ar(M_U(V_j^N \times \mathbb{R}^+)) &= (2\pi)^{\frac{d+1}{2}}\mu_U(V_j^N \times \mathbb{R}^+) \\
&= (2\pi)^{\frac{d}{2}+1}\int_{V_j^N} \frac{\delta_{c|\xi|} + \delta_{-c|\xi|}}{2}(\mathbb{R}^+)d\mu_{U_S}(\xi) \\
&= (2\pi)^{\frac{d}{2}+1}\left[\int_{V_j^N\setminus\{0\}} \frac{\delta_{c|\xi|} + \delta_{-c|\xi|}}{2}(\mathbb{R}^+)d\mu_{U_S}(\xi) + \int_{V_j^N\cap\{0\}} \frac{\delta_{c|\xi|} + \delta_{-c|\xi|}}{2}(\mathbb{R}^+)d\mu_{U_S}(\xi)\right] \\
&= (2\pi)^{\frac{d}{2}+1}\left[\frac{1}{2}\mu_{U_S}(V_j^N\setminus\{0\}) + \mu_{U_S}(V_j^N \cap \{0\})\right].
\end{aligned} \tag{6.62}$$

Following the same arguments one obtains

$$\mathbb{V}ar(M_U(V_j^N \times \mathbb{R}_*^-)) = \frac{(2\pi)^{\frac{d}{2}+1}}{2}\mu_{U_S}(V_j^N\setminus\{0\}). \tag{6.63}$$

We obtain thus the following results which are consequences of Proposition 6.2.3 and Theorem 6.2.1. We omit the proofs.

Proposition 6.2.5. *Let U be a real stationary GeRF over $\mathbb{R}^d \times \mathbb{R}$ which is solution to the homogeneous Wave equation* (6.56)*. Let M_U be its spatio-temporal Fourier Transform. Let $(U_N)_{N\in\mathbb{N}_*}$ be the sequence of Random Functions over $\mathbb{R}^d \times \mathbb{R}$ defined through* (6.60) *for an arbitrary Riemann sequence of partitions growing to $\mathbb{R}^d$, $(V_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$ and for arbitrary tag points $(\xi_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$. Then, $(U_N)_{N\in\mathbb{N}_*}$ converges to U in a mean-square-$\mathscr{S}'(\mathbb{R}^d \times \mathbb{R})$-weak-$*$ sense, that is,*

$$\mathbb{E}\left(|\langle U, \psi\rangle - \langle U_N, \psi\rangle|^2\right) \to 0, \quad \text{as } N \to \infty, \forall \psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R}). \tag{6.64}$$

For the next result, we remark that if U is a real stationary solution to the homogeneous Wave equation such that the spectral measure describing its spatial behaviour μ_{U_S} is finite, we obtain immediately from (6.57) that its spatio-temporal spectral measure μ_U is also finite, hence U is a continuous stationary Random Function.

Theorem 6.2.3. *Let U be a real stationary GeRF over $\mathbb{R}^d \times \mathbb{R}$ which is solution to the homogenoeus Wave equation* (6.56)*, and such that the spectral measure describing its spatial behaviour μ_{U_S} is finite. Let M_U be its spatio-temporal Fourier Transform. Let $(U_N)_{N\in\mathbb{N}_*}$ be the sequence of Random Functions over $\mathbb{R}^d \times \mathbb{R}$*

defined through (6.60) *for an arbitrary Riemann sequence of partitions growing to* $\mathbb{R}^d$, $(V_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$ *and for an arbitrary choice of tag points* $(\xi_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$. *Then,* $(U_N)_{N\in\mathbb{N}_*}$ *converges to* U, *which is a continuous Random Function, in a mean-square-uniformly on compacts sense, that is,*

$$\sup_{x\in K}\mathbb{E}\left(|U(x)-U_N(x)|^2\right)\to 0,\quad \text{as } N\to\infty, \forall K\subset\mathbb{R}^d\times\mathbb{R} \text{ compact}. \tag{6.65}$$

Remark 6.2.10. Concerning the explicit computation of the variances (6.62) and (6.63) for a particular spatial spectral measure μ_{U_S}, we have the same issues as in Remark 6.2.7. If we know how to calculate them explicitly, we can apply Theorem 6.2.3 and obtain an approximation which converges mean-square uniformly on compacts. However, if we rather do an approximation, similarly as in Eq. (6.30), some extra conditions over this spatial spectral measure must be required in order to justify the convergence. For instance, if we suppose the spatial traces of U to satisfy an equation of the form (6.18), and if we apply the procedures of section 6.2.2 to approximate U, the conditions on Theorem 6.2.2 are required.

To conclude, we remark that in this case we have a similarity with the approach followed in section 6.2.3 for first order evolution models: the approximation is only spatial. Temporally, our approximation U_N solves the homogeneous Wave equation strictly.

6.3 Implementation

The results presented in Section 6.2 are now applied to simulate approximations of some GeRFs and solutions of SPDEs. We will present examples in a spatial context with a study of the convergence of the approximations in the case of the Matérn model. We will illustrate other models with interesting properties in the case of dimension $d = 2$. In the spatio-temporal context, we will illustrate first order evolution models and Waving models.

The simulations are done in a Gaussian framework. Hence, all the non-correlated real random variables involved are independent. Before concerning about the simulation itself, we will point out some necessary technical specifications about the computational adaptation of our results and how to relate them to a classical Fourier Analysis computational problem, obtaining expressions associated to Discrete Fourier Transforms. The simulations are performed in R.

The Riemann *multi*-sequence of partitions

In this implementation section we will always use the same Riemann sequence of partitions growing to $\mathbb{R}^d$ (or to $\mathbb{R}^d\times\mathbb{R}$). The collection of sets of the form $(V_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$ will be now re-indexed in a convenient way. The approximation order will not be described by an integer $N\in\mathbb{N}_*$ but by a multi-index of positive

integers, which will be written in bold letters $\mathbf{N} \in \mathbb{N}_*^d$, and we will further require all of its components to grow together to ∞, which we denote by $\mathbf{N} \to \infty$. We denote by $[\![-(\mathbf{N}-1), \mathbf{N}-1]\!] := \{\mathbf{j} \in \mathbb{Z}^d \,\big|\, -(\mathbf{N}-1) \leqslant \mathbf{j} \leqslant \mathbf{N}-1\}$, where the inequality relation $-(\mathbf{N}-1) \leqslant \mathbf{j} \leqslant \mathbf{N}-1$ is taken component-wise. The collection of partitions will be described through a *Riemann multi-sequence of partitions growing to* $\mathbb{R}^d$, determined for every $\mathbf{j} \in [\![-(\mathbf{N}-1), \mathbf{N}-1]\!]$ and for every $\mathbf{N} \in \mathbb{N}_*^d$ through:

$$V_{\mathbf{j}}^{\mathbf{N}} = a_{\mathbf{N}} \left[j_1 - \frac{1}{2}, j_1 + \frac{1}{2} \right) \times \left[j_2 - \frac{1}{2}, j_2 + \frac{1}{2} \right) \times ... \times \left[j_d - \frac{1}{2}, j_d + \frac{1}{2} \right), \tag{6.66}$$

where $(a_{\mathbf{N}})_{\mathbf{N} \in \mathbb{N}_*^d}$ is a multi-sequence of positive numbers such that $a_{\mathbf{N}} \to 0$ as $\mathbf{N} \to \infty$ and such that $a_{\mathbf{N}}\mathbf{N} \to \infty$ as $\mathbf{N} \to \infty$. It is clear that for a fixed $\mathbf{N} \in \mathbb{N}_*^d$, the collection $(V_{\mathbf{j}}^{\mathbf{N}})_{\mathbf{j} \in [\![(-\mathbf{N}-1), \mathbf{N}-1]\!]}$ forms a partition of its reunion. Such a reunion is given by

$$\bigcup_{\mathbf{j} \in [\![-(\mathbf{N}-1), \mathbf{N}+1]\!]} V_{\mathbf{j}}^{\mathbf{N}} = a_{\mathbf{N}} \left[N_1 - \frac{1}{2}, N_1 + \frac{1}{2} \right) \times \left[N_2 - \frac{1}{2}, N_2 + \frac{1}{2} \right) \times ... \times \left[N_d - \frac{1}{2}, N_d + \frac{1}{2} \right), \tag{6.67}$$

which we may denote informally by

$$a_{\mathbf{N}} \left[\mathbf{N} - \frac{1}{2}, \mathbf{N} + \frac{1}{2} \right). \tag{6.68}$$

Since $a_{\mathbf{N}}\mathbf{N} \to \infty$ as $\mathbf{N} \to \infty$, the reunion (6.67) grows to the whole space $\mathbb{R}^d$ as $\mathbf{N}$ grows. In addition, all of the members of the partition $(V_{\mathbf{j}}^{\mathbf{N}})_{\mathbf{j} \in [\![-(\mathbf{N}-1), \mathbf{N}-1]\!]}$ have the same diameter for a fixed $\mathbf{N} \in \mathbb{N}_*^d$, given by

$$\operatorname{diam}(V_{\mathbf{j}}^{\mathbf{N}}) = a_{\mathbf{N}}\sqrt{d}, \tag{6.69}$$

which goes to 0 as $\mathbf{N}$ grows. Hence, we can properly say that (6.66) defines a *Riemann* multi-sequence of partitions growing to $\mathbb{R}^d$.

Finally, we choose the associated collection of tag points, which will be given simply by

$$\xi_{\mathbf{j}}^{\mathbf{N}} = a_{\mathbf{N}}\mathbf{j}, \tag{6.70}$$

for all $\mathbf{j} \in [\![-(\mathbf{N}-1), \mathbf{N}-1]\!]$ and for all $\mathbf{N} \in \mathbb{N}_*^d$.

This selection of Riemann multi-sequence of partitions and tag points is preferred for many reasons. We first remark that it is the most intuitive and classical *Riemann* partition: we made a partition given by rectangles and the selected tag points are the middle points. In addition, we remark that we have the following condition

$$-\left(V_{\mathbf{j}}^{\mathbf{N}}\right)^{\mathrm{o}} = \left(V_{-\mathbf{j}}^{\mathbf{N}}\right)^{\mathrm{o}} \quad \text{and} \quad -\xi_{\mathbf{j}}^{\mathbf{N}} = \xi_{-\mathbf{j}}^{\mathbf{N}}, \quad \forall \mathbf{j} \in [\![-(\mathbf{N}-1), \mathbf{N}-1]\!], \forall \mathbf{N} \in \mathbb{N}_*^d, \tag{6.71}$$

where $\left(V_{\mathbf{j}}^{\mathbf{N}}\right)^{\mathrm{o}}$ denotes the interior of the set $V_{\mathbf{j}}^{\mathbf{N}}$. This condition will be beneficial when considering ap-

proximations of Hermitian Random Measures, in order to identify which random variables are conjugate one of the other when simulating a real stationary GeRF. Finally, this selection is, as it will be made explicit below, closely related to the traditional form of a discrete Fourier Transform, for which there are well-known algorithms of fast computation such as the FFT (Cooley & Tukey, 1965).

We remark that, in any case, this choice is of course not the only possible one, and that other options may be more adapted for particular cases. For instance, if the target GeRF Z has an isotropic stationary covariance, a partition grounded on *pizza slices* centred at the origin, rather than rectangles, may be more comfortable for analytic computations, although the FFT algorithm is not immediately adapted.

Expression as a Discrete Fourier Transform

Let Z be a real stationary GeRF over $\mathbb{R}^d$. For simplicity, we will suppose that it is a continuous Random Function over $\mathbb{R}^d$ and we will simulate approximations of Z over a fixed domain of $\mathbb{R}^d$, say $[0, L]^d \subset \mathbb{R}^d$, with $L > 0$. The principle is to approach Z by the corresponding analogue approximation (6.4), which is a stationary Random Function. In the case of the Riemann multi-sequence of partitions considered here, $Z_{\mathbf{N}}$ is given by

$$Z_{\mathbf{N}}(x) = \frac{1}{(2\pi)^{\frac{d}{2}}} \sum_{\mathbf{j}=-(\mathbf{N}-1)}^{\mathbf{N}-1} M_Z(V_{\mathbf{j}}^{\mathbf{N}}) e^{i a_{\mathbf{N}} x^T \mathbf{j}} = \mathscr{F}^{-1}\left(\sum_{\mathbf{j}=-(\mathbf{N}-1)}^{\mathbf{N}-1} M_Z(V_{\mathbf{j}}^{\mathbf{N}}) \delta_{a_{\mathbf{N}}\mathbf{j}} \right)(x), \tag{6.72}$$

where $M_Z = \mathscr{F}(Z)$, which is Hermitian since Z is real. The random variables $(M_Z(V_{\mathbf{j}^{\mathbf{N}}}))_{\mathbf{j} \in [\![-(\mathbf{N}-1), \mathbf{N}-1]\!]}$ for a fixed $\mathbf{N} \in \mathbb{N}_*^d$ are all uncorrelated, although non-independent due to the Hermitianity condition. Let us assume, for simplicity, that the spectral measure of Z, μ_Z, satisfies that $\mu_Z(\partial V_{\mathbf{j}^{\mathbf{N}}}) = 0$, that is, that the boundary of the sets of the form (6.66) are μ_Z-null sets, which holds for example when μ_Z has a density with respect to the Lebesgue measure. In such a case one has that $M_Z(V_{\mathbf{j}}^{\mathbf{N}}) = M_Z((V_{\mathbf{j}}^{\mathbf{N}})^{\circ})$, and since M_Z is Hermitian, by (6.71) one has

$$M_Z(V_{\mathbf{j}}^{\mathbf{N}}) = \overline{M_Z(V_{-\mathbf{j}}^{\mathbf{N}})}. \tag{6.73}$$

We can conclude that the orthogonal Random Measure

$$M_{Z_{\mathbf{N}}} = \sum_{\mathbf{j}=-(\mathbf{N}-1)}^{\mathbf{N}-1} M_Z(V_{\mathbf{j}}^{\mathbf{N}}) \delta_{a_{\mathbf{N}}\mathbf{j}} \tag{6.74}$$

is Hermitian. Indeed, if $A \in \mathcal{B}_B(\mathbb{R}^d)$, then

$$\overline{M_{Z_{\mathbf{N}}}(A)} = \sum_{\mathbf{j}=-(\mathbf{N}-1)}^{\mathbf{N}-1} \overline{M_Z(V_{\mathbf{j}}^{\mathbf{N}})} \delta_{a_{\mathbf{N}}\mathbf{j}}(A) = \sum_{\mathbf{j}=-(\mathbf{N}-1)}^{\mathbf{N}-1} M_Z(V_{-\mathbf{j}}^{\mathbf{N}}) \delta_{-a_{\mathbf{N}}\mathbf{j}}(-A) = M_{Z_{\mathbf{N}}}(-A), \tag{6.75}$$

where we have used a simple rearrangement of the sum. Hence, Z_N is a real stationary continuous Random Function.

We recall that from Remark 6.2.1, the Random Function $Z_{\mathbf{N}}$ is periodic. If we want to simulate its values over the domain $[0, L]^d$ we need to ensure that no undesired periodical behaviour will be present. From expression (6.72), we see that the frequencies along the canonical directions are multiples of $a_{\mathbf{N}}$, hence it is through the control of this multi-sequence that we can control the periodicity of Z_N over the desired domain. The lower frequency in expression (6.72) in the direction $k \in \{1, ..., d\}$ is given by $j_k = 1$. In such a case, the associated function $e^{ia_{\mathbf{N}}x_k}$ has period $\frac{2\pi}{a_{\mathbf{N}}}$. From this, we obtain that if we want to simulate $Z_{\mathbf{N}}$ at points belonging to $[0, L]^d$, we must take $a_{\mathbf{N}} \leqslant \frac{2\pi}{L}$. In order to avoid a *reflective behaviour*, that is, not necessarily periodic but rather having undesirable cases such as $Z_{\mathbf{N}}(x) = -Z_{\mathbf{N}}(y)$ for two different points x, y, we must impose

$$a_{\mathbf{N}} \leqslant \frac{\pi}{L}. \tag{6.76}$$

With these considerations we can already simulate the Random Function $Z_{\mathbf{N}}$ over an arbitrary finite collection of points belonging to the domain $[0, L]^d$. Let us suppose that there are $M \in \mathbb{N}$ points in this domain where we want to simulate $Z_{\mathbf{N}}$. Supposing that the computations of the variance and the simulation of every random variable $M_Z(V_{\mathbf{j}}^{\mathbf{N}})$ are of order $\mathcal{O}(1)$, a direct computation of the sum (6.72) involves a quantity of operations with complexity $\mathcal{O}(MN_1N_2...N_d)$. The FFT algorithm allows to reduce the "M" part of this complexity to a $\log(M)$. However, in order to apply it, the points of evaluation in the domain $[0, L]^d$ must be on a regular grid. Let us fix the regular grid which must be used. We recall that the classical expression of a Discrete Fourier Transform in a uni-dimensional case of a vector of complex numbers $(X_0, ..., X_{N-1})$ is given by the vector $(Y_0, ..., Y_{N-1})$ determined by

$$Y_k = \sum_{j=0}^{N-1} X_j e^{-i\frac{2\pi}{N}kj}, \quad \forall k \in \{0, ..., N-1\}. \tag{6.77}$$

The minus sign in the exponentials in (6.77) can be removed without technical issues, the result being a non-normalized Discrete Inverse Fourier Transform, for which the FFT algorithm also applies. The adaptation to a form of the multiple sum (6.72) is done using algebraical rearrangements, for which the final expression is given by a combination of Discrete Fourier Transforms and non-normalized Discrete Inverse Fourier Transforms. What really lacks is the presence of an expression of the form $i\frac{2\pi}{N}kj$ in the exponentials in Eq. (6.72), whose multi-variate version consists in expressions of the form $i2\pi\big(\frac{x_1j_1}{N_1} + \frac{x_2j_2}{N_2} + ... + \frac{x_dj_d}{N_d}\big)$, for $x = (x_1, ..., x_d) \in \mathbb{R}^d$. In order to obtain such an expression, we consider for any fixed $\mathbf{N} \in \mathbb{N}_*^d$ and for any $\mathbf{k} \in \mathbb{N}^d$, the point in $\mathbb{R}^d$:

$$x_{\mathbf{k}}^{\mathbf{N}} = \frac{2\pi}{a_{\mathbf{N}}} \left(\frac{k_1}{N_1}, ..., \frac{k_d}{N_d} \right). \tag{6.78}$$

These points are set on a regular grid over $[0,\infty)^d$. If we evaluate $Z_{\mathbf{N}}$ on one of them we obtain

$$Z_{\mathbf{N}}(x_{\mathbf{k}}^{\mathbf{N}}) = \frac{1}{(2\pi)^{\frac{d}{2}}} \sum_{\mathbf{j}=-(\mathbf{N}-1)}^{\mathbf{N}-1} M_Z(V_{\mathbf{j}}^{\mathbf{N}}) e^{i2\pi\left(\frac{k_1 j_1}{N_1}+\ldots+\frac{k_d j_d}{N_d}\right)}, \quad \forall \mathbf{k} \in \mathbb{N}^d. \tag{6.79}$$

Hence, we obtain an expression which can be easily related to a multivariate form of combinations of Discrete Fourier Transform and Inverse Discrete Fourier Transforms. Such an adaptation must be done by rearranging the sum (6.79) into multi-sums over $[\![\mathbf{0}, \mathbf{N}-1]\!]$.

Methods based on the FFT algorithm produce arrays of the same quantity as the array which is transformed, and hence we will obtain values associated to the points $x_{\mathbf{k}}^{\mathbf{N}}$ such that $\mathbf{k} \in [\![\mathbf{0}, \mathbf{N}-1]\!]$. We remark that just some of these points are in the desired domain $[0, L]^d$. A point $x_{\mathbf{k}}^{\mathbf{N}}$ is in such a domain if and only if for the associated multi-index $\mathbf{k} = (k_1, ..., k_d) \in \mathbb{N}^d$ we have

$$k_l \leqslant \frac{a_{\mathbf{N}} N_l}{2\pi} L, \quad \forall l \in \{1, ..., d\}. \tag{6.80}$$

We remark that from condition (6.76), this implies that $k_1 \leqslant \frac{N_l}{2}$ for all $l \in \{1, ..., d\}$, hence we will only retain a small part of the values obtained when applying the FFT algorithm (less than the half for $d = 1$, for instance). We remark also that since $a_{\mathbf{N}}\mathbf{N} \to \infty$ as $\mathbf{N} \to \infty$, the quantity of points included in $[0, L]^d$ also grows, obtaining a finer simulation grid as we take a higher order of approximation $\mathbf{N}$. The total quantity of evaluation points included in the domain $[0, L]^d$ is

$$\prod_{l=1}^{d} \left(\left\lfloor \frac{a_{\mathbf{N}} N_l}{2\pi} L \right\rfloor + 1 \right). \tag{6.81}$$

The complexity of the algorithm is $\mathcal{O}(N_1 ... N_d \log(N_1 ... N_d))$.

The simulation of the random variables involved in the approximation

Consider the problem of simulating the complex random variables $(M_Z(V_{\mathbf{j}}^{\mathbf{N}}))_{\mathbf{j} \in [\![-(\mathbf{N}-1), \mathbf{N}-1]\!]}$ in expression (6.4). These complex random variables are non-correlated but they are not all mutually independent since they are related through Hermitianity conditions. The variances of each one of these random variables are given by

$$\mathbb{V}ar(M_Z(V_{\mathbf{j}}^{\mathbf{N}})) = (2\pi)^{\frac{d}{2}} \mu_Z(V_{\mathbf{j}}^{\mathbf{N}}). \tag{6.82}$$

We split $M_Z(V_{\mathbf{j}}^{\mathbf{N}})$ in its real and imaginary parts

$$M_Z(V_{\mathbf{j}}^{\mathbf{N}}) = M_Z^R(V_{\mathbf{j}}^{\mathbf{N}}) + i M_Z^I(V_{\mathbf{j}}^{\mathbf{N}}). \tag{6.83}$$

Since M_Z is orthogonal and Hermitian, from Proposition 3.3.3, M_Z^R and M_Z^I are non-correlated real Random Measures, and hence independent in this Gaussian framework. Recalling relation (6.71) and that we have supposed that M_Z is null over the boundaries of the sets $(V_{\mathbf{j}}^{\mathbf{N}})_{\mathbf{j}\in[\![-(\mathbf{N}-1),\mathbf{N}-1]\!],\mathbf{N}\in\mathbb{N}_*^d}$, we obtain the variances

$$\begin{aligned}\mathbb{V}ar(M_Z^R(V_{\mathbf{j}}^{\mathbf{N}})) &= \frac{(2\pi)^{\frac{d}{2}}}{2}\left(\mu_Z(V_{\mathbf{j}}^{\mathbf{N}}) + \mu_Z(V_{\mathbf{j}}^{\mathbf{N}} \cap V_{-\mathbf{j}}^{\mathbf{N}})\right)\\ \mathbb{V}ar(M_Z^I(V_{\mathbf{j}}^{\mathbf{N}})) &= \frac{(2\pi)^{\frac{d}{2}}}{2}\left(\mu_Z(V_{\mathbf{j}}^{\mathbf{N}}) - \mu_Z(V_{\mathbf{j}}^{\mathbf{N}} \cap V_{-\mathbf{j}}^{\mathbf{N}})\right).\end{aligned} \tag{6.84}$$

From relation (6.73) one gets that it is just necessary to simulate a part of these random variables, the rest of them being determined by the already simulated ones through a complex conjugation. The procedure is quite immediate in the case $d = 1$, where in the sum (6.72) we have to just to simulate the cases of index $j \geqslant 0$ and then obtain the case of the negative ones through the Hermitian condition. In higher dimensions the problem is a little bit more complicated.

- We first consider the case where $\mathbf{j}$ is such that $j_1 \geqslant 1$ and its reflections. In such a case, $V_{\mathbf{j}}^{\mathbf{N}} \cap V_{-\mathbf{j}} = \varnothing$ and hence the random variables $M_Z^R(V_{\mathbf{j}}^{\mathbf{N}})$ and $M_Z^I(V_{\mathbf{j}}^{\mathbf{N}})$ are independent with same variance equal to $(2\pi)^{\frac{d}{2}}\frac{\mu_Z(V_{\mathbf{j}}^{\mathbf{N}})}{2}$ (Eq. (6.84)). We simulate these independent real random variables and then we set $M_Z(V_{\mathbf{j}}^{\mathbf{N}}) = M_Z^R(V_{\mathbf{j}}^{\mathbf{N}}) + iM_Z^I(V_{\mathbf{j}}^{\mathbf{N}})$. We then set $M_Z(V_{-\mathbf{j}}^{\mathbf{N}}) = \overline{M_Z(V_{\mathbf{j}}^{\mathbf{N}})}$. The total number of independent real random variables simulated is $2(N_1 - 1)\prod_{l=2}^d(2N_l - 1)$.

- We then consider the case where $j_1 = 0$ and $j_2 \geqslant 1$. We apply an analogue procedure since we still having $V_{\mathbf{j}}^{\mathbf{N}} \cap V_{-\mathbf{j}}^{\mathbf{N}} = \varnothing$. The total number of real independent random variables simulated is $2(N_2 - 1)\prod_{l=3}^d(2N_l - 1)$.

- We apply recursively the same principle until the case $j_1 = j_2 = ... = j_{d-1} = 0$ and $j_d \geqslant 1$, where the total number of real independent random variables simulated is $2(N_d - 1)$.

- Finally, the random variable $M_Z(V_{\mathbf{0}}^{\mathbf{N}})$, where $\mathbf{0} = (0, ..., 0) \in \mathbb{N}^d$, is a real random variable since M_Z is Hermitian and $V_{\mathbf{0}}^{\mathbf{N}}$ is a symmetric set ($V_{\mathbf{0}}^{\mathbf{N}} = -V_{\mathbf{0}}^{\mathbf{N}}$). The variance of this real random variable is $(2\pi)^{\frac{d}{2}}\mu_Z(V_{\mathbf{0}}^{\mathbf{N}})$.

In this procedure, the final total number of real independent random variables simulated is

$$\prod_{l=1}^{d}(2N_l - 1), \tag{6.85}$$

which is the same number of sets in the partition $(V_{\mathbf{j}}^{\mathbf{N}})_{\mathbf{j}\in[\![-(\mathbf{N}-1),\mathbf{N}-1]\!]}$. We remark that this procedure gets highly memory-consuming for big values of $\mathbf{N}$.

6.3.1 Implementation in the case $d = 2$

In this section we apply the method to cases in dimension $d = 2$. We consider the case of SPDEs of the form

$$\mathcal{L}_g U = W, \tag{6.86}$$

restricting ourselves to the case where the solution U is a continuous Random Function, i.e. when $|g|^{-2}$ is integrable. We follow the approach described in Section 6.2.2, and hence we simulate the solution to the approximated problem

$$\mathcal{L}_g U_{\mathbf{N}} = W_{\mathbf{N}}, \tag{6.87}$$

where $W_{\mathbf{N}}$ is an approximation of the White Noise using expression (6.72). Since the Fourier Transform of a White Noise is a White Noise, we have the convenient simple expression of the variances:

$$\mathbb{V}ar(M_W(V_{\mathbf{j}}^{\mathbf{N}})) = \mathrm{Leb}(M_W(V_{\mathbf{j}}^{\mathbf{N}})) = a_{\mathbf{N}}^d, \tag{6.88}$$

which does not depend on $\mathbf{j}$. The expression for the solution $U_{\mathbf{N}}$ to (6.87) is given by

$$U_{\mathbf{N}}(x_{\mathbf{k}}^{\mathbf{N}}) = \frac{1}{(2\pi)^{\frac{d}{2}}} \sum_{\mathbf{j}=-(\mathbf{N}-1)}^{\mathbf{N}-1} \frac{M_W(V_{\mathbf{j}}^{\mathbf{N}})}{g(\xi_{\mathbf{j}}^{\mathbf{N}})} e^{i2\pi\left(\frac{k_1 j_1}{N_1}+\ldots+\frac{k_d j_d}{N_d}\right)}, \quad \forall \mathbf{k} \in \mathbb{N}^d, \tag{6.89}$$

for every point $x_{\mathbf{k}}^{\mathbf{N}}$ on the evaluation grid.

We set the simulation domain to be $[0, L] \times [0, L] \subset \mathbb{R}^2$, for $L = 100$. We also set from now on:

$$a_{\mathbf{N}} = \frac{\pi}{L} \frac{1}{\log_{10}(\max_{l \in \{1,\ldots,d\}} N_l)}, \tag{6.90}$$

which has been chosen arbitrarily. For simplicity, we will always work with multi-indices of the form $\mathbf{N} = (N, ..., N) \in \mathbb{N}_*^d$ for some $N \in \mathbb{N}_*$ which will determine the approximation order. Under these conditions, the total number of points in the simulation grid is given by (Eq. (6.81)):

$$\left(\left\lfloor \frac{N}{2\log_{10}(N)} \right\rfloor + 1\right) \times \left(\left\lfloor \frac{N}{2\log_{10}(N)} \right\rfloor + 1\right) = \left(\left\lfloor \frac{N}{2\log_{10}(N)} \right\rfloor + 1\right)^2. \tag{6.91}$$

For instance, for $N = 2^{10}$ the grid has 170×170 evaluation points, for $N = 2^{11}$ it has 310×310 points, and for $N = 2^{12}$, it has 567×567 points. From expression (6.78) it follows that the step between two neighbour points in the grid along an arbitrary canonical direction is given by $\frac{2L\log_{10}(N)}{N}$.

The variance of the approximation $U_{\mathbf{N}}$, is given by

$$\sigma^2_{U_{\mathbf{N}}} = \frac{a^d_{\mathbf{N}}}{(2\pi)^d} \sum_{\mathbf{j}=-(\mathbf{N}-1)}^{\mathbf{N}-1} \frac{1}{|g(\xi^{\mathbf{N}}_{\mathbf{j}})|^2}, \tag{6.92}$$

which can be obtained from the evaluation of Eq. (6.24) at $h = 0$ (we recall $d\mu_W(\xi) = (2\pi)^{-\frac{d}{2}} d\xi$). This is used in order to obtain normalized simulations, which are then approximations to the normalized version of the solution to (6.86).

Matérn model: illustrations and qualitative Error Analysis

We begin by considering the case of the well-known Matérn model. We consider then the equation

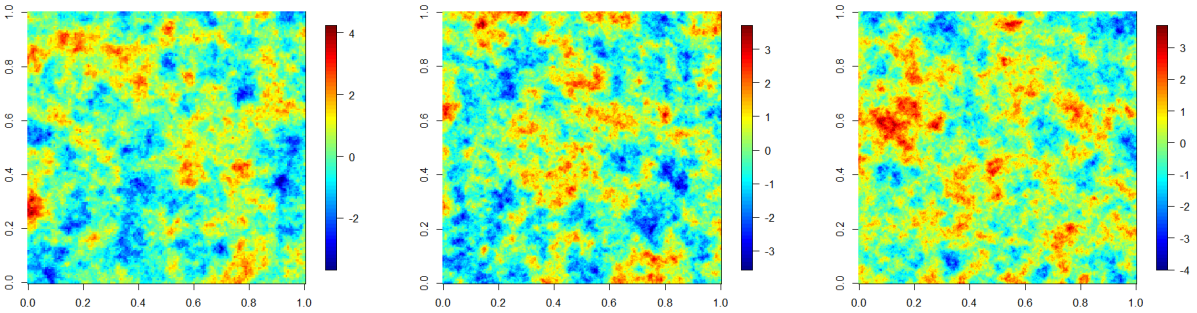
$$(\kappa^2 - \Delta)^{\frac{\alpha}{2}} U = W, \tag{6.93}$$

for some $\kappa > 0$ and $\alpha > \frac{d}{2}$. Hence $g(\xi) = (\kappa^2 + |\xi|^2)^{\frac{\alpha}{2}}$. We consider simulations of the normalized approximated version. Figure 6.1 shows illustrations of such simulations for different approximation orders N and for different regularity parameters α, maintaining the same scale parameter κ.

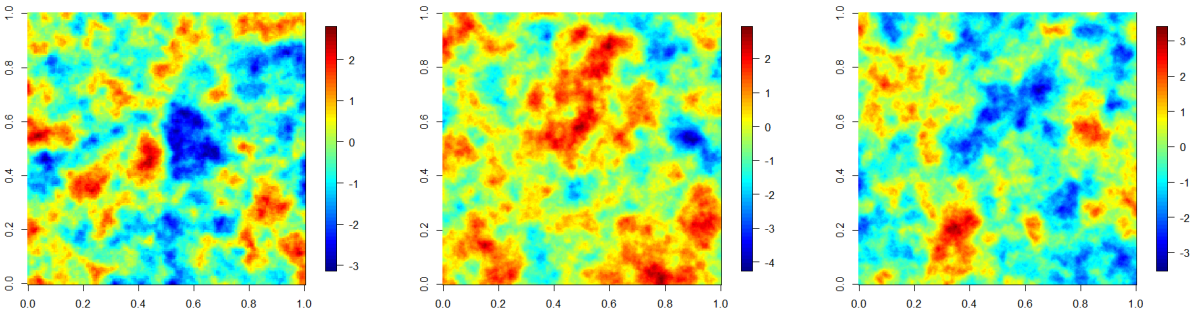
In order to analyse the quality of this approximate procedure, we computed the average experimental variograms of 50 independent simulations and compared it to the theoretical variogram of the Matérn model. The experimental variograms are considered using 20 separation distance bins of width 3, hence considering points separated up to a distance of 60 units (we recall that $L = 100$). The comparison is done for different orders of approximation, regularities and scale parameters. The cases of scale parameter $\kappa = \frac{1}{5}$ are presented in Figure 6.2. The cases of scale parameter $\kappa = \frac{1}{10}$ are presented in Figure 6.3.

In Figures 6.2 and 6.3 it can be appreciated that when the approximation order N grows, the average experimental variograms get closer to the theoretical variogram near to the origin. This tendency is stronger in the cases with high regularity and small practical range. In the sense of the behaviour at the origin, the method produces better approximations under higher regularities, which is not surprising since, as mentioned before, the theoretical approximation $U_{\mathbf{N}}$ is a smooth Random Function. In broad terms, the coincidence between the mean of the empirical mean of the variograms and the theoretical variogram is not ideal. The particular case $\alpha = 4$ and $\kappa = \frac{1}{10}$ presents more difficulties, probably related to the high practical range of this case.

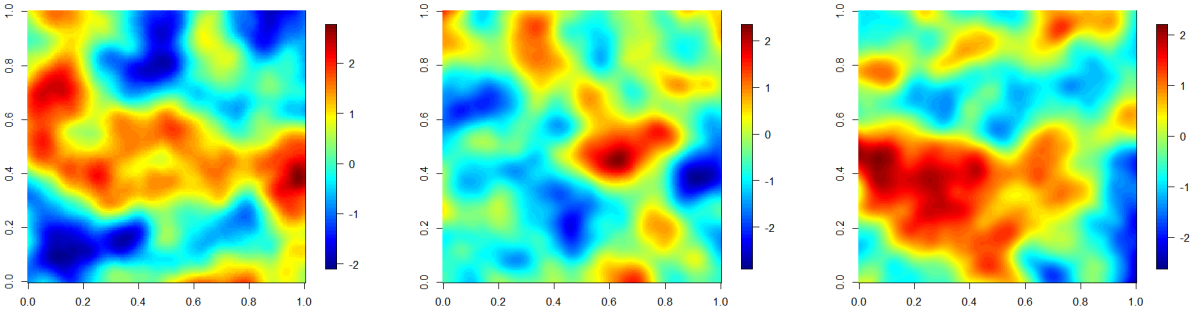
While it is theoretically proven that the approximation gets better as the approximation order grows (Theorem 6.2.2 applies in the case of the Matérn model), it is also true that in such case both the computation time of the method and the memory required to perform it grow. Hence, it is not an issue which may be easily tackled in practice without producing extra computational problems.



(a) $\alpha = 1.5$, $N = 2^{10}$ (b) $\alpha = 1.5$, $N = 2^{11}$ (c) $\alpha = 1.5$, $N = 2^{12}$

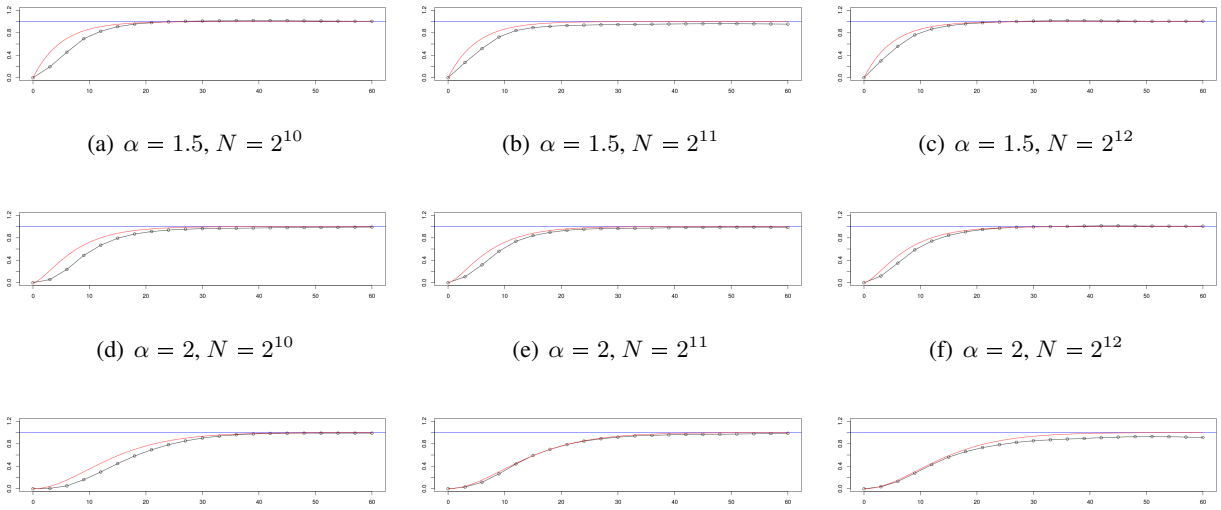


(d) $\alpha = 2$, $N = 2^{10}$ (e) $\alpha = 2$, $N = 2^{11}$ (f) $\alpha = 2$, $N = 2^{12}$



(g) $\alpha = 4$, $N = 2^{10}$ (h) $\alpha = 4$, $N = 2^{11}$ (i) $\alpha = 4$, $N = 2^{12}$

FIGURE 6.1: ILLUSTRATION OF APPROXIMATIONS OF THE MATÉRN MODEL FOR DIFFERENT ORDERS OF APPROXIMATION N AND DIFFERENT REGULARITIES. THE SCALE PARAMETER IS SET AT $\kappa = \frac{1}{5}$. THE VARIANCE IS NORMALIZED.

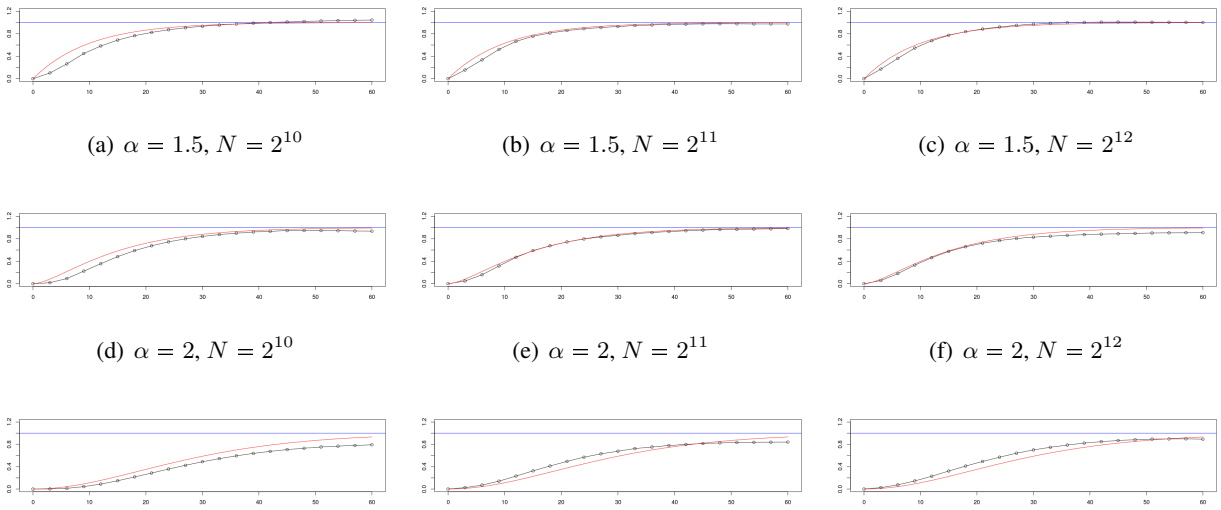


(a) $\alpha = 1.5$, $N = 2^{10}$ (b) $\alpha = 1.5$, $N = 2^{11}$ (c) $\alpha = 1.5$, $N = 2^{12}$

(d) $\alpha = 2$, $N = 2^{10}$ (e) $\alpha = 2$, $N = 2^{11}$ (f) $\alpha = 2$, $N = 2^{12}$

(g) $\alpha = 4$, $N = 2^{10}$ (h) $\alpha = 4$, $N = 2^{11}$ (i) $\alpha = 4$, $N = 2^{12}$

FIGURE 6.2: COMPARISON BETWEEN THE AVERAGE EXPERIMENTAL VARIOGRAMS OF 50 INDEPENDENT SIMULATIONS OF APPROXIMATIONS OF A MATÉRN MODEL (IN BLACK) AND THE THEORETICAL VARIOGRAM OF THE MATÉRN MODEL (IN RED). THE SCALE PARAMETER IS SET AT $\kappa = \frac{1}{5}$. NORMALIZED, THE VARIANCE IS REPRESENTED IN BLUE.



(a) $\alpha = 1.5$, $N = 2^{10}$ (b) $\alpha = 1.5$, $N = 2^{11}$ (c) $\alpha = 1.5$, $N = 2^{12}$

(d) $\alpha = 2$, $N = 2^{10}$ (e) $\alpha = 2$, $N = 2^{11}$ (f) $\alpha = 2$, $N = 2^{12}$

(g) $\alpha = 4$, $N = 2^{10}$ (h) $\alpha = 4$, $N = 2^{11}$ (i) $\alpha = 4$, $N = 2^{12}$

FIGURE 6.3: COMPARISON BETWEEN THE AVERAGE EXPERIMENTAL VARIOGRAMS OF 50 INDEPENDENT SIMULATIONS OF APPROXIMATIONS OF A MATÉRN MODEL (IN BLACK) AND THE THEORETICAL VARIOGRAM OF THE MATÉRN MODEL (IN RED). THE SCALE PARAMETER IS SET AT $\kappa = \frac{1}{10}$. NORMALIZED, THE VARIANCE IS REPRESENTED IN BLUE.

Illustrations of miscellaneous models

One real advantage of this simulation procedure is its generality: we can now easily simulate approximations of a large variety of models presenting different interesting behaviours and being related to different SPDEs. The key issue is the control of the symbol function g, which in the case of equation (6.86) determines the spectral measure of the solution, which has a density. Taking advantage of this condition, we show illustrations of a big variety of models related to different kinds of SPDEs.

All the simulations are normalized. The approximation order is set at $N = 2^{12}$.

Lim-Teo generalization of the Matérn model. We consider the SPDE

$$(\kappa^2 + (-\Delta)^{\alpha})^{\frac{\gamma}{2}} U = W, \tag{6.94}$$

for $\kappa > 0$, $\alpha > 0$ and $\gamma > 0$ such that $\alpha\gamma > \frac{d}{2}$. The unique stationary solution of this SPDE follows a generalization of the Matérn model which was studied in Lim & Teo (2009). The symbol function defining the operator in (6.94) is given by

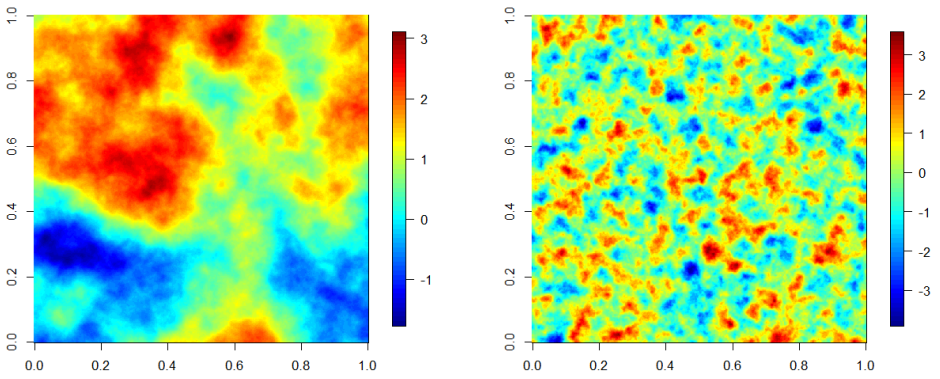
$$g(\xi) = (\kappa^2 + |\xi|^{2\alpha})^{\frac{\gamma}{2}}. \tag{6.95}$$

The mean-square regularity of this model is determined by the product $\alpha\gamma$. Since in this case we can dissociate this product by controlling the parameters α and γ separately, we are able to control the practical range of the model without changing the parameter κ, which now plays the role of a regularising parameter which guarantees the positivity of g and hence the existence of a unique stationary solution. We present in Figure 6.4 two illustrations of approximations of this model using our method. In both of them we have that $\alpha\gamma = 2$, hence the regularity, that is, the mean-square differentiability order is the same, and the parameter κ is also set at $\kappa = \frac{1}{5}$. We illustrate that through the change in the parameters α and γ we are able to control then the practical range.

Advections and other asymmetries. We present examples which consider a symbol function with non-null imaginary part. As seen in Section 5.2, such a condition produces an asymmetric behaviour of the variability of the Random Function with respect to the two spatial components (we may consider it as a non-symmetric anisotropy). The simplest way of doing this is by considering a vector $v \in \mathbb{R}^d$ and considering symbol functions of the form $g(\xi) = g_R + iv^T\xi$. The associated SPDE is then of the form

$$\mathcal{L}_{g_R} U + v^T \nabla U = W, \tag{6.96}$$

which is related to transport phenomena. We take advantage of the generality of our method to consider other non-typical asymmetries induced by particular symbol functions. The imaginary part g_I must be odd.

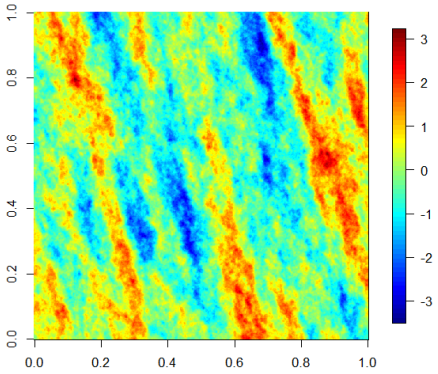


(a) $\alpha = 0.5, \gamma = 4$ (b) $\alpha = 3, \gamma = \frac{2}{3}$

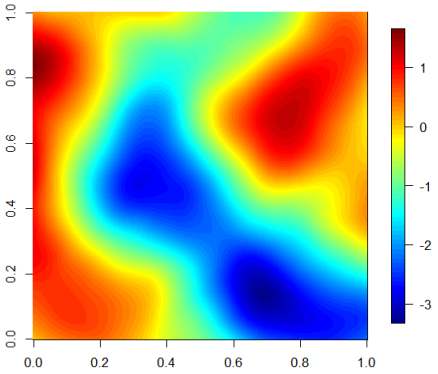
FIGURE 6.4: SIMULATIONS OF APPROXIMATIONS OF THE LIM-TEO GENERALIZATION OF THE MATÉRN MODEL. THE PARAMETER κ IS FIXED AT $\kappa = \frac{1}{5}$. NORMALIZED. WE APPRECIATE DIFFERENT PRACTICAL RANGES WHEN COMPARING BOTH SIMULATIONS, WHILE THE REGULARITIES ARE THE SAME.

We can obtain a large variety of them by considering any continuous polynomially bounded odd function f over $\mathbb{R}$ and then setting $g_I(\xi) = f(v^T\xi)$. It is then expected that some behaviour similar to the simple case of advection (6.96) will be present if we consider the same vector v. We consider the odd functions x^3, $\arctan$ and $\sin$. For the case of x^3 a classical differential operator of third order can be obtained. For the cases of functions $\arctan$ and $\sin$, we do not know if there exists a widely used operator associated to symbol functions involving them.

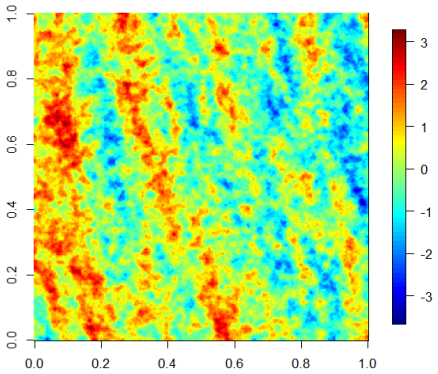
The results are illustrated in Figure 6.5. In all cases we consider the real part of g to be the one associated to the Matérn model, $g_R(\xi) = (\kappa^2 + |\xi|^2)^{\frac{\alpha}{2}}$, with $\kappa = \frac{1}{5}$ and $\alpha = 2$. We set $v = (-1, 4)$. In the basic case of $g_I(\xi) = v^T\xi$ an anisotropy is clearly apparent, which presents higher correlations at same distances in the direction of the vector v with respect to other directions. A similar anisotropy is clearly present in the cases of the functions $\sin$ and $\arctan$. In both cases extra small-range correlations are visible along the direction orthogonal to v. In the case of the $\sin$ function, a slight periodic behaviour along the direction of v can be perceived. In the case of the function x^3 the anisotropy is less clear. This can be explained from the extra regularity that the term $(v^T\xi)^3$ imposes to the Random Function. Since it is a high order polynomial, the associated spectral density of the solution, which is the inverse of a strictly positive polynomial of degree 6, is integrable with respect to high order polynomials, hence it is a more regular covariance model. The simulation shows a significant increase of the practical range with respect to the other cases.



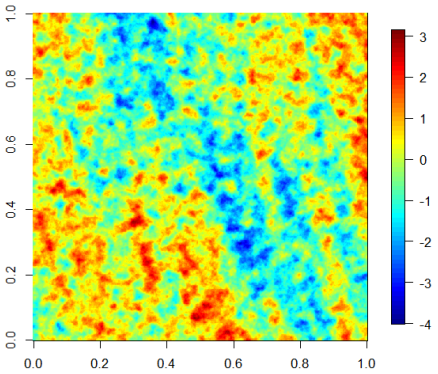
(a) $g(\xi) = (\kappa^2 + |\xi|^2)^{\frac{\alpha}{2}} + iv^T\xi$



(b) $g(\xi) = (\kappa^2 + |\xi|^2)^{\frac{\alpha}{2}} + i(v^T\xi)^3$



(c) $g(\xi) = (\kappa^2 + |\xi|^2)^{\frac{\alpha}{2}} + i\arctan(v^T\xi)$



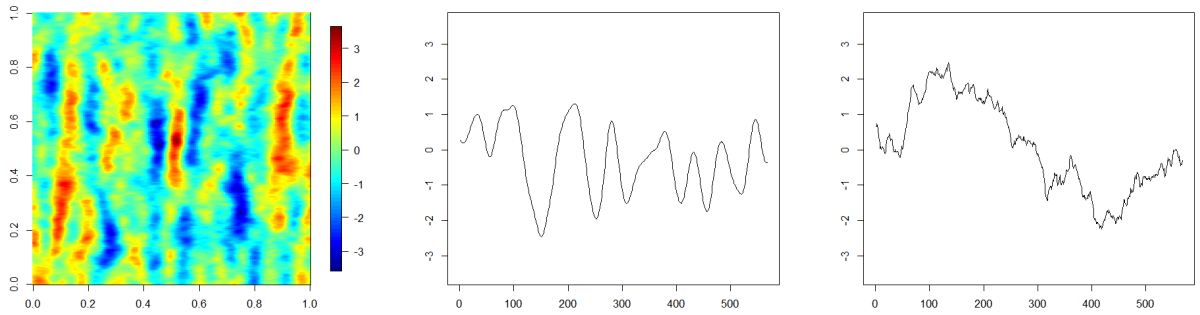
(d) $g(\xi) = (\kappa^2 + |\xi|^2)^{\frac{\alpha}{2}} + i\sin(v^T\xi)$

FIGURE 6.5: REALISATIONS OF MODELS WITH NON-SYMMETRIC BEHAVIOUR IN BOTH COMPONENTS, INDUCED BY AN ADVECTION VECTOR SET AT $v = (-1, 4)$. THE PARAMETERS OF THE REAL PART OF THE SYMBOLS FUNCTIONS ARE ALL SAME, SET AT $\kappa = \frac{1}{5}$ AND $\alpha = 2$.

Models with different regularities along the axes. Within this approach it is easy to construct models presenting different regularities along the axes. The idea is quite similar to the case of the Stein model in a spatio-temporal context (Example 4.5.5). We consider, for instance, a symbol function of the form $g(\xi) = \kappa^2 + |\xi_1|^{2\alpha_1} + |\xi_2|^{2\alpha_2}$, for $\kappa > 0$ and $\alpha_1, \alpha_2 > \frac{1}{2}$. The associated SPDE can be written as

$$\left(\kappa^2 + \left(-\frac{\partial^2}{\partial x_1^2}\right)^{\alpha_1} + \left(-\frac{\partial^2}{\partial x_2^2}\right)^{\alpha_2}\right) U = W. \tag{6.97}$$

Since α_1 (resp. α_2) controls the integrability of the spectral density with respect to the first (resp. second) frequency component, both parameters can control separately the regularities along the axes. In Figure 6.6 we present an illustration of such a model, for which we also show the behaviour along one axis when the other component is fixed at a particular position. As expected, a more regular behaviour along the component with the higher $\alpha-$value associated is clearly visible.



(a) $\alpha_1 = 3, \alpha_2 = 0.7$ (b) Trace at a fixed x_2 (c) Trace at a fixed x_1

FIGURE 6.6: SIMULATION OF AN APPROXIMATION OF A MODEL WITH DIFFERENT REGULARITIES ALLONG THE CANONICAL AXES SATISFYING SPDE (6.97). NORMALIZED. $\kappa = \frac{1}{5}$. AT THE CENTER AND AT THE RIGHT, WE PRESENT AN EXAMPLE OF TRACES ALONG THE AXES, OBTAINED BY FIXING ONE COMPONENT TO A PARTICULAR VALUE. THE FIXED VALUE IN BOTH CASES CORRESPONDS TO THE 50TH POSITION IN THE REGULAR SIMULATION GRID (EQUALS TO $8.642853 \in [0, 100]$ IN THIS CASE).

The model associated to Eq. (6.97) is symmetric in the sense of Section 5.1.2 (we can interpret, for instance, the second component as a time component). Hence, it is interesting to illustrate the behaviour of a non-symmetric case, as presented in Section 5.2. We will thus consider solutions to SPDEs of the form

$$\left(\kappa^2 - \frac{\partial^2}{\partial x_1^2}\right)^{\frac{\alpha}{2}} U + \frac{\partial^\beta U}{\partial x_2^\beta} = W, \tag{6.98}$$

for $\kappa > 0$, $\alpha > \frac{1}{2}$ and $\beta > \frac{1}{2}$. The associated symbol function is then

$$g(\xi) = (\kappa^2 + |\xi_1|^2)^{\frac{\alpha}{2}} + (i\xi_2)^{\beta}. \tag{6.99}$$

As we mentioned in Section 5.2, these models present a non-symmetric behaviour which in addition allows to control the regularity along the axes. In Figure 6.7, illustrations of approximations of these models are presented, with different combinations of the regularity parameters α and β. The parameter κ is set at $\kappa = \frac{1}{5}$. The regularity along the axes is studied similarly to the case of the symmetric model (6.97). A corresponding change in the regularities along the axes when changing the parameters α and β is observed, as expected.

6.3.2 First order evolution models: asymptotic convergence

We illustrate a simulation of a first order evolution model. We follow the approach proposed in Section 6.2.3. We consider the case where the spatial symbol function is the one associated to the Matérn model, $g(\xi) = (\kappa^2 + |\xi|^2)^{\frac{\alpha}{2}}$. The source term is supposed to follow a coloured in space and white in time noise, which is spatially approximated as explained in Section 6.2.3: $X^S_{\mathbf{N}} \boxtimes W_T$, with $X^S_{\mathbf{N}}$ representing the approximation of a Matérn model with parameters κ_{X_S} and α_{X_S}. We also consider a null initial condition.

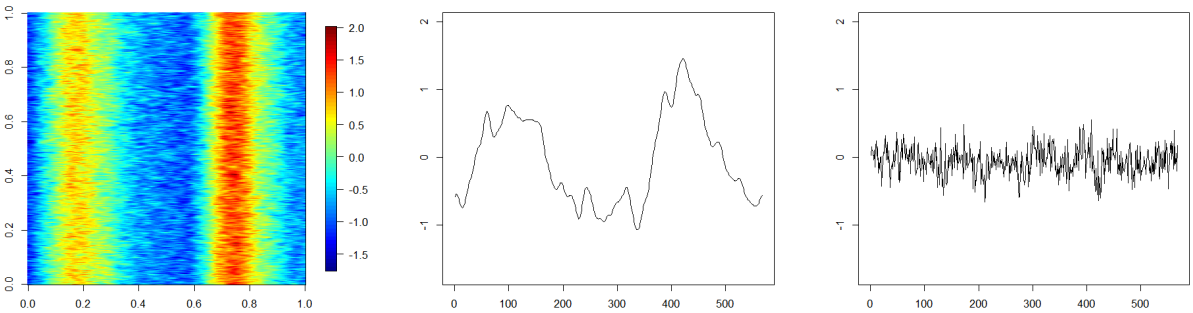
We simulate hence a GeRF $U_{\mathbf{N}}$ which satisfies the Cauchy problem:

$$\begin{cases} \dfrac{\partial U_{\mathbf{N}}}{\partial t} + (\kappa^2 - \Delta)^{\frac{\alpha}{2}} U_{\mathbf{N}} = X^S_{\mathbf{N}} \boxtimes W_T & \text{over } \mathbb{R}^d \times \mathbb{R}^+ \\ U_{\mathbf{N}}\big|_{t=0} = 0 \end{cases} . \tag{6.100}$$

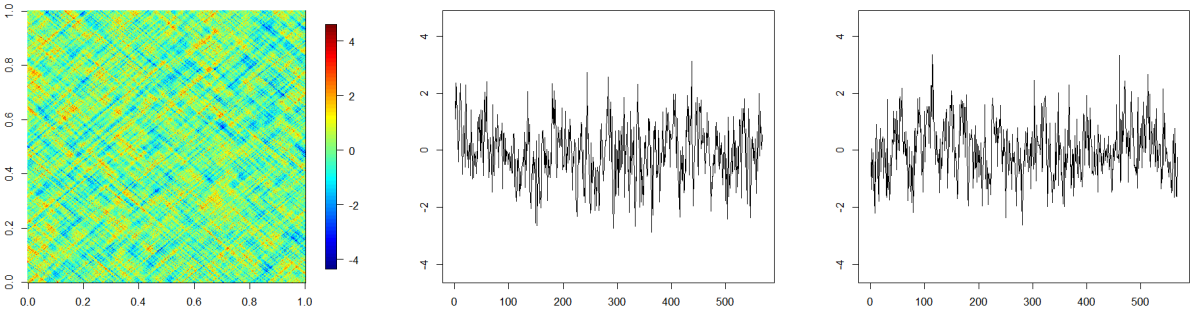
We set the parameters to $\kappa = \frac{1}{5}$, $\alpha = 3.12$, $\kappa_{X_S} = \frac{1}{5}$, and $\alpha_{X_S} = 0.65$. The spatial domain, as well as the other parameters of the approximation method are set as in section 6.3.1. The order of approximation N is set at $N = 2^{10}$. The time simulation is done over a regular temporal grid starting at $t = 0$ and with step $\Delta t = 0.1$.

Some images related to particular time locations are illustrated in Figure 6.8. At every fixed time location, we calculate an experimental spatial variogram following the same conventions as in Section 6.3.1, and we compare them with the spatial stationary variogram which the solution should follow once the time has flown long enough, as it was stated in Section 5.3.4. In this case, the theoretical limit spatial variogram is the one of a Matérn model, with scale parameter $\kappa_{U_S^{stat}} = \frac{1}{5}$, regularity parameter $\alpha_{U_S^{stat}} = \frac{\alpha}{2} + \alpha_{X_S} = 2.21$, and with variance (sill) equal to

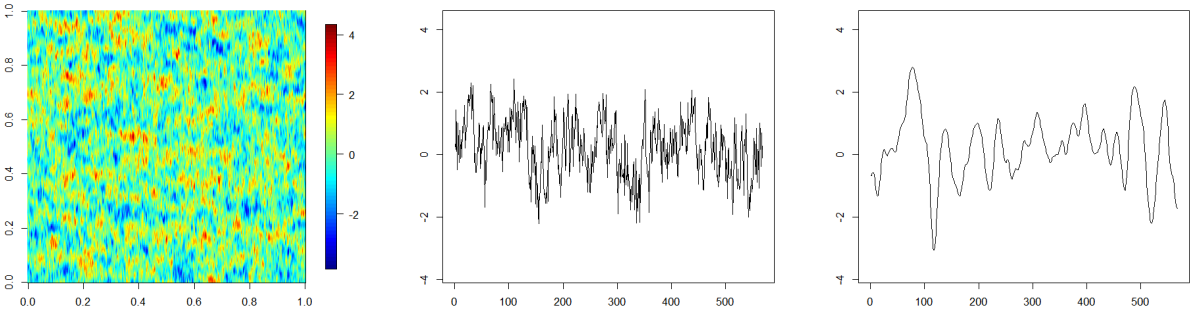
$$\sigma_{U_S^{stat}} = \frac{1}{2} \frac{1}{4\pi(\alpha_{U_S^{stat}} - 1)\kappa_{U_S^{stat}}^{2(\alpha_{U_S^{stat}} - 1)}} \approx 1.61615.$$



(a) $\alpha = 4, \beta = 0.7$ (b) Trace at a fixed x_2 (c) Trace at a fixed x_1



(d) $\alpha = 2, \beta = 1.8$ (e) Trace at a fixed x_2 (f) Trace at a fixed x_1



(g) $\alpha = 0.7, \beta = 4$ (h) Trace at a fixed x_2 (i) Trace at a fixed x_1

FIGURE 6.7: SIMULATION OF APPROXIMATIONS OF THE SOLUTION TO EQ. (6.98) FOR DIFFERENT COMBINATION OF REGULARITIES. NORMALIZED. $\kappa = \frac{1}{5}$. AT THE CENTER AND AT THE RIGHT, WE PRESENT AN EXAMPLE OF TRACES ALONG THE AXES, OBTAINED BY FIXING ONE COMPONENT TO A PARTICULAR VALUE. THE FIXED VALUE IS THE SAME AS IN FIGURE 6.6.

The convergence to the stationary model can be observed in this simulation.

6.3.3 Waving models

We finish this illustration section with simulations of a Waving model. We follow the approach proposed in Section 6.2.4. The spatial behaviour is set to be a normalized approximation of a Matérn model, obtaining thus Waving Matérn models (Section 5.2.4). Hence, the approximation $U_{\mathbf{N}}$ satisfies the system of SPDEs:

$$\begin{cases} \dfrac{\partial^2 U_{\mathbf{N}}}{\partial t^2} - c^2 \Delta U_{\mathbf{N}} = 0 & \text{over } \mathbb{R}^d \times \mathbb{R} \\ (\kappa^2 - \Delta)^{\frac{\alpha}{2}} U_{\mathbf{N},S} \overset{2nd\ o.}{=} \frac{1}{a} W_{S,\mathbf{N}} & \text{over } \mathbb{R}^d \end{cases}, \tag{6.101}$$

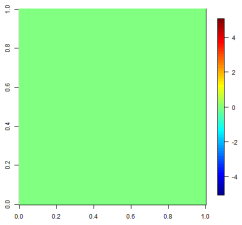
where $W_{\mathbf{N},S}$ denotes an approximation of a spatial White Noise following our method. $a > 0$ is a constant normalizing the variance of the solution.

We set the parameters of the spatial model to $\kappa = \frac{1}{5}$ and $\alpha = 2$. The wave propagation velocity c is set at $c = 8$. The spatial domain, as well as the other parameters of the approximation method are set as in section 6.3.1. The order of approximation N is set at $N = 2^{10}$. The time simulation is done over a regular temporal grid starting at $t = 0$ and with step $\Delta t = 0.1$. Some images related to particular time locations are illustrated in Figure 6.9. At every fixed time location, we calculate an spatial experimental variogram following the same conventions as in Section 6.3.1, and we compare them with the theoretical spatial variogram which is the one of a Matérn model. The resulting experimental variograms oscillate around the theoretical one, which is the expected behaviour due to the statistical variability of a experimental variogram with respect to its theoretical counterpart.

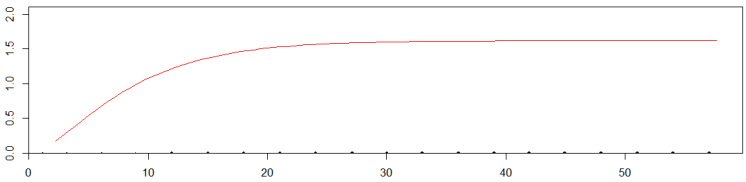
6.4 Discussion

The method discussed in this chapter allows us to obtain simulations of approximations of stationary models with great generality. The versatility of the method has allowed us to illustrate solutions to many SPDEs, with associated covariance models having non-trivial properties which are easily controllable by suitable parameters. It provides hence a quite general method to visualize the behaviour of new stationary models associated to SPDEs without the restrictions on the associated operator which are present in the case of the FEM. It also allows to simulate over a spatial regular grid with a large quantity of points.

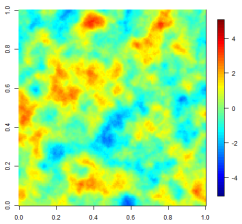
Within the framework worked out in this dissertation, we were able to theoretically prove the convergence of this method to the target model in weak and strong senses under suitable conditions. This is also a contribution to the development of this method, for which its main principles and applications have been already proposed in the literature, for example in Pardo-Iguzquiza & Chica-Olmo (1993) in the case of a general



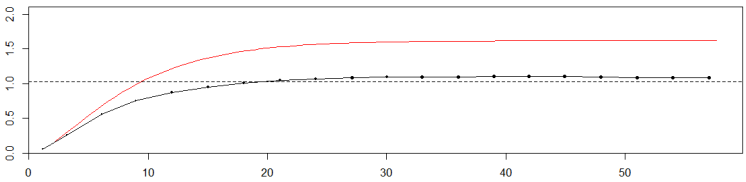
(a) $t = 0$



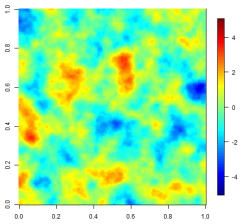
(b) Spatial experimental variogram



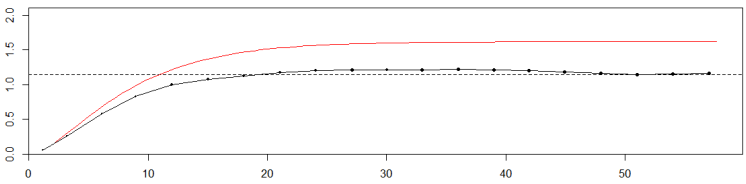
(c) $t = 30$



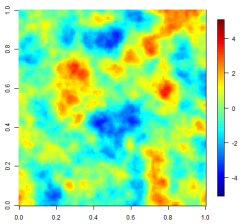
(d) Spatial experimental variogram



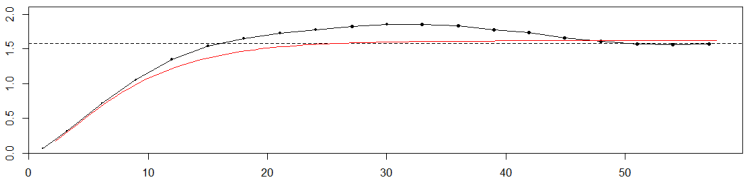
(e) $t = 60$



(f) Spatial experimental variogram

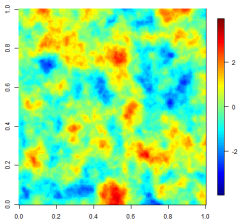


(g) $t = 90$

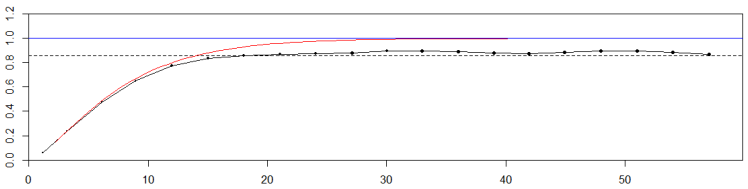


(h) Spatial experimental variogram

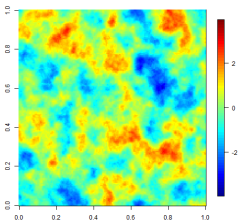
FIGURE 6.8: ILLUSTRATION OF THE SIMULATION OF A FIRST ORDER EVOLUTION MODEL CONVERGING ASYMPTOTICALLY SPATIO-TEMPORALLY TO ITS STATIONARY SOLUTION FOLLOWING SPATIALLY A MATÉRN MODEL. ILLUSTRATIONS AT DIFFERENT TIME LOCATIONS. AT THE RIGHT, THE ASSOCIATED SPATIAL EXPERIMENTAL VARIOGRAM (IN BLACK), COMPARED WITH THE THEORETICAL SPATIAL VARIOGRAM (IN RED) OF THE LIMIT MODEL. IN DOTTED BLACK, AN ESTIMATED SILL.



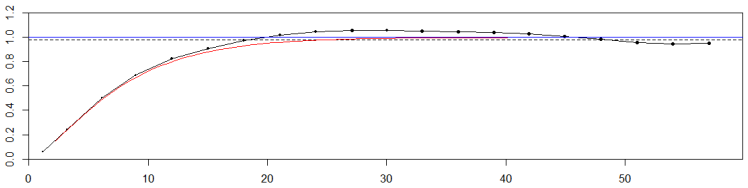
(a) $t = 1$



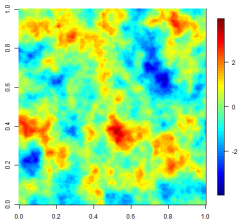
(b) Spatial experimental variogram



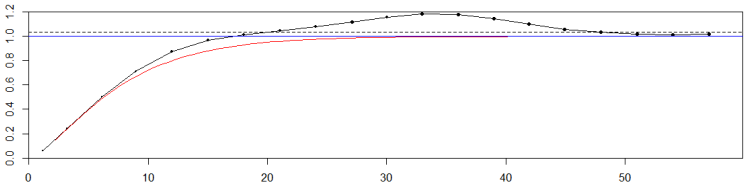
(c) $t = 1.5$



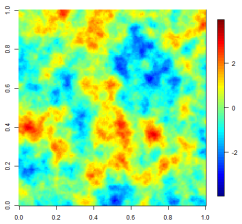
(d) Spatial experimental variogram



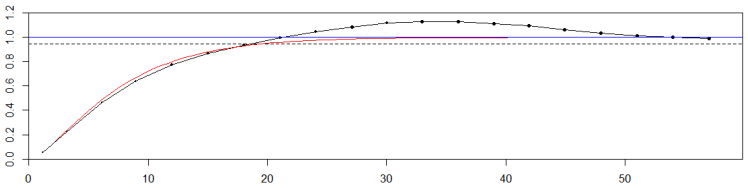
(e) $t = 2$



(f) Spatial experimental variogram



(g) $t = 2.5$



(h) Spatial experimental variogram

FIGURE 6.9: ILLUSTRATION OF THE SIMULATION OF A WAVING MODEL WITH NORMALIZED APPROXIMATED MATÉRN SPATIAL COVARIANCE, WITH $\kappa = \frac{1}{5}$ AND $\alpha = 2$. WAVE PROPAGATION VELOCITY $c = 8$. ILLUSTRATIONS AT DIFFERENT TIME LOCATIONS. AT THE RIGHT, THE ASSOCIATED SPATIAL EXPERIMENTAL VARIOGRAM (IN BLACK), COMPARED WITH THE THEORETICAL SPATIAL VARIOGRAM (IN RED). IN BLUE IT IS REMARKED THE UNITARY VARIANCE. IN DOTTED BLACK, AN ESTIMATED SILL.

stationary Random Function (as stated in Section 6.2.1) and in Lang & Potthoff (2011) for its application to SPDEs (as stated in Section 6.2.2). The uniform convergence on compacts sense of the covariance provides a framework where the variograhic behaviour of the approximated simulated models is similar enough to the target model when the approximation order is high enough. However, other criteria of *good approximation* can be stated for these simulations based on other ways of convergence. For instance, rather than focusing on a particular theoretical convergence of the approximation Random Fields to the target Random Field, one may define a good approximation criterion using statistical approaches, and arguing that the approximation is good enough if after some approximation order the model passes suitable statistical tests grounded on the target model. This approach is for example the one applied in Lang (2007), and the one proposed in Pereira & Desassis (2018) under other context of approximations of Gaussian Random Fields.

The use of the FFT algorithm allows us to simulate efficiently the approximations of the model over a regular grid on the desired domain. As the approximation order $\mathbf{N}$ grows, this method provides a more accurate simulation, both in the sense of a better approximation of the target Random Function and in the sense of the quantity of simulation points. Its complexity increases dramatically with the spatial dimension. The memory problem can be tackled using different approaches, see for instance the Singleton's algorithm presented in Teukolsky et al. (1992, Section 12.6). However, if we are interested in simulating an approximation of the model over a non-regular grid, or over a quantity of points which is much smaller than (6.81), the direct computation of the sum (6.72) is more convenient.

In a general and qualitative sense, the approximation method behaves relatively well in regular cases with limited practical range. In case with less regularity, the method provides approximations which may fail to recreate the regularity at low approximation orders. Although we have proved a uniform on compacts convergence of the covariance of the approximation to the target covariance model, we have not provided a vanishing bound to the general case of a resolution of the SPDE (6.18). Hence we have not been able to theoretically indicate at which approximation order this method provides an approximation similar enough to the target model. This increases the difficulties in the implementation of the method, since an augmentation on the approximation order increases the computational time and the memory storage of the method, specially for high dimensions.

Chapter 7

Conclusion and Perspectives

In this dissertation we have entered into the theoretical details of the current application of the SPDE approach in Geostatistics. We have set a particular framework of Stochastic Analysis, namely, the framework of Generalized Random Fields within the mean-square theory. We have exposed this framework and contrasted its differences with other approaches to Stochastic Calculus, and we concluded that it is the most adapted for the needs of geostatistical analysis. Within this framework where Random Functions, Measures and Distributions cohabit together, we have been able to give a rigorous notion of a linear SPDE and we have explained how such an equation determines the covariance structure of its potential solutions. We have been able to obtain criteria of existence and uniqueness of stationary solutions of linear SPDEs involving operators defined through a symbol. This has allowed us to recover many already known theoretical relationships between SPDEs and covariance models, and to obtain new relationships. These developments have also allowed us to construct spatio-temporal geostatistical models presenting non-trivial properties, allowing to control easily the spatio-temporal symmetry and separated regularity. We have been able to obtain interesting results concerning stationary solutions for physically-driven SPDEs, and we have entered into the details of the initial value problem relating SPDEs with a first-order temporal derivative operator involved. Concerning simulation methods, we have studied a particular already known spectral method which was perfectly adapted to the framework of this dissertation, and for which we have given rigorous mathematical proofs of its performance. This method has allowed us to easily illustrate models presenting non-trivial properties and to visualize the models we have developed in this dissertation.

As mentioned in the introductory chapter, the SPDE approach is a vast theoretical and practical field which requires, and will keep requiring, further research work and developments. We expect that the results presented in this dissertation will enlighten some ideas and motivate future research questions, but there are still many important points within this framework which have not been worked out in this dissertation. We present here our perspectives of future work for the points which we consider are the most important

ones. They are grounded on questions and approaches which arise immediately from the developments presented in this work. These perspectives concern three major issues: the theoretical question concerning the possibility of relating an arbitrary and well-known covariance model with a convenient SPDE (which is basically the inverse problem of what we have done in Chapters 4 and 5), the development of non-stationary models taking advantage of the SPDE approach, and the development of inference methods adapted to this framework and to the models developed in this dissertation.

A SPDE for a generic model

The third question presented in the introductory Section 1.3 was not addressed explicitly with whole generality. By following the approach presented in Chapters 4 and 5, we know that we can easily present a SPDE for a stationary model whose spectral measure has a density with respect to another spectral measure, relating the density to a convenient symbol function. We can also apply this idea for stationary models whose spectral measures are concentrated on suitable subsets of the frequency space, obtaining a homogeneous SPDE for the model to satisfy. With this aim we have been able, for example, to propose SPDEs for the Stein model (Example 4.5.5) and the $J-$Bessel model (Example 4.5.4). However, there are still many other models, even in a stationary framework, for which we do no not know how to relate them to a convenient SPDE, particularly if we do not know the spectral measure associated to the model. For instance, a convenient SPDE for the Gneiting class of covariance functions (5.4) would produce a huge interest within the spatio-temporal statistical community, since such a model is popular for its flexibility and capacity to obtain non-separable models, and for which we do not have an explicit spectral measure associated.

The direct general question "*how can we relate a covariance model to a convenient SPDE*" has then not been fully answered in this work. Well understood, any GeRF over $\mathbb{R}^d$, say Z, can be trivially related to a SPDE involving any operator $\mathcal{L}$ for which its application to Z is well-defined. But in such a case, the SPDE may not be *convenient* since on the right side of the equation we may obtain a rather complicated GeRF for which the facilities within the SPDE approach may not be immediate to apply. A more interesting question is, for example, the following one: *given a GeRF Z over $\mathbb{R}^d$ following a particular covariance structure, does it exist an operator $\mathcal{L} : \mathscr{S}'(\mathbb{R}^d) \to \mathscr{S}'(\mathbb{R}^d)$ such that*

$$\mathcal{L}Z = W? \tag{7.1}$$

Here W denotes, as usual, a White Noise over $\mathbb{R}^d$. This more precise question is more relevant for many reasons. For instance, results presented in Section 4.4 have shown the importance of the case with a White Noise source term. In addition, a White Noise is a sort of *simple model*, in the sense that, due to all of its properties presented in this dissertation, it is easy to treat, simulate and analyse.

We have actually advanced in this question during the PhD period. Although we do not have a formal

result, we have many indices as to how to answer this question theoretically. A first simple analysis shows that the solution is not unique: different operators may lead to the same covariance structure. For example, consider the SDEs over $\mathbb{R}$:

$$\left(\kappa + \frac{d}{dx}\right) U = \sigma W \quad ; \quad \left(\kappa - \frac{d}{dx}\right) U = \sigma W \quad ; \quad \left(\kappa^2 - \frac{d^2}{dx^2}\right)^{\frac{1}{2}} U = \sigma W, \tag{7.2}$$

where $\kappa, \sigma > 0$ and W is a White Noise. Using our typical approach with symbol functions, it is not difficult to prove that all these equations have unique stationary solutions, and that they all have the same spectral measure $d\mu_U(\xi) = \frac{\sigma^2 d\xi}{\kappa^2 + |\xi|^2}$, which corresponds to an exponential model (Eq. (3.5)). More generally, when facing equations of the form $\mathcal{L}_g U = X$ as seen in Chapter 4, we recall that the spectral measure of the potential solution(s) depends on g only through $|g|$, and hence different symbol functions with the same modulus generate the same covariance model. Hence, a criterion for selecting one of these solutions should be proposed. An example is, for instance, to require that the operator is self-adjoint and positive-definite, that is, such that $\langle \mathcal{L}(\varphi), \phi \rangle = \langle \varphi, \mathcal{L}(\phi) \rangle$ and that $\langle \mathcal{L}(\varphi), \overline{\varphi} \rangle \geqslant 0$ for all $\varphi, \phi \in \mathscr{S}(\mathbb{R}^d)$. For instance, in the case of SDEs (7.2) the only operator satisfying these properties among the presented operators is $\left(\kappa^2 - \frac{d^2}{dx^2}\right)^{\frac{1}{2}}$. We remark that, at any state, the solutions to equations (7.2) are different strictly speaking: they only coincide in the covariance structure that they follow.

More generally and regardless of this *uniqueness* issue, the existence of an operator $\mathcal{L}$ such that (7.1) holds can be studied following a different approach based on convenient spectral decompositions of Z and the White Noise W. The idea is actually quite simple. Let us suppose that we are in a Gaussian framework. It can be proven that any GeRF over $\mathbb{R}^d$ can be completely determined by an at most countable quantity of independent random variables. This may be astonishing at first: one may have the idea that, for instance, a Random Function $(Z(x))_{x \in \mathbb{R}^d}$ is determined by a non-countable quantity of random variables. While this is true in general, when imposing regularity conditions such as continuity it is easy to verify that Z is completely determined if we just specify the evaluations at a countable dense subset of $\mathbb{R}^d$. Although those evaluations are not necessarily independent, a similar but more technical analysis can be done to obtain an at most countable family of independent random variables which determines completely the Random Function Z. An approach using a Karhunen-Loève expansion, for instance, may provide such a family of random variables. In the framework of GeRF, we conjecture that an analogue procedure can be done. This arises from the fact that the Schwartz space $\mathscr{S}(\mathbb{R}^d)$ is separable (Reed & Simon, 1980, Corollary 2 to Theorem V.14). Hence, by defining the random variables $\langle Z, \varphi \rangle$ for all the functions φ belonging to a suitable countable dense subset of $\mathscr{S}(\mathbb{R}^d)$, we have determined completely the GeRF Z. We conjecture that a *generalized Karhunen-Loève expansion* can be constructed, and that it is possible to obtain an at most countable basis of *functions on* $\mathscr{S}(\mathbb{R}^d)$ such that the development of Z on this basis involves mutually independent random variables with the same variance. Such a basis can be obtained through the use of the covariance distribution of Z, $C_Z \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^d)$, and taking advantage of its positive-definite Kernel

structure in order to construct a pre-inner product over a suitable subspace of $\mathscr{S}(\mathbb{R}^d)$, from where the basis functions can be obtained as a result of an orthogonalisation process. We give as example the case of the White Noise W, for which we can use, for instance, the *Hermite functions* as a basis of functions in $\mathscr{S}(\mathbb{R}^d)$ (see Reed & Simon, 1980, Appendix to V.3). Let us denote by $(\mathfrak{h}_\beta)_{\beta\in\mathbb{N}^d}$ the collection of Hermite functions over $\mathbb{R}^d$. It is known that this countable collection of functions is included in $\mathscr{S}(\mathbb{R}^d)$, that they form an orthonormal system of $L^2(\mathbb{R}^d)$, and that *every test-function in $\mathscr{S}(\mathbb{R}^d)$ and every tempered distribution* can be decomposed in a formal series based on this system of functions (Reed & Simon, 1980, Theorems V.14 and V.15). In the stochastic framework, one has for the White Noise:

$$W = \sum_{\beta\in\mathbb{N}^d} \langle W, \mathfrak{h}_\beta\rangle \mathfrak{h}_\beta, \tag{7.3}$$

where the multi-series is considered in the sense of a mean-square-$\mathscr{S}'(\mathbb{R}^d)$-weak-$*$ sense. The orthonormality of the Hermite functions allows to conclude that the random variables $(\langle W, \mathfrak{h}_\beta\rangle)_{\beta\in\mathbb{N}^d}$ are mutually independent with same variance. For a general GeRF Z, the formal series

$$Z = \sum_{\beta\in\mathbb{N}^d} \langle Z, \mathfrak{h}_\beta\rangle \mathfrak{h}_\beta \tag{7.4}$$

also holds, the random variables $(\langle Z, \mathfrak{h}_\beta\rangle)_{\beta\in\mathbb{N}^d}$ being not necessarily independent. We conclude that, intuitively speaking, *any GeRF can be identified with a countably-infinite dimensional random vector*. A linear operator $\mathcal{L} : \mathscr{S}'(\mathbb{R}^d) \to \mathscr{S}'(\mathbb{R}^d)$ defined through an adjoint can then be identified with a *countably-infinite matrix*, determined by its action on the Hermite functions. The idea is then, to *construct this matrix determining the operator $\mathcal{L}$ in such a way that we could obtain a vector of independent random variables from the collection* $(\langle Z, \mathfrak{h}_\beta\rangle)_{\beta\in\mathbb{N}^d}$. This procedure could be done, for instance, through a generalization of the Cholesky factorisation to an infinite-dimensional case. With these developments, we conjecture then the following results:

- Let Z be a real GeRF over $\mathbb{R}^d$. Then, there exists a real linear operator $\mathcal{L}_1 : \mathscr{S}'(\mathbb{R}^d) \to \mathscr{S}'(\mathbb{R}^d)$ and a real White Noise W_1 such that
$$Z = \mathcal{L}_1 W_1. \tag{7.5}$$

- Let Z be a real GeRF over $\mathbb{R}^d$ with covariance distribution C_Z. Let us denote by $\mathrm{Ker}(Z) = \{\varphi \in \mathscr{S}(\mathbb{R}^d) \parallel \langle Z, \varphi\rangle \overset{a.s.}{=} 0\}$. If the quotient space $\mathscr{S}(\mathbb{R}^d)/\mathrm{Ker}(Z)$ has an infinite absolute basis, then there exists a real linear operator $\mathcal{L}_2 : \mathscr{S}'(\mathbb{R}^d) \to \mathscr{S}'(\mathbb{R}^d)$ such that
$$\mathcal{L}_2 Z = W_2, \tag{7.6}$$
where W_2 is a White Noise.

For these conjectures the intuition is actually simple. For the first one, we relate it to the discrete result of determining a finite random vector with a desired covariance matrix starting with a vector of independent random variables, the typical method being a Cholesky decomposition of the covariance matrix. For the second one, we follow the intuition that we cannot obtain countably many independent random variables with the same strictly positive variance starting from a random vector which is only determined by a finite quantity of independent (non-constant) random variables. Hence, to construct a White Noise starting from Z, we necessarily need that Z must be determined by an infinite quantity of independent random variables. The operators $\mathcal{L}_1$ and $\mathcal{L}_2$ are not the only ones satisfying these properties, but as stated before, extra requirements may provide a unique particular selection, such as requiring self-adjointness and positive-definiteness.

These developments, if achieved, would have a huge theoretical value which will improve our understanding of GeRFs and the relationship between SPDEs and covariance structures. We remark that we have not supposed stationarity here, so the conjectures involve quite general covariance structures. However, it is not immediate to apply them in a practical context. The construction of the operator $\mathcal{L}$ is rather abstract and it is not clear in general how to relate it to classical differential operators or operators defined through a symbol. A study on the action of these well-known operators on the Hermite functions may provide conditions that $\mathcal{L}$ must satisfy in order to identify it as an operator belonging to an already known class of operators. Simulation and inference methods associated to this development may be achieved by studying the suitability of a spectral method based on the Hermite basis.

Non-stationarity

The construction of non-stationary geostatistical models presents special issues. Concerning the positive-definiteness condition that a covariance Kernel must satisfy, constructing valid and flexible covariance models is specially intricate, since we do not even count on Bochner's Theorem to easily provide a positive-definite structure as in the stationary case. Many approaches of development on stationary models have been done. We refer to Fouedjio (2014) for a presentation and developments of different methodologies in this aim.

The SPDE approach is then one interesting proposition to construct non-stationary models in a simple way without the technical requirements on the positive-definite Kernel structure. Indeed, if the Random Field is the unique and well-defined solution to a SPDE which presents parameters varying spatially or spatio-temporally, then its covariance structure presents a non-stationary behaviour, which is in addition easily controllable and parametrizable. We do not even need to know the covariance itself, the SPDE makes all the job. This is the approach, for example, in Fuglstad et al. (2013), where a non-stationary anisotropical diffusion matrix is added on the Matérn equation (1.1).

This dissertation did not deal explicitly with this issue. In Chapter 5 we deal mainly with stationary

models. In Section 5.3 we construct solutions to first order evolution models which are in general non-stationary models, although the most interesting properties are rather present in the cases where stationarity holds. The simulation technique proposed in Chapter 6 works also, theoretically, for non-stationary models, provided that the Random Fields are such that their Fourier Transforms are slow-growing Random Measures. However, even in such a case, the Random Measure is not orthogonal, and since one of the particularities of this simulation method is that it takes advantage of this orthogonality, we conclude that this simulation method proposes no essential advantages at all: the random variables of the form $M_Z(V_j^N)$ in (6.4) are correlated and hence they are as difficult to simulate as the evaluation of the Random Field at arbitrary points in the space. Besides those facts, the subject of non-stationary models has not been properly treated. The main reason is actually that the stationary framework provided an easier context where to obtain new results, and hence was more attractive to attack in a first sight.

Many questions for future research follow. For instance, the application of the ideas presented in Fuglstad et al. (2013) into a spatio-temporal context would provide interesting spatio-temporal covariance model with non-trivial properties. This would be one possible approach for solving the SPDE (3.91). We have already mentioned some theoretical issues related to this equation in Section 3.8. Studying its deterministic counterpart may provide a potential enlightenment for the stochastic case under the right conditions, similarly to the case of first order evolution models presented in Section 5.3. The Lax-Milgram theorem may provide a theoretical justification of a solution to such equation (see for instance Clément & Martin, 2016) . We remark that the main theoretical difference to the stationary case is that when stationarity is supposed, the studied Random Fields behave as members of the space $\mathscr{V}'(\mathbb{R}^d)$ (Eq. (4.1)). Hence, it is easier to restrict our space of possible solutions and to determine the kinds of operations that can be applied on its members. In a non-stationary context, there is no particular special subspace (besides $\mathscr{S}'(\mathbb{R}^d)$ itself) on which to restrict our work: such a space will depend on the particular cases depending on the behaviour of the parameters of the model along the space-time. We remark, however, that a numerical method such as the FEM can be applied in a first tentative of studying equations such as Eq. (3.91). Even if a theoretical justification for its resolution is not provided, we may been able to study this equation in a qualitative way through the behaviour of empirical try-outs using the FEM.

Inference

In this dissertation we have not worked on inference methods for the models developed here or for more general contexts within the SPDE approach. We give, however, some indices that may be interesting in this aim. Some typical statistical inference methods such as likelihood based methods and Bayesian techniques are often based on a suitable manipulation of the precision matrices involved. It is in this point where the SPDE approach may provide a particular contribution. The conception of inference methods can be

accomplished in this framework following for example Lindgren et al. (2011), where the knowledge of the precision matrices associated to evaluations of the Random Field over the nodes in the discretization mesh is exploited. The general principle is the following: consider the SPDE, discretize it using the FEM, and work with the matrix associated to the FEM discretization to construct the precision matrix, regardless of the covariance matrix itself or to the explicit covariance function. The parameters to be estimated are the parameters defining the discretization matrix, in other words, the parameters of the SPDE. The advantages of such an approach depends on the SPDE considered and on the discretization or numerical method used to solve the SPDE selected.

Let us consider, for example, the simulation method presented in Chapter 6. When solving a SPDE such as in section 6.2.2, we have been able to determine properly the covariance function of the approximation U_N, given by (6.24). With this covariance function, we are actually able to provide an explicit expression for the covariance matrix of a random vector of the form $(U_N(x_k))_{k\in\{1,...,M\}}$, where $(x_1, ..., x_M)$ is an arbitrary finite collection of points in the working domain. Then, it is immediate that the covariance matrix, say Σ, can be factorised as $\Sigma = VD\overline{V^T}$, with two matrices $V \in \mathbb{C}^{M\times N}$ and $D \in \mathbb{C}^{N\times N}$ given by

$$V = \begin{bmatrix} e^{ix_1^T\xi_1^N} & e^{ix_1^T\xi_2^N} & \dots & e^{ix_1^T\xi_N^N} \\ e^{ix_2^T\xi_1^N} & e^{ix_2^T\xi_2^N} & \dots & e^{ix_2^T\xi_N^N} \\ \vdots & \vdots & \ddots & \vdots \\ e^{ix_M^T\xi_1^N} & e^{ix_M^T\xi_2^N} & \dots & e^{ix_M^T\xi_N^N} \end{bmatrix}, \quad D = \frac{1}{(2\pi)^d} \begin{bmatrix} \frac{1}{|g(\xi_1^N)|^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{|g(\xi_2^N)|^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \vdots & \frac{1}{|g(\xi_N^N)|^2} \end{bmatrix}. \tag{7.7}$$

We remark that the matrix V is neither square nor unitary in general, and it is not immediate to interpret the factorisation of Σ as a typical eigenvalue decomposition. However, this factorisation still provides an immediate expression for the covariance matrix which may be practical for some techniques. We have no general expression for the , which may not exist in some cases: we claim that the covariance function (6.24) is not strictly positive definite. It is not clear neither in which cases the precision matrix, when it exists, is sparse. We remark that when both the evaluation points $(x_1, ..., x_M)$ and the tag points $(\xi_j^N)_{j\in\{1,...,N\},N\in\mathbb{N}_*}$ are on a regular grid, the matrices are intimately related to classical Fourier matrices which appear in the analysis of Discrete Fourier Transforms. Hence, it is expected that a typical analysis technique from such a framework can be applied within our context. For instance, when $M = N$ and the regular grids are set conveniently, the matrices involved are exactly those of typical Discrete Fourier Transforms, and the factorisation $\Sigma = VD\overline{V^T}$ can be interpreted as an eigenvalue decomposition. The precision matrix is given simply by $VD^{-1}\overline{V^T}$ up to a multiplicative normalising constant. It is then possible to conceive inference methods based on Fourier Analysis methods, which are now interpreted within the SPDE approach, taking advantage of a convenient expression of the precision matrices.

We expect that inference will be one of the most important research directions within the SPDE approach

in a near future. After all, the SPDE approach may also inspire a new methodological paradigm to select geostatistical models for determined situations: under some phenomena theoretically submitted to physically driven PDEs, the geostatistician may now be *guided* by the physical consideration and select covariance models where the parameters carry some physical meaning. Then, inference techniques adapted to these kinds of models will be applied. Hence, the physical knowledge of the phenomena may now help the geostatistical practice. This is maybe a somewhat *naive dream* about how things will work in practice[1], but still deserves attention and try-outs within the future of Geostatistics.

[1] Reality almost never works as we want it to do...

Appendix

Appendix A

Proofs

A.1 Proof of Proposition 2.1.1

The necessity is an immediate conclusion from Rudin (1987, Theorem 6.4), from which we conclude that $|\mu|(\mathbb{R}^d) < \infty$. We prove the sufficiency. Let $\mu \in \mathscr{M}(\mathbb{R}^d)$. μ is by definition a function from $\mathcal{B}_B(\mathbb{R}^d)$ to $\mathbb{C}$ and we need to prove that we can extend its domain to $\mathcal{B}(\mathbb{R}^d)$ obtaining finite complex values and maintaining the $\sigma-$additivity property. We first prove that we can define $\mu(A)$ for all $A \in \mathcal{B}(\mathbb{R}^d)$. Let $A \in \mathcal{B}(\mathbb{R}^d)$. A particular property of the Euclidean space is that A can be partitioned in a countable family of mutually disjoint bounded Borel subsets: $A = \bigcup_{n\in\mathbb{N}} B_n$, $(B_n)_{n\in\mathbb{N}} \subset \mathcal{B}_B(\mathbb{R}^d)$, $B_n \cap B_m = \varnothing$ if $m \neq n$. As $|\mu| \in \mathscr{M}_F^+(\mathbb{R}^d)$, we consider that

$$\left|\sum_{n\in\mathbb{N}} \mu(B_n)\right| \leqslant \sum_{n\in\mathbb{N}} |\mu(B_n)| \leqslant \sum_{n\in\mathbb{N}} |\mu|(B_n) = |\mu|(A) < \infty. \tag{A.1}$$

Thus the series $\sum_{n\in\mathbb{N}} \mu(B_n)$ is absolutely convergent. Using the fact that a finite union of bounded set is also bounded, we define the complex number $b_A = \lim_{N\to\infty} \sum_{n\leqslant N} \mu(B_n) = \lim_{N\to\infty} \mu\left(\bigcup_{n\leqslant N} B_n\right)$. Let us prove that this limit does not depend on the collection of bounded Borel sets used as partition of A. Let $(C_n)_{n\in\mathbb{N}} \subset \mathcal{B}_B(\mathbb{R}^d)$ be a collection of mutually disjoint bounded Borel sets such that $A = \bigcup_{n\in\mathbb{N}} C_n$. By the same previous arguments, the series $\sum_{n\in\mathbb{N}} \mu(C_n)$ is absolutely convergent to a limit which is a complex number c_A. By triangular inequality one obtains

$$|b_A - c_A| \leqslant \left|b_A - \mu\left(\bigcup_{n\leqslant N} B_n\right)\right| + \left|\mu\left(\bigcup_{n\leqslant N} B_n\right) - \mu\left(\bigcup_{m\leqslant N} C_m\right)\right| + \left|c_A - \mu\left(\bigcup_{m\leqslant N} C_m\right)\right|. \tag{A.2}$$

Both terms $\left|b_A - \mu\left(\bigcup_{n\leqslant N} B_n\right)\right|$ and $\left|c_A - \mu\left(\bigcup_{m\leqslant N} C_m\right)\right|$ converge to 0 as N grows. Considering that $\bigcup_{n\leqslant N} B_n = \bigcup_{n\leqslant N, m\in\mathbb{N}} B_n \cap C_m$, we argue that $\mu(\bigcup_{n\leqslant N, m\in\mathbb{N}} B_n \cap C_m) = \sum_{n\leqslant N, m\in\mathbb{N}} \mu(B_n \cap C_m)$ since all the sets involved are bounded and their reunion too. Using the same argument as in Eq. (A.1), we can argue that

$$\sum_{n\leqslant N, m\in\mathbb{N}} \mu(B_n \cap C_m) + \sum_{n>N, m\in\mathbb{N}} \mu(B_n \cap C_m) = \sum_{n,m\in\mathbb{N}} \mu(B_n \cap C_m), \tag{A.3}$$

with all the series in (A.3) being absolutely convergent. Applying the same procedure to $\bigcup_{m\leqslant N} C_n = \bigcup_{n\in\mathbb{N}, m\leqslant N} B_n \cap C_m$ we obtain

$$\begin{aligned}
\left|\mu\left(\bigcup_{n\leqslant N} B_n\right) - \mu\left(\bigcup_{m\leqslant N} C_m\right)\right| &= \Bigg| \sum_{n,m\in\mathbb{N}} \mu(B_n \cap C_m) - \sum_{n>N, m\in\mathbb{N}} \mu(B_n \cap C_m) \\
&\quad - \sum_{n,m\in\mathbb{N}} \mu(B_n \cap C_m) + \sum_{n\in\mathbb{N}, m>N} \mu(B_n \cap C_m)\Bigg| \\
&\leqslant \left|\sum_{n>N, m\in\mathbb{N}} \mu(B_n \cap C_m)\right| + \left|\sum_{n\in\mathbb{N}, m>N} \mu(B_n \cap C_m)\right| \\
&\leqslant \underbrace{|\mu|\left(\bigcup_{n>N} B_n\right) + |\mu|\left(\bigcup_{m>N} C_m\right)}_{\to 0 \text{ as } N\to\infty \text{ since } |\mu|\in\mathscr{M}_F^+(\mathbb{R}^d)}.
\end{aligned} \tag{A.4}$$

We conclude that $|b_A - c_A|$ can be bounded by any arbitrarily small positive number and thus $b_A = c_A$, and thus $\sum_{n\in\mathbb{N}} \mu(B_n)$ converges to the same limit regardless of the partition on bounded Borel sets selected. We call then $\mu(A) := b_A \in \mathbb{C}$, and we have thus extended the domain of μ to all $\mathcal{B}(\mathbb{R}^d)$. The $\sigma-$additivity property still holds if the partition of the set A is made of bounded Borel sets. We still need to prove that it holds for an arbitrary countable partition in Borel sets $(A_n)_{n\in\mathbb{N}} \subset \mathcal{B}(\mathbb{R}^d)$, but this follows immediately: the series $\sum_{n\in\mathbb{N}} \mu(A_n)$ is absolutely convergent, which can be seen by applying the same argument as in Eq. (A.1), considering that the terms $\mu(A_n)$ for unbounded sets A_n are now well-defined. Every set A_n can be expressed as a union of mutually disjoint bounded Borel sets, for example of the form $A_n = \bigcup_{m\in\mathbb{N}} A_n \cap B_m$. From this we have

$$\sum_{n\in\mathbb{N}} \mu(A_n) = \sum_{n\in\mathbb{N}} \mu\left(\bigcup_{m\in\mathbb{N}} A_n \cap B_m\right) = \sum_{n\in\mathbb{N}} \sum_{m\in\mathbb{N}} \mu(A_n \cap B_m) = \mu(A). \tag{A.5}$$

Here the last equality is justified as the collection $(A_n \cap B_m)_{n,m\in\mathbb{N}}$ is a countable and mutually disjoint collection of bounded Borel sets whose reunion is A. This proves the $\sigma-$additivity condition and thus that μ is a well-defined finite measure. ■

A.2 Proofs of the variants of the Riesz Representation Theorem

Before proving both theorems let us describe the continuous linear functionals over $C(\mathbb{R}^d)$ and $C_{FD}(\mathbb{R}^d)$. Both are complete metric spaces, but from a more general point of view, they are locally convex topological vector spaces. Hence, their topologies are described by families of semi-norms (see Appendix D). Continuous linear functionals and then described by the corresponding families of semi-norms, following criteria D.0.1. Let us apply this principle to the cases of $C(\mathbb{R}^d)$ and $C_{FD}(\mathbb{R}^d)$.

In the case of the space $C(\mathbb{R}^d)$, a family of directed semi-norms defining the topology can be the norms $\|\cdot\|_{K,\infty}$ for every compact set $K \subset \mathbb{R}^d$.[1] We conclude then that a linear functional $T : C(\mathbb{R}^d) \to \mathbb{C}$ is continuous if and only if there exists $C > 0$ and there exists a compact set $K \subset \mathbb{R}^d$ such that

$$|\langle T, \varphi\rangle| \leqslant C\|\varphi\|_{\infty,K}, \quad \forall \varphi \in C(\mathbb{R}^d). \tag{A.6}$$

In the case of the space $C_{FD}(\mathbb{R}^d)$, a directed family of semi-norms inducing its topology is the family $(\|(1+|x|^2)^N(\cdot)\|_\infty)_{N\in\mathbb{N}}$. We conclude that a linear functional $T : C_{FD}(\mathbb{R}^d) \to \mathbb{C}$ is continuous if and only if there exists $C > 0$ and there exists $N \in \mathbb{N}$ such that

$$|\langle T, \varphi\rangle| \leqslant C\|(1+|x|^2)^N\varphi\|_\infty, \quad \forall \varphi \in C_{FD}(\mathbb{R}^d). \tag{A.7}$$

We are going to use the next three Lemmas, which are typical results from Analysis and Measure Theory.

Lemma A.2.1. *$C_c(\mathbb{R}^d)$ is a dense subspace of both $C_0(\mathbb{R}^d)$ and $C(\mathbb{R}^d)$ with their respective topologies.*

Proof: Consider a function $\varphi \in C(\mathbb{R}^d)$. For every $n \in \mathbb{N}$, consider a continuous function $\phi_n : \mathbb{R}^d \to [0,1]$ such that $\phi_n = 1$ over $\overline{B_n(0)}$ and $\phi_n = 0$ over $B_{n+1}(0)^c$, which can be constructed thanks to Urysohn's Lemma. Then, the sequence of functions $(\varphi_n)_{n\in\mathbb{N}}$ defined through $\varphi_n = \phi_n\varphi$ is in $C_c(\mathbb{R}^d)$ and it is clear that $\varphi_n \xrightarrow{C} \varphi$ since for every compact set $K \subset \mathbb{R}^d$ there exists a large enough $n_0 \in \mathbb{N}$ such that $\varphi = \phi_n\varphi$ over K for all $n \geqslant n_0$. If $\varphi \in C_0(\mathbb{R}^d)$, then for every $\epsilon > 0$ there exists a large enough $m \in \mathbb{N}$ for which $\|\varphi\|_{\infty,B_m(0)^c} < \epsilon$. For $n \geqslant m$, $\varphi_n = \varphi$ over $\overline{B_m}$, and then $\|\varphi - \varphi_n\|_\infty = \|(\phi_n - 1)\varphi\|_{\infty,B_m(0)^c} \leqslant \|\varphi\|_{\infty,B_m(0)^c} < \epsilon$. This proves that $\varphi_n \xrightarrow{C_0} \varphi$. ■

Lemma A.2.2. *Let $O \subset \mathbb{R}^d$ be an open set. Then, there exists a sequence of positive functions $(\varphi_n)_{n\in\mathbb{N}} \subset C_c(\mathbb{R}^d)$ such that $(\varphi_n)_{n\in\mathbb{N}}$ converges monotonically increasing point-wise to $\mathbf{1}_O$.*

Proof: We use the fact that every open set of the Euclidean space can be expressed as a union of a countable quantity of bounded open rectangles. This fact can be proven using the density of the rational

[1] We could have taken a countable family of these norms by considering only the compact sets of the form $\overline{B_N(0)}$ with $N \in \mathbb{N}$ as in metric (2.26), but this actually will not help us very much.

numbers. If O is an open set, we consider a countable collection of bounded open rectangles $(R_k)_{k\in\mathbb{N}}$ such that $O = \bigcup_{k\in\mathbb{N}} R_k$. We write these rectangles as

$$R_k = (a_1^k, b_1^k) \times \cdots \times (a_d^k, b_d^k), \quad k \in \mathbb{N}, \tag{A.8}$$

having of course $-\infty < a_j^k < b_j^k < \infty$ for all $j = 1, ..., d$ and for all $k \in \mathbb{N}$. For each $n \in \mathbb{N}$, $k \in \mathbb{N}$ and $j \in \{1, ..., d\}$, we consider the piece-wise linear function $\varphi_n^{(j,k)} : \mathbb{R} \to [0, 1]$ which takes the value 1 over the closed interval $[a_j^k + \frac{b_j^k - a_j^k}{2(n+1)}\ ,\ b_j^k - \frac{b_j^k - a_j^k}{2(n+1)}]$ and the value 0 outside the open interval $(a_j^k + \frac{b_j^k - a_j^k}{4(n+1)}\ ,\ b_j^k - \frac{b_j^k - a_j^k}{4(n+1)})$, its graph forming a trapezium for $n \geqslant 1$. For every $k \in \mathbb{N}$ and $n \in \mathbb{N}$ we define $\varphi_n^{(k)} : \mathbb{R}^d \to [0, 1]$ as the tensor product $\varphi_n^{(k)} = \varphi_n^{(1,k)} \otimes ... \otimes \varphi_n^{(d,k)}$, that is, the function determined by $\varphi_n^{(k)}(x) = \prod_{j=1}^d \varphi_n^{(j,k)}(x_j)$ for every $x = (x_1, ..., x_d) \in \mathbb{R}^d$. Finally, we define for each $n \in \mathbb{N}$ the function $\varphi_n : \mathbb{R}^d \to [0, 1]$ as $\varphi_n = \max_{k\leqslant n} \varphi_n^{(k)}$. It follows then that the sequence $(\varphi_n)_{n\in\mathbb{N}}$ is in $C_c(\mathbb{R}^d)$ since the maximum of a finite quantity of continuous function with compact support is continuous with compact support. We also have that $\mathrm{supp}(\varphi_n) \subset O$ for all n and that $0 \leqslant \varphi_n \leqslant \varphi_{n+1}$, so the sequence is monotonically increasing. Finally if $x \in O$, there is a rectangle R_k such that $x \in R_k$ and since R_k is open, it is clear that for a large enough $n_0 \in \mathbb{N}$ it will holds that $\varphi_n(x) = 1$ for every $n \geqslant n_0$. Since $\mathrm{supp}(\varphi_n) \subset O$, it follows that $\varphi_n(x) = 0$ for all $x \in O^c$ and for all $n \in \mathbb{N}$. This proves that $\varphi_n \to \mathbf{1}_O$ monotonically increasing and point-wise. ■

Lemma A.2.3. *Let $\mu \in \mathscr{M}(\mathbb{R}^d)$ and let $O \subset \mathbb{R}^d$ be an open set. Suppose that for every $\varphi \in C_c(\mathbb{R}^d)$ such that* $\mathrm{supp}(\varphi) \subset O$*, we have that $\int_{\mathbb{R}^d} \varphi(x) d\mu(x) = 0$. Then,* $\mathrm{supp}(\mu) \subset O^c$.

Proof: Let $\mu \in \mathscr{M}(\mathbb{R}^d)$ and $O \subset \mathbb{R}^d$ an open set with the specified condition. For simplicity we will first suppose that O is bounded. Consider the decomposition of μ in four positive measures $\mu = \mu_R^+ - \mu_R^- + i(\mu_I^+ - \mu_I^-)$. Let $A \subset O$ be any open subset of O. Consider a monotonically increasing sequence of functions $(\varphi_n)_{n\in\mathbb{N}} \subset C_c(\mathbb{R}^d)$ approaching $\mathbf{1}_A$ as in Lemma A.2.2. We have $\mathrm{supp}(\varphi_n) \subset A \subset O$, and hence $\int_{\mathbb{R}^d} \varphi_n d\mu(x) = 0$ for all $n \in \mathbb{N}$. This implies that

$$\int_{\mathbb{R}^d} \varphi_n d\mu_R^+ = \int_{\mathbb{R}^d} \varphi_n d\mu_R^- \quad ; \quad \int_{\mathbb{R}^d} \varphi_n d\mu_I^+ = \int_{\mathbb{R}^d} \varphi_n d\mu_I^-, \quad \forall n \in \mathbb{N}. \tag{A.9}$$

Using the Monotone Convergence Theorem 2.11, we take the limit when $n \to \infty$ to obtain

$$\mu_R^+(A) = \mu_R^-(A) \quad ; \quad \mu_I^+(A) = \mu_I^-(A), \quad \forall A \subset O, \text{ open.} \tag{A.10}$$

Thus, the positive measures μ_R^+ and μ_I^+ coincide respectively with the positive measures μ_R^- and μ_I^- over every open subset of O. Since the collection of open subsets of O is the collection which engenders the $\sigma-$algebra of Borel subsets of O, denoted by $\mathcal{B}(O)$, a typical result in Measure Theory (see for example Williams, 1990, Lemma 1.6, applicable since $\mu(O) < \infty$) guarantees that (A.10) holds for every $A \in \mathcal{B}(O)$. This proves that $\mu(A) = \mu_R^+(A) - \mu_R^-(A) + i(\mu_I^+(A) - \mu_I^-(A)) = 0$ for all $A \in \mathcal{B}(\mathbb{R}^d)$ such that $A \subset O$.

By definition of the total variation measure (see Eq. (2.2)), it follows that $|\mu|(O) = 0$. Since $\mathrm{supp}(\mu)$ is the complementary of the largest open set where $|\mu|$ is null, it follows that $\mathrm{supp}\,\mu \subset O^c$.

If O is an unbounded open set, it can be expressed as a countable union of bounded open sets, $O = \bigcup_{n\in\mathbb{N}} A_n$. Applying the previous argument for the sets $(A_n)_{n\in\mathbb{N}}$, it follows that

$$|\mu|(O) \leqslant \sum_{n\in\mathbb{N}} |\mu|(A_n) = 0, \tag{A.11}$$

from which it follows that $\mathrm{supp}(\mu) \subset O^c$. ■

A.2.1 Proof of Theorem 2.1.5

If $\mu \in \mathscr{M}_c(\mathbb{R}^d)$, it follows immediately from Eq. (2.27) that the integral with respect to μ defines a continuous linear functional on $C(\mathbb{R}^d)$ by setting $C = |\mu|(\mathrm{supp}(\mu))$ in (A.6).

We prove now the converse. Suppose $T : C(\mathbb{R}^d) \to \mathbb{C}$ is linear and continuous. Let $K \subset \mathbb{R}^d$ be a compact set and let $C > 0$ be such that $|\langle T, \varphi\rangle| \leqslant C\|\varphi\|_{\infty,K}$ for all $\varphi \in C(\mathbb{R}^d)$. In particular, this holds for every $\varphi \in C_0(\mathbb{R}^d)$, for which we also have $|\langle T, \varphi\rangle| \leqslant C\|\varphi\|_{\infty,K} \leqslant C\|\varphi\|_{\infty}$. This proves that T is also a continuous linear functional over $C_0(\mathbb{R}^d)$. By the Riesz Representation Theorem for finite measures 2.1.4, there exists a unique finite measure $\mu \in \mathscr{M}_F(\mathbb{R}^d)$ such that (2.25) holds. Consider now any function $\varphi \in C_c(\mathbb{R}^d)$ such that $\mathrm{supp}(\varphi) \subset K^c$. From Eq. (A.6) it follows that $\langle T, \varphi\rangle = \int_{\mathbb{R}^d} \varphi(x)d\mu(x) = 0$. Since K^c is open, we obtain from Lemma A.2.3 that $\mathrm{supp}(\mu) \subset K$. Hence, μ has compact support. Since $\langle T, \varphi\rangle = \int_{\mathbb{R}^d} \varphi(x)d\mu(x)$ holds for all $\varphi \in C_c(\mathbb{R}^d)$, the integral with respect to μ coincides with T as a continuous linear functional in a dense subspace of C (Lemma A.2.1). It follows that $\langle T, \varphi\rangle = \int_{\mathbb{R}^d} \varphi(x)d\mu(x)$ for all $\varphi \in C(\mathbb{R}^d)$. ■

A.2.2 Proof of Theorem 2.1.6

Proof: Let $\mu \in \mathscr{M}_{SG}(\mathbb{R}^d)$. From equation (2.30) it follows immediately that the integral with respect to μ defines a continuous linear functional on $C_{FD}(\mathbb{R}^d)$. Indeed, set $N \in \mathbb{N}$ such that $(1 + |x|^2)^{-N}|\mu|$ is finite and $C = \left|(1 + |x|^2)^{-N}\mu\right|(\mathbb{R}^d)$ in (A.7).

Let us prove the converse. Let $T \in C'_{FD}(\mathbb{R}^d)$ and let $C > 0$ and $N \in \mathbb{N}$ such that (A.7) holds. Let us define the linear functional $(1 + |x|^2)^{-N}T : C_{FD}(\mathbb{R}^d) \to \mathbb{C}$ by

$$(1 + |x|^2)^{-N}T(\varphi) := \langle T, (1 + |x|^2)^{-N}\varphi\rangle. \tag{A.12}$$

Since for all $\varphi \in C_{FD}(\mathbb{R}^d)$ the function $(1 + |x|^2)^{-N}\varphi$ is also in $C_{FD}(\mathbb{R}^d)$, this functional is well-defined.

In addition,

$$|(1+|x|^2)^{-N}T(\varphi)| = |\langle T,(1+|x|^2)^{-N}\varphi\rangle| \leqslant C\|(1+|x|^2)^N(1+|x|^2)^{-N}\varphi\|_\infty = C\|\varphi\|_\infty, \quad \text{(A.13)}$$

hence, T is continuous. Expression (A.13) holds in particular for every $\varphi \in C_c(\mathbb{R}^d) \subset C_{FD}(\mathbb{R}^d)$. Hence $(1+|x|^2)^{-N}T$ is a bounded linear functional in the sense of the supremum norm on $C_c(\mathbb{R}^d)$. By Hahn-Banach extension Theorem (Reed & Simon, 1980, Theorem III.5 or Theorem V.3), $(1+|x|^2)^{-N}T$ can be extended to a continuous linear functional over $C_0(\mathbb{R}^d)$, and since $C_c(\mathbb{R}^d)$ is dense in $C_0(\mathbb{R}^d)$ by Lemma A.2.1, the extension is unique and Eq. (A.13) holds for every $\varphi \in C_0(\mathbb{R}^d)$. By Riesz Representation Theorem for finite measures 2.1.4, we conclude that $(1+|x|^2)^{-N}T$ is identified with a unique finite measure $\nu \in \mathscr{M}_F(\mathbb{R})$. Consider then the multiplication measure $\mu = (1+|x|^2)^N\nu$, which is in $\mathscr{M}_{SG}(\mathbb{R}^d)$. We conclude that for every $\varphi \in C_{FD}(\mathbb{R}^d)$ we have

$$\langle T,\varphi\rangle = \langle T,\frac{(1+|x|^2)^N}{(1+|x|^2)^N}\varphi\rangle = \langle (1+|x|^2)^{-N}T,(1+|x|^2)^N\varphi\rangle = \int_{\mathbb{R}^d}\varphi(x)(1+|x|^2)^N d\nu(x) = \int_{\mathbb{R}^d}\varphi(x)d\mu(x). \quad \text{(A.14)}$$

This completes the proof. ■

A.3 Proofs of Propositions 2.1.2 and 2.1.3

A.3.1 Proof of Proposition 2.1.2

Since for all the cases for $j \in \{\text{“}c\text{”},\text{“}F\text{”},\text{“}SG\text{”}\}$, $\mu \in \mathscr{M}_j(\mathbb{R}^d)$ if and only if $|\mu| \in \mathscr{M}_j^+(\mathbb{R}^d)$, it is sufficient to prove this claim for positive measures. The case $j =$ “c” is straightforward from (2.39) since if $\mathrm{supp}(\nu) = K$ it follows that $\mathrm{supp}(\mu\delta^{\{y=x\}}) \subset K \times K$. Conversely, if $\mathrm{supp}(\mu\delta^{\{y=x\}}) = K_2 \subset \mathbb{R}^d \times \mathbb{R}^d$, K_2 compact, then there exists a large enough compact set $K \subset \mathbb{R}^d$ such that $K_2 \subset K \times K$, for which we obtain that $\mathrm{supp}(\mu) \subset K$. When $j =$ “F” it is also straightforward from (2.39) since $\mu(\mathbb{R}^d) = \mu\delta^{\{y=x\}}(\mathbb{R}^d \times \mathbb{R}^d)$. Finally, for $j =$ “SG”, from Eq. (2.38) we have that for every $N \in \mathbb{N}$,

$$\int_{\mathbb{R}^d\times\mathbb{R}^d}\frac{d(\mu\delta^{\{y=x\}})(x,y)}{(1+|x|^2+|y|^2)^N} = \int_{\mathbb{R}^d}\frac{d\mu(x)}{(1+2|x|^2)^N}. \quad \text{(A.15)}$$

It follows that if one of the two integrals is finite for some $N \in \mathbb{N}$ then the other is finite. Hence, we can find a strictly positive polynomial such that the multiplication between its reciprocal and the corresponding measure is a finite measure. ■

A.3.2 Proof of Proposition 2.1.3

As in Proposition 2.1.2, we prove this just for positive measures. The case $j =$ "F" follows immediately from (2.44) since $\mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}(\mathbb{R}^d \times \mathbb{R}) = \mu_1(\mathbb{R}^d) + \mu_2(\mathbb{R}^d \backslash \{0\})$. We remark that $\mu_2(\{0\}) < \infty$ since we suppose $\mu_2 \in \mathscr{M}(\mathbb{R}^d)$.

For $j =$ "c", if $\mathrm{supp}(\mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}) \subset K_d \times K_1$, with $K_d \subset \mathbb{R}^d$ and $K_1 \subset \mathbb{R}$ compact, then it follows that

$$\mu_1(K_d^c) = \mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}(K_d^c \times \mathbb{R}^+) = 0 = \mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}(K_d^c \times \mathbb{R}^-_*) = \mu_2(K_d^c), \tag{A.16}$$

which proves that $\mathrm{supp}(\mu_1) \cup \mathrm{supp}(\mu_2) \subset K_d$. Conversely, suppose $\mathrm{supp}(\mu_1) \cup \mathrm{supp}(\mu_2) \subset K_d$ for some $K_d \subset \mathbb{R}^d$ compact. This implies that both μ_1 and μ_2 are finite so $\mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}$ is finite. Consider the compact subset of $\mathbb{R}$, $K_1 = \{a|x| \in \mathbb{R} \;\big|\; -c \leqslant a \leqslant c,\ x \in K_d\}$. Let us evaluate $\mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}((K_d \times K_1)^c)$. Using the $\sigma-$additivity and the finiteness of $\mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}((K_d \times K_1)^c)$, we obtain

$$\mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}((K_d \times K_1)^c) = \mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}(K_d^c \times K_1^c) + \mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}(K_d^c \times K_1) + \mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}(K_d \times K_1^c). \tag{A.17}$$

Since $\mathrm{supp}(\mu_1) \cup \mathrm{supp}(\mu_2) \subset K_d$, it is immediate from Eq. (2.44) that $\mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}(K_d^c \times K_1^c) = \mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}(K_d^c \times K_1) = 0$. From Eq. (2.44) we also obtain

$$\mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}(K_d \times K_1^c) = \int_{K_d} \delta_{c|x|}(K_1^c) d\mu_1(x) + \int_{K_d \backslash \{0\}} \delta_{-c|x|}(K_1^c) d\mu_2(x). \tag{A.18}$$

Since for $x \in K_d$ we have $\pm c|x| \in K_1$, both expressions $\delta_{c|x|}(K_1^c)$ and $\delta_{-c|x|}(K_1^c)$ are null, and hence $\mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}(K_d \times K_1^c) = 0$. We conclude that $\mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}((K_d \times K_1)^c) = 0$ and therefore $\mathrm{supp}(\mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}) \subset K_d \times K_1$, so $\mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)} \in \mathscr{M}_c(\mathbb{R}^d \times \mathbb{R})$.

Finally, when $j =$ "SG", we consider that for all $N \in \mathbb{N}$ it holds that

$$\int_{\mathbb{R}^d \times \mathbb{R}} \frac{d\mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}(x,t)}{(1 + |x|^2 + |t|^2)^N} = \int_{\mathbb{R}^d} \frac{d\mu_1(x)}{(1 + (1+c^2)|x|^2)^N} + \int_{\mathbb{R}^d \backslash \{0\}} \frac{d\mu_2(x)}{(1 + (1+c^2)|x|^2)^N}. \tag{A.19}$$

Hence, if N is such that the integral on the left side of (A.19) is finite, the multiplication between the reciprocal of the polynomial $x \mapsto (1 + (1+c^2)|x|^2)^N$ and μ_1 is a finite measure, as well as its multiplication with μ_2. Hence both measures are slow-growing. Conversely, supposing that μ_1 and μ_2 are slow-growing, we can find a number $N \in \mathbb{N}$ such that both integrals on the right side of (A.19) are finite, and hence $\mu^{\mathcal{C}^c}_{(\mu_1,\mu_2)}$ is slow-growing. ■

A.4 Proof of Propositions about tensor products (Section 2.2.3)

A.4.1 Proof of Proposition 2.2.1

We consider the next Lemma. It can be found in the context of complex finite measures over abstract measure spaces in Rudin (1987, Theorem 6.12).

Lemma A.4.1. *Let $\mu \in \mathscr{M}(\mathbb{R}^d)$. Then, there exists a complex measurable function $f_\mu : \mathbb{R}^d \to \mathbb{C}$ such that $|f_\mu(x)| = 1$ for all $x \in \mathbb{R}^d$ and such that $\mu = f_\mu|\mu|$.*

Proof: See Rudin (1987, Theorem 6.12) for the case of finite measures. The case of not finite measures is straightforward by restricting the analysis to a collection of disjoint bounded Borel sets whose union is the whole space $\mathbb{R}^d$ and defining the function f_μ as the sum of the corresponding functions restricted to the corresponding bounded Borel sets. ■

Proof of Proposition 2.2.1:

Let $\mu \in \mathscr{M}(\mathbb{R}^d)$ and $\nu \in \mathscr{M}(\mathbb{R}^m)$. Let $f_\mu : \mathbb{R}^d \to \mathbb{C}$ and $f_\nu : \mathbb{R}^m \to \mathbb{C}$ be two measurable complex functions obtained as in Lemma A.4.1. Consider $O \in \mathcal{B}_B(\mathbb{R}^d \times \mathbb{R}^m)$. By definition of the total variation measure, we have that

$$\begin{aligned}
|\mu \otimes \nu|(O) &= \sup\left\{ \sum_{n\in\mathbb{N}} |(\mu\otimes\nu)(O_n)| \;\middle|\; (O_n)_{n\in\mathbb{N}} \subset \mathcal{B}(\mathbb{R}^d\times\mathbb{R}^m) \text{ partition of } O \right\} \\
&= \sup\left\{ \sum_{n\in\mathbb{N}} \left| \int_{O_n} f_\mu(x) f_{\nu(y)} d|\mu|(x) d|\nu|(y)\right| \;\middle|\; (O_n)_{n\in\mathbb{N}} \subset \mathcal{B}(\mathbb{R}^d\times\mathbb{R}^m) \text{ partition of } O \right\} \\
&\leqslant \sup\left\{ \sum_{n\in\mathbb{N}} (|\mu|\otimes|\nu|)(O_n) \;\middle|\; (O_n)_{n\in\mathbb{N}} \subset \mathcal{B}(\mathbb{R}^d\times\mathbb{R}^m) \text{ partition of } O \right\} \\
&= (|\mu|\otimes|\nu|)(O).
\end{aligned} \tag{A.20}$$

Let us now consider the case where O is of the form $O = A \times B$, with $A \in \mathcal{B}_B(\mathbb{R}^d)$ and $B \in \mathcal{B}_B(\mathbb{R}^m)$. Since the definition of the total variation measure over a set uses the supremum over all partitions of the set, it holds in particular that

$$(|\mu|\otimes|\nu|)(A\times B) \geqslant |\mu\otimes\nu|(A\times B) \geqslant \sup\left\{ \sum_{n,m\in\mathbb{N}} |\mu\otimes\nu(A_n\times B_m)| \right\} = \sup\left\{ \sum_{n\in\mathbb{N}} |\mu(A_n)| \sum_{m\in\mathbb{N}} |\nu(B_m)| \right\}, \tag{A.21}$$

where the supremum is taken over all the possible collections $(A_n \times B_m)_{n,m\in\mathbb{N}} \subset \mathcal{B}(\mathbb{R}^d \times \mathbb{R}^m)$ which satisfies that $(A_n) \subset \mathcal{B}(\mathbb{R}^d)$ is a partition of A, $(B_m)_{m\in\mathbb{N}} \subset \mathcal{B}(\mathbb{R}^m)$ is a partition of B. For every such a

partition of $A \times B$, it holds that

$$\left| |\mu|(A)|\nu|(B) - \sum_{n\in\mathbb{N}} |\mu(A_n)| \sum_{m\in\mathbb{N}} |\nu(B_m)| \right| \leqslant \underbrace{\left| |\mu|(A) - \sum_{n\in\mathbb{N}} |\mu(A_n)| \right|}_{\text{arbitrarily small}} \underbrace{\sum_{m\in\mathbb{N}} |\nu(B_m)|}_{\leqslant |\nu|(B)} + \underbrace{\left| |\nu|(B) - \sum_{m\in\mathbb{N}} |\nu(B_m)| \right|}_{\text{arbitrarily small}} |\mu|(A). \tag{A.22}$$

Hence, we can find such a partition $(A_n \times B_m)_{n,m\in\mathbb{N}}$ which approaches $(|\mu| \otimes |\nu|)(A \times B)$. This proves that the inequalities in Eq. (A.21) are, in fact, equalities. Hence

$$(|\mu| \otimes |\nu|)(A \times B) = |\mu \otimes \nu|(A \times B) \quad \forall A \in \mathcal{B}_B(\mathbb{R}^d), B \in \mathcal{B}_B(\mathbb{R}^m). \tag{A.23}$$

The equality of both measures in the whole system of rectangles guarantees the equality for every set in $\mathcal{B}_B(\mathbb{R}^d \times \mathbb{R}^m)$. ■

A.4.2 Proof of Proposition 2.2.2

To establish this Proposition, we need a few Lemmas describing some basic behaviour of some objects in Distribution Theory.

Lemma A.4.2. *Let $\psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m)$. Let $(y_n)_{n\in\mathbb{N}} \subset \mathbb{R}^m$ be a sequence such that $y_n \to y_0 \in \mathbb{R}^m$. Then, $\psi(\cdot, y_n) \overset{\mathscr{S}(\mathbb{R}^d)}{\longrightarrow} \psi(\cdot, y_0)$.*

Proof: Let $\psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m)$. For $N \in \mathbb{N}$ and $\alpha \in \mathbb{N}^d$, let us consider the function $\Psi_{N,\alpha} : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{C}$ defined through $\Psi_{N,\alpha}(x, y) = (1 + |x|^2)^N D^{(\alpha,0_m)}\psi(x, y)$ for every $(x, y) \in \mathbb{R}^d \times \mathbb{R}^m$. Here $0_m = (0, ..., 0) \in \mathbb{N}^m$ denotes the multi-index with m null components and $(\alpha, 0_m) \in \mathbb{N}^d \times \mathbb{N}^m$ denotes the concatenation between α and 0_m. Clearly $\Psi_{N,\alpha} \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m)$. Let $(y_n)_{n\in\mathbb{N}} \subset \mathbb{R}^m$ such that $y_n \to y_0 \in \mathbb{R}^m$. We consider then that

$$\sup_{x\in\mathbb{R}^d} |\Psi_{N,\alpha}(x, y_n) - \Psi_{N,\alpha}(x, y_0)| \leqslant \| \, |\nabla\Psi_{N,\alpha}| \, \|_\infty |y_n - y_0| \to 0, \tag{A.24}$$

where $\nabla\Psi_{N,\alpha}$ denotes the gradient of the function $\Psi_{N,\alpha}$. Here $\| \, |\nabla\Psi_{N,\alpha}| \, \|_\infty = \sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^m} |\nabla\Psi_{N,\alpha}(x, y)|$ is a finite number since $\Psi_{N,\alpha} \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m)$ and hence all of its derivatives are bounded. Since N and α were arbitrary, this proves the convergence $\psi(\cdot, y_n) \overset{\mathscr{S}(\mathbb{R}^d)}{\longrightarrow} \psi(\cdot, y_0)$. ■

Lemma A.4.3. *Let $\varphi \in \mathscr{S}(\mathbb{R}^d)$. Let $(t_n)_{n\in\mathbb{N}} \subset \mathbb{R}_*$ be a sequence such that $t_n \to 0$ as $n \to \infty$. Let $e_j \in \mathbb{R}^d$ be the canonical vector in the direction $j \in \{1, ..., d\}$. Then, $\frac{\varphi(\cdot+t_n e_j)-\varphi(\cdot)}{t_n} \overset{\mathscr{S}(\mathbb{R}^d)}{\longrightarrow} \frac{\partial\varphi}{\partial x_j}$ as $n \to \infty$.*

Proof: Let $\varphi \in \mathscr{S}(\mathbb{R}^d)$ and let $\alpha \in \mathbb{N}^d$. Let $x \in \mathbb{R}^d$. We consider the Taylor's formula for $D^\alpha\varphi$ at x

with integral form of the reminder

$$D^\alpha\varphi(x+t_ne_j) = D^\alpha\varphi(x) + t_n\frac{\partial D^\alpha\varphi}{\partial x_j}(x) + t_n^2\int_0^1(1-t)\frac{\partial^2 D^\alpha\varphi}{\partial x_j^2}(x+t_ne_jt)dt. \tag{A.25}$$

Let $N\in\mathbb{N}$. We obtain thus

$$(1+|x|^2)^N\left(\frac{D^\alpha\varphi(x+t_ne_j)-D^\alpha\varphi}{t_n} - \frac{\partial D^\alpha\varphi}{\partial x_j}(x)\right) = t_n\int_0^1(1-t)(1+|x|^2)^N\frac{\partial^2 D^\alpha\varphi}{\partial x_j^2}(x+t_ne_jt)dt. \tag{A.26}$$

Using the convexity of the functions $x\in\mathbb{R}^d\mapsto|x|^2$ and $x\in\mathbb{R}\mapsto|x|^N$, one can verify that for every $x,y\in\mathbb{R}^d$ it holds that $(1+|x|^2)^N\leqslant 2^{N-1}(1+2|x-y|^2)^N+2^{2N-1}|y|^2$. Applying this idea with $y=-t_ne_jt$ in (A.26), it follows that the integral in this equation can be bounded in the following way:

$$\begin{aligned}\left|\int_0^1(1-t)(1+|x|^2)^N\frac{\partial^2 D^\alpha\varphi}{\partial x_j^2}(x+t_ne_jt)dt\right| &\leqslant 2^{N-1}\int_0^1(1-t)(1+2|x+t_ne_jt|^2)^N\left|\frac{\partial^2 D^\alpha\varphi}{\partial x_j^2}(x+t_ne_jt)\right|dt\\ &\quad+2^{2N-1}|t_n|^{2N}\int_0^1(1-t)|t|^{2N}\left|\frac{\partial^2 D^\alpha\varphi}{\partial x_j^2}(x+t_ne_jt)\right|dt\\ &\leqslant 2^{N-2}\underbrace{\sup_{x\in\mathbb{R}^d}\left|(1+2|x|^2)^N\frac{\partial^2 D^\alpha\varphi}{\partial x_j^2}(x)\right|}_{<\infty\text{ since }\frac{\partial^2 D^\alpha\varphi}{\partial x_j^2}\in\mathscr{S}(\mathbb{R}^d)}+2^{2N-1}|t_n|^{2N}\left\|\frac{\partial^2 D^\alpha\varphi}{\partial x_j^2}\right\|_\infty.\end{aligned} \tag{A.27}$$

It follows that

$$\begin{aligned}\sup_{x\in\mathbb{R}^d}\left|(1+|x|^2)^N\left(\frac{D^\alpha\varphi(x+t_ne_j)-D^\alpha\varphi(x)}{t_n}-\frac{\partial D^\alpha\varphi}{\partial x_j}(x)\right)\right| &\leqslant |t_n|2^{N-2}\left\|(1+2|x|^2)^N\frac{\partial^2 D^\alpha\varphi}{\partial x_j^2}\right\|_\infty\\ &\quad+|t_n|^{2N+1}2^{2N-1}\left\|\frac{\partial^2 D^\alpha\varphi}{\partial x_j^2}\right\|_\infty\\ &\to 0\quad\text{as } n\to\infty.\end{aligned} \tag{A.28}$$

Since $D^\alpha(\frac{\partial\varphi}{\partial x_j}) = \frac{\partial}{\partial x_j}(D^\alpha\varphi)$ and since $D^\alpha\left(\frac{\varphi(\cdot+t_ne_j)-\varphi(\cdot)}{t_n}\right) = \frac{D^\alpha\varphi(\cdot+t_ne_j)-D^\alpha\varphi(\cdot)}{t_n}$, this proves that

$$\frac{\varphi(\cdot+t_ne_j)-\varphi(\cdot)}{t_n}\xrightarrow{\mathscr{S}(\mathbb{R}^d)}\frac{\partial\varphi}{\partial x_j}.\quad\blacksquare$$

Proof of Proposition 2.2.2:

Let $\mathcal{L}_1:\mathscr{S}(\mathbb{R}^d)\to\mathscr{S}(\mathbb{R}^d)$ be linear and continuous. Let $\psi\in\mathscr{S}(\mathbb{R}^d\times\mathbb{R}^m)$. We need to verify that $(\mathcal{L}_1\otimes\mathcal{I}_m)(\psi)$ defined as in 2.2.8 is in $\mathscr{S}(\mathbb{R}^d\times\mathbb{R}^m)$ and that the operator defines a continuous mapping. Note that for any fixed $y\in\mathbb{R}^m$, the function $\psi(\cdot,y)$ is in $\mathscr{S}(\mathbb{R}^d)$. Hence, for every y the function $x\mapsto\mathcal{L}_1(\psi(\cdot,y))(x)$ is a well-defined function in $\mathscr{S}(\mathbb{R}^d)$. This proves that $\mathcal{L}_1\otimes\mathcal{I}_m$ is a well-defined mapping

from $\mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m)$ to the space $\mathbb{C}^{\mathbb{R}^d \times \mathbb{R}^m}$ of all complex functions defined over $\mathbb{R}^d \times \mathbb{R}^m$. It is straightforward that this mapping is linear. For simplicity, we will denote by $\Psi = (\mathcal{L}_1 \otimes \mathcal{I}_m)(\psi) \in \mathbb{C}^{\mathbb{R}^d \times \mathbb{R}^m}$.

We need to prove that Ψ is in $\mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m)$. We begin by proving that it is continuous. Let $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}^d$ and $(y_n)_{n\in\mathbb{N}} \subset \mathbb{R}^m$ be sequences such that $(x_n, y_n) \to (x_0, y_0) \in \mathbb{R}^d \times \mathbb{R}^m$. It holds then that

$$\begin{aligned} |\Psi(x_n, y_n) - \Psi(x_0, y_0)| &\leqslant |\Psi(x_n, y_n) - \Psi(x_n, y_0)| + |\Psi(x_n, y_0) - \Psi(x_0, y_0)| \\ &= |\mathcal{L}_1(\psi(\cdot, y_n) - \psi(\cdot, y_0))(x_n)| + |\mathcal{L}_1(\psi(\cdot, y_0))(x_n) - \mathcal{L}_1(\psi(\cdot, y_0))(x_0)| \\ &\leqslant \sup_{u\in\mathbb{R}^d} |\mathcal{L}_1(\psi(\cdot, y_n) - \psi(\cdot, y_0))(u)| + |\mathcal{L}_1(\psi(\cdot, y_0))(x_n) - \mathcal{L}_1(\psi(\cdot, y_0))(x_0)|. \end{aligned} \tag{A.29}$$

The term $\sup_{u\in\mathbb{R}^d} |\mathcal{L}_1(\psi(\cdot, y_n) - \psi(\cdot, y_0))(u)|$ goes to zero since by Lemma A.4.2 we have that $\psi(\cdot, y_n) \xrightarrow{\mathscr{S}(\mathbb{R}^d)} \psi(\cdot, y_0)$ and $\mathcal{L}_1$ is continuous. The term $|\mathcal{L}_1(\psi(\cdot, y_0)(x_n) - \mathcal{L}_1(\psi(\cdot, y_0))(x_0)|$ goes to zero since $x \mapsto \mathcal{L}_1(\psi(\cdot, y_0))(x) \in \mathscr{S}(\mathbb{R}^d)$ and hence it is continuous at x_0. This proves that Ψ is continuous in $\mathbb{R}^d \times \mathbb{R}^m$.

Let us now prove that Ψ is differentiable. The differentiability of Ψ with respect to the *first* components is immediate since $\Psi(\cdot, y) \in \mathscr{S}(\mathbb{R}^d)$ for every $y \in \mathbb{R}^m$. Hence, for every $\alpha \in \mathbb{N}^d$ the function $(x, y) \mapsto D^{(\alpha,0_m)}\Psi(x, y)$ is well-defined. Its continuity can be verified using the same arguments used to prove the continuity of Ψ. Let us now prove the differentiability with respect to the *second* components. Let $e_j \in \mathbb{R}^m$ be the canonical vector in direction $j \in \{1, ..., m\}$. Let $t \neq 0$ and consider the expression of the form $\frac{\Psi(x,y+te_j)-\Psi(x,y)}{t}$ for a fixed $(x, y) \in \mathbb{R}^d \times \mathbb{R}^m$. The linearity of $\mathcal{L}_1$ guarantees that

$$\frac{\Psi(x, y + te_j) - \Psi(x, y)}{t} = \frac{\mathcal{L}_1(\psi(\cdot, y + te_j))(x) - \mathcal{L}_1(\psi(\cdot, y))(x)}{t} = \mathcal{L}_1\left(\frac{\psi(\cdot, y + te_j) - \psi(\cdot, y)}{t}\right)(x). \tag{A.30}$$

A slightly different but valid interpretation of Lemma A.4.3 allows to conclude that $\frac{\psi(\cdot,y+te_j)-\psi(\cdot,y)}{t} \xrightarrow{\mathscr{S}(\mathbb{R}^d)} \frac{\partial\psi}{\partial y_j}(\cdot, y)$ for all $y \in \mathbb{R}^m$ as $t \to 0$. Since $\mathcal{L}_1$ is continuous it follows that $\mathcal{L}_1\left(\frac{\psi(\cdot,y+te_j)-\psi(\cdot,y)}{t}\right)(x) \to \mathcal{L}_1\left(\frac{\partial\psi}{\partial y_j}(\cdot, y)\right)(x)$ as $t \to 0$. Hence,

$$\frac{\Psi(x, y + te_j) - \Psi(x, y)}{t} \to \mathcal{L}_1\left(\frac{\partial\psi}{\partial y_j}(\cdot, y)\right)(x) \quad \text{as } t \to 0. \tag{A.31}$$

Since this is valid for every $(x, y) \in \mathbb{R}^d \times \mathbb{R}^m$, this proves that Ψ is also differentiable at every point with respect to the *second* components, and that $\frac{\partial\Psi}{\partial y_j}(x, y) = \mathcal{L}_1\left(\frac{\partial\psi}{\partial y_j}(\cdot, y)\right)(x)$. In addition, since $\frac{\partial\psi}{\partial y_j} \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m)$, the same argument used to prove the continuity of Ψ can be used to prove the continuity of $\frac{\partial\Psi}{\partial y_j}$. Since this procedure can be repeated for any derivative of any order with respect to the *second* components of Ψ, it follows that all the derivatives of Ψ exist and are continuous. This proves that $\Psi \in C^\infty(\mathbb{R}^d \times \mathbb{R}^m)$. We remark that we have implicitly proved that $(\mathcal{L}_1 \otimes \mathcal{I}_m)\left(\frac{\partial\psi}{\partial y_j}\right) = \frac{\partial}{\partial y_j}((\mathcal{L}_1 \otimes \mathcal{I}_m)(\psi))$, and this is also

valid for differential operators with respect to the *second* components of bigger order:

$$D^{(0_d,\beta)}((\mathcal{L}_1 \otimes \mathcal{I}_m)(\psi)) = (\mathcal{L}_1 \otimes \mathcal{I}_m)(D^{(0_d,\beta)}\psi), \tag{A.32}$$

for any $\psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m)$ and any $\beta \in \mathbb{N}^m$. Here $0_d \in \mathbb{N}^d$ denotes the multi-index with d null components.

We will prove at the same time that the function Ψ and its derivatives are fast-decreasing and that $\mathcal{L}_1 \otimes \mathcal{I}_m$ is continuous. For this we will recall the criterion of a continuous linear operator over $\mathscr{S}(\mathbb{R}^d)$ given in Eq. (2.53). We will use this criterion for both cases $n = d$ and $n = d + m$, separating the multi-index notation using concatenations. Let $\alpha, \gamma \in \mathbb{N}^d$ and let $\beta, \theta \in \mathbb{N}^m$. We need to study the expression

$$x^\alpha y^\beta D^{(\gamma,\theta)}\Psi(x,y), \quad (x,y) \in \mathbb{R}^d \times \mathbb{R}^m.$$

Using the linearity of $\mathcal{L}_1$ and the exchange between differential operators with respect to *second* components (A.32), we conclude that for all $(x,y) \in \mathbb{R}^d \times \mathbb{R}^m$ it holds that

$$x^\alpha y^\beta D^{(\gamma,\theta)}\Psi(x,y) = x^\alpha y^\beta D^{(\gamma,0_m)}\mathcal{L}_1(D^{(0_d,\theta)}\psi(\cdot,y))(x)) = x^\alpha D^{(\gamma,0_m)}\mathcal{L}_1\left(y^\beta D^{(0_d,\theta)}\psi(\cdot,y)\right)(x). \tag{A.33}$$

Using then the continuity of $\mathcal{L}_1$ over $\mathscr{S}(\mathbb{R}^d)$, it follows that there exist $C > 0$ and $N \in \mathbb{N}$ such that

$$\begin{aligned}
\sup_{x\in\mathbb{R}^d}\left|x^\alpha y^\beta D^{(\gamma,\theta)}\Psi(x,y)\right| &= \sup_{x\in\mathbb{R}^d}\left|x^\alpha D^{(\gamma,0_m)}\mathcal{L}_1\left(y^\beta D^{(0_d,\theta)}\psi(\cdot,y)\right)(x)\right| \\
&\leqslant C \sum_{\substack{\alpha',\gamma'\in\mathbb{N}^d \\ |\alpha'|,|\gamma'|\leqslant N}} \sup_{x\in\mathbb{R}^d}\left|x^{\alpha'} D^{(\gamma',0_m)}\left(y^\beta D^{(0_d,\theta)}\psi\right)(x,y)\right| \\
&= C \sum_{\substack{\alpha',\gamma'\in\mathbb{N}^d \\ |\alpha'|,|\gamma'|\leqslant N}} \sup_{x\in\mathbb{R}^d}\left|x^{\alpha'} y^\beta D^{(\gamma',\theta)}(\psi)(x,y)\right|.
\end{aligned} \tag{A.34}$$

By taking the supremum over all possible $y \in \mathbb{R}^m$ we finally obtain

$$\begin{aligned}
\sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^m}\left|x^\alpha y^\beta D^{(\gamma,\theta)}(\mathcal{L}_1 \otimes \mathcal{I}_m)(\psi)(x,y)\right| &= \sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^m}\left|x^\alpha y^\beta D^{(\gamma,\theta)}\Psi(x,y)\right| \\
&\leqslant C \sum_{\substack{\alpha',\gamma'\in\mathbb{N}^d \\ |\alpha'|,|\gamma'|\leqslant N}} \sup_{y\in\mathbb{R}^m}\sup_{x\in\mathbb{R}^d}\left|x^{\alpha'} y^\beta D^{(\gamma',\theta)}(\psi)(x,y)\right| \\
&\leqslant C \sum_{\substack{\alpha',\gamma'\in\mathbb{N}^d \\ \beta',\theta'\in\mathbb{N}^m \\ |\alpha'|,|\beta'|,|\gamma'|,|\theta'|\leqslant N+|\beta|+|\theta|}} \sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^m}\left|x^{\alpha'} y^{\beta'} D^{(\gamma',\theta')}(\psi)(x,y)\right|.
\end{aligned} \tag{A.35}$$

Here we have used that for any bounded function $f : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{C}$ it holds that $\sup_{y\in\mathbb{R}^m}\sup_{x\in\mathbb{R}^d}|f(x,y)| = \sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^m}|f(x,y)|$. Since $\psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m)$ is arbitrary, the expression (A.35) proves two things:

first that all the derivatives of $(\mathcal{L}_1 \otimes \mathcal{I}_m)(\psi)$ have fast decreasing behaviour and thus $(\mathcal{L}_1 \otimes \mathcal{I}_m)(\psi) \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^m)$, and that the operator $\mathcal{L}_1 \otimes \mathcal{I}_m$ is continuous, which follows from criterion (2.53). This proves Proposition 2.2.2. ■

A.5 Proof of Proposition 3.2.1

The proof of this Proposition follows standard arguments. We have decided to specify them in the setting of this Proposition. We first prove the deterministic analogue, which is not obvious.

Lemma A.5.1. *Let $f \in C(\mathbb{R}^d)$, $\mu \in \mathscr{M}(\mathbb{R}^d)$ and $A \in \mathcal{B}_B(\mathbb{R}^d)$. Then, for every Riemann sequence of partitions of A, $(V_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$, and for every choice of tag points $x_j^N \in V_j^N$, it holds that*

$$\int_A f(x)d\mu(x) = \lim_{N\to\infty} \sum_{j=1}^{N} \mu(V_j^N) f(x_j^N), \tag{A.36}$$

where $\int_A f(x)d\mu(x)$ denotes the Lebesgue integral of f with respect to μ over A.

Proof: We consider that for a fixed $N \in \mathbb{N}_*$,

$$\sum_{j=1}^{N} \mu(V_j^N) f(x_j^N) = \sum_{j=1}^{N} \int_A f(x_j^N) \mathbf{1}_{V_j^N}(x) d\mu(x) = \int_A \left(\sum_{j=1}^{N} f(x_j^N) \mathbf{1}_{V_j^N}(x) \right) d\mu(x). \tag{A.37}$$

Consider for every $N \in \mathbb{N}_*$ the measurable function $f_N = \sum_{j=1}^{N} f(x_j^N)\mathbf{1}_{V_j^N}(x)$. From Eq. (A.37), the Riemann sum in (A.36) is just $\int_A f_N(x)d\mu(x)$. Let $\epsilon > 0$. Since A is bounded and f is continuous, f is uniformly continuous on A. Hence, there exists $\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ for every $x, y \in A$ such that $|x - y| < \delta$. Let $N_0 \in \mathbb{N}_*$ be large enough such that $\max\limits_{j\in\{1,\dots,N\}} \operatorname{diam}(V_j^N) < \delta$ for all $N \geqslant N_0$. Hence, if $x, y \in V_j^N$, then $|f(x) - f(y)| < \epsilon$. It follows that if $N \geqslant N_0$, then

$$\sup_{x\in A} |f_N(x) - f(x)| \leqslant \max_{j\in\{1,\dots,N\}} \sup_{x\in V_j^N} |f_N(x) - f(x)| = \max_{j\in\{1,\dots,N\}} \sup_{x\in V_j^N} |f(x_j^N) - f(x)| \leqslant \epsilon. \tag{A.38}$$

This proves that $f_N \to f$ uniformly on A. Hence,

$$\left| \int_A f_N(x) - f(x) d\mu(x) \right| \leqslant |\mu|(A) \|f_N - f\|_{\infty,A} \to 0, \tag{A.39}$$

which proves the convergence (A.36). ■

The following Lemma extends in some sense the result of Lemma A.5.1 to the case of *double dimension* in a particularly convenient way.

Lemma A.5.2. *Let $F \in C(\mathbb{R}^d \times \mathbb{R}^m)$, $A \in \mathcal{B}_B(\mathbb{R}^d), B \in \mathcal{B}_B(\mathbb{R}^m)$, and $\Lambda \in \mathscr{M}(\mathbb{R}^d \times \mathbb{R}^m)$. Then, for every Riemann sequence of partitions of A, $(V_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*} \subset \mathcal{B}_B(\mathbb{R}^d)$, for every Riemann sequence of partitions of B, $(U_k^M)_{k\in\{1,\dots,M\},M\in\mathbb{N}_*} \subset \mathcal{B}_B(\mathbb{R}^d \times \mathbb{R}^m)$, and for every choice of tag points $x_j^N \in V_j^N$ and $y_k^M \in U_k^M$, the following double limits hold:*

$$\begin{aligned}\int_{A\times B} F(x,y)d\Lambda(x,y) &= \lim_{N,M\to\infty} \sum_{j=1}^{N}\sum_{k=1}^{M} F(x_j^N, y_k^N)\Lambda(V_j^N \times U_k^M) \\ &= \lim_{N\to\infty}\lim_{M\to\infty} \sum_{j=1}^{N}\sum_{k=1}^{M} F(x_j^N, y_k^N)\Lambda(V_j^N \times U_k^M) \\ &= \lim_{M\to\infty}\lim_{N\to\infty} \sum_{j=1}^{N}\sum_{k=1}^{M} F(x_j^N, y_k^N)\Lambda(V_j^N \times U_k^M).\end{aligned} \tag{A.40}$$

Proof: We remark, first of all, that the limit $\lim_{N,M\to\infty}$ means the limit when both N, M grow to ∞ together[2]. For every $N, M \in \mathbb{N}_*$ let $F_{N,M}$ be the measurable function defined by $F_{N,M}(x,y) = \sum_{j=1}^{N}\sum_{k=1}^{M} F(x_j^N, y_k^N)\mathbf{1}_{V_j^N}(x)\mathbf{1}_{U_k^M}(y)$. The double sums in (A.40) are then the Lebesgue integral of $F_{N,M}$ with respect to Λ. Let us analyse the class of subsets of $\mathbb{R}^d \times \mathbb{R}^m$, $(V_j^N \times U_k^M)_{(j,k)\in\{1,\dots,N\}\times\{1,\dots,M\},N,M\in\mathbb{N}_*}$. This class forms a sort of Riemann *double-sequence* partition of $A \times B$. More precisely, it is immediate that $A \times B = \bigcup_{(j,k)\in\{1,\dots,N\}\times\{1,\dots,M\}} V_j^N \times U_k^M$ for all $N, M \in \mathbb{N}_*$ and that $(V_{j_1}^N \times U_{k_1}^M) \cap (V_{j_2}^N \times U_{k_2}^M) = \varnothing$ if $(j_1, k_1) \neq (j_2, k_2)$. Since for every vector $(x, y) \in \mathbb{R}^d \times \mathbb{R}^m$ it holds that $|(x,y)|_{d+m}^2 = |x|_d^2 + |y|_m^2$, with $|\cdot|_n$ being the Euclidean norm in $\mathbb{R}^n$, we conclude that $\operatorname{diam}(V_j^N \times U_k^M) = \sqrt{\operatorname{diam}(V_j^N)^2 + \operatorname{diam}(U_k^M)^2}$. Hence,

$$\lim_{N,M\to\infty} \max_{(j,k)\in\{1,\dots,N\}\times\{1,\dots,M\}} \operatorname{diam}(V_j^N \times U_k^M) = 0,$$

and the same holds for the corresponding iterated limits $\lim_{N\to\infty}\lim_{M\to\infty}$ and $\lim_{M\to\infty}\lim_{N\to\infty}$. We argue then that the same arguments used in the proof of Lemma A.5.1 can be used in this case. Indeed, since F is continuous and $A \times B$ is bounded, F is uniformly continuous on $A \times B$. Let $\epsilon > 0$. Then, there exists $\delta > 0$ such that for every $(x, y), (u, v) \in A \times B$ such that $|(x, y) - (u, v)| < \delta$, it holds that $|F(x, y) - F(u, v)| \leqslant \epsilon$. Take $N_0 \in \mathbb{N}_*$ big enough such that for every $N, M \geqslant N_0$ we have that $\max_{(j,k)\in\{1,\dots,N\}\times\{1,\dots,M\}} \operatorname{diam}(V_j^N \times U_k^M) < \delta$. Then, it holds that

$$\sup_{(x,y)\in A\times B} |F_{N,M}(x,y) - F(x,y)| \leqslant \max_{(j,k)\in\{1,\dots,N\}\times\{1,\dots,M\}} \sup_{(x,y)\in V_j^N\times U_k^M} |F_{N,M}(x,y) - F(x,y)| \leqslant \epsilon. \tag{A.41}$$

We conclude that $\lim_{N,M\to\infty} F_{N,M} = F$ uniformly on $A \times B$. Since $A \times B$ is bounded and $\Lambda \in \mathscr{M}(\mathbb{R}^d \times$

[2] If $(a_{n,m})_{n,m\in\mathbb{M}}$ is a *double-sequence* of complex numbers, we say that $\lim_{n,m\to\infty} a_{n,m} = a$, for a complex number a, if for every $\epsilon > 0$ there exists $N_0 \in \mathbb{N}_*$ such that for every $n, m \geqslant N_0$ we have $|a_{n,m} - a| < \epsilon$. It does not hold, in general, that $\lim_{n,m\to\infty} a_{n,m} = \lim_{n\to\infty}\lim_{m\to\infty} a_{n,m} = \lim_{m\to\infty}\lim_{n\to\infty} a_{n,m}$. See Habil (2016).

$\mathbb{R}^m$), then $|\Lambda|(A \times B) < \infty$. This proves that

$$\left| \int_{A\times B} F_{N,M}(x,y) - F(x,y) d\Lambda(x,y) \right| \leqslant |\Lambda|(A \times B) \|F_{N,M} - F\|_{\infty, A\times B} \to 0, \tag{A.42}$$

as N, M grow together. We conclude that the double limit in (A.40) holds.

Consider now the iterative limits in (A.40). For every $N \in \mathbb{N}_*$ and for every $j \in \{1, ..., N\}$, let us define the set application $\lambda_{V_j^N} : \mathcal{B}_B(\mathbb{R}^m) \to \mathbb{C}$ through $\lambda_{V_j^N}(U) = \Lambda(V_j^N \times U)$ for every $U \in \mathcal{B}_B(\mathbb{R}^m)$. Clearly, $\lambda_{V_j^N}$ is a well-defined complex measure in $\mathscr{M}(\mathbb{R}^m)$. Hence, from Lemma A.5.1 it holds that

$$\lim_{M\to\infty} \sum_{k=1}^{M} F(x, y_k^N) \Lambda(V_j^N \times U_k^M) = \lim_{M\to\infty} \sum_{k=1}^{M} F(x, y_k^N) \lambda_{V_j^N}(U_k^M) = \int_B F(x,y) d\lambda_{V_j^N}(y). \tag{A.43}$$

The integral in (A.43) is thus well-defined for every $N \in \mathbb{N}_*$. An analogue argument can be used to prove the existence of the partial limit with $\lim_{N\to\infty}$. By a known property of double-sequences (Habil, 2016, Theorem 2.13), this proves that the iterative limits in (A.40) converge to the same limit which is equal the double limit. Hence, the double sums converge, in all the senses, to the Lebesgue integral of F with respect to Λ over $A \times B$. ■

Proof of Proposition 3.2.1:

Let us prove that the integral exists as a square-integrable random variable for a fixed arbitrary Riemann sequence of partitions of A, $(V_j^N)_{j\in\{1,...,N\}, N\in\mathbb{N}_*}$, and for fixed tag points $x_j^N \in V_j^N$. Consider the sequence of squared-integrable random variables $\left(\sum_{j=1}^n \mu(V_j^n) Z(x_j^n)\right)_{n\in\mathbb{N}_*}$. Each random variable of this sequence is in $L^2(\Omega, \mathcal{A}, \mathbb{P})$ since it is a finite linear combination of random variables in $L^2(\Omega, \mathcal{A}, \mathbb{P})$. We remark that $\mu(V_j^N) \in \mathbb{C}$ since $\mu \in \mathscr{M}(\mathbb{R}^d)$. Let us prove that it is a Cauchy sequence. For $n, m \in \mathbb{N}_*$, consider the expression

$$\mathbb{E}\left(\left| \sum_{j=1}^{n} \mu(V_j^n) Z(x_j^n) - \sum_{k=1}^{m} \mu(V_k^m) Z(x_j^m) \right|^2 \right). \tag{A.44}$$

We split it and we apply the linearity of the expectation to obtain,

$$\begin{aligned}
&\mathbb{E}\Big(\sum_{j=1}^{n}\sum_{k=1}^{n} Z(x_j^n)Z(x_k^n)\mu(V_j^n)\overline{\mu(V_k^n)} + \sum_{j=1}^{m}\sum_{k=1}^{m} Z(x_j^m)Z(x_k^m)\mu(V_j^m)\overline{\mu(V_k^m)} \\
&- \sum_{j=1}^{n}\sum_{k=1}^{m} Z(x_j^n)Z(x_k^m)\mu(V_j^n)\overline{\mu(V_k^m)} - \sum_{j=1}^{m}\sum_{k=1}^{n} Z(x_j^m)Z(x_k^n)\mu(V_j^m)\overline{\mu(V_k^n)} \Big) \\
&= \Big(\sum_{j=1}^{n}\sum_{k=1}^{n} \big(C_Z(x_j^n,x_k^n) + m_Z(x_j^n)m_Z(x_k^n)\big)\mu(V_j^n)\overline{\mu(V_k^n)} \\
&+ \sum_{j=1}^{m}\sum_{k=1}^{m} \big(C_Z(x_j^m,x_k^m) + m_Z(x_j^m)m_Z(x_k^m)\big)\mu(V_j^m)\overline{\mu(V_k^m)} \\
&- \sum_{j=1}^{n}\sum_{k=1}^{m} \big(C_Z(x_j^n,x_k^m) + m_Z(x_j^n)m_Z(x_k^m)\big)\mu(V_j^n)\overline{\mu(V_k^m)} \\
&- \sum_{j=1}^{m}\sum_{k=1}^{n} \big(C_Z(x_j^m,x_k^n) + m_Z(x_j^m)m_Z(x_k^n)\big)\mu(V_j^m)\overline{\mu(V_k^n)} \Big).
\end{aligned} \tag{A.45}$$

This is equal to

$$\begin{aligned}
&\Big\{ \sum_{j=1}^{n}\sum_{k=1}^{n} C_Z(x_j^n,x_k^n)\mu(V_j^n)\overline{\mu(V_k^n)} + \Big|\sum_{j=1}^{n} m_Z(x_j^n)\mu(V_j^n)\Big|^2 \\
&+ \sum_{j=1}^{m}\sum_{k=1}^{m} C_Z(x_j^m,x_k^m)\mu(V_j^m)\overline{\mu(V_k^m)} + \Big|\sum_{j=1}^{m} m_Z(x_j^m)\mu(V_j^m)\Big|^2 \\
&- \sum_{j=1}^{n}\sum_{k=1}^{m} C_Z(x_j^n,x_k^m)\mu(V_j^n)\overline{\mu(V_k^m)} - \sum_{j=1}^{n} m_Z(x_j^n)\mu(V_j^n)\overline{\sum_{k=1}^{m} m_Z(x_k^m)\mu(V_k^m)} \\
&- \sum_{j=1}^{m}\sum_{k=1}^{n} C_Z(x_j^m,x_k^n)\mu(V_j^m)\overline{\mu(V_k^n)} - \sum_{j=1}^{m} m_Z(x_j^m)\mu(V_j^m)\overline{\sum_{k=1}^{n} m_Z(x_k^n)\mu(V_k^n)} \Big\}.
\end{aligned} \tag{A.46}$$

We remark that, from Lemma A.5.1, the sums of the form $\sum_{j=1}^{N} m_Z(x_j^N)\mu(V_j^N)$ converge to $\int_A m_Z(x)d\mu(x)$ as $N \to \infty$. For the double sums involving the covariance, we apply Lemma A.5.2 to conclude that sums of the form $\sum_{j=1}^{N}\sum_{k=1}^{M} C_Z(x_j^N,x_k^M)\mu(V_j^N)\overline{\mu(U_k^M)}$ converge to $\int_{A\times A} C_Z(x,y)d(\mu\otimes\overline{\mu})(x,y)$ as N, M grow together, and also for the corresponding iterative limits. From this we obtain that (A.44) can be made arbitrarily close, as n, m grow together, to

$$\begin{aligned}
&2\Big(\int_{A\times A} C_Z(x,y)d(\mu\otimes\overline{\mu})(x,y) + \Big|\int_A m_Z(x)d\mu(x)\Big|^2 - \int_{A\times A} C_Z(x,y)d(\mu\otimes\overline{\mu})(x,y) - \Big|\int_A m_Z(x)d\mu(x)\Big|^2 \Big) \\
&= 0.
\end{aligned} \tag{A.47}$$

This proves that the sequence $\left(\sum_{j=1}^{n}\mu(V_j^n)Z(x_j^n)\right)_{n\in\mathbb{N}_*}$ is a Cauchy sequence in $L^2(\Omega,\mathcal{A},\mathbb{P})$, which is a Hilbert space. Hence, it converges to a well-defined square-integrable random variable which we will note for now I_V.

Let us prove that this limit does not depend on the Riemann sequence of partitions chosen. Let us then consider $(U_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$ be another Riemann sequence of partitions of A, and let us fix any arbitrary collection of tag points $y_j^N\in U_j^N$. The sequence $\left(\sum_{j=1}^{n}Z(y_j^n)\mu(U_j^n)\right)_{n\in\mathbb{N}_*}$ converges to a square-integrable random variable that we will call I_U. We have then that

$$I_V - I_U = \lim_{n\to\infty}\sum_{j=1}^{n} Z(x_j^n)\mu(V_j^n) - Z(y_j^n)\mu(U_j^n), \tag{A.48}$$

where the limit is taken in the sense of $L^2(\Omega,\mathcal{A},\mathbb{P})$. It follows that

$$\mathbb{E}(|I_V - I_U|^2) = \lim_{n\to\infty}\mathbb{E}\left(\left|\sum_{j=1}^{n} Z(x_j^n)\mu(V_j^n) - Z(y_j^n)\mu(U_j^n)\right|^2\right).$$

If we compute the expression $\mathbb{E}(|I_V - I_U|^2) = \lim_{n\to\infty}\mathbb{E}\left(\left|\sum_{j=1}^{n} Z(x_j^n)\mu(V_j^n) - Z(y_j^n)\mu(U_j^n)\right|^2\right)$, one obtains similar expressions as in Eq. (A.45). By applying Lemmas A.5.1 and A.5.2, one obtains that this expression converges to (A.47) as n grows, hence it vanishes. This proves that $I_U = I_V$ in $L^2(\Omega,\mathcal{A},\mathbb{P})$, and hence the limit is unique and it does not depend on the Riemann sequence of partitions of A selected. The limit I_V will be denoted then by $\int_A Z(x)d\mu(x)$.

We finally prove the formulas of the mean (3.10) and the covariance (3.11). Let $(V_j^N)_{j\in\{1,\dots,N\},N\in\mathbb{N}_*}$ be a Riemann sequence of partitions of A, with tag points $x_j^N\in V_j^N$. The formula of the mean is immediate from Lemma A.5.1 since

$$\mathbb{E}\left(\int_A Z(x)d\mu(x)\right) = \mathbb{E}\left(\lim_{N\to\infty}\sum_{j=1}^{N} Z(x_j^N)\mu(V_j^N)\right) = \lim_{N\to\infty}\sum_{j=1}^{N} m_Z(x_j^N)\mu(V_j^N) = \int_A m_Z(x)d\mu(x). \tag{A.49}$$

Finally, if $(U_k^M)_{k\in\{1,\dots,M\},M\in\mathbb{N}_*}$ is any Riemann sequence of partitions of B with its associated points $y_k^M\in U_k^M$, then applying Lemma A.5.2 to the case $\Lambda=\mu\otimes\overline{\nu}$, one obtains

$$\begin{aligned}
\mathbb{C}ov\left(\int_A Z(x)d\mu(x)\,,\int_B Z(x)d\nu(x)\right) &= \mathbb{C}ov\left(\lim_{N\to\infty}\sum_{j=1}^{N} Z(x_j^N)\mu(V_j^N)\,,\lim_{M\to\infty}\sum_{k=1}^{M} Z(y_k^M)\nu(U_k^M)\right)\\
&= \lim_{N\to\infty}\lim_{M\to\infty}\mathbb{C}ov\left(\sum_{j=1}^{N} Z(x_j^N)\mu(V_j^N)\,,\sum_{k=1}^{M} Z(y_k^M)\nu(U_k^M)\right)\\
&= \lim_{N\to\infty}\lim_{M\to\infty}\sum_{j=1}^{N}\sum_{k=1}^{M} C_Z(x_j^N,y_k^M)\mu(V_j^N)\overline{\nu(U_k^M)}\\
&= \int_A\int_B C_Z(x,y)d\mu(x)d\overline{\nu}(y),
\end{aligned} \tag{A.50}$$

where we have used the sesquilinearity of the covariance and the convergence of covariances of double-sequences of square integrable random variables[3]. This proves the desired result. ■

A.6 Proof of Proposition 3.3.1

Let M be a Random Measure over $\mathbb{R}^d$ and let $f:\mathbb{R}^d\to\mathbb{C}$ be a measurable function such that $f\in\mathscr{L}^1(\mathbb{R}^d,m_M)$ and $f\otimes\overline{f}\in\mathscr{L}^1(\mathbb{R}^d\times\mathbb{R}^d,C_M)$. In order to prove that f is integrable with respect to M, we need to prove, following the definition of the integral (3.24), that for any sequence of simple functions $(f_n)_{n\in\mathbb{N}}$ converging point-wise to f and such that $|f_n|$ converges point-wise monotonically increasing to $|f|$, the sequence $(\int_{\mathbb{R}^d} f_n(x)dM(x))_{n\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\Omega,\mathcal{A},\mathbb{P})$. Let us consider $(f_n)_{n\in\mathbb{N}}$ such a sequence. We have then for $m,n\in\mathbb{N}$,

$$\begin{aligned}
\mathbb{E}\left(\left|\int_{\mathbb{R}^d} f_n(x)dM(x)-\int_{\mathbb{R}^d} f_m(x)dM(x)\right|^2\right) &= \mathbb{E}\left(\left|\int_{\mathbb{R}^d} f_n(x)-f_m(x)dM(x)\right|^2\right)\quad\text{(by linearity)}\\
&= \mathbb{V}ar\left(\int_{\mathbb{R}^d} f_n(x)-f_m(x)dM(x)\right)\\
&\quad+|\mathbb{E}\left(\int_{\mathbb{R}^d} f_n(x)-f_m(x)dM(x)\right)|^2\\
&= \int_{\mathbb{R}^d\times\mathbb{R}^d}(f_n(x)-f_m(x))\overline{(f_n(y)-f_m(y))}dC_M(x,y)\\
&\quad+\left|\int_{\mathbb{R}^d}(f_n(x)-f_m(x))dm_M(x)\right|.
\end{aligned} \tag{A.51}$$

[3] If $(X_n)_{n\in\mathbb{N}_*}$ and $(Y_m)_{m\in\mathbb{N}_*}$ are two sequences of square integrable random variables converging in $L^2(\Omega,\mathcal{A},\mathbb{P})$ to X and Y respectively, then $\lim_{n,m\to\infty}\mathbb{C}ov(X_n,Y_m)=\lim_{m\to\infty}\lim_{n\to\infty}\mathbb{C}ov(X_n,Y_m)=\lim_{n\to\infty}\lim_{m\to\infty}\mathbb{C}ov(X_n,Y_m)=\mathbb{C}ov(X,Y)$. This can be proven using the Chauchy-Schwarz inequality.

The convergence to 0 of the final expression in (A.51) as $n, m \to \infty$ is guaranteed from the facts that

$$\int_{\mathbb{R}^d} f_n(x) dm_M(x) \to \int_{\mathbb{R}^d} f(x) dm_M(x)$$

as $n \to \infty$ and

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} f_n(x)\overline{f_m}(y) dC_M(x,y) \to \int_{\mathbb{R}^d \times \mathbb{R}^d} (f \otimes \overline{f})(x,y) dC_M(x,y)$$

as $n, m \to \infty$. These last convergences are guaranteed by the integrability of the involved functions with respect to the involved measures and by Dominated Convergence Theorem 2.1.2. This proves that f is integrable with respect to M. The formula of the mean 3.25 and covariance 3.26 hold evidently when f and g are simple functions, and with the Dominated Convergence Theorem, one proves that the same holds when f and g satisfy the required conditions. ■

A.7 Proof of Proposition 3.3.2 and its variants

Let M be a Random Measure over $\mathbb{R}^d$, and consider the random variables $(\langle M, \varphi \rangle)_{\varphi \in C_c(\mathbb{R}^d)}$ defined through the integrals of φ with respect to M. The linear functional $\varphi \in C_c(\mathbb{R}^d) \mapsto \langle M, \varphi \rangle \in L^2(\Omega, \mathcal{A}, \mathbb{P})$ is continuous if and only if for all compact $K \subset \mathbb{R}^d$ there exists $C_K > 0$ such that

$$\|\langle M, \varphi \rangle\|_{L^2(\Omega,\mathcal{A},\mathbb{P})} \leqslant C_K \|\varphi\|_\infty, \quad \forall \varphi \in C_c(\mathbb{R}^d) \text{ such that } \operatorname{supp}(\varphi) \subset K. \tag{A.52}$$

This comes from the typical criterion of continuity of lineal operators between locally convex vector spaces (Theorem D.0.1). Thus, let $K \subset \mathbb{R}^d$ be a compact set and let $\varphi \in C_c(\mathbb{R}^d)$ such that $\operatorname{supp}(\varphi) \subset K$. Considering that $m_M \in \mathscr{M}(\mathbb{R}^d) = C_c'(\mathbb{R}^d)$ and $C_M \in \mathscr{M}(\mathbb{R}^d \times \mathbb{R}^d) = C_c'(\mathbb{R}^d \times \mathbb{R}^d)$ (Riesz Representation Theorem 2.1.3), it follows from criterion 2.22 that there exist $C_K^{m_M} > 0$ and $C_{K\times K}^{C_M} > 0$ such that

$$|\langle m_M, \varphi \rangle| \leqslant C_K^{m_M} \|\varphi\|_\infty \quad ; \quad |\langle C_M, \varphi \otimes \overline{\varphi} \rangle| \leqslant C_{K\times K}^{C_M} \|\varphi\|_\infty^2, \qquad \forall \varphi \in C_c(\mathbb{R}^d) \text{ such that } \operatorname{supp}(\varphi) \subset K. \tag{A.53}$$

Hence,

$$\begin{aligned} \|\langle M, \varphi \rangle\|_{L^2(\Omega,\mathcal{A},\mathbb{P})} &= \sqrt{\mathbb{E}\left(|\langle M, \varphi \rangle|^2\right)} \\ &= \sqrt{\mathbb{V}ar(\langle M, \varphi \rangle) + \left|\mathbb{E}\left(\langle M, \varphi \rangle\right)\right|^2} \\ &= \sqrt{\langle C_M, \varphi \otimes \overline{\varphi} \rangle + |\langle m_M, \varphi \rangle|^2} \\ &\leqslant \sqrt{C_{K\times K}^{C_M} \|\varphi\|_\infty^2 + (C_K^{m_M})^2 \|\varphi\|_\infty^2} \\ &= \sqrt{C_{K\times K}^{C_M} + (C_K^{m_M})^2} \|\varphi\|_\infty. \end{aligned} \tag{A.54}$$

Hence, proves that M defines a continuous linear functional.

The proof for the cases of the spaces $C_{FD}(\mathbb{R}^d), C_0(\mathbb{R}^d)$ and $C(\mathbb{R}^d)$ is done in complete analogy, con-

sidering criteria of continuity of m_M, C_M and M according to the topology associated to each space. ■

A.8 Proof of Proposition 3.3.3

Let M be an Hermitian orthogonal Random Measure over $\mathbb{R}^d$, with weight ν_M. The parity of ν_M comes immediately from

$$\nu_M(A) = \mathbb{V}ar(M(A)) = \mathbb{V}ar(\overline{M(-A)}) = \mathbb{V}ar(M(-A)) = \nu_M(-A), \quad \forall A \in \mathcal{B}_B(\mathbb{R}^d). \tag{A.55}$$

Here we have used that the variance of a complex random variable equals the variance of its conjugate. Let M_R and M_I be the real and imaginary parts of M respectively. Let $A, B \in \mathcal{B}_B(\mathbb{R}^d)$. Using the Hermitianity of M and the parity of ν_M, we obtain that

$$\begin{aligned}
\mathbb{C}ov(M_R(A), M_I(B)) &= \mathbb{E}\left(M_R(A)M_I(B)\right) \\
&= \mathbb{E}\left(\frac{M(A) + \overline{M(A)}}{2}\frac{M(B) - \overline{M(B)}}{2i}\right) \\
&= \frac{-i}{4}\mathbb{E}\left(M(A)\overline{M(-B)} - M(A)\overline{M(B)} + M(-A)\overline{M(-B)} - M(-A)\overline{M(B)}\right) \\
&= \frac{-i}{4}\Big(\mathbb{C}ov(M(A), M(-B)) - \mathbb{C}ov(M(A), M(B)) \\
&\qquad + \mathbb{C}ov(M(-A), M(-B)) - \mathbb{C}ov(M(-A), M(B))\Big) \\
&= \frac{-i}{4}\Big(\nu_M(A \cap (-B)) - \nu_M(A \cap B) + \underbrace{\nu_M((-A) \cap (-B))}_{=\nu_M(A \cap B)} - \underbrace{\nu_M((-A) \cap B)}_{=\nu_M(A \cap (-B))}\Big) \\
&= 0.
\end{aligned} \tag{A.56}$$

Hence, M_R and M_I are non-correlated Random Measures. The expressions for the covariance measures of both M_R and M_I are obtained following the same principles as in (A.56). ■

A.9 Proof of Proposition 4.2.1

Let Z be a real stationary GeRF over $\mathbb{R}^d$ with covariance distribution $C_Z \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^d)$, stationary covariance distribution $\rho_Z \in \mathscr{S}'(\mathbb{R}^d)$ and spectral measure $\mu_Z \in \mathscr{M}_{SG}^+(\mathbb{R}^d)$. Let g be a symbol function over $\mathbb{R}^d$, with $\mathcal{L}_g$ being its associated operator. Since Z is real and stationary, its Fourier Transform is an Hermitian slow-growing orthogonal Random Measure (Theorem 3.4.2). Since g is measurable and polynomially bounded, it is clear that $g \otimes \overline{g}$ is locally integrable with respect to $C_{\mathscr{F}(Z)}$ and that the measure $g \otimes \overline{g} C_{\mathscr{F}(Z)}$ is in $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^d)$ (see Section 2.1.3). Hence, the multiplication $g\mathscr{F}(Z)$ is a well-defined Random Measure over $\mathbb{R}^d$ (see Section 3.3.3), and since $C_{g\mathscr{F}(Z)} = g \otimes \overline{g} C_{\mathscr{F}(Z)} = |g|^2 (2\pi)^{\frac{d}{2}} \mu_Z \delta^{\{y=x\}}$, it follows that it is a

slow-growing orthogonal Random Measure. It is also Hermitian since g is Hermitian. Finally, using again Theorem 3.4.2, it follows that it Inverse Fourier Transform $\mathscr{F}^{-1}(g\mathscr{F}(Z))$, which is then $\mathcal{L}_g Z$, is a real stationary GeRF over $\mathbb{R}^d$, with spectral measure $\mu_{\mathcal{L}_g Z} = |g|^2 \mu_Z$. The expression for the stationary covariance $\rho_{\mathcal{L}_g Z}$ is immediate since $\mu_Z = \mathscr{F}(\rho_Z)$, and hence $\rho_{\mathcal{L}_g Z} = \mathscr{F}^{-1}(|g|^2\mu_Z) = \mathscr{F}^{-1}(|g|^2\mathscr{F}(\rho_Z)) = \mathcal{L}_{|g|^2}\rho_Z$. ■

A.10 Proof of Theorem 4.3.1

We will need the following intuitive but not-so-obvious Lemma.

Lemma A.10.1. *Let M be a Random Measure over $\mathbb{R}^d$ with mean measure m_M and covariance measure C_M. Let $f : \mathbb{R}^d \to \mathbb{C}$ be a measurable function such that $\frac{1}{f}$ is locally integrable with respect to m_M and $\frac{1}{f} \otimes \frac{1}{f}$ is locally integrable with respect to C_M. Then, $f(\frac{1}{f}M) = M$, that is, $f(\frac{1}{f}M)$ is a well-defined Random Measure which is a modification of M.*

Proof: We first recall that the condition of modification means that for all $A \in \mathcal{B}_B(\mathbb{R}^d)$, $(f\frac{1}{f}M)(A) \overset{a.s.}{=} M(A)$. When all the involved measures are slow-growing, this implies an equality in the sense of modification between GeRFs when interpreting the measures as GeRFs. This is implied from the definition of integrals of slow-growing measures with respect to functions in the Schwartz space (Example 3.4.2).

The local integrability conditions on $\frac{1}{f}$ and on $\frac{1}{f} \otimes \frac{1}{f}$ imply two things: first that $\frac{1}{f}$ is locally integrable with respect to M, and hence that the Random Measure $\frac{1}{f}M$ is well-defined (see Section 3.3.3), and second that $|m_M|(f^{-1}(\{0\})) = 0 = |C_M|(f^{-1}(\{0\}) \times f^{-1}(\{0\}))$. This implies that $M(f^{-1}(\{0\})) \overset{a.s.}{=} 0$. In addition, following Proposition 3.3.1 and Eq. (3.28), we conclude that f is locally integrable with respect to $\frac{1}{f}M$, hence the multiplication Random Measure $f(\frac{1}{f}M)$ is well-defined. Using the $\sigma-$additivity of M one obtains for $A \in \mathcal{B}_B(\mathbb{R}^d)$,

$$M(A) \overset{a.s.}{=} M(A \cap \{f \neq 0\}) \overset{a.s.}{=} \int_{A\cap\{f\neq 0\}} \frac{f(x)}{f(x)} dM(x) \overset{a.s.}{=} \int_{A\cap\{f\neq 0\}} f(x) d(\frac{1}{f}M)(x) \overset{a.s.}{=} \int_A f(x) d(\frac{1}{f}M)(x). \tag{A.57}$$

This proves that $M(A) \overset{a.s.}{=} (f\frac{1}{f}M)(A)$, and hence $f\frac{1}{f}M$ is a modification of M. ■

Proof of Theorem 4.3.1:

Let X be a real stationary GeRF over $\mathbb{R}^d$ with spectral measure μ_X. Let g be a symbol function over $\mathbb{R}^d$ and let $\mathcal{L}_g$ be its associated operator. We start by proving the existence criterion. Let us prove the necessity. Suppose there exists a real stationary GeRF over $\mathbb{R}^d$, say U, satisfying (4.3). Let μ_U be its spectral measure. Proposition 4.2.1, implies that $|g|^2\mu_U = \mu_X$. This implies in particular that $\mu_X(g^{-1}(\{0\})) = 0$. Since $\mu_U \in \mathscr{M}^+_{SG}(\mathbb{R}^d)$, we can take $N \in \mathbb{N}$ such that $(1+|\xi|^2)^{-N}\mu_U$ is finite. We have then that

$$\int_{\mathbb{R}^d} \frac{d\mu_X(\xi)}{(1+|\xi|^2)^N |g(\xi)|^2} = \int_{\{g\neq 0\}} \frac{|g(\xi)|^2}{(1+|\xi|^2)^N} \frac{d\mu_U(\xi)}{|g(\xi)|^2} = \int_{\{g\neq 0\}} \frac{d\mu_U(\xi)}{(1+|\xi|^2)^N} \leqslant \int_{\mathbb{R}^d} \frac{d\mu_U(\xi)}{(1+|\xi|^2)^N} < \infty. \tag{A.58}$$

Let us prove the sufficiency. Since X is a real stationary GeRF, $\mathscr{F}(X)$ is an Hermitian slow-growing orthogonal Random Measure, with covariance measure $C_{\mathscr{F}(X)} = (2\pi)^{\frac{d}{2}}\mu_X\delta^{\{y=x\}}$ (Eq. (3.68) applied to the Fourier Transform). Following the developments in Section 3.3.3, let us verify that the multiplication $\frac{1}{g}\mathscr{F}(X)$ is a well-defined Random Measure. Let $A \in \mathcal{B}_B(\mathbb{R}^d)$. The function $\frac{1}{g}\mathbf{1}_A$ is integrable with respect to $\mathscr{F}(X)$ if $\frac{1}{g}\mathbf{1}_A \otimes \overline{\frac{1}{g}\mathbf{1}_A} \in \mathscr{L}^1(\mathbb{R}^d \times \mathbb{R}^d, C_{\mathscr{F}(X)})$. We have that,

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \left|\frac{1}{g(\xi)\overline{g(\eta)}}\right| \mathbf{1}_{A\times A}(\xi,\eta)d|C_{\mathscr{F}(X)}|(\xi,\eta) = \int_A \frac{(2\pi)^{\frac{d}{2}}}{|g(\xi)|^2}d\mu_X(\xi). \tag{A.59}$$

Since condition (4.7) implies in particular that the measure $|g|^{-2}\mu_X$ is locally finite (the polynomial controlling only its growing behaviour), the expression (A.59) is finite since A is bounded. Hence, $\frac{1}{g}\mathscr{F}(X)$ is a well-defined Random Measure. In addition, its covariance measure is given by

$$C_{\frac{1}{g}\mathscr{F}(X)} = \frac{1}{g}\otimes\frac{1}{\overline{g}}C_{\mathscr{F}(X)} = (2\pi)^{\frac{d}{2}}(|g|^{-2}\mu_X)\delta^{\{y=x\}} \tag{A.60}$$

Hence, $\frac{1}{g}\mathscr{F}(X)$ is a well-defined orthogonal Random Measure with weight $\nu_{\frac{1}{g}\mathscr{F}(X)} = (2\pi)^{\frac{d}{2}}|g|^{-2}\mu_X$. It is in addition Hermitian since both $\frac{1}{g}$ and $\mathscr{F}(X)$ are Hermitian. Finally, from condition (4.7), it follows that the weight measure $\nu_{\frac{1}{g}\mathscr{F}(X)}$ is slow-growing and hence, by the arguments developed in Section 3.3.6, $\frac{1}{g}\mathscr{F}(X)$ is an Hermitian slow-growing orthogonal Random Measure. Its Inverse Fourier Transform is then well-defined, and from Theorem 3.4.2, $U = \mathscr{F}^{-1}(\frac{1}{g}\mathscr{F}(X))$ is a real stationary GeRF over $\mathbb{R}^d$. From Lemma A.10.1 it follows that $g\frac{1}{g}\mathscr{F}(X) = \mathscr{F}(X)$, and hence

$$\mathscr{F}^{-1}(g\mathscr{F}(U)) = \mathscr{F}^{-1}(\frac{1}{g}g\mathscr{F}(X)) = X. \tag{A.61}$$

Hence $\mathcal{L}_g U = X$, so the existences of a strict stationary solution to Eq. 4.3 is proven.

Let us now prove the uniqueness criterion. Let us prove the necessity. Let us suppose that $g^{-1}(\{0\}) \neq \varnothing$. Consider μ_{U_H} be an even positive slow-growing measure concentrated on $g^{-1}(\{0\})$. We can take, for instance, $\mu_{U_H} = \delta_{\xi_0} + \delta_{-\xi_0}$ for any $\xi_0 \in \mathbb{R}^d$ such that $g(\xi_0) = 0$ (we remark that in such a case $g(-\xi_0) = 0$ since g is Hermitian). It follows that $|g|^2\mu_{U_H} = 0$. Let U be a real stationary solution to 4.3 and let U_H be a real stationary GeRF with spectral measure μ_{U_H} independent of U. It follows that $\mathcal{L}_g(U_H) = 0$, since by Proposition 4.2.1, $\mu_{\mathcal{L}_g U_H} = |g|^2\mu_{U_H} = 0$. Hence $\mathcal{L}_g(U + U_H) = \mathcal{L}_g U = X$. Since the addition of two independent stationary GeRFs is a stationary GeRF, it follows that the solution is not unique. This proves the necessity.

Let us finally prove the sufficiency. Let us suppose that there is no uniqueness, and hence there are two different stationary solutions U_1 and U_2 to (4.3), that is, that one is not the modification of the other. Hence,

$\mathcal{L}_g U_1 = X = \mathcal{L}_g U_2$. Taking Fourier Transform, one obtains

$$g\mathscr{F}(U_1) = X = g\mathscr{F}(U_2). \tag{A.62}$$

Let us call $f = \frac{1}{g}$. Since $|g| > 0$, f is measurable taking finite complex values. Eq. (A.62) is then

$$\frac{1}{f}\mathscr{F}(U_1) = \mathscr{F}(X) = \frac{1}{f}\mathscr{F}(U_2). \tag{A.63}$$

This implies that $|f|^{-2}\mu_{U_1} = \mu_X \in \mathscr{M}_{SG}^{+}(\mathbb{R}^d)$, which in particular implies that $(\frac{1}{f} \otimes \frac{1}{\bar{f}})$ is locally integrable with respect to $C_{\mathscr{F}(U_1)}$, since

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} \left| \frac{1}{f(\xi)\overline{f(\eta)}} \right| \mathbf{1}_{A\times A}(\xi,\eta) d|C_{\mathscr{F}(U_1)}|(\xi,\eta) = \int_A \frac{(2\pi)^{\frac{d}{2}}}{|f(\xi)|^2} d\mu_{U_1}(\xi) = (2\pi)^{\frac{d}{2}}\mu_X(A) < \infty, \tag{A.64}$$

for all $A \in \mathcal{B}_B(\mathbb{R}^d)$. The same can be stated for the local integrability of $(\frac{1}{f} \otimes \frac{1}{\bar{f}})$ with respect to $C_{\mathscr{F}(U_2)}$. By Lemma A.10.1, it follows that multiplying (A.63) by f we obtain

$$\mathscr{F}(U_1) = f\frac{1}{f}\mathscr{F}(U_1) = f\frac{1}{f}\mathscr{F}(U_2) = \mathscr{F}(U_2), \tag{A.65}$$

where the equality means that the involved Random Measures are a modification one of another. Taking Inverse Fourier Transform, we finally obtain

$$U_1 = U_2. \tag{A.66}$$

Hence, U_1 is a modification of U_2. The contradiction allows us to conclude that the stationary solution to (4.3) is unique up to a modification.

It follows that the solution $U = \mathscr{F}^{-1}(\frac{1}{g}\mathscr{F}(X))$, is the only stationary solution that satisfies (4.3) up to a modification. From this, the fact that the spectral measure of the unique stationary solution to (4.3) is of the form (4.8) comes immediately. ■

A.11 Proof of Theorem 4.4.1

For the proof of this Theorem we need two Lemmas. The first one is a straightforward result that we just mention in order to make it explicit. The second one is a known result of convolutions between some kinds of distributions.

Lemma A.11.1. *Let f be a function defining a tempered distribution over $\mathbb{R}^d$ such that $\mathscr{F}(f) \in L^1(\mathbb{R}^d)$. Let*

$\mu \in \mathscr{M}_F(\mathbb{R}^d)$. Then $\mathscr{F}(f)$ and $\mathscr{F}(\mu)$ are convolvable and the exchange formula of the Fourier Transform holds: $\mathscr{F}(f\mu) = (2\pi)^{-\frac{d}{2}}\mathscr{F}(f) * \mathscr{F}(\mu)$.

Proof: The convolvability is straightforward since if μ is finite, $\mathscr{F}(\mu) \in C_B(\mathbb{R}^d)$, and every bounded measurable function is convolvable with an integrable one. We remark that $f \in C_0(\mathbb{R}^d)$ by Riemann-Lebesgue Lemma (Theorem 2.2.1). Since μ is finite and f is continuous and bounded, the multiplication $f\mu$ is in $\mathscr{M}_F(\mathbb{R}^d)$, hence $\mathscr{F}(f\mu)$ is a continuous and bounded function. The Exchange Formula must be verified in the classical sense of continuous functions. Let us call $\phi = \mathscr{F}(f)$. Using the formula of the Inverse Fourier Transform and Fubini's Theorem, we obtain

$$\begin{aligned}
\mathscr{F}(f\mu)(\xi) &= \frac{1}{(2\pi)^{\frac{d}{2}}}\int_{\mathbb{R}^d} e^{-ix^T\xi} f(x)d\mu(x) \\
&= \frac{1}{(2\pi)^{\frac{d}{2}}}\int_{\mathbb{R}^d} e^{-ix^T\xi}\frac{1}{(2\pi)^{\frac{d}{2}}}\int_{\mathbb{R}^d} e^{ix^T\eta}\phi(\eta)d\eta d\mu(x) \\
&= \frac{1}{(2\pi)^{\frac{d}{2}}}\int_{\mathbb{R}^d}\frac{1}{(2\pi)^{\frac{d}{2}}}\int_{\mathbb{R}^d} e^{i(\xi-\eta)^Tx}d\mu(x)\phi(\eta)d\eta \\
&= \frac{1}{(2\pi)^{\frac{d}{2}}}\int_{\mathbb{R}^d}\mathscr{F}(\mu)(\xi-\eta)\mathscr{F}(f)(\eta)d\eta = (2\pi)^{-\frac{d}{2}}(\mathscr{F}(\mu) * \mathscr{F}(f))(\xi).\blacksquare
\end{aligned} \tag{A.67}$$

Lemma A.11.2. *Let $f_1 \in L^1(\mathbb{R}^d)$ and $f_2 \in L^\infty(\mathbb{R}^d)$. Then, for every $\alpha, \beta \in \mathbb{N}^d$, $D^\alpha f_1$ and $D^\beta f_2$ are $\mathscr{S}'-$convolvable. The next equality holds in the sense of distributions for every $\alpha, \beta \in \mathbb{N}^d$:*

$$D^\alpha f_1 * D^\beta f_2 = D^{\alpha+\beta}(f_1 * f_2) = D^\alpha(f_1 * D^\beta f_2) = D^\beta(D^\alpha f_1 * f_2). \tag{A.68}$$

Proof: f_1 and f_2 are convolvable in the classical sense of functions and their convolution is a continuous and bounded function. Hence $f_1 * f_2 \in \mathscr{S}'(\mathbb{R}^d)$, and the derivative $D^{\alpha+\beta}(f_1 * f_2)$ is a well-defined element of $\mathscr{S}'(\mathbb{R}^d)$.

Since in this case we are not in the framework explained in Section 2.2.2, we rather use the definition through the tensor product presented in Section 2.2.3. Hence, we need to verify if the corresponding derivatives of f_1 and f_2 are $\mathscr{S}'-$convolvable.

Consider then $f_1 \in L^1(\mathbb{R}^d)$ and $f_2 \in L^\infty(\mathbb{R}^d)$. The tensor product $f_1 \otimes f_2$ is simply the function $(x, y) \mapsto f_1(x)f_2(y)$. A typical argument using a change of variables and Fubini's Theorem guarantees that the convolution $f_1 * f_2$ is well-defined in the sense of distributions, obtaining

$$\langle f_1 * f_2, \varphi\rangle = \langle f_1 \otimes f_2, (x, y) \mapsto \varphi(x + y)\rangle, \quad \forall\varphi \in \mathscr{S}(\mathbb{R}^d). \tag{A.69}$$

Let us consider right now the derivatives in distributional sense $D^\alpha f_1$ and $D^\beta f_2$. Their tensor product is then $D^\alpha f_1 \otimes D^\beta f_2$, and it holds that (Schwartz, 1966, Theorem VII, Chapter IV) $D^\alpha f_1 \otimes D^\beta f_2 = D^{(\alpha,\beta)}(f_1 \otimes f_2)$.

By definition of the derivative,

$$\langle D^{(\alpha,\beta)}(f_1 \otimes f_2), \psi\rangle := (-1)^{|\alpha|+|\beta|}\langle f_1 \otimes f_2, D^{(\alpha,\beta)}\psi\rangle, \quad \forall \psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R}^d).$$

Since $D^{(\alpha,\beta)}((x,y) \mapsto \varphi(x+y)) = (x,y) \mapsto D^{\alpha+\beta}\varphi(x+y)$, and $D^{\alpha+\beta}\varphi \in \mathscr{S}(\mathbb{R}^d)$ for every $\varphi \in \mathscr{S}(\mathbb{R}^d)$, it follows that

$$\langle D^{\alpha} f_1 \otimes D^{\beta} f_2, (x,y) \mapsto \varphi(x+y)\rangle = (-1)^{|\alpha|+|\beta|}\langle f_1 \otimes f_2, (x,y) \mapsto D^{\alpha+\beta}\varphi(x+y)\rangle = (-1)^{|\alpha|+|\beta|}\langle f_1 * f_2, D^{\alpha+\beta}\varphi\rangle, \tag{A.70}$$

for all $\varphi \in \mathscr{S}(\mathbb{R}^d)$. This proves two things: first $D^{\alpha}f_1$ and $D^{\beta}f_2$ are $\mathscr{S}'-$convolvable, and second $D^{\alpha+\beta}(f_1 * f_2) = D^{\alpha}f_1 * D^{\beta}f_2$, which follows from an immediate application of the definition of the derivative in (A.69). Applying the same principles for the convolutions $D^{\alpha}f_1 * f_2$ and $f_1 * D^{\beta}f_2$, one obtains equality (A.68) . ■

Proof of Theorem 4.4.1:

We first prove the case of condition 1, which is actually immediate within the framework of tempered distributions. Indeed, if $\frac{1}{g} \in \mathcal{O}_M(\mathbb{R}^d)$, then it is obvious that $|g| > 0$ and that the PBR condition holds, hence there exists a unique stationary solution to (4.3) (and to (4.12) too). Its spectral measure is then given by the multiplication $\mu_U = |g|^{-2}\mu_X$. The condition $\frac{1}{g} \in \mathcal{O}_M(\mathbb{R}^d)$ also implies that $|g|^{-2} \in \mathcal{O}_M(\mathbb{R}^d)$. Since $\mu_X \in \mathscr{M}^+_{SG}(\mathbb{R}^d) \subset \mathscr{S}'(\mathbb{R}^d)$, the multiplication measure $|g|^{-2}\mu_X$ is simply the multiplication between a multiplicator of the Schwartz space (Section 2.2.2) and a tempered distribution. Hence, the Exchange Formula for the Fourier Transform (Eq. (2.65)) holds, obtaining:

$$\rho_U = \mathscr{F}(\mu_U) = \mathscr{F}(|g|^{-2}\mu_X) = (2\pi)^{-\frac{d}{2}}\mathscr{F}(|g|^{-2}) * \mathscr{F}(\mu_X) = \mathscr{F}((2\pi)^{-\frac{d}{2}}|g|^{-2}) * \rho_X = \mathscr{F}(\mu_U^W) * \rho_X = \rho_U^W * \rho_X. \tag{A.71}$$

The case of condition 2 is more sophisticated. We will use Lemmas A.11.1 and A.11.2. Let us first prove that the solution to (4.3) exists and that it is unique. If g satisfies condition 2, then the function $(1+|x|^2)^{-N}|g|^{-2}$ is the Inverse Fourier Transform of an integrable function, and hence it is in $C_0(\mathbb{R}^d)$ by Riemann-Lebesgue Lemma (Theorem 2.2.1). Hence, $|g|^{-2} \leqslant \|(1+|x|^2)^{-N}|g|^{-2}\|_{\infty}(1+|x|^2)^N$ and so $|g|^{-2}$ is polynomially bounded. We conclude that the PBR condition holds, hence there exists a unique stationary solution to (4.3) with spectral measure $\mu_U = |g|^{-2}\mu_X$.

Lemma A.11.1 gives us the proof for the case when $N = 0$ in condition 2 and μ_X is a finite measure. Let us prove the general case. Consider g satisfying condition 2 and $\mu_X \in \mathscr{M}^+_{SG}(\mathbb{R}^d)$ any arbitrary spectral measure. Let $N_g \in \mathbb{N}$ such that condition 2 holds. Let $N_{\mu_X} \in \mathbb{N}$ such that $(1+|x|^2)^{-N_{\mu_X}}\mu_X$ is a finite measure. We consider that

$$|g|^{-2}\mu_X = (1+|x|^2)^{N_g+N_{\mu_X}} \underbrace{\frac{|g|^{-2}}{(1+|x|^2)^{N_g}}}_{\in \mathscr{F}^{-1}(L^1(\mathbb{R}^d))} \underbrace{\frac{\mu_X}{(1+|x|^2)^{N_{\mu_X}}}}_{\in \mathscr{M}^+_F(\mathbb{R}^d)}. \tag{A.72}$$

Considering the exchange formula for the case of Lemma A.11.1 and the property of the Fourier Transform with respect to multiplications by polynomials, $\mathscr{F}((1+|x|^2)^M T) = (1-\Delta)^{(M)} T$ for $M \in \mathbb{N}$ and $T \in \mathscr{S}'(\mathbb{R}^d)$, we obtain

$$\mathscr{F}(|g|^{-2}\mu_X) = (2\pi)^{-\frac{d}{2}}(1-\Delta)^{N_g+N_{\mu_C}} \left(\underbrace{\mathscr{F}\left(\frac{|g|^{-2}}{(1+|x|^2)^{N_g}}\right)}_{\in L^1(\mathbb{R}^d)} * \underbrace{\mathscr{F}\left(\frac{\mu_X}{(1+|x|^2)^{N_{\mu_X}}})\right)}_{\in C_B(\mathbb{R}^d)\subset L^\infty(\mathbb{R}^d)} \right). \tag{A.73}$$

We then apply Lemma A.11.2 to split conveniently the differential operators between the two convolving functions. We conclude that

$$\begin{aligned}
\mathscr{F}(|g|^{-2}\mu_X) &= (2\pi)^{-\frac{d}{2}} \left[(1-\Delta)^{N_g}\mathscr{F}\left(\frac{|g|^{-2}}{(1+|x|^2)^{N_g}}\right)\right] * \left[(1-\Delta)^{N_{\mu_X}}\mathscr{F}\left(\frac{\mu_X}{(1+|x|^2)^{N_{\mu_X}}}\right)\right] \\
&= (2\pi)^{-\frac{d}{2}}\mathscr{F}\left((1+|x|^2)^{N_g}\frac{|g|^{-2}}{(1+|x|^2)^{N_g}}\right) * \mathscr{F}\left((1+|x|^2)^{N_{\mu_X}}\frac{\mu_X}{(1+|x|^2)^{N_{\mu_X}}}\right) \\
&= (2\pi)^{-\frac{d}{2}}\mathscr{F}(|g|^{-2}) * \mathscr{F}(\mu_X) \\
&= \mathscr{F}(\mu_U^W) * \mathscr{F}(\mu_X) = \rho_U^W * \rho_X . \blacksquare
\end{aligned} \tag{A.74}$$

A.12 Proof of Proposition 4.6.1

Let $f \in \mathscr{V}'(\mathbb{R}^d)$. Let g be a symbol function over $\mathbb{R}^d$ and let $\mathcal{L}_g$ be its associated operator. Let us prove the necessity. Let $u \in \mathscr{V}'(\mathbb{R}^d)$ be such that $\mathcal{L}_g(u) = f$. Hence, $g\mathscr{F}(u) = \mathscr{F}(f) \in \mathscr{M}_{SG}(\mathbb{R}^d)$. Thus, there exists $N \in \mathbb{N}$ such that $(1+|\xi|^2)^{-N} g\mathscr{F}(u)$ is finite. It also follows that $|\mathscr{F}(f)|(g^{-1}(\{0\})) = 0$. Hence,

$$\int_{\mathbb{R}^d} \frac{d|\mathscr{F}(f)|(\xi)}{(1+|\xi|^2)^N |g(\xi)|} = \int_{\{g\neq 0\}} \frac{|g(\xi)|}{(1+|\xi|^2)^N} \frac{d|\mathscr{F}(u)|(\xi)}{|g(\xi)|} = \int_{\{g\neq 0\}} \frac{d|\mathscr{F}(u)|(\xi)}{(1+|\xi|^2)^N} \leqslant \int_{\mathbb{R}^d} \frac{d|\mathscr{F}(u)|(\xi)}{(1+|\xi|^2)^N} < \infty. \tag{A.75}$$

This proves that $\frac{1}{g}\mathscr{F}(f)$ is slow-growing, and in particular $|g|^{-1}$ is locally integrable with respect to $\mathscr{F}(f)$.

Let us prove the sufficiency. If $\frac{1}{g}$ is locally integrable with respect to $\mathscr{F}(f)$, then the multiplication $\frac{1}{g}\mathscr{F}(f)$ defines a measure in $\mathscr{M}(\mathbb{R}^d)$. Since we have supposed in addition that $\frac{1}{g}\mathscr{F}(f) \in \mathscr{M}_{SG}(\mathbb{R}^d)$, then its inverse Fourier Transform, $u = \mathscr{F}^{-1}(\frac{1}{g}\mathscr{F}(f))$ is well-defined in distributional sense, and it is an element of $\mathscr{V}'(\mathbb{R}^d)$. Let us verify that it solves (4.29). It is clear that $\mathscr{F}(u) = \frac{1}{g}\mathscr{F}(f)$. Since $\frac{1}{g}$ is locally integrable

with respect to the measure $\mathscr{F}(f)$, it follows that $|\mathscr{F}(f)|\left(g^{-1}\left(\{0\}\right)\right) = 0$. Hence,

$$\mathscr{F}(f)(A) = \int_{A\cap\{g\neq 0\}} \frac{g(x)}{g(x)} d\mathscr{F}(f)(x) = \int_{A\cap\{g\neq 0\}} g(x) d\left(\frac{1}{g}\mathscr{F}(f)\right)(x) = \int_A g(x) d\left(\mathscr{F}(u)\right)(x) = \left(g\mathscr{F}(u)\right)(A). \tag{A.76}$$

We conclude that $g\mathscr{F}(u) = \mathscr{F}(f) \in \mathscr{M}_{SG}(\mathbb{R}^d)$. Taking Inverse Fourier Transform, it follows that $\mathcal{L}_g(u) = f$, hence the existence is proven.

Let us prove the necessity of the uniqueness criterion. If we suppose that $g^{-1}(\{0\}) \neq \varnothing$, then any slow-growing measure μ concentrated on $g^{-1}(\{0\})$ (for instance, $\mu = \delta_{\xi_0}$ for any $\xi_0 \in g^{-1}(\{0\})$), satisfies $g\mu = 0$. Hence, setting $u_H = \mathscr{F}^{-1}(\mu)$, one gets that $\mathcal{L}_g u_H = 0$. If u is any solution to (4.29), then $u + u_H$ is also a solution to (4.29), which implies non-uniqueness.

Let us now prove the sufficiency of uniqueness criterion. Let us suppose $|g| > 0$. This implies that $\frac{1}{g}$ is a complex measurable function. It is straightforward hence that for any measure $\mu \in \mathscr{M}(\mathbb{R}^d)$, $\mu = g\frac{1}{g}\mu = \frac{1}{g}g\mu$. If there are two solutions $u_1, u_2 \in \mathscr{V}'(\mathbb{R}^d)$ to (4.29), then $g\mathscr{F}(u_1) = g\mathscr{F}(u_2) = \mathscr{F}(f) \in \mathscr{M}_{SG}(\mathbb{R}^d)$. Multiplying by $\frac{1}{g}$, one obtains that $\mathscr{F}(u_1) = \mathscr{F}(u_2)$, and hence $u_1 = u_2$. The solution is then unique. ■

A.13 Proof of Proposition 5.1.1

Let Z be a real stationary GeRF over $\mathbb{R}^d \times \mathbb{R}$ with temporally integrable spectral measure $\mu_Z \in \mathscr{M}^+_{SG}(\mathbb{R}^d \times \mathbb{R})$. As stated in Section 5.1.2, the covariance distribution $\rho_Z \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R})$ has a continuous-in-time representation $(\rho_Z^u)_{u\in\mathbb{R}} \subset \mathscr{S}'(\mathbb{R}^d)$. The objective is to define the random variables of the form $\langle Z_t, \varphi\rangle$, with $\varphi \in \mathscr{S}(\mathbb{R}^d)$ and $t \in \mathbb{R}$.

Let $t \in \mathbb{R}$ and let $(\theta_n^t)_{n\in\mathbb{N}}$ be a sequence of positive functions in $\mathscr{S}(\mathbb{R})$ converging to δ_t in the following sense: $\int_{\mathbb{R}} \theta_n^t(u)du = 1$ for all $n \in \mathbb{N}$, and for every $f \in C(\mathbb{R})$ polynomially bounded, one has that

$$\langle f, \theta_n^t\rangle \to \langle \delta_t, f\rangle = f(t), \quad \text{as } n \to \infty. \tag{A.77}$$

An example of such a sequence is given by a sequence of Gaussian probability density functions centred at t with the variance decreasing to 0 as n grows. It follows that

$$\mathscr{F}_T(\theta_n^t)(\omega) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\omega u}\theta_n^t(u)du \to \frac{1}{\sqrt{2\pi}} e^{-i\omega t}, \quad \text{as } n \to \infty, \forall \omega \in \mathbb{R}. \tag{A.78}$$

Hence, the temporal Fourier Transforms of the functions in $(\theta_n^t)_{n\in\mathbb{N}}$ converge point-wise to the function $\omega \mapsto \frac{1}{\sqrt{2\pi}} e^{-i\omega t}$.

Let us fix $\varphi \in \mathscr{S}(\mathbb{R}^d)$ and $t \in \mathbb{R}$ and let us consider the sequence of square-integrable random variables $(\langle Z, \varphi \boxtimes \theta_n^t\rangle)_{n\in\mathbb{N}}$. Let us prove that it converges in $L^2(\Omega, \mathcal{A}, \mathbb{P})$. Let $n, m \in \mathbb{N}$. Using the linearity and that

$C_{\mathscr{F}^{-1}(Z)} = (2\pi)^{\frac{d+1}{2}} \mu_Z \delta^{\{x=y\}}$, we obtain that

$$\begin{aligned}\mathbb{E}\left(\left|\langle Z, \varphi \boxtimes \phi_n^t\rangle - \langle Z, \varphi \boxtimes \theta_m^t\rangle\right|^2\right) &= \mathbb{E}\left(\left|\langle Z, \varphi \boxtimes (\theta_n^t - \theta_m^t)\rangle\right|^2\right) \\ &= \mathbb{E}\left(\left|\langle \mathscr{F}^{-1}(Z), \mathscr{F}_S(\varphi) \boxtimes \mathscr{F}_T(\theta_n^t - \theta_m^t)\rangle\right|^2\right) \\ &= (2\pi)^{\frac{d+1}{2}} \int_{\mathbb{R}^d \times \mathbb{R}} |\mathscr{F}_S(\varphi)(\xi)|^2 \left|\mathscr{F}_T(\theta_n^t - \theta_m^t)(\omega)\right|^2 d\mu_Z(\xi, \omega).\end{aligned} \tag{A.79}$$

The point-wise convergence of $\mathscr{F}_T(\theta_n^t)$ to the function $\omega \mapsto \frac{1}{\sqrt{2\pi}} e^{-i\omega t}$ implies that

$$\begin{aligned}\left|\mathscr{F}_T(\theta_n^t - \theta_m^t)(\omega)\right|^2 &= \left|\mathscr{F}_T(\theta_n^t)(\omega)\right|^2 + \left|\mathscr{F}_T(\theta_m^t)(\omega)\right|^2 - \left(\mathscr{F}_T(\theta_n^t)(\omega)\overline{\mathscr{F}_T(\theta_m^t)(\omega)} + \mathscr{F}_T(\theta_n^t)(\omega)\overline{\mathscr{F}_T(\theta_m^t)(\omega)}\right) \\ &\xrightarrow[n,m\to\infty]{} \left|\frac{1}{\sqrt{2\pi}} e^{-i\omega t}\right|^2 + \left|\frac{1}{\sqrt{2\pi}} e^{-i\omega t}\right|^2 - \left(\frac{1}{\sqrt{2\pi}} e^{-i\omega t}\overline{\frac{1}{\sqrt{2\pi}} e^{-i\omega t}} + \overline{\frac{1}{\sqrt{2\pi}} e^{-i\omega t}} \frac{1}{\sqrt{2\pi}} e^{-i\omega t}\right) \\ &= 0.\end{aligned} \tag{A.80}$$

Hence, $\left|\mathscr{F}_T(\theta_n^t - \theta_m^t)\right|^2 \to 0$ point-wise as $n, m \to \infty$. In addition we have that

$$\begin{aligned}\left|\mathscr{F}_T(\theta_n^t - \theta_m^t)(\omega)\right|^2 &= \left|\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\omega u}(\theta_n^t - \theta_m^t)(u) du\right|^2 \\ &\leqslant \left(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} |\theta_n^t - \theta_m^t|(u) du\right)^2 \\ &\leqslant \left(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (\theta_n^t + \theta_m^t)(u) du\right)^2 \\ &\leqslant \frac{2}{\pi}.\end{aligned} \tag{A.81}$$

Hence, $\left|\mathscr{F}_T(\theta_n^t - \theta_m^t)\right|^2 \to 0$ point-wise as $n, m \to \infty$ and in addition dominated by $\frac{2}{\pi}$. Since μ_Z is temporally integrable, one has

$$(2\pi)^{\frac{d+1}{2}} \int_{\mathbb{R}^d \times \mathbb{R}} |\mathscr{F}_S(\varphi)(\xi)|^2 \frac{\pi}{2} d\mu_Z(\xi, \omega) < \infty. \tag{A.82}$$

It follows from Dominated Convergence Theorem that

$$(2\pi)^{\frac{d+1}{2}} \int_{\mathbb{R}^d \times \mathbb{R}} |\mathscr{F}_S(\varphi)(\xi)|^2 \left|\mathscr{F}_T(\theta_n^t - \theta_m^t)(\omega)\right|^2 d\mu_Z(\xi, \omega) \to 0, \text{ as } n, m \to \infty. \tag{A.83}$$

Hence, the sequence of random variables $(\langle Z, \varphi \boxtimes \theta_n^t\rangle)_{n\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\Omega, \mathcal{A}, \mathbb{P})$, and hence it is convergent to a unique square-integrable random variable, which we will denote by $\langle Z_t, \varphi\rangle$. This limit does not depend on the sequence $(\theta_n^t)_{n\in\mathbb{N}}$. This can be concluded by considering another sequence of temporal test-functions, say $(\vartheta_n^t)_{n\in\mathbb{N}}$, converging δ_t in the same sense as $(\theta_n^t)_{n\in\mathbb{N}}$, and considering the

sequence of random variables $(\langle Z, \varphi \boxtimes (\theta_n^t - \vartheta_n^t)\rangle)_{n\in\mathbb{N}}$. Following similar arguments as those exposed in this proof, one proves that $\langle Z, \varphi \boxtimes (\theta_n^t - \vartheta_n^t)\rangle \to 0$ in $L^2(\Omega, \mathcal{A}, \mathbb{P})$ as $n \to \infty$, and hence

$$\langle Z_t, \varphi\rangle = \lim_{n\to\infty} \langle Z, \varphi \boxtimes \theta_n^t\rangle = \lim_{n\to\infty} \langle Z, \varphi \boxtimes \vartheta_n^t\rangle. \tag{A.84}$$

Let us now fix two time coordinates $t, s \in \mathbb{R}$. Consider two sequences of temporal test-functions $(\theta_n^t)_{n\in\mathbb{N}}$ and $(\theta_n^s)_{n\in\mathbb{N}}$ converging respectively to δ_t and to δ_s in the sense given above. Let $f : \mathbb{R} \to \mathbb{C}$ be a continuous and bounded function. Then, Fubini's Theorem, a change of variable and a passage to limit, we conclude that

$$\langle f, \theta_n^t * \check{\overline{\theta_n^s}}\rangle = \int_{\mathbb{R}} f(u) \int_{\mathbb{R}} \theta_n^t(u-v)\overline{\theta_n^s}(-v)dvdu = \int_{\mathbb{R}\times\mathbb{R}} f(u-v)\theta_n^t(u)\overline{\theta_n^s}(v)d(u,v) \to f(t-s), \quad \text{as } n \to \infty. \tag{A.85}$$

Let us then consider two spatial test-functions $\varphi, \phi \in \mathscr{S}(\mathbb{R}^d)$. Then, one has

$$\begin{aligned}
\mathbb{C}ov(\langle Z_t, \varphi\rangle, \langle Z_s, \phi\rangle) &= \lim_{n\to\infty} \mathbb{C}ov(\langle Z, \varphi \boxtimes \theta_n^t\rangle, \langle Z_s, \phi \boxtimes \theta_n^2\rangle) \\
&= \lim_{n\to\infty} \langle \rho_Z, (\varphi * \check{\overline{\phi}}) \boxtimes (\theta_n^t * \check{\overline{\theta_n^s}})\rangle \\
&= \lim_{n\to\infty} \int_{\mathbb{R}} \langle \rho_Z^u, \varphi * \check{\overline{\phi}}\rangle (\theta_n^t * \check{\overline{\theta_n^s}})(u)du \\
&= \langle \rho_Z^{t-s}, \varphi * \check{\overline{\phi}}\rangle,
\end{aligned} \tag{A.86}$$

where we have used the continuity of the function $u \in \mathbb{R} \mapsto \langle \rho_Z^u, \varphi * \check{\overline{\phi}}\rangle$ and the limit expression (A.85).

Using the continuity-in-time structure of the family $(\rho_Z^u)_{u\in\mathbb{R}}$, one concludes immediately that for every $\varphi \in \mathscr{S}(\mathbb{R}^d)$, the Random Function $t \mapsto \langle Z_t, \varphi\rangle$ is a continuous Random Function, which is in addition stationary. On the spatial dimension, using the tempered structure in space of the family $(\rho_Z^u)_{u\in\mathbb{R}} \subset \mathscr{S}'(\mathbb{R}^d)$, one obtains that for every $t \in \mathbb{R}$, the family of random variables $(\langle Z_t, \varphi\rangle)_{\varphi\in\mathscr{S}(\mathbb{R}^d)}$ satisfies the linearity and continuity conditions to be a well-defined real spatial GeRF, which is also stationary. This completes the proof. ■

A.14 Proof of Proposition 5.2.1

Let $\beta > 0$. For an arbitrary symbol function $g : \mathbb{R}^d \to \mathbb{R}$, with g_R and g_I being its real and imaginary parts respectively, we denote by f_{g_R,g_I} the function over $\mathbb{R}^d \times \mathbb{R}$

$$f_{g_R,g_I}(\xi,\omega) = (i\omega)^\beta + g(\xi) = |\omega|^\beta \cos\left(\frac{\beta\pi}{2}\right) + g_R(\xi) + i\left(\mathrm{sgn}(\omega)|\omega|^\beta \sin\left(\frac{\beta\pi}{2}\right) + g_I(\xi)\right). \tag{A.87}$$

Let us first prove the sufficiency of our claim. Let $g_R : \mathbb{R}^d \to \mathbb{R}$ satisfying the PBR condition and such that $g_R \cos(\frac{\beta\pi}{2}) \geqslant 0$. Let $p : \mathbb{R}^d \to \mathbb{R}_*^+$ be a strictly positive polynomial such that $|g_R| \geqslant 1/p$. When β is an odd integer, we get $\cos(\frac{\beta\pi}{2}) = 0$, and it is thus straightforward that $|f_{g_R,g_I}|^2 \geqslant g_R^2 \geqslant 1/p^2$, from which we obtain that f_{g_R,g_I} has PBR for any chosen g_I. When β is not and odd integer, the choice of the sign of g_R is made in order to make that both $\cos(\frac{\beta\pi}{2})$ and g_R have the same sign, and hence $|g_R \cos(\frac{\beta\pi}{2})| = |g_R||\cos(\beta\frac{\pi}{2})|$. Thus, for all $(\xi, \omega) \in \mathbb{R}^d \times \mathbb{R}$ we have

$$|f_{g_R,g_I}(\xi,\omega)|^2 \geqslant \left(|\omega|^\beta \cos(\frac{\beta\pi}{2}) + g_R(\xi)\right)^2 = \left(|\omega|^\beta |\cos(\frac{\beta\pi}{2})| + |g_R|(\xi)\right)^2 \geqslant |g_R(\xi)|^2 \geqslant \frac{1}{p^2(\xi)}. \tag{A.88}$$

Hence, f_{g_R,g_I} satisfies the PBR condition for any chosen g_I.

Let us now prove the necessity. Suppose that for every g_I there exists a strictly positive polynomial $q_{g_I} : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}_*^+$ such that $|f_{g_R,g_I}| \geqslant \frac{1}{q_{g_I}}$, which is equivalent to say that f_{g_R,g_I} satisfies the PBR condition for any g_I. Then, in particular for $g_I = 0$ and evaluating at $\omega = 0$, we get $|f_{g_R,0}(\xi,0)|^2 = g_R^2(\xi) \geqslant q_0(\xi,0)^{-2}$ from which we obtain that g_R satisfies the PBR condition. Let β be such that $\cos(\frac{\beta\pi}{2}) < 0$. Since g_R has PBR, it cannot take the value 0. Suppose there exists $\xi_1 \in \mathbb{R}^d$ such that $g_R(\xi_1) > 0$. If we consider $\omega_{\xi_1} = \left(-g_R(\xi_1)/\cos(\frac{\beta\pi}{2})\right)^{\frac{1}{\beta}}$, we obtain that for every g_I,

$$f_{g_R,g_I}(\xi_1,\omega_{\xi_1}) = -g_R(\xi_1) + g_R(\xi_1) + i\left(-g_R(\xi_1)\tan\left(\frac{\beta\pi}{2}\right) + g_I(\xi_1)\right). \tag{A.89}$$

It suffices then to take a particular measurable polynomially bounded odd function g_I such that $g_I(\xi_1) = g_R(\xi_1)\tan(\frac{\beta\pi}{2})$, to obtain $f_{g_R,g_I}(\xi_1,\omega_{\xi_1}) = 0$. This proves that f_{g_R,g_I} does not satisfies the PBR condition. The contradiction proves that g_R must be a negative function. An analogue argument is used to prove that g_R must be a positive function when β is such that $\cos(\beta\pi/2) > 0$.■

A.15 Proofs regarding the stochastic Heat equation (Section 5.2.3)

A.15.1 Existence of stationary solutions

According to Theorem 4.3.1, there exists a stationary solution to the stochastic Heat equation with White Noise source term (5.56) if and only if the spatio-temporal measure

$$(\omega^2 + a^2|\xi|^4)^{-1} d\xi d\omega$$

is in $\mathscr{M}_{SG}^+(\mathbb{R}^d \times \mathbb{R})$. This would hold if the function $(\xi,\omega) \mapsto (\omega^2 + a^2|\xi|^4)^{-1}$ is locally integrable, the slow-growing behaviour being provided by the fact that this function is bounded outside every neighbourhood of

the origin. It suffices thus to study the integrability over subsets of $\mathbb{R}^d \times \mathbb{R}$ of the form $B_R^{(d)}(0) \times [-M, M]$ for $R, M > 0$, where $B_R^{(d)}(0) \subset \mathbb{R}^d$ is the ball of radius R centered in 0. Using integration with polar coordinates in the spatial domain and the symmetry in the time interval, we obtain

$$\int_{B_R^{(d)}(0)\times[-M,M]} \frac{1}{\omega^2 + a^2|\xi|^4} d(\xi,\omega) = C \int_0^R \arctan\left(\frac{M}{ar^2}\right) r^{d-3} dr \tag{A.90}$$

for some positive constant C. Since we have that $\arctan(\frac{M}{aR^2}) \leqslant \arctan(\frac{M}{ar^2}) \leqslant \frac{\pi}{2}$ for all $r \in [0, R]$, we conclude that the integral (A.90) is finite only for $d > 2$. We conclude that there exist stationary solutions to the SPDE (5.56) only for spatial dimensions $d \geqslant 3$. In these cases, the stationary solutions would be continuous Random Functions if the measure $(\omega^2 + a^2|\xi|^4)^{-1} d\xi d\omega$ was finite, which would hold if the limit when M and R go to ∞ would exist and was finite. However, by seeing that $\int_0^R \arctan(\frac{M}{ar^2}) r^{d-3} dr \geqslant \arctan(\frac{M}{aR^2}) \frac{R^{d-2}}{d-2}$, and by letting $M \to \infty$ first and $R \to \infty$ second, one gets that the limit is not finite. Hence, the stationary solutions to (5.56) in spatial dimensions higher that 2 have a meaning as GeRFs and not as continuous Random Functions.

A.15.2 Covariance structure for $d = 3$

The covariance distribution (5.57) is the Fourier Transform of the spatio-temporal spectral measure

$$d\mu_U(\xi,\omega) = (2\pi)^{-\frac{d+1}{2}} (\omega^2 + a^2|\xi|^4)^{-1} d\xi d\omega$$

for $d = 3$. This measure is not finite. The computation of the Fourier Transform $\rho_U = \mathscr{F}(\mu_U)$ is obtained as the limit in a distributional sense of continuous functions. Let us be precise.

Let $R > 0$ and let us denote by μ_U^R the restriction of the measure μ_U to the subset $B_R^{(3)}(0) \times \mathbb{R} \subset \mathbb{R}^3 \times \mathbb{R}$, i.e.

$$d\mu_U^R(\xi,\omega) = (2\pi)^{-\frac{d+1}{2}} (\omega^2 + a^2|\xi|^4)^{-1} \mathbf{1}_{B_R^{(3)}(0)}(\xi) d\xi d\omega.$$

This measure is even, positive and finite, so $\rho_U^R = \mathscr{F}(\mu_U^R)$ is a continuous positive-definite function over $\mathbb{R}^d \times \mathbb{R}$. Since for every $\psi \in \mathscr{S}(\mathbb{R}^3 \times \mathbb{R})$ we have $\langle \mu_U^R, \psi \rangle = \langle \mu_U, \psi \mathbf{1}_{B_R^{(3)}(0)\times\mathbb{R}} \rangle$, and that $\psi \mathbf{1}_{B_R^{(3)}\times\mathbb{R}} \to \psi$ point-wise and dominated by $|\psi| \in \mathscr{L}^1(\mathbb{R}^3 \times \mathbb{R}, \mu_U)$ as $R \to \infty$, we obtain by Dominated Convergence Theorem that

$$\langle \mu_U^R, \psi \rangle \to \langle \mu_U, \psi \rangle \quad \text{as } R \to \infty. \tag{A.91}$$

Hence, $\mu_U^R \to \mu_U$ as $R \to \infty$ in a $\mathscr{S}'(\mathbb{R}^3 \times \mathbb{R})-$weak-$*$ sense. By continuity of the Fourier Transform, we

have that $\rho_R^U \overset{\mathscr{S}'}{\to} \rho_U$. Let us calculate $\rho_U^R(h,u)$ for $(h,u) \in \mathbb{R}^3 \times \mathbb{R}$.

$$\begin{aligned}
\rho_U^R(h,u) &= \frac{1}{(2\pi)^4} \int_{\mathbb{B}_R^{(3)}(0)} \int_{\mathbb{R}} \frac{e^{-iu\omega - ih^T\xi}}{\omega^2 + a^2|\xi|^4} d\omega d\xi \\
&= \frac{1}{(2\pi)^3} \frac{1}{2a} \int_{B_R^{(3)}(0)} e^{-ih^T\xi} \frac{e^{-a|\xi|^2|u|}}{a|\xi|^2} d\xi \\
&= \frac{1}{(2\pi)^{\frac{3}{2}}} \frac{1}{2a} \sqrt{\frac{2}{\pi}} \int_0^R \frac{J_{\frac{1}{2}}(|h|r)}{\sqrt{r|h|}} e^{-a|u|r^2} dr \\
&= \frac{1}{(2\pi)^2} \frac{1}{a|h|} \int_0^R \frac{\sin(|h|r)}{r} e^{-a|u|r^2} dr. \qquad \text{(A.92)}
\end{aligned}$$

Here we have used the expression of the Fourier Transform of radial functions (Donoghue, 1969, Chapter 41). Let us evaluate the limit of $\rho_U^R(h,u)$ when $R \to \infty$ for $|h| \neq 0 \neq |u|$. Consider the function $f_R : \mathbb{R}^+ \to \mathbb{R}$ defined by $f_R(\lambda) = \int_0^R \frac{\sin(\lambda r)}{r} e^{-a|u|r^2} dr$ for $\lambda > 0$, and $f_R(0) = 0$. A typical application of the Dominated Convergence Theorem proves that f_R is continuous over $\mathbb{R}^+$ and differentiable over $\mathbb{R}_*^+$. Differentiating under the integral, we have that $f_R'(\lambda) = \int_0^R \cos(\lambda r) e^{-a|u|r^2} dr$ for $\lambda > 0$. Using the expression of the Fourier Transform of a Gaussian function, one proves that

$$\lim_{R\to\infty} f_R'(\lambda) = \sqrt{\frac{\pi}{4a|u|}} e^{-\frac{\lambda^2}{4a|u|}}, \qquad \text{(A.93)}$$

for every $\lambda > 0$. Using $f_R(\lambda) = \int_0^\lambda f_R'(s) ds$ and again the Dominated Convergence Theorem, we obtain

$$\lim_{R\to\infty} f_R(\lambda) = \int_0^\lambda \sqrt{\frac{\pi}{4a|u|}} e^{-\frac{s^2}{4a|u|}} ds = \frac{\pi}{2} \operatorname{erf}\left(\frac{\lambda}{2\sqrt{a|u|}}\right). \qquad \text{(A.94)}$$

Using this result in (A.92) with $\lambda = |h|$ and $R \to \infty$, we finally obtain the distribution *associated to the function*

$$\rho_U(h,u) = \frac{1}{(2\pi)^2} \frac{\pi}{2a|h|} \operatorname{erf}\left(\frac{|h|}{2\sqrt{a|u|}}\right)., \qquad \text{(A.95)}$$

which is the expression in (5.57).

It is worth emphasizing that this expression is only valid in a distributional sense. The distribution ρ_U is only meaningful when applied to test-functions, satisfying

$$\langle \rho_U, \psi \rangle = \lim_{R\to\infty} \langle \rho_U^R, \psi \rangle, \quad \forall \psi \in \mathscr{S}(\mathbb{R}^3 \times \mathbb{R}). \qquad \text{(A.96)}$$

The expression *associated to the function* (A.95) refers to the fact that for every test-function ψ *such that its*

support does not contain the origin, we have

$$\langle \rho_U, \psi \rangle = \int_{\mathbb{R}^3 \times \mathbb{R}} \frac{1}{(2\pi)^2} \frac{\pi}{2a|h|} \operatorname{erf}\left(\frac{|h|}{2\sqrt{a|u|}} \right) \psi(h,u) dh du. \tag{A.97}$$

This expression does not hold for a general test-function $\psi \in \mathscr{S}(\mathbb{R}^3 \times \mathbb{R})$.

A.16 Proof of Proposition 6.2.1

Let $\varphi \in \mathscr{S}(\mathbb{R}^d)$. By linearity, the expression $\langle Z, \varphi \rangle - \langle Z_N, \varphi \rangle$ is simply $\langle Z - Z_N, \varphi \rangle$. Setting $\phi = \mathscr{F}^{-1}(\varphi) \in \mathscr{S}(\mathbb{R}^d)$, we have

$$\langle Z - Z_N, \varphi \rangle = \langle \mathscr{F}(Z - Z_N), \phi \rangle.$$

We have that

$$\mathscr{F}(Z - Z_N) = M_Z - M_{Z_N}. \tag{A.98}$$

From the definition of M_{Z_N} one obtains that

$$\langle M_{Z_N}, \phi \rangle = \langle \sum_{j=1}^{N} M_Z(V_j^N) \delta_{\xi_j^N}, \phi \rangle = \langle M_Z, \sum_{j=1}^{N} \phi(\xi_j^N) \mathbf{1}_{V_j^N} \rangle. \tag{A.99}$$

Hence,

$$\begin{aligned} \mathbb{E}\left(|\langle Z_N - Z, \varphi \rangle|^2 \right) &= \mathbb{E}\left(|\langle M_{Z_N} - M_Z, \phi \rangle|^2 \right) \\ &= \mathbb{E}\left(\left| \langle M_Z, \phi - \sum_{j=1}^{N} \phi(\xi_j^N) \mathbf{1}_{V_j^N} \rangle \right|^2 \right) \\ &= (2\pi)^{\frac{d}{2}} \int_{\mathbb{R}^d} \left| \phi(\xi) - \sum_{j=1}^{N} \phi(\xi_j^N) \mathbf{1}_{V_j^N}(\xi) \right|^2 d\mu_Z(\xi). \end{aligned} \tag{A.100}$$

Consider the sequence of functions $\phi_N = \sum_{j=1}^{N} \phi(\xi_j^N) \mathbf{1}_{V_j^N}$. Let us verify that it converges point-wise to ϕ. Let $\xi \in \mathbb{R}^d$. Since the union of the Riemann sequence of partitions $(V_j^N)_{j \in \{1,\dots,N\}, N \in \mathbb{N}_*}$ grows to $\mathbb{R}^d$ as N grows, there exists $N_0 \in \mathbb{N}_*$ such that for all $N \geqslant N_0$, $\xi \in \bigcup_{j=1}^{N} V_j^N$, and it belongs, of course, to just one of the V_j^N's. Since ϕ is continuous at ξ, for every $\epsilon > 0$ we can find a $\delta > 0$ such that $|\phi(\xi) - \phi(\eta)| < \epsilon$ if $|\xi - \eta| < \delta$. By taking $N_1 \in \mathbb{N}$ large enough such that if $N \geqslant N_1$, $\ell_N < \delta$, one gets that for $N \geqslant \max\{N_0, N_1\}$,

$$|\phi_N(\xi) - \phi(\xi)| = |\phi(\xi) - \phi(\xi_j^N)| < \epsilon, \tag{A.101}$$

where ξ_j^N is the tag point of the set V_j^N which contains ξ.

Consider now $M \in \mathbb{N}$ such that $(1+|\xi|^2)^{-2M}\mu_Z \in \mathscr{M}_F^+(\mathbb{R}^d)$. Expression (A.100) equals to

$$(2\pi)^{\frac{d}{2}} \int_{\mathbb{R}^d} \left|(1+|\xi|^2)^M \left(\phi(\xi) - \phi_N(\xi)\right)\right|^2 \frac{d\mu_Z(\xi)}{(1+|\xi|^2)^{2M}}. \tag{A.102}$$

Consider the sequence of functions defined through $f_N(\xi) = \left|(1+|\xi|^2)^M \left(\phi(\xi) - \phi_N(\xi)\right)\right|^2$. Since $\phi_N \to \phi$ point-wise, $f_N \to 0$ point-wise. It is clear that

$$|f_N(\xi)| \leqslant \underbrace{\sup_{\eta\in\mathbb{R}^d} \left|(1+|\eta|^2)^M \phi(\eta)\right|^2}_{<\infty \text{ since } \phi\in\mathscr{S}(\mathbb{R}^d)} + \sup_{\eta\in\mathbb{R}^d} \left|(1+|\eta|^2)^M \phi_N(\eta)\right|^2, \quad \forall \xi \in \mathbb{R}^d. \tag{A.103}$$

By convexity of the function $\xi \in \mathbb{R}^d \mapsto |\xi|^2$, it holds that

$$(1+|\xi|^2)^M \leqslant (1+2|\xi-\xi_j^N|^2 + 2|\xi_j^N|^2)^M \leqslant 2^M (1+|\xi-\xi_j^N|^2 + |\xi_j^N|^2)^M.$$

Using a binomial expansion, we have that for every $\xi \in \mathbb{R}^d$,

$$\begin{aligned}
(1+|\xi|^2)^M |\phi_N(\xi)| &= \sum_{j=1}^N (1+|\xi|^2)^M |\phi(\xi_j^N)| \mathbf{1}_{V_j^N}(\xi) \\
&\leqslant \sum_{j=1}^N 2^M (1+|\xi_j^N|^2 + |\xi-\xi_j^N|^2)^M |\phi(\xi_j^N)| \mathbf{1}_{V_j^N}(\xi) \\
&= 2^M \sum_{j=1}^N \sum_{k=0}^M \binom{M}{k} \underbrace{(1+|\xi_j^N|^2)^{M-k} |\phi(\xi_j^N)|}_{\leqslant \sup_{\eta\in\mathbb{R}^d} |(1+|\eta|^2)^{M-k}\phi(\eta)|} \underbrace{|\xi-\xi_j^N|^{2k}}_{\leqslant \ell_N^{2k}} \mathbf{1}_{V_j^N}(\xi) \\
&\leqslant 2^M \sum_{k=0}^M \sup_{\eta\in\mathbb{R}^d} \left|(1+|\eta|^2)^{M-k}\phi(\eta)\right| \ell_N^{2k}
\end{aligned} \tag{A.104}$$

Since $\phi \in \mathscr{S}(\mathbb{R}^d)$ and $\ell_N \to 0$ this expression is bounded by a constant which does neither depend on N nor on ξ, which we will denote by $C_{\phi,M} > 0$. It follows that

$$|f_N| \leqslant \sup_{\eta\in\mathbb{R}^d} \left|(1+|\eta|^2)^M \phi(\eta)\right|_\infty^2 + C_{\phi,M}^2. \tag{A.105}$$

Hence, the sequence $(f_N)_{N\in\mathbb{N}_*}$ is dominated by a constant, which is of course integrable with respect to the finite measure $(1+|\xi|^2)^{-2M}\mu_Z$. By Dominated Convergence Theorem, expression (A.102) goes to 0 as N grows, which proves the result. ■

A.17 Proof of Theorem 6.2.1

We consider the expressions

$$Z(x) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{ix^T\xi} dM_Z(\xi) \quad ; \quad Z_N(x) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{ix^T\xi} dM_{Z_N}(\xi). \tag{A.106}$$

Considering that

$$\int_{\mathbb{R}^d} e^{ix^T\xi} dM_{Z_N}(\xi) = \sum_{j=1}^{N} e^{ix^T\xi_j^N} M_Z(V_j^N) = \int_{\mathbb{R}^d} \sum_{j=1}^{N} e^{ix^T\xi_j^N} \mathbf{1}_{V_j^N}(\xi) dM_Z(\xi), \tag{A.107}$$

and using the simple fact $e^{ix^T\xi} = \sum_{j=1}^{N} e^{ix^T\xi} \mathbf{1}_{V_j^N}(\xi) + e^{ix^T\xi} \mathbf{1}_{D_N}(\xi)$ for all $x \in \mathbb{R}^d$ and $\xi \in \mathbb{R}^d$, we obtain

$$\begin{aligned}
\mathbb{E}\left(|Z(x) - Z_N(x)|^2\right) &= \frac{1}{(2\pi)^d} \mathbb{E}\left(\left| \int_{\mathbb{R}^d} e^{ix^T\xi} dM_Z(\xi) - \int_{\mathbb{R}^d} e^{ix^T\xi} dM_{Z_N}(\xi) \right|^2 \right) \\
&= \frac{1}{(2\pi)^d} \mathbb{E}\left(\left| \int_{\mathbb{R}^d} e^{ix^T\xi} - \sum_{j=1}^{N} e^{ix^T\xi_j^N} \mathbf{1}_{V_j^N}(\xi) dM_Z(\xi) \right|^2 \right) \\
&= \frac{1}{(2\pi)^d} \mathbb{E}\left(\left| \int_{\mathbb{R}^d} \sum_{j=1}^{N} (e^{ix^T\xi} - e^{ix^t\xi_j^N}) \mathbf{1}_{V_j^N}(\xi) + e^{ix^T\xi} \mathbf{1}_{D_N}(\xi) dM_Z(\xi) \right|^2 \right) \\
&= \frac{1}{(2\pi)^d} \mathbb{E}\left(\left| \sum_{j=1}^{N} \int_{V_j^N} e^{ix^T\xi} - e^{ix^T\xi_j^N} dM_Z(\xi) + \int_{D_N} e^{ix^T\xi} dM_Z(\xi) \right|^2 \right).
\end{aligned} \tag{A.108}$$

Since M_Z is orthogonal, the stochastic integrals of the form $\int_{V_j^N} e^{ix^T\xi} - e^{ix^T\xi_j^N} dM_Z(\xi)$ and $\int_{D_N} e^{ix^\xi} dM_Z(\xi)$ are mutually non-correlated for all $j \in \{1, ..., N\}$ and $N \in \mathbb{N}_*$. Hence, expression (A.108) equals to

$$\begin{aligned}
&\frac{1}{(2\pi)^d} \left[\sum_{j=1}^{N} \mathbb{E}\left(\left| \int_{V_j^N} e^{ix^T\xi} - e^{ix^T\xi_j^N} dM_Z(\xi) \right|^2 \right) + \mathbb{E}\left(\left| \int_{D_N} e^{ix^T\xi} dM_Z(\xi) \right|^2 \right) \right] \\
&= \frac{1}{(2\pi)^d} \left(\sum_{j=1}^{N} (2\pi)^{\frac{d}{2}} \int_{V_j^N} \left| e^{ix^T\xi} - e^{ix^T\xi_j^N} \right|^2 d\mu_Z(\xi) + (2\pi)^{\frac{d}{2}} \int_{D_N} |e^{ix^T\xi}|^2 d\mu_Z(\xi) \right) \\
&= \frac{1}{(2\pi)^{\frac{d}{2}}} \left(\sum_{j=1}^{N} \int_{V_j^N} \left| e^{ix^T\xi} - e^{ix^T\xi_j^N} \right|^2 d\mu_Z(\xi) + \mu_Z(D_N) \right).
\end{aligned} \tag{A.109}$$

Using a typical application of Taylor's Theorem, one obtains that

$$|e^{ix^T\xi} - e^{ix^T\xi_j^N}| \leqslant 2|x||\xi - \xi_j^N| \leqslant 2|x|\ell_N, \quad \forall \xi \in V_j^N. \tag{A.110}$$

Applying this to expression (A.109), we obtain

$$\begin{aligned}
\mathbb{E}\left(|Z(x) - Z_N(x)|^2\right) &= \frac{1}{(2\pi)^{\frac{d}{2}}}\left(\int_{V_j^N}\left|e^{ix^T\xi} - e^{ix^T\xi_j^N}\right|^2 d\mu_Z(\xi) + \mu_Z(D_N)\right) \\
&\leqslant \frac{1}{(2\pi)^{\frac{d}{2}}}\left[4|x|^2\ell_N^2\mu_Z\left(\bigcup_{j=1}^N V_j^N\right) + \mu_Z(D_N)\right] \\
&\leqslant \frac{1}{(2\pi)^{\frac{d}{2}}}\left[4|x|^2\ell_N^2\mu_Z(\mathbb{R}^d) + \mu_Z(D_N)\right].
\end{aligned} \tag{A.111}$$

Since μ_Z is a positive finite measure, we conclude that for any $K \subset \mathbb{R}^d$ compact,

$$\sup_{x\in K}\mathbb{E}\left(|Z(x) - Z_N(x)|^2\right) \leqslant \frac{1}{(2\pi)^{\frac{d}{2}}}\Big[4\underbrace{\ell_N^2}_{\to 0 \text{ as } N\to\infty}\underbrace{\mu_Z(\mathbb{R}^d)\sup_{x\in K}|x|^2}_{<\infty} + \underbrace{\mu_Z(D_N)}_{\to 0 \text{ as } N\to\infty}\Big] \to 0. \tag{A.112}$$

This proves the mean-square-uniformly on compacts convergence and gives a bound for analysing the rate of convergence. ■

A.18 Proof of Proposition 6.2.2

Let us consider the case with the addition of the Random Function R_N. By definition of R_N, one has that

$$R_N(x) = \frac{1}{(2\pi)^{\frac{d}{2}}}\int_{D_N} e^{ix^T d_N} dM_Z(\xi). \tag{A.113}$$

Following the same procedure and arguments as in (A.108) and (A.109), one gets that for all $x \in \mathbb{R}^d$

$$\begin{aligned}
\mathbb{E}\left(|Z(x) - (Z_N(x) + R_N(x))|^2\right) &= \frac{1}{(2\pi)^d}\mathbb{E}\left(\left|\sum_{j=1}^N\int_{V_j^N} e^{ix^T\xi} - e^{ix^T\xi_j^N} dM_Z(\xi) + \int_{D_N} e^{ix^T\xi} - e^{ix^T d_N} dM_Z(\xi)\right|^2\right) \\
&= \frac{1}{(2\pi)^{\frac{d}{2}}}\left(\sum_{j=1}^N\int_{V_j^N}\left|e^{ix^T\xi} - e^{ix^T\xi_j^N}\right|^2 d\mu_Z(\xi) + \int_{D_N}|e^{ix^T\xi} - e^{ix^T d_N}|^2 d\mu_Z(\xi)\right) \\
&\leqslant \frac{1}{(2\pi)^{\frac{d}{2}}}\Big[4\ell_N^2\mu_Z(\mathbb{R}^d)|x|^2 + 2\mu_Z(D_N)\Big].
\end{aligned} \tag{A.114}$$

Hence,

$$\sup_{x\in K}\mathbb{E}\left(|Z(x)-(Z_N(x)+R_N(x))|^2\right)\leqslant\frac{2}{(2\pi)^{\frac{d}{2}}}\left[2\ell_N^2\mu_Z(\mathbb{R}^d)\sup_{x\in K}|x|^2+\mu_Z(D_N)\right]\to 0, \tag{A.115}$$

which proves that Z_N+R_N converges mean-square-uniformly on compacts to Z, and also proves the bound (6.13).

The convergence of $\frac{\sigma_Z}{\sigma_{Z_N}}Z_N$ to Z using the triangular inequality:

$$\begin{aligned}\mathbb{E}\left(\left|\frac{\sigma_Z}{\sigma_{Z_N}}Z_N(x)-Z(x)\right|^2\right)&=\frac{1}{\sigma_{Z_N}^2}\mathbb{E}\left(|(\sigma_Z-\sigma_{Z_N})Z_N(x)+(Z_N(x)-Z(x))\sigma_{Z_N}|^2\right)\\&\leqslant\frac{2}{\sigma_{Z_N}^2}\left[\mathbb{E}\left(|(\sigma_Z-\sigma_{Z_N})Z_N(x)|^2\right)+\mathbb{E}\left(|\sigma_{Z_N}(Z_N(x)-Z(x))|^2\right)\right]\\&=\frac{2}{\sigma_{Z_N}^2}\left(|\sigma_Z-\sigma_{Z_N}|^2\sigma_{Z_N^2}+\sigma_{Z_N^2}\mathbb{E}\left(|Z_N(x)-Z(x)|^2\right)\right)\\&=2|\sigma_Z-\sigma_{Z_N}|^2+2\mathbb{E}\left(|Z_N(x)-Z(x)|^2\right).\end{aligned} \tag{A.116}$$

Since $\sigma_{Z_N}\to\sigma_Z$ and from Theorem 6.2.1 the convergence in the sense of mean-square-uniformly on compact sets follows. The bound (6.14) also follows immediately.

The equality of the variances between $\frac{\sigma_Z}{\sigma_Z}Z_N$ is straightforward. The equality between the variances of Z_N+R_N and Z follows from

$$\begin{aligned}\mathbb{V}ar(Z(x))=\mathbb{V}ar(Z(0))=\mathbb{V}ar\left(\frac{1}{(2\pi)^{\frac{d}{2}}}M_Z(\mathbb{R}^d)\right)=\mathbb{V}ar\left(\frac{1}{(2\pi)^{\frac{d}{2}}}(M_{Z_N}(\mathbb{R}^d)+M_{R_N}(\mathbb{R}^d))\right)\\=\mathbb{V}ar(Z_N(0)+R_N(0))=\mathbb{V}ar(Z_N(x)+R_N(x)).\end{aligned} \tag{A.117}$$

Here we have used that R_N and Z_N are non-correlated stationary random Functions and that

$$M_{Z_N}(\mathbb{R}^d)+M_{R_N}(\mathbb{R}^d)=M_Z(\mathbb{R}^d), \tag{A.118}$$

which is easy to conclude from the definition of R_N. ■

A.19 Proof of Proposition 6.2.3

We follow the same arguments as in the proof of Proposition 6.2.1. We consider $\varphi \in \mathscr{S}(\mathbb{R}^d)$ and we set $\phi = \mathscr{F}^{-1}(\varphi)$. We call

$$M_{X_N} := \mathscr{F}(X_N) = \sum_{j=1}^{N} M_X(V_j^N)\delta_{\xi_j^N}. \tag{A.119}$$

Using that

$$\langle M_{X_N}, \frac{1}{g}\phi\rangle = \sum_{j=1}^{N} M_X(V_j^N)\frac{\phi(\xi_j^N)}{g(\xi_j^N)} = \langle M_X, \sum_{j=1}^{N}\frac{\phi(\xi_j^N)}{g(\xi_j^N)}\mathbf{1}_{V_j^N}\rangle, \tag{A.120}$$

we conclude that

$$\begin{aligned}
\mathbb{E}\left(|\langle U,\varphi\rangle - \langle U_N,\varphi\rangle|^2\right) &= \mathbb{E}\left(\left|\langle \mathcal{L}_{\frac{1}{g}}X,\varphi\rangle - \langle \mathcal{L}_{\frac{1}{g}}X_N,\varphi\rangle\right|^2\right)\\
&= \mathbb{E}\left(\left|\langle M_X, \frac{1}{g}\phi\rangle - \langle M_{X_N}, \frac{1}{g}\phi\rangle\right|^2\right)\\
&= \mathbb{E}\left(\left|\langle M_X, \frac{1}{g}\phi - \sum_{j=1}^{N}\frac{\phi(\xi_j^N)}{g(\xi_j^N)}\mathbf{1}_{V_j^N}\rangle\right|^2\right)\\
&= (2\pi)^{\frac{d}{2}}\int_{\mathbb{R}^d}\left|\frac{\phi(\xi)}{g(\xi)} - \sum_{j=1}^{N}\frac{\phi(\xi_j^N)}{g(\xi_j^N)}\mathbf{1}_{V_j^N}(\xi)\right|^2 d\mu_X(\xi)\\
&= (2\pi)^{\frac{d}{2}}\left[\int_{\mathbb{R}^d}\left|\frac{\phi(\xi)}{g(\xi)} - \sum_{j=1}^{N}\frac{\phi(\xi_j^N)}{g(\xi_j^N)}\mathbf{1}_{V_j^N}(\xi)\right|^2 d\mu_X(\xi) + \int_{D_N}\left|\frac{\phi(\xi)}{g(\xi)}\right|^2 d\mu_X(\xi)\right].
\end{aligned} \tag{A.121}$$

The convergence to 0 of the first integral is concluded using Dominated Convergence Theorem following the same arguments as in the proof of Proposition 6.2.1. Namely, using the continuity of $\frac{\phi}{g}$ one proves that

$$\left|\frac{\phi(\xi)}{g(\xi)} - \sum_{j=1}^{N}\frac{\phi(\xi_j^N)}{g(\xi_j^N)}\mathbf{1}_{V_j^N}(\xi)\right|^2 \to 0 \tag{A.122}$$

as N grows for any $\xi \in \mathbb{R}^d$. Using that g satisfies the PBR condition, there exists $c > 0$ and $m \in \mathbb{N}$ such that $|g(\xi)| \geqslant \frac{c}{(1+|\xi|^2)^m}$ for all $\xi \in \mathbb{R}^d$, and hence $|\frac{\phi(\xi)}{g(\xi)}| \leqslant \frac{1}{c}|(1+|\xi|^2)^m\phi(\xi)|$ for all $\xi \in \mathbb{R}^d$. We apply then the same procedures as in equations (A.102), (A.103), and (A.104) to prove that the convergence is dominated,

using M such that $(1+|\xi|^2)^{-2M}\mu_Z$ is finite, replacing ϕ_N by $\sum_{j=1}^{N}\frac{\phi(\xi_j^N)}{g(\xi_j^N)}\mathbf{1}_{V_j^N}(\xi)$ and $f_N(\xi)$ by

$$\left|(1+|\xi|^2)^M\left(\frac{\phi(\xi)}{g(\xi)}-\sum_{j=1}^{N}\frac{\phi(\xi_j^N)}{g(\xi_j^N)}\mathbf{1}_{V_j^N}(\xi)\right)\right|^2,$$

and taking advantage from the inequalities $|\frac{\phi(\xi_j^N)}{g(\xi_j^N)}|\leqslant\frac{1}{c}|(1+|\xi_j^N|^2)^m\phi(\xi_j^N)|$ in (A.104).

Finally, the integral

$$\int_{D_N}\left|\frac{\phi(\xi)}{g(\xi)}\right|^2 d\mu_X(\xi) \tag{A.123}$$

vanishes as $N\to 0$ since $|\frac{\phi}{g}|^2$ is integrable with respect to μ_X and D_N decreases to $\varnothing$. ■

A.20 Proof of Theorem 6.2.2

We follow a similar approach as in Theorem 6.2.2. Since $|g|^{-2}$ is integrable with respect to μ_X we have that $\frac{1}{g}$ is integrable with respect to the Random Measure M_X (Section 3.3.6). We recall that we have supposed that g satisfies the PBR condition, hence $|g|>0$. We consider that the unique stationary solution U to (6.18) can be written as

$$U(x)=\frac{1}{(2\pi)^{\frac{d}{2}}}\int_{\mathbb{R}^d}\frac{e^{ix^T\xi}}{g(\xi)}dM_X(\xi). \tag{A.124}$$

And the approximation U_N can be written as

$$U_N(x)=\frac{1}{(2\pi)^{\frac{d}{2}}}\int_{\mathbb{R}^d}\frac{e^{ix^T\xi}}{g(\xi)}dM_{X_N}(\xi)=\frac{1}{(2\pi)^{\frac{d}{2}}}\int_{\mathbb{R}^d}\sum_{j=1}^{N}\frac{e^{ix^T\xi_j^N}}{g(\xi_j^N)}\mathbf{1}_{V_j^N}(\xi)dM_X(\xi). \tag{A.125}$$

Using the orthogonality of M_X we obtain that

$$\begin{aligned}
\mathbb{E}\left(|U(x)-U_N(x)|^2\right) &= \frac{1}{(2\pi)^d}\mathbb{E}\left(\left|\int_{\mathbb{R}^d}\frac{e^{ix^T\xi}}{g(\xi)}-\sum_{j=1}^N\frac{e^{ix^T\xi_j^N}}{g(\xi_j^N)}\mathbf{1}_{V_j^N}(\xi)dM_X(\xi)\right|^2\right)\\
&= \frac{1}{(2\pi)^d}\mathbb{E}\left(\left|\int_{\mathbb{R}^d}\sum_{j=1}^N\left(\frac{e^{ix^T\xi}}{g(\xi)}-\frac{e^{ix^T\xi_j^N}}{g(\xi_j^N)}\right)\mathbf{1}_{V_j^N}(\xi)dM_X(\xi)+\int_{D_N}\frac{e^{ix^T\xi}}{g(\xi)}dM_X(\xi)\right|^2\right)\\
&= \frac{1}{(2\pi)^{\frac{d}{2}}}\left[\int_{\mathbb{R}^d}\sum_{j=1}^N\left|\frac{e^{ix^T\xi}}{g(\xi)}-\frac{e^{ix^T\xi_j^N}}{g(\xi_j^N)}\right|^2\mathbf{1}_{V_j^N}(\xi)d\mu_X(\xi)+\int_{D_N}\left|\frac{e^{ix^T\xi}}{g(\xi)}\right|^2d\mu_X(\xi)\right]\\
&= \frac{1}{(2\pi)^{\frac{d}{2}}}\left[\int_{\mathbb{R}^d}\sum_{j=1}^N\left|\frac{e^{ix^T\xi}}{g(\xi)}-\frac{e^{ix^T\xi_j^N}}{g(\xi)}+\frac{e^{ix^T\xi_j^N}}{g(\xi)}-\frac{e^{ix^T\xi_j^N}}{g(\xi_j^N)}\right|^2\mathbf{1}_{V_j^N}(\xi)d\mu_X(\xi)+\int_{D_N}\frac{d\mu_X(\xi)}{|g(\xi)|^2}\right]\\
&\leqslant \frac{1}{(2\pi)^{\frac{d}{2}}}\left\{2\int_{\mathbb{R}^d}\sum_{j=1}^N\left|e^{ix^T\xi}-e^{ix^T\xi_j^N}\right|^2\mathbf{1}_{V_j^N}(\xi)\frac{d\mu_X(\xi)}{|g(\xi)|^2}\right.\\
&\qquad\left.+2\int_{\mathbb{R}^d}\sum_{j=1}^N\left|\frac{1}{g(\xi)}-\frac{1}{g(\xi_j^N)}\right|^2\mathbf{1}_{V_j^N}(\xi)d\mu_X(\xi)+\int_{D_N}\frac{d\mu_X(\xi)}{|g(\xi)|^2}\right\}\\
&\leqslant \frac{1}{(2\pi)^{\frac{d}{2}}}\left\{\underbrace{8|x|^2\ell_N^2\int_{\mathbb{R}^d}\frac{d\mu_X(\xi)}{|g(\xi)|^2}}_{\to 0}+2\int_{\mathbb{R}^d}\sum_{j=1}^N\left|1-\frac{g(\xi)}{g(\xi_j^N)}\right|^2\mathbf{1}_{V_j^N}(\xi)\frac{d\mu_X(\xi)}{|g(\xi)|^2}+\underbrace{\int_{D_N}\frac{d\mu_X(\xi)}{|g(\xi)|^2}}_{\to 0}\right\}.
\end{aligned}\tag{A.126}$$

Here we have used an inequality of the form $|x+y|^2\leqslant 2(|x|^2+|y|^2)$ and the bound (A.110). The vanishing integrals are justified by the integrability of $|g|^{-2}$ with respect to μ_X and the fact that D_N decreases to $\varnothing$. For every $K\subset\mathbb{R}^d$ compact, we can take the supremum over all $x\in K$ in procedure (A.126) and the vanishing integrals still going to zero. We need then to verify the convergence to 0 of the integral

$$\int_{\mathbb{R}^d}\sum_{j=1}^N\left|1-\frac{g(\xi)}{g(\xi_j^N)}\right|^2\mathbf{1}_{V_j^N}(\xi)\frac{d\mu_X(\xi)}{|g(\xi)|^2}.\tag{A.127}$$

From the continuity of g and using the same arguments which concluded (A.101) in the proof of Proposition 6.2.1, on proves the point-wise convergence

$$\sum_{j=1}^N\left|1-\frac{g(\xi)}{g(\xi_j^N)}\right|^2\mathbf{1}_{V_j^N}(\xi)\to 0,\quad \text{as } N\to\infty,\forall\xi\in\mathbb{R}^d.\tag{A.128}$$

Let us prove that the convergence is dominated. For this, we use condition (6.27) to conclude that

$$\left|1-\frac{g(\xi)}{g(\xi_j^N)}\right|^2\leqslant 1+\left|\frac{g(\xi)}{g(\xi_j^N)}\right|^2\leqslant 1+\frac{C_2^2}{C_1^2}\left(\frac{1+|\xi|^2}{1+|\xi_j^N|^2}\right)^\alpha,\quad \forall\xi\in\mathbb{R}^d,\forall j\in\{1,...,N\}.\tag{A.129}$$

From the mean value theorem one can conclude the inequality:

$$\left|\frac{\log(1+u^2)-\log(1+v^2)}{u-v}\right| \leqslant \sup_{w\geqslant 0}\left|\frac{2w}{1+w^2}\right| \leqslant 1, \quad \forall u,v \geqslant 0, \tag{A.130}$$

from where we can conclude that

$$\left|\log\left(\frac{1+|\xi|^2}{1+|\xi_j^N|^2}\right)\right| \leqslant \big||\xi|-|\xi_j^N|\big| \leqslant |\xi-\xi_j^N|. \tag{A.131}$$

Using the monotony of the logarithm, one obtains ($\wedge$ denotes minimum and $\vee$ maximum)

$$\log\left(\frac{1+(|\xi|\vee|\xi_j^N|)^2}{1+(|\xi|\wedge|\xi_j^N|)^2}\right) = \left|\log\left(\frac{1+|\xi|^2}{1+|\xi_j^N|^2}\right)\right| \leqslant |\xi-\xi_j^N|, \tag{A.132}$$

and hence

$$\frac{1+(|\xi|\vee|\xi_j^N|)^2}{1+(|\xi|\wedge|\xi_j^N|)^2} \leqslant e^{|\xi-\xi_j^N|}. \tag{A.133}$$

We obtain thus that for $\xi \in V_j^N$:

$$\left|1-\frac{g(\xi)}{g(\xi_j^N)}\right|^2 \leqslant 1+\frac{C_2^2}{C_1^2}\left(\frac{1+|\xi|^2}{1+|\xi_j^N|^2}\right)^{\alpha} \leqslant 1+\frac{C_2^2}{C_1^2}\left(\frac{1+(|\xi|\vee|\xi_j^N|)^2}{1+(|\xi|\wedge|\xi_j^N|)^2}\right)^{\alpha} \leqslant 1+\frac{C_2^2}{C_1^2}e^{\alpha|\xi-\xi_j^N|} \leqslant 1+\frac{C_2^2}{C_1^2}e^{\alpha\ell_N}. \tag{A.134}$$

Since $\ell_N \to 0$ and $N \to \infty$, we obtain thus that the convergence (A.128) is dominated by the constant $1+\frac{C_2^2}{C_1^2}e^{\alpha\sup_{N\in\mathbb{N}_*}\ell_N}$. Since $|g|^{-2}$ is integrable with respect to μ_X, from Dominated Convergence Theorem we conclude that the integral (A.127) vanishes as $N \to \infty$. This proves the mean-square-uniformly on compacts convergence. ■

A.21 Proof of Proposition 6.2.4

Let $Y = \mathscr{F}_S(X_S) \boxtimes X_T$ (symbolically). Since X is stationary and its temporal trace X_T is a Random Measure, we have that Y is a Random Measure over $\mathbb{R}^d \times \mathbb{R}$. Let $C_{X_T} \in \mathscr{M}_{SG}(\mathbb{R}\times\mathbb{R})$ be the covariance measure of X_T, and let μ_{X_S} be the spectral measure of the spatial trace X_S. The covariance measure of Y is then determined by

$$\langle C_Y, \psi\rangle = (2\pi)^{\frac{d}{2}}\int_{\mathbb{R}\times\mathbb{R}}\int_{\mathbb{R}^d}\psi((\xi,t),(\xi,s))d\mu_{X_S}(\xi)dC_{X_T}(t,s), \quad \forall\psi\in\mathscr{S}((\mathbb{R}^d\times\mathbb{R})\times(\mathbb{R}^d\times\mathbb{R})), \tag{A.135}$$

form which we remark that Y has the structure of an orthogonal Random Measure in space, but not necessarily on time.

We set as usual $V_0 = \mathscr{F}_S(U_0)$ and $V_{0,N}$ its approximation given by (6.35). We consider the solution to the transformed problem (6.36), expressed through its càdlàg-in-time representation $(V_{N,t})_{t\in\mathbb{R}^+}$ given by equation (6.43). Let V be the solution to (6.31), and let $(V_t)_{t\in\mathbb{R}^+}$ be its càdlàg-in-time representation. It is clear that (6.45) holds if and only if it also holds for the spatial Fourier Transforms V_N and V of U_N and U respectively, hence it suffices to analyse the mean-square-$\mathscr{S}'(\mathbb{R}^d)$-weak-$*$ in space and point-wise in time convergence of $(V_N)_{N\in\mathbb{N}_*}$ to V. Let $\varphi \in \mathscr{S}(\mathbb{R}^d)$. Using the independence of V_0 and Y we have

$$\mathbb{E}\left(\left|\langle V_t - V_{N,t}, \varphi\rangle\right|^2\right) = \mathbb{E}\left(\left|\langle V_0 - V_{0,N}, e^{-tg}\varphi\rangle\right|^2\right) + \mathbb{E}\left(\left|\int_{\mathbb{R}^d\times[0,t]} e^{-(t-s)g(\xi)}\varphi(\xi)d(Y-Y_N)(\xi,s)\right|^2\right). \tag{A.136}$$

For the first expectation we consider that

$$\langle V_{0,N}, e^{-tg}\varphi\rangle = \langle V_0, \sum_{j=1}^N e^{-tg(\xi_j^N)}\varphi(\xi_j^N)\mathbf{1}_{V_j^N}\rangle, \tag{A.137}$$

and hence

$$\begin{aligned}
\mathbb{E}\left(\left|\langle V_0 - V_{0,N}, e^{-tg}\varphi\rangle\right|^2\right) &= \mathbb{E}\left(\left|\left\langle V_0, e^{-tg}\varphi - \sum_{j=1}^N e^{-tg(\xi_j^N)}\varphi(\xi_j^N)\mathbf{1}_{V_j^N}\right\rangle\right|^2\right) \\
&= (2\pi)^{\frac{d}{2}}\int_{\mathbb{R}^d}\left|e^{-tg(\xi)}\varphi(\xi) - \sum_{j=1}^N e^{-tg(\xi_j^N)}\varphi(\xi_j^N)\mathbf{1}_{V_j^N}\right|^2 d\mu_{U_0}(\xi) \\
&= (2\pi)^{\frac{d}{2}}\int_{\mathbb{R}^d}\sum_{j=1}^N\left|e^{-tg(\xi)}\varphi(\xi) - e^{-tg(\xi_j^N)}\varphi(\xi_j^N)\right|^2\mathbf{1}_{V_j^N}(\xi)d\mu_{U_0}(\xi) \\
&\quad + (2\pi)^{\frac{d}{2}}\int_{D_N}|e^{-tg(\xi)}\varphi(\xi)|^2 d\mu_{U_0}(\xi).
\end{aligned} \tag{A.138}$$

Using the same arguments using Dominated Convergence Theorem as in the proof the Proposition 6.2.1. and considering that $g_R \geqslant 0$ and hence $|e^{-tg}| \leqslant 1$, we obtain that (A.138) vanishes as $N \to \infty$.

Let us now bound the second expectation in the right side of (A.136). Using that

$$\int_{\mathbb{R}^d\times[0,t]} e^{-(t-s)g(\xi)}\varphi(\xi)dY_N(\xi,s) = \int_{\mathbb{R}^d\times[0,t]}\sum_{j=1}^N e^{-(t-s)g(\xi_j^N)}\varphi(\xi_j^N)\mathbf{1}_{V_j^N}(\xi)dY(\xi,s), \tag{A.139}$$

we obtain that

$$
\begin{aligned}
&\mathbb{E}\left(\left|\int_{\mathbb{R}^d\times[0,t]} e^{-(t-s)g(\xi)}\varphi(\xi)d(Y-Y_N)(\xi,s)\right|^2\right)\\
&=\mathbb{E}\left(\left|\int_{\mathbb{R}^d\times[0,t]} e^{-(t-s)g(\xi)}\varphi(\xi)-\sum_{j=1}^N e^{-(t-s)g(\xi_j^N)}\varphi(\xi_j^N)\mathbf{1}_{V_j^N}(\xi)dY(\xi,s)\right|^2\right)\\
&=\mathbb{E}\left(\left|\int_{\mathbb{R}^d\times[0,t]}\sum_{j=1}^N\left(e^{-(t-s)g(\xi)}\varphi(\xi)-e^{-(t-s)g(\xi_j^N)}\varphi(\xi_j^N)\right)\mathbf{1}_{V_j^N}(\xi)dY(\xi,s)+\int_{D_N\times[0,t]}e^{-(t-s)g(\xi)}dY(\xi,s)\right|^2\right)\\
&=(2\pi)^{\frac{d}{2}}\int_{\mathbb{R}^d}\int_{[0,t]\times[0,s]}\Bigg\{\\
&\sum_{j=1}^N\left(e^{-(t-u)g(\xi)}\varphi(\xi)-e^{-(t-u)g(\xi_j^N)}\varphi(\xi_j^N)\right)\left(e^{-(s-v)\overline{g}(\xi)}\overline{\varphi}(\xi)-e^{-(t-u)\overline{g}(\xi_j^N)}\overline{\varphi}(\xi_j^N)\right)\mathbf{1}_{V_j^N}(\xi)\Bigg\}dC_{X_T}(u,v)d\mu_{X_S}(\xi)\\
&+(2\pi)^{\frac{d}{2}}\int_{D_N}\int_{[0,t]\times[0,s]}e^{-(t-u)g(\xi)-(s-v)\overline{g}(\xi)}|\varphi(\xi)|^2dC_{X_T}(u,v)d\mu_{X_S}(\xi).
\end{aligned}
\tag{A.140}
$$

Following the approach of Proposition 6.2.1, one proves that for all $\xi\in\mathbb{R}^d$,

$$
\sum_{j=1}^N\left(e^{-(t-u)g(\xi)}\varphi(\xi)-e^{-(t-u)g(\xi_j^N)}\varphi(\xi_j^N)\right)\left(e^{-(s-v)\overline{g}(\xi)}\overline{\varphi}(\xi)-e^{-(t-u)\overline{g}(\xi_j^N)}\overline{\varphi}(\xi_j^N)\right)\mathbf{1}_{V_j^N}(\xi)\to 0,
\tag{A.141}
$$

as $N\to\infty$, and using again that $|e^{-tg(\xi)}|\leqslant 1$ for every $t\in\mathbb{R}^+$, one proves that the convergence is dominated. Finally, since $|\varphi|^2$ is integrable with respect to $\mu_{X_S}\in\mathscr{M}_{SG}^+(\mathbb{R}^d)$ and D_N decreases to $\varnothing$, then

$$
\left|\int_{D_N}\int_{[0,t]\times[0,s]}e^{-(t-u)g(\xi)-(s-v)\overline{g}(\xi)}|\varphi(\xi)|^2dC_{X_T}(u,v)d\mu_{X_S}(\xi)\right|\leqslant\underbrace{|C_{X_T}|([0,t]\times[0,s])}_{<\infty}\underbrace{\int_{D_N}|\varphi(\xi)|^2d\mu_{X_S}(\xi)}_{\to 0\text{ as }N\to\infty}.
\tag{A.142}
$$

This proves that convergence of (A.136) to 0. ■

Appendix B

Formal Construction of Generalized Random Fields

In this Appendix we present a way of constructing (or rather, we show the existence of) Generalized Random Fields as stated in Definition 3.4.1, following a wide-range of possible laws, not-necessarily Gaussian, and following desired mean and covariance structures. It is based on the classical Kolmogorov's Theorem of existence of Stochastic Processes with finite-dimensional evaluations following a family of compatible laws. This Theorem will be sufficient for our developments. This construction is completely analogue to the one described in Ma (2009) in the case of continuous Random Functions, where the construction is simply done by multiplying a Gaussian Random Function with a suitable independent positive random variable. Here we present the general case for Generalized Random Fields describing explicitly the multi-dimensional laws of the constructed stochastic process.

In Section 3.4 we have defined a real GeRF Z as a real and continuous linear mapping from $\mathscr{S}(\mathbb{R}^d)$ to $L^2(\Omega, \mathcal{A}, \mathbb{P})$. Then, we have concluded that a mean distribution $m_Z \in \mathscr{S}'(\mathbb{R}^d)$ and a covariance distribution $C_Z \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^d)$ exist. This definition of a GeRF is quite useful for the understanding of this mathematical object and for practical applications based on the analysis of the mean and covariance structures. Nevertheless, the formal mathematical proof of the existence of such a mathematical object is usually done *backwards*, that is, we *choose* an arbitrary distribution $m_Z \in \mathscr{S}'(\mathbb{R}^d)$ which will be our mean distribution, and we also *choose* a distribution of two variables $C_Z \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^d)$ defining a positive-definite Kernel which will be our covariance distribution. Starting from these deterministic objects we construct a stochastic process indexed by $\mathscr{S}(\mathbb{R}^d)$ following a particular law with its mean and covariance structures being characterised by m_Z and C_Z respectively.

We will make explicit this construction, for which we will use a *Fourier-Transformed* version of Kolmogorov's Theorem.

B.1 Reminders on Kolmogorov's Theorem

Kolmogorov's Theorem is the basic tool used to construct stochastic processes indexed by arbitrary sets. Although this approach could produce some technical problems regarding the regularity of the sample-paths of the processes as mentioned in Section 3.7, this framework will be sufficient for us.

Theorem B.1.1 (Kolmogorov, real values case). *Let T be a non-empty set. For every finite vector of elements in T, $(t_1, ..., t_N) \in T^N$, with $N \in \mathbb{N}_*$, we consider a probability measure $\mu_{(t_1,...,t_N)}$ over $\mathbb{R}^N$. Suppose that the so-defined family of probability measures $(\mu_{(t_1,...,t_N)})_{(t_1,...,t_N)\in T^N, N\in\mathbb{N}_*}$ satisfies the following compatibility conditions:*

- ***Permutability:*** *$\mu_{(t_1,...,t_N)}(A_1 \times ... \times A_N) = \mu_{(t_{\sigma(1)},...,t_{\sigma(N)})}(A_{\sigma(1)} \times ... \times A_{\sigma(N)})$, for any collection of Borel sets $A_1, ..., A_N \in \mathcal{B}(\mathbb{R})$ and for any permutation $\sigma : \{1, ..., N\} \to \{1, ..., N\}$, for all $N \in \mathbb{N}_*$.*
- ***Projectivity:*** *$\mu_{(t_1,...,t_N,t_{N+1})}(A_1 \times ... \times A_N \times \mathbb{R}) = \mu_{(t_1,...,t_N)}(A_1 \times ... \times A_N)$ for any collection of Borel sets $A_1, ..., A_N \in \mathcal{B}(\mathbb{R})$, for all $N \in \mathbb{N}_*$.*

Then, there exists a probability space where a family of real random variables $(X_t)_{t\in T}$ such that the law of an arbitrary finite vector of this family $(X_{t_1}, ..., X_{t_N})$ is $\mu_{(t_1,...,t_N)}$ can be well-defined.

To be more precise about what the statement "*there exists a probability space where...*" means, Kolmogorov's Theorem actually states that there exists a unique probability measure over the space $\mathbb{R}^T$ of all mappings from T to $\mathbb{R}$ equipped with the cylinders $\sigma-$algebra such that the random vectors constructed through the evaluation of these functions over a finite quantity of points of T have the corresponding probability law in the family $(\mu_{(t_1,...,t_N)})_{(t_1,...,t_N)\in T^N, N\in\mathbb{N}_*}$. We are not going to enter into these details, and we will just use Kolmogorov's Theorem to assure that a real stochastic process indexed by an arbitrary set exists provided that we have a compatible family of probability measures describing the laws of the finite vector valued sub-families. See Kolmogorov (1956, Section 4, Chapter III) for a statement of this theorem and a proof. See also the development in Dellacherie & Meyer (1978, Chapter III, N$^{\text{o}}$ 50 to 52), applicable when $\mathbb{R}$ is replaced by a complete metrizable space. Another source exposing this Theorem with additional conditions concerning *compact classes* associated to each point $t \in T$ is the classical Neveu (1970, Chapter III, §3).

Let $\vec{X}$ be a random vector taking values in $\mathbb{R}^N$ with $N \in \mathbb{N}_*$ and let $\mu_{\vec{X}}$ be its probability law over $\mathbb{R}^N$. As we know, this probability measure can be completely described by its characteristic function

$$\varphi_{\vec{X}}(\xi) = \mathbb{E}(e^{-i\xi^T\vec{X}}) = \int_{\mathbb{R}^N} e^{-i\xi^T x} d\mu_{\vec{X}}(x), \quad \xi \in \mathbb{R}^N. \tag{B.1}$$

Hence, defining a compatible family of probability measures is equivalent to defining a *compatible family of characteristic functions*. We recall that from Bochner's Theorem 3.4.1 it can be concluded that necessary

and sufficient conditions for a function $\varphi : \mathbb{R}^N \to \mathbb{C}$ to be a characteristic function is to be continuous, to satisfy $\varphi(0) = 1$, and to be a positive-definite function (see Eq. (3.2)). Thus, we are going to re-write Kolmogorov's Theorem in terms of these characteristic functions.

Theorem B.1.2 (Kolmogorov, real values case, characteristic functions version). *Let T be a non-empty set. For every vector of elements in T, $(t_1, ..., t_N) \in T^N$, with $N \in \mathbb{N}_*$, we consider a characteristic function $\varphi_{(t_1,...,t_N)} : \mathbb{R}^N \to \mathbb{C}$. Suppose that the so-defined family $(\varphi_{(t_1,...,t_N)})_{(t_1,...,t_N)\in T^N, N\in\mathbb{N}_*}$ satisfies the following compatibility conditions:*

- ***Permutability:*** *$\varphi_{(t_1,...,t_N)}(\xi_1, ..., \xi_N) = \varphi_{(t_{\sigma(1)},...,t_{\sigma(N)})}(\xi_{\sigma(1)}, ..., \xi_{\sigma(N)})$, for all vector $\vec{\xi} = (\xi_1, ..., \xi_N) \in \mathbb{R}^N$ and for all permutation $\sigma : \{1, ..., N\} \to \{1, ..., N\}$, for all $N \in \mathbb{N}_*$.*
- ***Projectivity:*** *$\varphi_{(t_1,...,t_N,t_{N+1})}(\xi_1, ..., \xi_N, 0) = \varphi_{(t_1,...,t_N)}(\xi_1, ..., \xi_N)$, for all vector $\vec{\xi} = (\xi_1, ..., \xi_N) \in \mathbb{R}^N$, for all $N \in \mathbb{N}_*$.*

Then, there exists a probability space where a family of real random variables $(X_t)_{t\in T}$ such that the law of an arbitrary finite vector of this family $(X_{t_1}, ..., X_{t_N})$ is the law associated to the characteristic function $\varphi_{(t_1,...,t_N)}$ can be well-defined.

We will see that it is much easier to work in the context of this transformed theorem. In particular for some laws of square-integrable random variables, it will be quite easy to see *where* must the covariance and mean structure act.

B.2 Schoenberg's Theorem and some characteristic functions

Defining a compatible family of characteristic functions satisfying our desired properties is easier than expected. For this, we need then to define characteristic functions over $\mathbb{R}^N$ with N taking different values. Our approach is to use a very useful theorem due to Schoenberg which describes isotropic positive-definite continuous functions in any dimension.

Theorem B.2.1 (Schoenberg). *Let $g : [0, \infty) \to \mathbb{R}$ be a continuous function. Then, g has the property that for every $N \in \mathbb{N}_*$ the function $\varphi : \mathbb{R}^N \to \mathbb{R}$ defined through*

$$\varphi(x) = g(\|x\|) \tag{B.2}$$

is a positive-definite function if and only if g is of the form:

$$g(t) = \int_{[0,\infty)} e^{-rt^2} d\nu(r), \tag{B.3}$$

with ν being a positive finite measure over $([0,\infty), \mathcal{B}([0,\infty)))$.

See (Donoghue, 1969, Chapter 41) for a proof. Thus, this theorem describes continuous positive-definite functions which are radial and valid for any arbitrary dimension. We see that if we ask the measure ν to satisfy $\nu([0,\infty)) = 1$, that is, to be a probability measure, the associated function φ satisfies $\varphi(0) = 1$, hence it is a characteristic function of the form

$$\varphi(\xi) = \int_{[0,\infty)} e^{-r|\xi|^2} d\nu(r), \quad \xi \in \mathbb{R}^N. \tag{B.4}$$

This is a valid characteristic function over $\mathbb{R}^N$ for any dimension N. We see that this function is always positive and it decreases as we advance in some particular direction in $\mathbb{R}^N$. This implies that the probability law over $\mathbb{R}^N$ associated to this characteristic function is invariant under rotations (thus symmetric with respect to 0) and its support is the whole space $\mathbb{R}^N$ (it is necessarily not bounded in every direction).

Since we work in a L^2 context, we also need φ to be twice continuously differentiable. By Dominated Convergence Theorem, this requires that the measure ν must satisfy

$$\int_{[0,\infty)} r^2 d\nu(r) < \infty. \tag{B.5}$$

We will suppose this holds. Hence, the law of the random vector associated to the characteristic function (B.4) has square-integrable components. In particular, every real random vector of dimension N having (B.4) as characteristic function have 0 mean and uncorrelated components (not necessarily independent). We are going to see that in order to construct vectors with some particular covariance structure (matrix), all we need to do is to conveniently insert an *anisotropy matrix*.

B.3 Construction

Let us fix once and for all a real mean distribution $m_Z \in \mathscr{S}'(\mathbb{R}^d)$ and a real covariance distribution $C_Z \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^d)$ defining a positive-definite Kernel. We are going to construct a real stochastic process indexed by *real* functions of the Schwartz space. This subspace of the Schwartz space will be denoted by $\mathscr{S}_{\mathbb{R}}(\mathbb{R}^d)$. Thus, we need to construct finite-dimensional laws or characteristic functions indexed by finite vectors of real test-functions. Let us fix a probability measure ν over $([0,\infty), \mathcal{B}([0,\infty))$ which satisfies (B.5).

For all $N \in \mathbb{N}_*$, we consider a finite vector of real test-functions $(\phi_1, ..., \phi_N)$, and we define the following characteristic function over $\mathbb{R}^N$:

$$\varphi_{(\phi_1,...,\phi_N)}(\xi_1, ..., \xi_N) = e^{-i\sum_{j=1}^N \xi_j \langle m_Z, \phi_j \rangle} \int_{[0,\infty)} e^{-\frac{r}{2\int_{[0,\infty)} t d\nu(t)} \sum_{j,k=1}^N \xi_j \xi_k \langle C_Z, \phi_j \otimes \phi_k \rangle} d\nu(r). \tag{B.6}$$

We verify that this function is well-defined as a characteristic function. For that, consider a function φ defined as in (B.4). Consider any symmetric positive-definite matrix $A \in \mathbb{R}^{N\times N}$, and let $\sqrt{A}$ be its symmetric positive-definite square root. Then, the function $\varphi \circ \sqrt{A} : \mathbb{R}^N \to \mathbb{R}$, is continuous and it is straightforward that it is positive-definite. In addition, its evaluation at 0 equals to 1. Hence, it is a well defined characteristic function, where $\sqrt{A}$ works as an *anisotropy matrix* for the isotropic function φ. If in addition we consider any vector $v \in \mathbb{R}^N$, then it is straightforward that the function $e^{-iv^T(\cdot)}\varphi(\sqrt{A}(\cdot))$ is also a characteristic function. Here the vector v acts as a *translation vector* for the original associated law.

In Eq. (B.6) we have then considered a mean vector $\mathbf{m}^{(\phi_1,...,\phi_N)} \in \mathbb{R}^N$ using m_Z through

$$\mathbf{m}^{(\phi_1,...,\phi_N)} := (\langle m_Z, \phi_1\rangle, ..., \langle m_Z, \phi_N\rangle), \tag{B.7}$$

and we have also considered a covariance matrix $\mathbf{C}^{(\phi_1,...,\phi_N)} = (\mathbf{C}^{(\phi_1,...,\phi_N)}_{j,k})_{j,k=1}^N \in \mathbb{R}^{N\times N}$, obtained using C_Z thorough

$$\mathbf{C}^{(\phi_1,...,\phi_N)}_{j,k} := \langle C_Z, \phi_j \otimes \phi_k\rangle, \quad \forall j,k \in \{1,...,N\}. \tag{B.8}$$

Since $C_Z \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^d)$ defines a positive-definite Kernel, the matrix (B.8) is positive-definite and symmetric. We express thus the function (B.6) more explicitly through this vector and this matrix:

$$\varphi_{(\phi_1,...,\phi_N)}(\vec{\xi}) = e^{-i\vec{\xi}^T\mathbf{m}^{(\phi_1,...,\phi_N)}} \int_{[0,\infty)} e^{-\frac{r}{2\int_{[0,\infty)} td\nu(t)}\vec{\xi}^T\mathbf{C}^{(\phi_1,...,\phi_N)}\vec{\xi}} d\nu(r) \tag{B.9}$$

The matrix $\frac{1}{2\int_{[0,\infty)} td\nu(t)}\mathbf{C}^{(\phi_1,...,\phi_N)}$ is positive-definite and we can take its square root as an anisotropy matrix. The vector $\mathbf{m}^{(\phi_1,...,\phi_N)}$ is acting as a translation vector. It turns out that the function (B.9) is a valid characteristic function for every vector $(\phi_1, ..., \phi_N) \in \mathscr{S}_{\mathbb{R}}(\mathbb{R}^d)^N$.

It is quite easy to verify that this so-defined family of characteristic functions form a compatible family according to Kolmogorov's Theorem B.1.2. Indeed, both the expressions $\vec{\xi}^T\mathbf{m}^{(\phi_1,...,\phi_N)}$ and $\vec{\xi}^T\mathbf{C}^{(\phi_1,...,\phi_N)}\vec{\xi}$ are stable under permutations of the respective components of the vectors and matrices involved, and if we evaluate any of the components of $\vec{\xi}$ in 0 we will obtain analogous expressions in dimension $N-1$, hence the projectivity condition is also satisfied. We conclude that there exists a real stochastic process $(\langle Z, \phi\rangle)_{\phi\in\mathscr{S}_{\mathbb{R}}(\mathbb{R}^d)}$ whose finite-dimensional laws are described by the characteristic functions (B.9). Now, for any arbitrary *complex* test-function $\phi \in \mathscr{S}(\mathbb{R}^d)$, we simply define the associated random variable

$$\langle Z, \phi\rangle := \langle Z, \phi_R\rangle + i\langle Z, \phi_I\rangle, \tag{B.10}$$

where ϕ_R and ϕ_I are the real and the imaginary parts of ϕ respectively. The resulting stochastic process $(\langle Z, \phi\rangle)_{\phi\in\mathscr{S}(\mathbb{R}^d)}$ is then well-defined and it has the property of producing real random variables when evaluated at real test-functions.

Differentiating (B.9) and evaluating at 0, we can conclude that for any vector of test-functions $(\phi_1, ..., \phi_N)$ it holds that

$$i\frac{\partial\varphi_{(\phi_1,...,\phi_N)}}{\partial\xi_j}(0) = \mathbb{E}(\langle Z, \phi_j\rangle) = \mathbf{m}_j^{(\phi_1,...,\phi_N)} = \langle m_Z, \phi_j\rangle. \tag{B.11}$$

And differentiating twice and evaluating at 0 , considering real test-functions $(\phi_1, ..., \phi_N)$, we obtain

$$\begin{aligned} -\frac{\partial^2\varphi_{(\phi_1,...,\phi_N)}}{\partial\xi_k\partial\xi_j}(0) &= \mathbb{E}(\langle Z, \phi_j\rangle\langle Z, \phi_k\rangle) = \mathbf{m}_j^{(\phi_1,...,\phi_N)}\mathbf{m}_k^{(\phi_1,...,\phi_N)} - \mathbf{C}_{j,k}^{(\phi_1,...,\phi_N)} \\ &= \langle m_Z, \phi_j\rangle\langle m_Z, \phi_k\rangle + \langle C_Z, \phi_j \otimes \phi_k\rangle. \end{aligned} \tag{B.12}$$

And thus we obtain for two arbitrary real test-functions ϕ and ψ:

$$\mathbb{E}(\langle Z, \phi\rangle) = \langle m_Z, \phi\rangle \tag{B.13}$$

$$\mathbb{C}ov(\langle Z, \phi\rangle, \langle Z, \psi\rangle) = \langle C_Z, \phi \otimes \psi\rangle \tag{B.14}$$

By linearity we see that formula (B.13) also holds for complex test-functions. The sesquilinearity of the covariance allows to conclude that for two complex test-functions ϕ and ψ it holds that

$$\mathbb{C}ov(\langle Z, \phi\rangle, \langle Z, \psi\rangle) = \langle C_Z, \phi \otimes \overline{\psi}\rangle, \tag{B.15}$$

which is what we expected to have. Hence, the distributions m_Z and C_Z do describe the moments of the stochastic process Z.

We still need to verify the linearity and continuity conditions, but this is straightforward in this squared-integrable context. Indeed, for the linearity consider $\phi, \psi \in \mathscr{S}(\mathbb{R}^d)$ and $\alpha, \beta \in \mathbb{C}$. A few calculations based on the linearity of m_Z and the bi-linearity of $(\phi, \psi) \mapsto \langle C_Z, \phi \otimes \psi\rangle$ of C_Z allow to show that

$$\mathbb{E}\left(|\langle Z, \alpha\phi + \beta\psi\rangle - (\alpha\langle Z, \phi\rangle + \beta\langle Z, \psi\rangle)|^2\right) = 0. \tag{B.16}$$

Hence, $\langle Z, \alpha\phi + \beta\psi\rangle \overset{a.s.}{=} \alpha\langle Z, \phi\rangle + \beta\langle Z, \psi\rangle$. To prove the continuity, let us consider a sequence of test-functions $(\phi_n)_{n\in\mathbb{N}} \subset \mathscr{S}(\mathbb{R}^d)$ such that $\phi_n \overset{\mathscr{S}}{\to} 0$. Considering that this implies that $\phi_n \otimes \overline{\phi_n} \overset{\mathscr{S}(\mathbb{R}^d\times\mathbb{R}^d)}{\longrightarrow} 0$ [1], we obtain from the continuity of m_Z and C_Z that

$$\mathbb{E}(|\langle Z, \phi_n\rangle|^2) = \mathbb{V}ar(\langle Z, \phi_n\rangle) + \mathbb{E}(\langle Z, \phi_n\rangle)\mathbb{E}(\overline{\langle Z, \phi_n\rangle}) = \langle C_Z, \phi_n \otimes \overline{\phi_n}\rangle + \langle m_Z, \phi_n\rangle\langle m_Z, \overline{\phi_n}\rangle \to 0. \tag{B.17}$$

This proves that the so constructed stochastic processes Z defines a continuous mapping in the mean-square

[1] This is concluded immediately from the fact that for every $\phi \in \mathscr{S}(\mathbb{R}^d)$, $\sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d} |\phi(x)\phi(y)| = \|\phi\|_\infty^2$, and applying this to every function of the form $x^\alpha D^\beta\phi$.

sense. Hence, we have proved the existence of a GeRF following the required mean and covariance structures and with finite-dimensional laws described by the family of characteristic functions determined by (B.9).

Remark B.3.1. Using ν of the form $\nu = \delta_t$ for some $t \in (0, \infty)$ in (B.9), we obtain a Gaussian process.

Remark B.3.2. An analogue construction can be done outside the framework of square-integrable random variables, but still using a "mean" and "covariance" structures as basis. Such a construction is done by considering $m_Z \in \mathscr{S}'(\mathbb{R}^d)$ and $C_Z \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^d)$ taken as always and considering the family of characteristic functions:

$$\varphi_{(\phi_1,\ldots,\phi_N)}(\vec{\xi}) = e^{-i\vec{\xi}^T \mathbf{m}^{(\phi_1,\ldots,\phi_N)}} \int_{[0,\infty)} e^{-r\vec{\xi}^T \mathbf{C}^{(\phi_1,\ldots,\phi_N)}\vec{\xi}} d\nu(r). \tag{B.18}$$

From the arguments exposed above, this defines a compatible family of characteristic functions. The generalization is done by considering a probability measure ν which does not satisfy (B.5). It is not even necessary to require $\int_{[0,\infty)} t d\nu(t) < \infty$. The only difference is that the continuity of the associated stochastic process $(\langle Z, \phi\rangle)_{\phi \in \mathscr{S}(\mathbb{R}^d)}$ must be interpreted *in probability* and not in the sense of $L^2(\Omega, \mathcal{A}, \mathbb{P})$. The arguments that we have used in this Appendix do not work to prove the linearity and continuity in such a case, since we have used the mean-square structure. However, both linearity and continuity can be proven by analysing directly the characteristic functions associated. Indeed, by doing some algebraic calculations using the linearity of m_Z and the bi-linearity of the Kernel associated to C_Z, one can prove that the characteristic function of a random variable of the form $\langle Z, \alpha\phi + \beta\psi\rangle - (\alpha\langle Z, \phi\rangle + \beta\langle Z, \psi\rangle)$, with $\alpha, \beta \in \mathbb{C}$ and $\phi, \psi \in \mathscr{S}(\mathbb{R}^d)$, is the constant function 1. Hence the random variable equals 0 almost-surely (its probability law is the Dirac measure). Analogously, if $\phi_n \xrightarrow{\mathscr{S}} 0$ one proves that the characteristic functions of the associated random variables $\langle Z, \phi_n\rangle$ converges point-wise to the function 1. From Lévy's Theorem one concludes that the sequence $(\langle Z, \phi_n\rangle)_{n\in\mathbb{N}}$ converges in law to 0, and since every sequence of random variables which converges in law to a constant also does it in probability, it follows that Z is continuous in probability. We omit the details of this procedure. We conclude that we can construct a GeRF Z continuous in probability with the same principles as done in the square-integrable case, using m_Z and C_Z as basic tools. Here the "mean" and "covariance" distributions do not determine, exactly, the first two moments of Z, since they do not necessarily exist. However, *they do describe the dependence structure between the random variables of the process*, and they do it in a quite analogous way as the mean and covariance do: through a *translation* and an *anisotropy*.

Remark B.3.3. Using another space of test-functions such as $\mathscr{D}(\mathbb{R}^d)$ or $\mathscr{E}(\mathbb{R}^d)$, with the mean and covariance belonging to the corresponding dual space (the covariance defining, of course, a positive-definite Kernel), the same procedure can be used to prove the existence of Random Distributions in the generic sense and of Random Distributions with compact support. The same idea also holds if we take as spaces of test-functions $C_c(\mathbb{R}^d)$, $C_{FD}(\mathbb{R}^d)$, $C_0(\mathbb{R}^d)$ and $C(\mathbb{R}^d)$ in order to construct Random Measures, slow-growing Random Measures, finite Random Measures and Random Measures of compact support respectively, by making the mean and the covariance being in suitable spaces of measures. Random Measures interpreted

as set-functions also enters in this framework, applying the same procedures by replacing the test-functions with indicator functions of bounded Borel sets. Finally, and getting back a little into a more basic framework, this procedure also shows the existence of Random Functions with any mean function and covariance functions. For that, replace simply the test-functions with Dirac measures at the corresponding points in the space, interpreting of course $\langle m_Z, \delta_x \rangle$ and $\langle C_Z, \delta_x \otimes \delta_y \rangle$ as $m_Z(x)$ and $C_Z(x, y)$ respectively.

Appendix C

Formal resolution of first order evolution equations.

In this Appendix we deal with the resolution of a spatio-temporal SPDE of the form

$$\frac{\partial U}{\partial t} + \mathcal{L}_g U = X, \tag{C.1}$$

where X and U are GeRFs over $\mathbb{R}^d \times \mathbb{R}^+$ having a suitable behaviour, and $\mathcal{L}_g$ is a *spatial* operator defined through a continuous symbol function $g : \mathbb{R}^d \mapsto \mathbb{C}$ (see Chapter 4) for which we will suppose that its real part satisfies $g_R \geqslant 0$. We will follow a *traditional* approach and obtain an existence and uniqueness result by fixing an *initial condition* that the solution must follow. This actually poses more theoretical problems than those which are simplified. When requiring, for example, that the *evaluation at* 0 of the solution must be equal to some particular spatial GeRF, we have already implicitly required that the solution must *have a functional meaning in time*, at least at $t = 0$ or at a neighbourhood of $t = 0$. This restricts the space of possible solutions to the problem and hence it also restricts, in principle, the type of operators we can apply over members of this space. However, we will fix a special space where the *functional meaning in time* of the solution can be guaranteed, without losing much generality. The approach is done by considering spaces of spatio-temporal distributions such that *their temporal derivatives are measures in time*. This selection has been done by following the fact that the distributional primitive of every measure over $\mathbb{R}$ can be identified with a *càdlàg* function.

We will first show how to solve the associated deterministic problem. The resolution of the stochastic version of the problem will follow analogously.

We will use the same notational conventions used in the spatio-temporal context presented in Chapter 5.

C.1 Solving the deterministic problem

C.1.1 Convenient spaces of test-functions and their dual spaces

In Chapter 4 we have introduced a convenient space of tempered distributions over $\mathbb{R}^d$, $\mathscr{V}'(\mathbb{R}^d)$, which is the space of all tempered distributions such that their Fourier Transforms belong to $\mathscr{M}_{SG}(\mathbb{R}^d)$. We have not justified the notation $\mathscr{V}'(\mathbb{R}^d)$ as the dual of a particular space, but using the Riesz Representation Theorem for slow-growing measures 2.1.6 this can be done easily. Let us consider the next space of test-functions:

$$\mathscr{V}(\mathbb{R}^d) = \{\varphi \in C_0^\infty(\mathbb{R}^d) \;\big|\; \exists \phi \in C_{FD}(\mathbb{R}^d) \text{ such that } \varphi = \mathscr{F}(\phi)\} = \mathscr{F}(C_{FD}(\mathbb{R}^d)). \tag{C.2}$$

We have defined $\mathscr{V}(\mathbb{R}^d)$ in such a way that $\mathscr{V}(\mathbb{R}^d) \subset C_0^\infty(\mathbb{R}^d)$, but this requirement actually follows from Riemann-Lebesgue Lemma (Theorem 2.2.1). Indeed, a function $\phi \in C_{FD}(\mathbb{R}^d)$ can be multiplied by any polynomial, the result being always integrable. Hence, the Fourier Transform of ϕ is smooth and all of its derivatives vanish at infinity. $\mathscr{V}(\mathbb{R}^d)$ is a strict subspace of $C_0^\infty(\mathbb{R}^d)$[1]. The space $\mathscr{V}(\mathbb{R}^d)$ can be defined equivalently as the space of functions in $C_0^\infty(\mathbb{R}^d)$ such that their Inverse Fourier Transforms are in $C_{FD}(\mathbb{R}^d)$. In such a case, the Inverse Fourier Transform must be interpreted in distributional sense, considering the members of $C_0^\infty(\mathbb{R}^d)$ as tempered distributions. Of course, if we use the Inverse Fourier Transform instead of the Fourier Transform in the definition (C.2) of $\mathscr{V}(\mathbb{R}^d)$, the space remains the same.

We endow $\mathscr{V}(\mathbb{R}^d)$ with the topology induced by the directed family of semi-norms

$$p_N(\varphi) = \sup_{\xi \in \mathbb{R}^d} \left|(1 + |\xi|^2)^N \mathscr{F}(\varphi)(\xi)\right|, \quad N \in \mathbb{N}. \tag{C.3}$$

This topology is equivalent to the one induced by the metric

$$(\varphi, \phi) \mapsto \sum_{N \in \mathbb{N}} \frac{1}{2^{N+1}} \frac{\sup_{\xi \in \mathbb{R}^d} \left|(1 + |\xi|^2)^N \mathscr{F}(\varphi - \phi)(\xi)\right|}{1 + \sup_{\xi \in \mathbb{R}^d} \left|(1 + |\xi|^2)^N \mathscr{F}(\varphi - \phi)(\xi)\right|}, \quad \forall \varphi, \phi \in \mathscr{V}(\mathbb{R}^d). \tag{C.4}$$

Hence, a sequence of functions $(\varphi_n)_{n \in \mathbb{N}} \subset \mathscr{V}(\mathbb{R}^d)$ converges to 0 as $n \to \infty$ on $\mathscr{V}(\mathbb{R}^d)$, denoted by $\varphi_n \xrightarrow{\mathscr{V}} 0$, if and only if $\mathscr{F}(\varphi_n) \xrightarrow{C_{FD}} 0$ (See Section 2.1.4 for a recall on the topology of $C_{FD}(\mathbb{R}^d)$).

Since $C_{FD}(\mathbb{R}^d) \subset \mathscr{L}^1(\mathbb{R}^d)$, the Fourier Transform of a fast-decreasing continuous function is defined classically as an integral. The same holds for the Inverse Fourier Transform. For the members in $\mathscr{V}(\mathbb{R}^d)$ these operations are algebraically defined in distributional sense, having a one-to-one correspondence between members in $\mathscr{V}(\mathbb{R}^d)$ and $C_{FD}(\mathbb{R}^d)$, and satisfying $\mathscr{F}^{-1}(\mathscr{F}(\varphi)) = \varphi$ for every $\varphi \in \mathscr{V}(\mathbb{R}^d)$ and also for every

[1] For $d = 1$, the function $\xi \mapsto \frac{2}{\sqrt{2\pi}} \frac{\sin(\xi)}{\xi}$ is in $C_0^\infty(\mathbb{R})$, but it is the Fourier Transform of $\mathbf{1}_{[-1,1]}$, which is not in $C_{FD}(\mathbb{R})$.

$\varphi \in C_{FD}(\mathbb{R}^d)$. From Theorem D.0.1, it follows immediately that the Fourier Transform and its inverse are continuous linear operators from $\mathscr{V}(\mathbb{R}^d)$ to $C_{FD}(\mathbb{R}^d)$ (or from $C_{FD}(\mathbb{R}^d)$ to $\mathscr{V}(\mathbb{R}^d)$, as pleasure). The completeness of $C_{FD}(\mathbb{R}^d)$ and the continuity of the Fourier Transform imply the completeness of $\mathscr{V}(\mathbb{R}^d)$. $\mathscr{V}(\mathbb{R}^d)$ is a Fréchet space.

Proposition C.1.1. *The Schwartz space satisfies $\mathscr{S}(\mathbb{R}^d) \subset \mathscr{V}(\mathbb{R}^d)$ and $\mathscr{S}(\mathbb{R}^d) \subset C_{FD}(\mathbb{R}^d)$ and it is a dense subspace of both spaces with their respective topologies.*

Proof: The inclusions are straightforward. We will just prove the density of $\mathscr{S}(\mathbb{R}^d)$ in $C_{FD}(\mathbb{R}^d)$. The density in $\mathscr{V}(\mathbb{R}^d)$ follows immediately from the continuity of the Fourier Transform.

We first prove that if $f \in C_{FD}(\mathbb{R}^d)$ and $\varphi \in \mathscr{S}(\mathbb{R}^d)$, then $f * \varphi \in \mathscr{S}(\mathbb{R}^d)$. It is clear that f is integrable and bounded, as well as φ which is in addition smooth. Thus $f * \varphi$ is a smooth integrable and bounded function, and its Fourier Transform satisfies $\mathscr{F}(f * \varphi) = (2\pi)^{\frac{d}{2}}\mathscr{F}(f)\mathscr{F}(\varphi)$. We have that $\mathscr{F}(\varphi) \in \mathscr{S}(\mathbb{R}^d)$ since $\mathscr{F}$ is a bijective endomorphism of $\mathscr{S}(\mathbb{R}^d)$. Since $f \in C_{FD}(\mathbb{R}^d)$, then $\mathscr{F}(f) \in \mathscr{V}(\mathbb{R}^d) \subset C_0^\infty(\mathbb{R}^d) \subset \mathcal{O}_M(\mathbb{R}^d)$. This implies that $(2\pi)^{\frac{d}{2}}\mathscr{F}(f)\mathscr{F}(\varphi) \in \mathscr{S}(\mathbb{R}^d)$. This proves that $f * \varphi = \mathscr{F}^{-1}\left((2\pi)^{\frac{d}{2}}\mathscr{F}(f)\mathscr{F}(\varphi)\right) \in \mathscr{S}(\mathbb{R}^d)$.

Let $(\phi_n)_{n\in\mathbb{N}} \subset \mathscr{S}(\mathbb{R}^d)$ be a regularizing sequence of positive compactly supported smooth functions, such that $\operatorname{supp}(\phi_n) = \overline{B_{\frac{1}{n}}(0)}$ and $\int_{\mathbb{R}^d}\phi_n(x)dx = 1$ for all $n \in \mathbb{N}$. We consider the sequence of functions $f_n = f * \phi_n$, which are all in $\mathscr{S}(\mathbb{R}^d)$. We will prove that $f_n \overset{C_{FD}}{\to} f$. Let $m \in \mathbb{N}$ be fixed. We must show that $\|(1+|x|^2)^m(f_n - f)\|_\infty \to 0$ as $n \to \infty$. Let $\epsilon > 0$. As $f \in C_{FD}(\mathbb{R}^d)$, we can take $R > 0$ large enough such that for every x such that $|x| > R-1$, $(1+2|x|^2)^m|f(x)| < \frac{\epsilon}{3(2^{m-1}+2^{2m-1})}$ holds. Notice that in this case, $(1+|x|^2)^m|f(x)| < \frac{\epsilon}{3}$. Since f is continuous, it is uniformly continuous over the compact set $\overline{B_{R+1}(0)}$. Thus, there exists $\delta > 0$ such that if $|x-y| < \delta$, then $|f(x) - f(y)| < \frac{\epsilon}{3(1+R^2)^m}$ for all $x, y \in \overline{B_{R+1}(0)}$. Consider $n_0 \in \mathbb{N}$ such that $\frac{1}{n_0} < \delta$. Then, for all $n \geqslant n_0$,

$$\begin{aligned}
\|(1+|x|^2)^m(f-f_n)\|_\infty &= \sup_{x\in\mathbb{R}^d}\left|\int_{B_{\frac{1}{n}}(0)}(1+|x|^2)^m(f(x)-f(x-y))\phi_n(y)dy\right| \\
&\leqslant \underbrace{\sup_{x\in\overline{B_R(0)}}\left|\int_{B_{\frac{1}{n}}(0)}(1+|x|^2)^m(f(x)-f(x-y))\phi_n(y)dy\right|}_{(a)} \\
&\quad + \underbrace{\sup_{x\in\overline{B_R(0)}^c}\left|\int_{B_{\frac{1}{n}}(0)}(1+|x|^2)^m(f(x)-f(x-y))\phi_n(y)dy\right|}_{(b)}.
\end{aligned} \tag{C.5}$$

For the first term (a), the uniform continuity of f implies

$$(a) = \sup_{x\in\overline{B_R(0)}} \left| \int_{B_{\frac{1}{n}}(0)} (1+|x|^2)^m (f(x)-f(x-y))\phi_n(y)dy \right| \leqslant \int_{B_{\frac{1}{n}}(0)} (1+R^2)^m \frac{\epsilon}{3(1+R^2)^m}\phi_n(y)dy = \frac{\epsilon}{3}. \tag{C.6}$$

Regarding the second term (b), the integral is split to obtain

$$(b) \leqslant \sup_{x\in\overline{B_R(0)}^c} \Big\{ \underbrace{\int_{B_{\frac{1}{n}}(0)} (1+|x|^2)^m |f(x)|\phi_n(y)dy}_{\leqslant \frac{\epsilon}{3}} + \int_{B_{\frac{1}{n}}(0)} (1+|x|^2)^m |f(x-y)|\phi_n(y)dy \Big\}. \tag{C.7}$$

When applying a convexity argument twice, one shows that $(1+|x|^2)^m \leqslant 2^{m-1}[(1+2|x-y|^2)^m + 2^m|y|^{2m}]$ for all x and y, and thus

$$\begin{aligned}\int_{B_{\frac{1}{n}}(0)} (1+|x|^2)^m |f(x-y)|\phi_n(y)dy \leqslant 2^{m-1}\Bigg[&\int_{B_{\frac{1}{n}}(0)} \underbrace{(1+2|x-y|^2)^m |f(x-y)|}_{<\frac{\epsilon}{3(2^{m-1}+2^{2m-1})} \text{ from } |x-y|>R-1} \phi_n(y)dy \\ &+ 2^m \int_{B_{\frac{1}{n}}(0)} \underbrace{|f(x-y)|}_{<\frac{\epsilon}{3(2^{m-1}+2^{2m-1})}} \underbrace{|y|^{2m}}_{\leqslant 1} \phi_n(y)dy \Bigg] \\ &< 2^{m-1}\left(\frac{\epsilon}{3(2^{m-1}+2^{2m-1})} + 2^m \frac{\epsilon}{3(2^{m-1}+2^{2m-1})} \right) = \frac{\epsilon}{3}.\end{aligned} \tag{C.8}$$

Hence considering (C.7) and (C.8) we finally obtain $(b) < \frac{2\epsilon}{3}$. Putting together this result and (C.6) on equation (C.5), we finally obtain that for all $n \geqslant n_0$,

$$\|(1+|x|^2)^m (f-f_n)\|_\infty < \epsilon, \tag{C.9}$$

hence $\|(1+|x|^2)^m(f-f_n)\|_\infty \to 0$. Since m was arbitrary, this result holds for all m. We therefore conclude that $f_n \overset{C_{FD}}{\to} f$. Since for any arbitrary $f \in C_{FD}(\mathbb{R}^d)$ we can find a sequence included in $\mathscr{S}(\mathbb{R}^d)$ which converges to f, we conclude that $\mathscr{S}(\mathbb{R}^d)$ is dense in $C_{FD}(\mathbb{R}^d)$. ■

The dual space of $\mathscr{V}(\mathbb{R}^d)$ is then denoted by $\mathscr{V}'(\mathbb{R}^d)$. From Theorem D.0.1, it follows that a linear functional $T : \mathscr{V}(\mathbb{R}^d) \to \mathbb{C}$ is in $\mathscr{V}'(\mathbb{R}^d)$ if and only if there exist $C > 0$ and $N \in \mathbb{N}$ such that

$$|\langle T, \varphi\rangle| \leqslant C \sup_{\xi\in\mathbb{R}^d} \left|(1+|\xi|^2)^N \mathscr{F}(\varphi)(\xi)\right|, \quad \forall \varphi \in \mathscr{V}(\mathbb{R}^d). \tag{C.10}$$

The density of $\mathscr{S}(\mathbb{R}^d)$ in $\mathscr{V}(\mathbb{R}^d)$ and the integrability of the functions in $C_{FD}(\mathbb{R}^d)$ allow to conclude the following inclusions:

$$\mathscr{S}(\mathbb{R}^d) \subset \mathscr{V}(\mathbb{R}^d) \subset \mathscr{V}'(\mathbb{R}^d) \subset \mathscr{S}'(\mathbb{R}^d). \tag{C.11}$$

The Fourier Transform over $\mathscr{V}'(\mathbb{R}^d)$ can be defined equivalently as the restriction over $\mathscr{V}'(\mathbb{R}^d)$ of the Fourier Transform on $\mathscr{S}'(\mathbb{R}^d)$, or as the transpose of the Fourier Transform $\mathscr{F} : C_{FD}(\mathbb{R}^d) \to \mathscr{V}(\mathbb{R}^d)$. Using this second option, it follows immediately that it is a sequentially continuous linear functional $\mathscr{F} : \mathscr{V}'(\mathbb{R}^d) \mapsto C'_{FD}(\mathbb{R}^d)$ in the sense of a weak-$*$ convergence. Since $C'_{FD}(\mathbb{R}^d) = \mathscr{M}_{SG}(\mathbb{R}^d)$ from Riesz Representation Theorem 2.1.6, we conclude that $\mathscr{V}'(\mathbb{R}^d)$ is indeed the space of tempered distributions such that their Fourier Transforms are slow-growing measures, as it was defined in Chapter 4.

Let us now introduce some special spaces of test-functions which consider that we work over the *positive time*, hence over the space $\mathbb{R}^d \times \mathbb{R}^+$ rather than over $\mathbb{R}^d \times \mathbb{R}$.

We define the next space of test-functions:

$$C_{FD}(\mathbb{R}^d \times \mathbb{R}^+) := \{\psi \in C(\mathbb{R}^d \times \mathbb{R}^+) \,\big|\, \sup_{(\xi,t)\in\mathbb{R}^d\times\mathbb{R}^+} \left|(1+|x|^2)^{n_S}(1+t^2)^{n_T}\psi(\xi,t)\right| < \infty, \forall n_S, n_T \in \mathbb{N}\}. \tag{C.12}$$

This space can be equivalently defined as the space of restrictions of functions in $C_{FD}(\mathbb{R}^d \times \mathbb{R})$ to the subset $\mathbb{R}^d \times \mathbb{R}^+$. $C_{FD}(\mathbb{R}^d \times \mathbb{R}^+)$ is endowed with the topology induced by the metric

$$(\psi_1, \psi_2) \mapsto \sum_{n_S, n_T \in \mathbb{N}} \frac{1}{2^{n_S+n_T+2}} \frac{\sup\limits_{(\xi,t)\in\mathbb{R}^d\times\mathbb{R}^+} \left|(1+|\xi|^2)^{n_S}(1+t^2)^{n_T}(\psi_1-\psi_2)(\xi,t)\right|}{1+\sup\limits_{(\xi,t)\in\mathbb{R}^d\times\mathbb{R}^+} \left|(1+|\xi|^2)^{n_S}(1+t^2)^{n_T}(\psi_1-\psi_2)(\xi,t)\right|}, \tag{C.13}$$

for all $\psi_1, \psi_2 \in C_{FD}(\mathbb{R}^d \times \mathbb{R}^+)$. $C_{FD}(\mathbb{R}^d \times \mathbb{R}^+)$ is a Fréchet space. We denote by $C'_{FD}(\mathbb{R}^d \times \mathbb{R}^+)$ its dual space. A linear functional $T : C_{FD}(\mathbb{R}^d \times \mathbb{R}^+) \to \mathbb{C}$ is continuous if and only if there exist $C > 0$ and $N_S, N_T \in \mathbb{N}$ such that

$$|\langle T, \psi\rangle| \leqslant C \sup_{(\xi,t)\in\mathbb{R}^d\times\mathbb{R}^+} \left|(1+|\xi|^2)^{N_S}(1+t^2)^{N_T}\psi(\xi,t)\right|, \quad \forall \psi \in C_{FD}(\mathbb{R}^d \times \mathbb{R}^+). \tag{C.14}$$

Let us now define the space $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$. We define it as the space of slow-growing complex measures such that their absolute variation over $\mathbb{R}^d \times \mathbb{R}^-_*$ is null:

$$\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+) := \{\mu \in \mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}) \,\big|\, |\mu|(\mathbb{R}^d \times \mathbb{R}^-_*) = 0\}. \tag{C.15}$$

We could have defined $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$ as a space of measures over $\mathbb{R}^d \times \mathbb{R}^+$ without concerning on what happens over the negative time, but it is actually easier to work with measures defined over the whole space $\mathbb{R}^d \times \mathbb{R}$ but for which we only look at their behaviours over the subset $\mathbb{R}^d \times \mathbb{R}^+$. $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$ is also equivalently defined as the space of slow-growing measures with support included in $\mathbb{R}^d \times \mathbb{R}^+$.

The next Proposition follows from Riesz Representation Theorem.

Proposition C.1.2. $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+) = C'_{FD}(\mathbb{R}^d \times \mathbb{R}^+)$.

Proof: The inclusion $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+) \subset C'_{FD}(\mathbb{R}^d \times \mathbb{R}^+)$ is immediate. Let $T \in C'_{FD}(\mathbb{R}^d \times \mathbb{R}^+)$. Let $\psi \in C_{FD}(\mathbb{R}^d \times \mathbb{R})$. We define the action of T over ψ as

$$\langle T, \psi \rangle := \langle T, \psi\big|_{\mathbb{R}^d \times \mathbb{R}^+} \rangle. \tag{C.16}$$

Since $T \in C_{FD}(\mathbb{R}^d \times \mathbb{R}^+)$, there exist $C > 0$ and $N_S, N_T \in \mathbb{N}$ such that

$$\begin{aligned} |\langle T, \psi \rangle| &= \left| \langle T, \psi\big|_{\mathbb{R}^d \times \mathbb{R}^+} \rangle \right| \\ &\leqslant C \sup_{(\xi,t) \in \mathbb{R}^d \times \mathbb{R}^+} \left| (1 + |\xi|^2)^{N_S} (1 + t^2)^{N_T} \psi(\xi, t) \right| \\ &\leqslant C \sup_{(\xi,t) \in \mathbb{R}^d \times \mathbb{R}} \left| (1 + |\xi|^2)^{N_S} (1 + t^2)^{N_T} \psi(\xi, t) \right|. \end{aligned} \tag{C.17}$$

This proves that the extension of T to the whole space $C_{FD}(\mathbb{R}^d \times \mathbb{R})$ is a continuous linear functional and hence there exists a slow-growing measure $\mu \in \mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R})$ such that

$$\langle T, \psi\big|_{\mathbb{R}^d \times \mathbb{R}^+} \rangle = \int_{\mathbb{R}^d \times \mathbb{R}} \psi(x, t) d\mu(x, t), \quad \forall \psi \in C_{FD}(\mathbb{R}^d \times \mathbb{R}). \tag{C.18}$$

It can be concluded either using the inequality (C.17) or the Lemma A.2.3 that the measure μ has its support in $\mathbb{R}^d \times \mathbb{R}^+$. This proves that $C'_{FD}(\mathbb{R}^d \times \mathbb{R}^+) = \mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$. ■

We also define the space $C_{FD}(\mathbb{R}^+)$ of continuous fast decreasing functions over $\mathbb{R}^+$, with an analogous topology which makes it a Fréchet space. Its dual can be represented as the space $\mathscr{M}_{SG}(\mathbb{R}^+)$ of measures over the positive-time, which can as well be identified with the space of measures in $\mathscr{M}_{SG}(\mathbb{R})$ which have their supports contained in $\mathbb{R}^+$. We omit the details for such claims since they are analogous to the analysis already done for the space $C_{FD}(\mathbb{R}^d \times \mathbb{R}^+)$.

We consider right now the following space of test-functions which are spatial Fourier Transforms of functions in $C_{FD}(\mathbb{R}^d \times \mathbb{R}^+)$:

$$\mathscr{V}(\mathbb{R}^d) \widehat{\boxtimes} C_{FD}(\mathbb{R}^+) := \{ \psi \in C(\mathbb{R}^d \times \mathbb{R}^+) \;\big|\; \exists \psi_2 \in C_{FD}(\mathbb{R}^d \times \mathbb{R}^+) \text{ such that } \psi = \mathscr{F}_S(\psi_2) \}. \tag{C.19}$$

Members of this class act as members of $\mathscr{V}(\mathbb{R}^d)$ spatially and as members of $C_{FD}(\mathbb{R}^+)$ temporally. That is, if $\psi \in \mathscr{V}(\mathbb{R}^d) \widehat{\boxtimes} C_{FD}(\mathbb{R}^+)$, then for every $x \in \mathbb{R}^d$, $\psi(x, \cdot) \in C_{FD}(\mathbb{R}^+)$ and for every $t \in \mathbb{R}^+$, $\psi(\cdot, t) \in \mathscr{V}(\mathbb{R}^d)$. It is immediate that this set of functions forms a complex vector space. Every function of the form $\varphi \boxtimes \phi$, with $\varphi \in \mathscr{V}(\mathbb{R}^d)$ and $\phi \in C_{FD}(\mathbb{R}^+)$ is a member of this class, as well as any finite linear

combination of functions of this form. The notation $\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)$ has a deep inspiration in the theory of Nuclear spaces: the notation $E\hat{\otimes}F$, when E and F are more general topological vector spaces, is used to represent a completition, under suitable topologies, of the space $E\otimes F$ of finite linear combinations of tensor products (see Trèves (1967, Part III) or Grothendieck (1955)). We will not enter in those details explicitly, and we will simply work with definition (C.19) and its notation [2].

The space $\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)$ is endowed with the topology induced by the metric:

$$(\psi_1,\psi_2)\mapsto \sum_{n_S,n_T\in\mathbb{N}} \frac{1}{2^{n_S+n_T+2}} \frac{\sup\limits_{(\xi,t)\in\mathbb{R}^d\times\mathbb{R}^+} \left|(1+|\xi|^2)^{n_S}(1+t^2)^{n_T}\mathscr{F}_S^{-1}(\psi_1-\psi_2)(\xi,t)\right|}{1+\sup\limits_{(\xi,t)\in\mathbb{R}^d\times\mathbb{R}^+} \left|(1+|\xi|^2)^{n_S}(1+t^2)^{n_T}\mathscr{F}_S^{-1}(\psi_1-\psi_2)(\xi,t)\right|}, \tag{C.20}$$

$\forall\psi_1,\psi_2\in\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)$. In virtue of Theorem D.0.1, the spatial Fourier Transform (and its inverse) defines a continuous linear functional from $C_{FD}(\mathbb{R}^d\times\mathbb{R}^+)$ to $\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)$, and it also does it in the reversed sense from $\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)$ to $C_{FD}(\mathbb{R}^d\times\mathbb{R}^+)$, with the spatial Fourier Transform being interpreted in a distributional sense. From the continuity of $\mathscr{F}_S$ and the completeness of $C_{FD}(\mathbb{R}^d\times\mathbb{R}^+)$, it can be concluded that $\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)$ is a Fréchet space.

The dual space of $\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)$ is denoted by $\left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'$. A linear functional

$$T:\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\to\mathbb{C}$$

is a member of $\left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'$ if and only if there exist $C>0$ and $N_S,N_T\in\mathbb{N}$ such that

$$|\langle T,\psi\rangle|\leqslant C\sup_{(\xi,t)\in\mathbb{R}^d\times\mathbb{R}^+}\left|(1+|\xi|^2)^{N_S}(1+t^2)^{N_T}\mathscr{F}_S^{-1}(\psi)(\xi,t)\right|,\quad\forall\psi\in\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+). \tag{C.21}$$

The spatial Fourier Transform $\mathscr{F}_S$ over $\left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'$ is defined as the adjoint of the Fourier Transform over $C_{FD}(\mathbb{R}^d\times\mathbb{R}^+)$, whose range is the space $\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)$. We obtain thus a sequentially continuous linear operator

$$\mathscr{F}_S:\left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'\to\mathscr{M}_{SG}(\mathbb{R}^d\times\mathbb{R}^+).$$

More explicitly, $\mathscr{F}_S(T)$ is the distribution in $C'_{FD}(\mathbb{R}^d\times\mathbb{R}^+)=\mathscr{M}_{SG}(\mathbb{R}^d\times\mathbb{R}^+)$ such that

$$\langle\mathscr{F}_S(T),\psi\rangle:=\langle T,\mathscr{F}_S(\psi)\rangle,\quad\forall\psi\in C_{FD}(\mathbb{R}^d\times\mathbb{R}^+). \tag{C.22}$$

The spatial Fourier Transform and its inverse interchange thus the spaces $\left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'$ and

[2] This notation would be fully justified if it turns out that the space $\mathscr{V}(\mathbb{R}^d)$ is nuclear. See Trèves (1967, Proposition 50.7).

$\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$.

We remark that both spaces $\left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'$ and $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$ are subspaces of $\mathscr{S}'(\mathbb{R}^d \times \mathbb{R})$, that is, their members are tempered distributions. To see this, we can consider that for every $\psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R})$, we can define the application

$$\langle T, \psi \rangle := \langle T, \psi\big|_{\mathbb{R}^d \times \mathbb{R}^+} \rangle, \tag{C.23}$$

for $T \in \mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$ or $T \in \left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'$. It is easy to see that for every $\psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R})$ the restriction $\psi\big|_{\mathbb{R}^d \times \mathbb{R}^+}$ is in both spaces $C_{FD}(\mathbb{R}^d \times \mathbb{R}^+)$ and $\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)$, hence this application is well-defined. It also defines a tempered distribution, which can be obtained easily following criterion (2.49) and using criteria (C.10) and (C.14). It turns out that every distribution in $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$ or in $\left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'$ can be differentiated any number of times, and that the spatial, temporal and spatio-temporal Fourier Transforms can be applied freely. We say thus, that $\left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'$ is the space of all **tempered distributions whose spatial Fourier Transforms are slow-growing measures over** $\mathbb{R}^d \times \mathbb{R}^+$.

C.1.2 Temporal Integration and càdlàg-in-time primitives

Let $T \in \left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'$. This distribution acts as a *measure in the positive time*. To see this, we consider the following fact:

- For every $\varphi \in \mathscr{V}(\mathbb{R}^d)$, $\theta \in C_{FD}(\mathbb{R}^+) \mapsto \langle T, \varphi \boxtimes \theta \rangle$ is a slow-growing measure over $\mathbb{R}^+$.
- For every $\theta \in C_{FD}(\mathbb{R}^+)$, $\varphi \in \mathscr{V}(\mathbb{R}^d) \mapsto \langle T, \varphi \boxtimes \theta \rangle$ is in $\mathscr{V}'(\mathbb{R}^d)$.

This comes immediately from criterion (C.21). Indeed, let us consider C, N_S, N_T as in (C.21). If $\theta \in C_{FD}(\mathbb{R}^+)$, set $C_\theta = C \sup_{t\in\mathbb{R}^+} \left|(1+t^2)^{N_T}\theta(t)\right|$. Then,

$$|\langle T, \varphi \boxtimes \theta \rangle| \leqslant C_\theta \sup_{\xi\in\mathbb{R}^d} \left|(1+|x|^2)^{N_S}\mathscr{F}_S(\varphi)(\xi)\right|, \quad \forall \varphi \in \mathscr{V}(\mathbb{R}^d), \tag{C.24}$$

which shows that $\langle T, \cdot \boxtimes \theta \rangle \in \mathscr{V}'(\mathbb{R}^d)$. On the other hand, by fixing $\varphi \in \mathscr{V}(\mathbb{R}^d)$ and setting $C_\varphi = C \sup_{\xi\in\mathbb{R}^d} \left|(1+|\xi|^2)^{n_S}\mathscr{F}_S(\varphi)(\xi)\right|$ we obtain

$$|\langle T, \varphi \boxtimes \theta \rangle| \leqslant C_\varphi \sup_{t\in\mathbb{R}^+} \left|(1+t^2)^{N_T}\theta(t)\right|, \quad \forall \theta \in C_{FD}(\mathbb{R}^+). \tag{C.25}$$

Hence $\langle T, \varphi \boxtimes \cdot \rangle$ is in $C'_{FD}(\mathbb{R}^+)$, and hence it is a slow-growing measure by Riesz Representation Theorem.

The fact that T acts temporally as a measure implies that we can construct integrals with respect to its time component. To be precise, for every $\varphi \in \mathscr{V}(\mathbb{R}^d)$, we can extend the domain of definition of the application $\theta \mapsto \langle T, \varphi \boxtimes \theta \rangle$ to every measurable function with fast decreasing behaviour over $\mathbb{R}^+$. In

particular, for any $A \subset \mathbb{R}^+$ Borel and bounded, we have the right to write:

$$\langle T, \varphi \boxtimes \mathbf{1}_A \rangle, \tag{C.26}$$

expression that can be defined through limiting arguments. In particular, when A is any bounded closed interval, the definition can be done by considering limits of functions in $C_c(\mathbb{R}^+)$ which converge point-wise monotonically-decreasing and bounded by 1 to the indicator function $\mathbf{1}_A$ (a sequence of functions whose graphs form trapeziums, for instance). Hence, we can consider for any $t \in \mathbb{R}^+$, the expression $\langle T, \varphi \boxtimes \mathbf{1}_{[0,t]} \rangle$. Since T is a measure in the temporal component, the function $t \in \mathbb{R}^+ \mapsto \langle T, \varphi \boxtimes \mathbf{1}_{[0,t]} \rangle$ is a càdlàg function.

Since $\left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)' \subset \mathscr{S}'(\mathbb{R}^d \times \mathbb{R})$, the distributions in $\left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'$ can be differentiated with respect to the temporal component in a distributional sense. With these considerations, let us describe the distributions in $\left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'$ whose temporal derivatives are also in the set $\left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'$. According to our intuition, such a distribution should follow a *càdlàg behaviour in time*, since their temporal derivatives behave like a measure in the time component. The next Proposition states that this intuition is correct.

Proposition C.1.3. *Let $T \in \left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'$ such that $\frac{\partial T}{\partial t} \in \left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'$. Then, there exists a family of spatial distributions $(T_t)_{t\in\mathbb{R}^+} \subset \mathscr{V}'(\mathbb{R}^d)$ such that*

- *the function $t \in \mathbb{R}^+ \mapsto \langle T_t, \varphi \rangle$ is a càdlàg function for every $\varphi \in \mathscr{V}(\mathbb{R}^d)$.*
- *$\langle T, \varphi \boxtimes \theta \rangle = \int_{\mathbb{R}^+} \langle T_t, \varphi \rangle \theta(t) dt$ for all $\varphi \in \mathscr{V}(\mathbb{R}^d)$ and for all $\theta \in C_{FD}(\mathbb{R}^+)$.*

Before proving this Proposition, let us prove the next Lemma:

Lemma C.1.1. *Let $f : \mathbb{R} \to \mathbb{C}$ be a càdlàg function. Let $t_0 \in \mathbb{R}$. Let $(\theta_n^{(t_0)})_{n\in\mathbb{N}}$ be a sequence of functions approaching δ_{t_0} from the right-side, that is, $(\theta_n^{(t_0)})_{n\in\mathbb{N}} \subset C_c^+(\mathbb{R})$ is such that for all $n \in \mathbb{N}$, $\operatorname{supp}(\theta_n^{(t_0)}) \subset \left[t_0, t_0 + \frac{1}{n+1}\right]$ and $\int_{\mathbb{R}} \theta_n^{(t_0)}(t) dt = 1$. Then,*

$$\lim_{n\to\infty} \int_{\mathbb{R}} f(t)\theta_n^{(t_0)}(t) dt = f(t_0). \tag{C.27}$$

Analogously, if $(\theta_n^{(t_0)})_{n\in\mathbb{N}}$ is a sequence of functions approaching δ_{t_0} from the left-side, that is, $(\theta_n^{(t_0)})_{n\in\mathbb{N}} \subset C_c^+(\mathbb{R})$ is such that for all $n \in \mathbb{N}$, $\operatorname{supp}(\theta_n^{(t_0)}) \subset \left[t_0 - \frac{1}{n+1}, t_0\right]$ and $\int_{\mathbb{R}} \theta_n^{(t_0)}(t) dt = 1$. Then,

$$\lim_{n\to\infty} \int_{\mathbb{R}} f(t)\theta_n^{(t_0)}(t) dt = f(t_0^-), \tag{C.28}$$

where $f(t_0^-) = \lim_{t\to t_0^-} f(t)$.

Proof: Let $f : \mathbb{R} \mapsto \mathbb{C}$ be a càdlàg function and let $(\theta_n^{(t_0)})_{n\in\mathbb{N}}$ be such a sequence of functions approaching δ_{t_0} by the right-side. Let $\epsilon > 0$. Since f is càdlàg, there exists $\delta > 0$ [3] such that $|f(t_0) - f(t)| < \epsilon$ for all t such that $0 \leqslant t - t_0 < \delta$. Considering $n_0 \in \mathbb{N}$ such that $\frac{1}{n_0+1} < \delta$, it follows that if $n > n_0$ then

$$\left|\int_{\mathbb{R}} f(t)\theta_n^{(t_0)}(t)dt - f(t_0)\right| = \left|\int_{t_0}^{t_0+\frac{1}{n+1}} (f(t) - f(t_0))\theta_n^{(t_0)}(t)dt\right| \leqslant \epsilon \int_{t_0}^{t_0+\frac{1}{n+1}} \theta_n^{(t_0)}(t)dt = \epsilon. \quad \text{(C.29)}$$

Which proves the convergence. The converge to the left-limit $f(t_0^-)$ is done analogously. ■

Proof of Proposition C.1.3: The proof consists in verifying that the family of (evidently linear) applications

$$\varphi \mapsto \langle T_t, \varphi\rangle := \langle \frac{\partial T}{\partial t}, \varphi \boxtimes \mathbf{1}_{[0,t]}\rangle, \quad t \in \mathbb{R}^+, \quad \text{(C.30)}$$

satisfies the required properties. Let us first prove that (C.30) does define a spatial distribution in $\mathscr{V}'(\mathbb{R}^d)$ for any fixed $t \in \mathbb{R}^+$. Since $\frac{\partial T}{\partial t} \in \left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'$, there exist $C > 0$ and $N_S, N_T \in \mathbb{N}$ such that (criterion (C.21))

$$|\langle \frac{\partial T}{\partial t}, \varphi \boxtimes \theta\rangle| \leqslant C \sup_{(\xi,t)\in\mathbb{R}^d\times\mathbb{R}^+} \left|(1+|\xi|^2)^{N_S}(1+t^2)^{N_T}\mathscr{F}_S(\phi)(\xi)\theta(t)\right|, \quad \forall\varphi \in \mathscr{V}(\mathbb{R}^d), \forall\theta \in C_{FD}(\mathbb{R}^+). \quad \text{(C.31)}$$

Let us approach the function $\mathbf{1}_{[0,t]}$ by a convenient sequence of functions. Let us consider a sequence of positive functions $(\theta_n^{(t)})_{n\in\mathbb{N}} \subset C_c(\mathbb{R}^+)$, continuously differentiable over $\mathbb{R}_*^+$ and such that $\theta_n^{(t)} = 1$ over $[0,t]$ and $\theta_n^{(t)} = 0$ over $\left[t + \frac{1}{n+1}, \infty\right)$, having a decreasing behaviour over the open interval $(t, t + \frac{1}{n+1})$. Then, the functions $(\theta_n^{(t)})_{n\in\mathbb{N}}$ converge point-wise to the function $\mathbf{1}_{[0,t]}$, and they are all bounded by $\mathbf{1}_{[0,t+1]}$. It follows from Dominated Convergence Theorem (we recall that $\frac{\partial T}{\partial t}$ acts as a measure on the temporal component) that

$$\begin{aligned}\left|\langle \frac{\partial T}{\partial t}, \varphi \boxtimes \mathbf{1}_{[0,t]}\rangle\right| &= \lim_{n\to\infty} \left|\langle \frac{\partial T}{\partial t}, \varphi \boxtimes \theta_n^{(t)}\rangle\right| \\ &\leqslant \lim_{n\to\infty} C \sup_{u\in\mathbb{R}^+} \left|(1+u^2)^{N_T}\theta_n^{(t)}(u)\right| \sup_{\xi\in\mathbb{R}^d} \left|(1+|\xi|^2)^{N_S}\mathscr{F}_S(\varphi)(\xi)\right| \quad \text{(C.32)} \\ &\leqslant C(1+(t+1)^2)^{N_T} \sup_{\xi\in\mathbb{R}^d} \left|(1+|\xi|^2)^{N_S}\mathscr{F}_S(\varphi)(\xi)\right|.\end{aligned}$$

This proves that the family (C.30) defines spatial distributions in $\mathscr{V}'(\mathbb{R}^d)$ for all $t \in \mathbb{R}^+$ (criterion (C.10)). Let us fix $\varphi \in \mathscr{V}(\mathbb{R}^d)$. Since $\frac{\partial T}{\partial t}$ behaves as a measure in the temporal component, the function $t \mapsto \langle \frac{\partial T}{\partial t}, \varphi \boxtimes \mathbf{1}_{[0,t]}\rangle$ is a càdlàg function, and from (C.32) it follows that it is also polynomially bounded. Hence,

[3] Do not confuse with the Dirac measure at 0.

for every $\theta \in C_{FD}(\mathbb{R}^+)$, the integral

$$\int_{\mathbb{R}^+} \langle \frac{\partial T}{\partial t}, \varphi \boxtimes \mathbf{1}_{[0,t]} \rangle \theta(t) dt$$

is well-defined. We remark that the sequence of differentiable functions $(\theta_n^{(t)})_{n\in\mathbb{N}}$ is such that the sequence of derivatives $(\frac{d\theta_n^{(t)}}{dt})_{n\in\mathbb{N}}$ satisfies that $\mathrm{supp}(\frac{d\theta_n^{(t)}}{dt}) \subset \left[t, t+\frac{1}{n+1}\right]$, $\frac{d\theta_n^{(t)}}{dt} \leqslant 0$ and $-\int_{\mathbb{R}^+} \frac{d\theta_n^{(t)}}{dt}(s)ds = 1$ for all $n \in \mathbb{N}$, hence they approach $-\delta_t$ from the right. Let $\mu^{T,\varphi}$ be the slow-growing measure over $\mathbb{R}^+$ such that $\mu^{T,\varphi}(A) = \langle T, \varphi \boxtimes \mathbf{1}_A \rangle$ for every bounded Borel set $A \subset \mathbb{R}^+$. From these considerations, it follows that

$$\begin{aligned}
\int_{\mathbb{R}^+} \langle \frac{\partial T}{\partial t}, \varphi \boxtimes \mathbf{1}_{[0,t]} \rangle \theta(t) dt &= \lim_{n\to\infty} \int_{\mathbb{R}^+} \langle \frac{\partial T}{\partial t}, \varphi \boxtimes \theta_n^{(t)} \rangle \theta(t) dt && \text{(since } \theta_n^{(t)} \to \mathbf{1}_{[0,t]} \text{ point-wise dominated)} \\
&= \lim_{n\to\infty} -\int_{\mathbb{R}^+} \langle T, \varphi \boxtimes \frac{d\theta_n^{(t)}}{dt} \rangle \theta(t) dt && \text{(Def. of derivative)} \\
&= \lim_{n\to\infty} \int_{\mathbb{R}^+} \int_{\mathbb{R}^+} -\frac{d\theta_n^{(t)}}{dt}(s) d\mu^{T,\varphi}(s) \theta(t) dt \\
&= \lim_{n\to\infty} \int_{\mathbb{R}^+} \int_{\mathbb{R}^+} -\frac{d\theta_n^{(t)}}{dt}(s) \theta(t) dt d\mu^{T,\varphi}(s) && \text{(Fubini's Theorem)} \\
&= \int_{\mathbb{R}^+} \theta(s) d\mu^{T,\varphi}(s) = \langle T, \varphi \boxtimes \phi \rangle. && \text{(Lemma C.1.1)}
\end{aligned} \tag{C.33}$$

Hence,

$$\int_{\mathbb{R}^+} \langle \frac{\partial T}{\partial t}, \varphi \boxtimes \mathbf{1}_{[0,t]} \rangle \theta(t) dt = \int_{\mathbb{R}^+} \langle T_t, \varphi \rangle \theta(t) dt = \langle T, \varphi \boxtimes \phi \rangle, \quad \forall \varphi \in \mathscr{V}(\mathbb{R}^d), \forall \theta \in C_{FD}(\mathbb{R}^+), \tag{C.34}$$

and thus the family (C.30) satisfies all the required conditions. ■

If $T \in \left(\mathscr{V}(\mathbb{R}^d) \widehat{\boxtimes} C_{FD}(\mathbb{R}^+)\right)'$ satisfies the conditions presented in Proposition C.1.3, we say that it has a $\mathscr{V}'(\mathbb{R}^d)$**-in-space and càdlàg-in-time** behaviour. We also say that T has a **càdlàg-in-time representation** and that the associated family of spatial distributions $(T_t)_{t\in\mathbb{R}^+} \subset \mathscr{V}'(\mathbb{R}^d)$ is the **càdlàg-in-time representation** of T.

We remark that the notion of *càdlàg-in-time representation* can be applied for many other distributions, not necessarily with a spatial behaviour in $\mathscr{V}'(\mathbb{R}^d)$. We remark in particular the case where $T \in \mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$, for which we say that it has a càdlàg-in-time representation if there exists a family of slow-growing spatial measures, $(T_t)_{t\in\mathbb{R}^+} \subset \mathscr{M}_{SG}(\mathbb{R}^d)$ such that

- the function $t \in \mathbb{R}^+ \mapsto \langle T_t, \varphi \rangle$ is a càdlàg function for every $\varphi \in C_{FD}(\mathbb{R}^d)$.
- $\langle T, \varphi \boxtimes \theta \rangle = \int_{\mathbb{R}^+} \langle T_t, \varphi \rangle \theta(t) dt$ for all $\varphi \in C_{FD}(\mathbb{R}^d)$ and for all $\theta \in C_{FD}(\mathbb{R}^+)$.

The next Proposition follows almost immediately.

Proposition C.1.4. *Let $T \in \mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$. Then, T has a càdlàg-in-time representation if and only if $\mathscr{F}_S(T) \in \mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)$ has a càdlàg-in-time representation.*

Proof: Since $\mathscr{F}_S(\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)) = \mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)$, it suffices to take $(\mathscr{F}_S(T_t))_{t\in\mathbb{R}^+}$ as the family defining the càdlàg-in-time representation of $\mathscr{F}_S(T)$. ■

C.1.3 The PDE

Let us now consider the PDE (C.1). In principle, we interpret this PDE in distributional sense considering $X \in \mathscr{S}'(\mathbb{R}^d \times \mathbb{R})$ and looking at for solutions in $\mathscr{S}'(\mathbb{R}^d \times \mathbb{R})$. This can be worked out for example when $g \in \mathcal{O}_M(\mathbb{R}^d)$, since for those distributions since both operators $\frac{\partial}{\partial t}$ and $\mathcal{L}_g$ are well-defined. However, since we also want to restrain the solution to obey a particular *initial condition*, more requirements on X, and hence on the solution must be added. Let us explain a method to do this. Such a method will allow us to properly speak about an initial condition and in addition to consider the cases where g is continuous and polynomially bounded, not necessarily in $\mathcal{O}_M(\mathbb{R}^d)$.

First of all, we restrain the space where X belongs to our space $\left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'$. Let us apply the spatial Fourier Transform to problem (C.1) to obtain the spatially-multiplicative problem:

$$\frac{\partial V}{\partial t} + gV = Y, \tag{C.35}$$

where $Y = \mathscr{F}_S(X)$ and $V = \mathscr{F}_S(U)$. From the developments presented in the previous section, the distribution Y is in $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$. We recall that we have supposed $g_R \geqslant 0$. Let us define the following operator, which will be called the ***Duhamel's Operator***. Consider $\mathcal{D}_g : C_{FD}(\mathbb{R}^d \times \mathbb{R}^+) \to C_{FD}(\mathbb{R}^d \times \mathbb{R}^+)$ defined through

$$\mathcal{D}_g(\psi)(\xi, t) := \int_t^\infty e^{-(s-t)g(\xi)}\psi(\xi, s)ds \quad (\xi, t) \in \mathbb{R}^d \times \mathbb{R}^+. \tag{C.36}$$

We remark that this operator is nothing but a *temporal convolution* with the function $s \in \mathbb{R} \mapsto e^{sg}\mathbf{1}_{(-\infty,0]}(s)$ (we may, for instance, extend the domain of ψ to $\mathbb{R}^d \times \mathbb{R}$ by making it null over $\mathbb{R}^d \times \mathbb{R}^-_*$, in order to properly define such a convolution). Let us verify that this operator is well-defined. The linearity of $\mathcal{D}_g$ is straightforward. If we consider a sequence $(\xi_n, t_n)_{n\in\mathbb{N}} \subset \mathbb{R}^d \times \mathbb{R}^+$ such that $(\xi_n, t_n) \to (\xi, t) \in \mathbb{R}^d \times \mathbb{R}^+$ as $n \to \infty$, it is immediate to verify using the continuity of g and ψ that

$$e^{-(s-t_n)g(\xi_n)}\psi(\xi_n, s)\mathbf{1}_{(t_n,\infty)}(s) \to e^{-(s-t)g(\xi)}\psi(\xi, s)\mathbf{1}_{(t,\infty)}(s), \quad \text{as } n \to \infty, \forall s \in \mathbb{R}^+. \tag{C.37}$$

Considering that $g_R \geqslant 0$, we also have that $|e^{-ag(\xi)}| \leqslant 1$ for every $a \geqslant 0$, hence

$$\left|e^{-(s-t_n)g(\xi_n)}\psi(\xi_n,s)\mathbf{1}_{(t_n,\infty)}(s)\right| \leqslant |\psi(\xi_n,s)| \leqslant \frac{|\psi(\xi_n,s)|}{(1+s^2)}(1+s^2) \leqslant \frac{\sup_{(\eta,u)\in\mathbb{R}^d\times\mathbb{R}^+}\left|\psi(\eta,s)(1+s^2)\right|}{1+s^2}. \tag{C.38}$$

We conclude that the convergence (C.37) is dominated by the integrable function

$$s \mapsto \frac{\sup_{(\eta,u)\in\mathbb{R}^d\times\mathbb{R}^+}\left|\psi(\eta,s)(1+s^2)\right|}{1+s^2}$$

(we have used that $\psi \in C_{FD}(\mathbb{R}^d\times\mathbb{R}^+)$). It follows from Dominated Convergence Theorem that the function $\mathcal{D}_g(\psi)$ is continuous. Let $n_S, n_T \in \mathbb{N}$. We consider that

$$\begin{aligned}
\sup_{(\xi,t)\in\mathbb{R}^d\times\mathbb{R}^+}\left|(1+|\xi|^2)^{n_S}(1+t^2)^{n_T}\mathcal{D}_g(\psi)(\xi,t)\right| &= \sup_{(\xi,t)\in\mathbb{R}^d\times\mathbb{R}^+}\left|(1+|\xi|^2)^{n_S}(1+t^2)^{n_T}\int_t^\infty e^{-(s-t)g(\xi)}\psi(\xi,s)ds\right| \\
&\leqslant \sup_{(\xi,t)\in\mathbb{R}^d\times\mathbb{R}^+}\left\{\int_t^\infty (1+|\xi|^2)^{n_S}(1+t^2)^{n_T}|\psi(\xi,s)|ds\right\} \\
&\leqslant \sup_{(\xi,t)\in\mathbb{R}^d\times\mathbb{R}^+}\left\{\int_t^\infty (1+|\xi|^2)^{n_S}(1+s^2)^{n_T}|\psi(\xi,s)|ds\right\} \\
&\leqslant \sup_{(\xi,t)\in\mathbb{R}^d\times\mathbb{R}^+}\left\{\int_{\mathbb{R}^+} (1+|\xi|^2)^{n_S}(1+s^2)^{n_T+1}|\psi(\xi,s)|\frac{ds}{(1+s^2)}\right\} \\
&\leqslant \frac{\pi}{4}\sup_{(\xi,s)\in\mathbb{R}^d\times\mathbb{R}^+}\left|(1+|\xi|^2)^{n_S}(1+s^2)^{n_T+1}\psi(\xi,s)\right|.
\end{aligned} \tag{C.39}$$

This proves that $\mathcal{D}_g(\psi)$ is a fast-decreasing function, hence $\mathcal{D}_g$ is well-defined. In addition, this also proves that $\mathcal{D}_g$ is a continuous linear operator from $C_{FD}(\mathbb{R}^d\times\mathbb{R}^+)$ to $C_{FD}(\mathbb{R}^d\times\mathbb{R}^+)$ (Theorem D.0.1). The adjoint operator of Duhamel's operator is denoted by $\mathcal{D}_g^*$ and it is a sequentially continuous linear operator from $\mathscr{M}_{SG}(\mathbb{R}^d\times\mathbb{R}^+)$ to $\mathscr{M}_{SG}(\mathbb{R}^d\times\mathbb{R}^+)$. Explicitly,

$$\langle\mathcal{D}_g^*(T),\psi\rangle := \langle T,\mathcal{D}_g(\psi)\rangle, \quad \forall T\in\mathscr{M}_{SG}(\mathbb{R}^d\times\mathbb{R}^+), \forall\psi\in C_{FD}(\mathbb{R}^d\times\mathbb{R}^+). \tag{C.40}$$

Proposition C.1.5. *Let $Y \in \mathscr{M}_{SG}(\mathbb{R}^d\times\mathbb{R}^+) \subset \mathscr{S}'(\mathbb{R}^d\times\mathbb{R})$. Then, $\mathcal{D}_g^*(Y)$ satisfies*

$$\frac{\partial\mathcal{D}_g^*(Y)}{\partial t} + g\mathcal{D}_g^*(Y) = Y, \tag{C.41}$$

in the sense of $\mathscr{S}'(\mathbb{R}^d\times\mathbb{R})$.

Proof: Let $\psi\in\mathscr{S}(\mathbb{R}^d\times\mathbb{R})$. We have that

$$\langle\frac{\partial\mathcal{D}_g^*(Y)}{\partial t},\psi\rangle = -\langle\mathcal{D}_g^*(Y),\frac{\partial\psi}{\partial t}\rangle = -\langle\mathcal{D}_g^*(Y),\frac{\partial\psi}{\partial t}\big|_{\mathbb{R}^d\times\mathbb{R}^+}\rangle = -\langle Y,\mathcal{D}_g^*\left(\frac{\partial\psi}{\partial t}\big|_{\mathbb{R}^d\times\mathbb{R}^+}\right)\rangle. \tag{C.42}$$

Now, we consider that for every $(\xi,t)\in\mathbb{R}^d\times\mathbb{R}^+$, we have by integration by parts that

$$\begin{aligned}
-\mathcal{D}_g\left(\frac{\partial\psi}{\partial t}\Big|_{\mathbb{R}^d\times\mathbb{R}^+}\right)(\xi,t) &= -\int_t^\infty e^{-(s-t)g(\xi)}\frac{\partial\psi}{\partial t}(\xi,s)ds \\
&= -\left[e^{-(s-t)g(\xi)}\psi(\xi,s)\big|_{s=t}^{s=\infty} - \int_t^\infty e^{-(s-t)g(\xi)}(-g(\xi))\psi(\xi,s)ds\right] \\
&= -\left[-\psi(\xi,t) + g(\xi)\int_t^\infty e^{-(s-t)g(\xi)}\psi(\xi,s)ds\right] \\
&= \psi(\xi,t) - g(\xi)\mathcal{D}_g(\psi)(\xi,t).
\end{aligned} \tag{C.43}$$

From the definition of Duhamel's Operator, it is immediate that $g\mathcal{D}_g(\psi)=\mathcal{D}_g(g\psi)$. We obtain thus that

$$\langle\frac{\partial\mathcal{D}_g^*(Y)}{\partial t},\psi\rangle = \langle Y,\psi\rangle - \langle Y,\mathcal{D}_g(g\psi)\rangle = \langle Y,\psi\rangle - \langle\mathcal{D}_g^*(Y),g\psi\rangle = \langle Y,\psi\rangle - \langle g\mathcal{D}_g^*(Y),\psi\rangle, \tag{C.44}$$

where the equality $\langle\mathcal{D}_g^*(Y),g\psi\rangle = \langle g\mathcal{D}_g^*(Y),\psi\rangle$ is justified since $\mathcal{D}_g^*(Y)$ is a slow-growing measure and g is a continuous polynomially bounded continuous function, hence the multiplication is well-defined as a measure in $\mathscr{M}_{SG}(\mathbb{R}^d\times\mathbb{R}^+)$. This proves that $\mathcal{D}_g^*(Y)$ satisfies the PDE (C.41) in the sense of $\mathscr{S}'(\mathbb{R}^d\times\mathbb{R})$. ■

Corollary C.1.1. *Let $Y\in\mathscr{M}_{SG}(\mathbb{R}^d\times\mathbb{R}^+)$. Then, $\frac{\partial\mathcal{D}_g^*(Y)}{\partial t}\in\mathscr{M}_{SG}(\mathbb{R}^d\times\mathbb{R}^+)$.*

Proof: $\frac{\partial\mathcal{D}_g^*(Y)}{\partial t} = Y - g\mathcal{D}_g^*(Y)\in\mathscr{M}_{SG}(\mathbb{R}^d\times\mathbb{R}^+)$ since $\mathcal{D}_g^*$ and the multiplication by g are operations with values in $\mathscr{M}_{SG}(\mathbb{R}^d\times\mathbb{R}^+)$. ■

From Propositions C.1.3 and C.1.4, the following Corollary follows immediately.

Corollary C.1.2. *Let $Y\in\mathscr{M}_{SG}(\mathbb{R}^d\times\mathbb{R}^+)$. Then, $\mathcal{D}_g^*(Y)$ has a càdlàg-in-time representation.*

We conclude then that we can always find a solution to the transformed problem (C.35) which has a càdlàg-in-time behaviour, and hence for which the notion of an *initial condition* makes sense. Nevertheless, the following result shows that we still have some difficulties if we want to consider any arbitrary initial condition to a Cauchy problem associated to equation (C.35).

Proposition C.1.6. *$\mathcal{D}_g^*(Y)$ is the unique possible solution in $\mathscr{M}_{SG}(\mathbb{R}^d\times\mathbb{R}^+)$ to equation* (C.35).

This result follows from a simple fact, which is actually a statement equivalent to Proposition C.1.6 *the homogeneous problem*

$$\frac{\partial V_H}{\partial t} + gV_H = 0 \tag{C.45}$$

has no non-trivial solutions in $\mathscr{M}_{SG}(\mathbb{R}^d\times\mathbb{R}^+)$.

Proof: Let us suppose there are two solutions in $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$, say V_1 and V_2. Then, by linearity of the equation, the difference $V_H = V_1 - V_2 \in \mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$ must satisfy the homogeneous problem (C.45). Let us see if such kinds of solutions exist. Let us look for solutions to the homogeneous problem in the bigger space of measures $\mathscr{M}(\mathbb{R}^d \times \mathbb{R})$. We recall that $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+) \subset \mathscr{M}(\mathbb{R}^d \times \mathbb{R}^+) \subset \mathscr{D}'(\mathbb{R}^d \times \mathbb{R})$. If $V_H \in \mathscr{M}(\mathbb{R}^d \times \mathbb{R})$ satisfies (C.45), we have that $\frac{\partial V_H}{\partial t} \in \mathscr{M}(\mathbb{R}^d \times \mathbb{R})$ since $gV_H \in \mathscr{M}(\mathbb{R}^d \times \mathbb{R})$. With a typical analysis we have that V_H satisfies

$$\frac{\partial}{\partial t}\left(e^{tg}V_H\right) = 0. \tag{C.46}$$

Hence,

$$e^{tg}V_H = S \boxtimes \mathbf{1} \tag{C.47}$$

for some $S \in \mathscr{D}'(\mathbb{R}^d \times \mathbb{R})$, and since we have required that V_H must be a measure, S must be in $\mathscr{M}(\mathbb{R}^d)$. It turns out that V_H is of the form

$$V_H = e^{-tg}\left(S \boxtimes \mathbf{1}\right). \tag{C.48}$$

Since we want V_H to be slow-growing, it is necessary to require $S \in \mathscr{M}_{SG}(\mathbb{R}^d)$. However, expression (C.48) does not provide a measure with support on $\mathbb{R}^d \times \mathbb{R}^+$ unless $S = 0$. Hence, there is no solution to (C.45) in $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$ besides the trivial solution. This completes the proof. ■

In the last proof we remark that if we consider the *restriction* of the measure V_H over $\mathbb{R}^d \times \mathbb{R}^+$, which is the measure which is null outside $\mathbb{R}^d \times \mathbb{R}^+$ an equals V_H over $\mathbb{R}^d \times \mathbb{R}^+$, denoted by $\mathbf{1}_{\mathbb{R}^d \times \mathbb{R}^+}V_H$, then we do obtain a measure in $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$ if $S \in \mathscr{M}_{SG}(\mathbb{R}^d)$. However, this measure does not satisfy (C.45) in the sense of $\mathscr{S}'(\mathbb{R}^d \times \mathbb{R})$.

C.1.4 The Cauchy Problem

Let us right now concentrate, once and for all on the Cauchy problem of the form

$$\begin{cases} \dfrac{\partial U}{\partial t} + \mathcal{L}_g U = X \\ U\big|_{t=0} = U_0 \end{cases} \tag{C.49}$$

for suitable distributions X and U_0.

We require $X \in \left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'$ and $U_0 \in \mathscr{V}'(\mathbb{R}^d)$. We apply the spatial Fourier Transform and we work with the transformed Cauchy problem

$$\begin{cases} \dfrac{\partial V}{\partial t} + gV = Y \\ V\big|_{t=0} = V_0 \end{cases}, \tag{C.50}$$

where $Y = \mathscr{F}_S(X) \in \mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$, $V_0 = \mathscr{F}_S(U_0) \in \mathscr{M}_{SG}(\mathbb{R}^d)$ and $V = \mathscr{F}_S(U)$ is the transformed unknown.

Proposition C.1.6 guarantees that if we look at for solutions in $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$ to the Cauchy problem (C.50), we will have just one possibility: $\mathcal{D}^*_g(Y)$. We recall that $\mathcal{D}^*_g(Y)$ has a càdlàg-in-time representation, which we will denote by $(\mathcal{D}^*_g(Y)_t)_{t\in\mathbb{R}^+} \subset \mathscr{M}_{SG}(\mathbb{R}^d)$. This measure is the solution to the Cauchy problem (C.50) if and only if it holds that $\mathcal{D}^*_g(Y)_0 = V_0$. Hence, we have no freedom at all to fix an arbitrary initial condition. The next result proves that, with just some arrangements, we can gain more freedom in the initial condition by requiring that the restriction to $\mathbb{R}^d \times \mathbb{R}^+$ of the solution is in $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$, rather than the solution itself.

Theorem C.1.1. *Let $Y \in \mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$. Let $g : \mathbb{R}^d \to \mathbb{C}$ be a continuous symbol function such that $g_R \geqslant 0$. Let $V_0 \in \mathscr{M}_{SG}(\mathbb{R}^d)$. Then, there exists a unique measure $V \in \mathscr{M}(\mathbb{R}^d \times \mathbb{R})$ such that*

- *$\frac{\partial V}{\partial t} + gV = Y$ in the sense of $\mathscr{D}'(\mathbb{R}^d \times \mathbb{R})$.*
- *Its restriction to $\mathbb{R}^d \times \mathbb{R}^+$ is in $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$ and it has a càdlàg-in-time representation whose evaluation at $t = 0$ is V_0.*

Proof: for the existence, consider simply the measure $V \in \mathscr{M}(\mathbb{R}^d \times \mathbb{R})$ defined by

$$V = e^{-tg}\left((V_0 - \mathcal{D}^*_g(Y)_0) \boxtimes \mathbf{1}\right) + \mathcal{D}^*_g(Y), \tag{C.51}$$

which is more explicitly expressed as (we recall that $\mathcal{D}^*_g(Y)$ is a measure over $\mathbb{R}^d \times \mathbb{R}^+$ and V_0 and $\mathcal{D}^*_g(Y)_0$ are measures over $\mathbb{R}^d$):

$$\langle V, \psi\rangle = \int_{\mathbb{R}}\int_{\mathbb{R}^d} e^{-tg(\xi)}\psi(\xi,t)d(V_0 - \mathcal{D}^*_g(Y)_0)(\xi)dt + \int_{\mathbb{R}^d\times\mathbb{R}^+} \psi(\xi,t)d(\mathcal{D}^*_g(Y))(\xi,t), \quad \forall \psi \in \mathscr{D}(\mathbb{R}^d \times \mathbb{R}). \tag{C.52}$$

The fact that V satisfies the PDE (C.35) arises from Proposition C.1.5 and from the fact that the measure $e^{-tg}\left((V_0 - \mathcal{D}^*_g(Y)_0) \boxtimes \mathbf{1}\right)$ satisfies the homogeneous equation (C.45), since it is of the form (C.47).

The restriction of V over $\mathbb{R}^d \times \mathbb{R}^+$ is given by

$$\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}V = e^{-tg}\left((V_0 - \mathcal{D}^*_g(Y)_0) \boxtimes \mathbf{1}_{\mathbb{R}^+}\right) + \mathcal{D}^*_g(Y). \tag{C.53}$$

It is immediate that $\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}V$ is in $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$ since $g_R \geqslant 0$. A typical computation using the derivative of the product and Proposition C.1.5 allow to conclude that

$$\frac{\partial}{\partial t}\left(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}V\right) = -ge^{-tg}\left(V_0 - \mathcal{D}^*_g(Y)\right) \boxtimes \mathbf{1}_{\mathbb{R}^+} + e^{-gt}\left(V_0 - \mathcal{D}^*_g(Y)_0\right) \boxtimes \delta_{0_T} - g\mathcal{D}^*_g(Y) + Y, \tag{C.54}$$

where δ_{0_T} is the Dirac measure at $0 \in \mathbb{R}^+$. We remark that

$$e^{-gt}\left(V_0 - \mathcal{D}_g^*(Y)_0\right) \boxtimes \delta_{0_T} = \left(V_0 - \mathcal{D}_g^*(Y)_0\right) \boxtimes \delta_{0_T}.$$

From this we obtain that the temporal derivative of $\mathbf{1}_{\mathbb{R}^d \times \mathbb{R}^+} V$ is a measure in $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$ since it is a sum of measures in $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$. From Propositions C.1.4 and C.1.3, it follows that $\mathbf{1}_{\mathbb{R}^d \times \mathbb{R}^+} V$ has a càdlàg-in-time representation. Applying expression (C.52) to test-functions of the form $\psi = \varphi \boxtimes \theta$ with $\varphi \in C_{FD}(\mathbb{R}^d)$ and $\theta \in C_{FD}(\mathbb{R}^+)$ (which can be done since the involved distributions are measures), one obtains that the càdlàg-in-time representation of V over $\mathbb{R}^d \times \mathbb{R}^+$ (which is evidently the same as the càdlàg representation of $\mathbf{1}_{\mathbb{R}^d \times \mathbb{R}^+} V$) is given by

$$V_t = e^{-tg}\left(V_0 - \mathcal{D}_g^*(Y)_0\right) + \mathcal{D}_g^*(Y)_t \in \mathscr{M}_{SG}(\mathbb{R}^d), \quad \forall t \in \mathbb{R}^+. \tag{C.55}$$

And from this, it is immediate that the evaluation at $t = 0$ if this càdlàg representation equals the desired initial condition V_0. This proves the existence.

If we suppose that there are two measures V_1 and V_2 satisfying the conditions in Theorem C.1.1, we consider then the difference $V_H = V_1 - V_2$ must satisfies the homogenoeus problem (C.45), and hence it must be of the form (C.47) for some $S \in \mathscr{M}_{SG}(\mathbb{R}^d)$. But this implies that the evaluation at 0 of its càdlàg-in-time representation is $V_{H,0} = S$. Since in addition, $V_1\big|_{t=0} = V_2\big|_{t=0}$, then it follows that $V_{H,0}$ must be null, and hence $S = 0$. It follows that $V_H = 0$. This proves that V is the unique solution to (C.50) satisfying the required properties. ■

As stated in the proof of Theorem C.1.1, the solution V has a càdlàg-in-time representation over $\mathbb{R}^+$. Let us describe the corresponding family $(V_t)_{t\in\mathbb{R}^+}$ which determines it. For this, let us first study the càdlàg representation of $\mathcal{D}_g^*(Y)$. Let $\varphi \in C_{FD}(\mathbb{R}^d)$ and $\theta \in C_{FD}(\mathbb{R}^+)$. We have thus that

$$\begin{aligned}
\langle \mathcal{D}_g^*(Y), \varphi \boxtimes \theta \rangle &= \langle Y, \mathcal{D}_g^*(\varphi \boxtimes \theta) \rangle \\
&= \int_{\mathbb{R}^d \times \mathbb{R}^+} \mathcal{D}_g(\varphi \boxtimes \theta)(\xi, s) dY(\xi, s) \\
&= \int_{\mathbb{R}^d \times \mathbb{R}^+} \int_s^\infty e^{-(t-s)g(\xi)} \varphi(\xi)\theta(t) dt dY(\xi, s) \\
&= \int_{\mathbb{R}^d \times \mathbb{R}^+} \int_{\mathbb{R}^+} e^{-(t-s)g(\xi)} \varphi(\xi) \mathbf{1}_{[s,\infty)}(t)\theta(t) dt dY(\xi, s) \\
&= \int_{\mathbb{R}^+} \int_{\mathbb{R}^d \times \mathbb{R}^+} e^{-(t-s)g(\xi)} \varphi(\xi) \mathbf{1}_{[0,t]}(s) dY(\xi, s)\theta(t) dt,
\end{aligned} \tag{C.56}$$

where we have used Fubini's Theorem. It follows that

$$\langle \mathcal{D}_g^*(Y)_t, \varphi \rangle = \int_{\mathbb{R}^d \times \mathbb{R}^+} e^{-(t-s)g(\xi)} \varphi(\xi) \mathbf{1}_{[0,t]}(s) dY(\xi, s), \quad \forall t \in \mathbb{R}^+. \tag{C.57}$$

We write then,

$$\mathcal{D}_g^*(Y)_t = \int_{[0,t]} e^{-(t-s)g(\cdot)} Y(\cdot \times ds) \in \mathscr{M}_{SG}(\mathbb{R}^d), \quad \forall t \in \mathbb{R}^+, \tag{C.58}$$

for which we mean the measure such that

$$\mathcal{D}_g^*(Y)_t(A) = \int_{A \times [0,t]} e^{-(t-s)g(\xi)} dY(\xi, s), \quad \forall A \in \mathcal{B}_B(\mathbb{R}^d). \tag{C.59}$$

We remark that $\mathcal{D}_g^*(Y)_0$ is simply

$$\mathcal{D}_g^*(Y)_0 = Y(\cdot \times \{0\}). \tag{C.60}$$

It follows that the càdlàg-in-time representation over $\mathbb{R}^+$ of the solution V determined in Theorem C.1.1 is given by the family $(V_t)_{t\in\mathbb{R}^+} \subset \mathscr{M}_{SG}(\mathbb{R}^d)$ defined by

$$V_t = e^{-tg}\left(V_0 - Y(\cdot \times \{0\})\right) + \int_{[0,t]} e^{-(t-s)g(\cdot)} Y(\cdot \times ds), \quad \forall t \in \mathbb{R}^+. \tag{C.61}$$

Theorem C.1.2. *Let $X \in \left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'$. Let $g : \mathbb{R}^d \to \mathbb{C}$ be a continuous spatial symbol function such that $g_R \geqslant 0$. Let $U_0 \in \mathscr{V}'(\mathbb{R}^d)$. Then, there exists a unique distribution $U \in \left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'$ such that*

- *It has a càdlàg-in-time representation whose evaluation at $t = 0$ is U_0.*

- *It satisfies*

$$\langle \frac{\partial U}{\partial t} + \mathcal{L}_g U, \psi\rangle = \langle X, \psi\rangle, \quad \forall \psi \in \mathscr{S}(\mathbb{R}^d \times \mathbb{R}) \text{ such that } \operatorname{supp}\psi \subset \mathbb{R}^d \times \mathbb{R}^+. \tag{C.62}$$

Proof: Let us prove the existence. Let $Y = \mathscr{F}_S(X) \in \mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$ and $V_0 = \mathscr{F}_S(U_0) \in \mathscr{M}_{SG}(\mathbb{R}^d)$. Let us then consider the solution V of the transformed problem (C.50) obtained from Theorem C.1.1 and which is given by (C.51). Let us consider its restriction to $\mathbb{R}^d \times \mathbb{R}^+$, $\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}V$, which is given by (C.53). We define then

$$U = \mathscr{F}_S^{-1}\left(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}V\right) = \mathscr{F}_S^{-1}\left(e^{-tg}\left((V_0 - \mathcal{D}_g^*(Y)_0)\boxtimes \mathbf{1}_{\mathbb{R}^+}\right) + \mathcal{D}_g^*(Y)\right). \tag{C.63}$$

Clearly $U \in \left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'$ since it is the spatial inverse Fourier Transform of a measure in $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$. Since $\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}$ has a càdlàg-in-time representation, it follows that U also has it (Proposition C.1.4), and that its evaluation at $t = 0$ is $U_0 = \mathscr{F}_S^{-1}(V_0)$.

Using the rule of the product of the derivative and Proposition C.1.5, we obtain that

$$\begin{aligned}\frac{\partial}{\partial t}\left(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}V\right) &= -g\left(e^{-tg}\left((V_0-\mathcal{D}_g^*(Y)_0)\boxtimes\mathbf{1}_{\mathbb{R}^+}\right)+\mathcal{D}_g^*(Y)\right)+Y-e^{-tg}\left((V_0-\mathcal{D}_g^*(Y)_0)\boxtimes\delta_{0_T}\right)\\ &= -g\left(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}V\right)+Y+(V_0-\mathcal{D}_g^*(Y)_0)\boxtimes\delta_{0_T}.\end{aligned} \tag{C.64}$$

Since $\frac{\partial}{\partial t}\circ\mathscr{F}_S^{-1}=\mathscr{F}_S^{-1}\circ\frac{\partial}{\partial t}$ over $\mathscr{S}'(\mathbb{R}^d\times\mathbb{R})$, it follows that U satisfies, in the sense of $\mathscr{S}'(\mathbb{R}^d\times\mathbb{R})$, the equation

$$\frac{\partial U}{\partial t}+\mathcal{L}_gU=X+\left(U_0-\mathscr{F}_S^{-1}\left(\mathcal{D}_g^*(Y)_0\right)\right)\boxtimes\delta_{0_T}. \tag{C.65}$$

If we restrain the space of test-functions to those in $\mathscr{S}(\mathbb{R}^d\times\mathbb{R})$ such that their supports are included in $\mathbb{R}^d\times\mathbb{R}^+$, then we will obtain condition (C.62) since for such kinds of test-functions we have $\psi(\cdot,0)=0$ and hence $\langle(U_0-\mathscr{F}_S^{-1}\left(\mathcal{D}_g^*(Y)_0\right))\boxtimes\delta_{0_T},\psi\rangle=0$. This proves the existence of such a solution.

The uniqueness is proven in a typical manner, by supposing that there are two different solutions satisfying the conditions and then taking the difference between the solutions. It follows that such difference must be of the form

$$U_H=\mathscr{F}_S^{-1}(e^{-tg}S\boxtimes\mathbf{1}_{\mathbb{R}^+}) \tag{C.66}$$

for some $S\in\mathscr{V}(\mathbb{R}^d)$. U_H has a càdlàg-in-time representation which must be null. It is then immediate to conclude that $S=0$, and hence there is a unique solution satisfying the desired conditions. ■

The solution $U\in\left(\mathscr{V}(\mathbb{R}^d)\hat{\boxtimes}C_{FD}(\mathbb{R}^+)\right)'$ satisfying (C.49) in the sense of Theorem C.1.2 can be then described through its càdlàg-in-time representation $(U_t)_{t\in\mathbb{R}^+}\subset\mathscr{V}'(\mathbb{R}^d)$, given by

$$U_t=\mathscr{F}_S^{-1}\left(e^{-tg}\left(\mathscr{F}_S(U_0)-\mathscr{F}_S(X)(\cdot\times\{0\})\right)+\int_{[0,t]}e^{-(t-s)g(\cdot)}\mathscr{F}_S(X)(\cdot\times ds)\right),\quad\forall t\in\mathbb{R}^+. \tag{C.67}$$

This is simply $U_t=\mathscr{F}_S^{-1}(V_t)$, with $(V_t)_{t\in\mathbb{R}^+}\subset\mathscr{M}_{SG}(\mathbb{R}^d\times\mathbb{R}^+)$ being the càdlàg-in-time representation of the solution to the transformed problem.

C.1.5 Asymptotic Analysis

Consider now the problem (C.49) with the slightly different conditions that both $Y:=\mathscr{F}_S(X)$ and $M_X:=\mathscr{F}(X)$ are in $\mathscr{M}_{SG}(\mathbb{R}^d\times\mathbb{R})$. Hence, $X\in\mathscr{V}'(\mathbb{R}^d\times\mathbb{R})$ and the restriction of $\mathscr{F}_S(X)$ to $\mathbb{R}^d\times\mathbb{R}^+$ is in $\mathscr{M}_{SG}(\mathbb{R}^d\times\mathbb{R}^+)$. In such a case we can analyse the equation

$$\frac{\partial U}{\partial t}+\mathcal{L}_gU=X \tag{C.68}$$

following the approach presented in Section 4.6. Hence, we can solve the problem by analysing the spatio-temporal Fourier Transformed equation

$$(i\omega + g(\xi))M_U = M_X, \tag{C.69}$$

where $M_U = \mathscr{F}(U)$. Let us suppose there exists $\kappa > 0$ such that $g_R \geqslant \kappa$. In such a case, Proposition 4.6.1 guarantees the existence of a unique solution in $\mathscr{V}'(\mathbb{R}^d \times \mathbb{R})$ given by

$$U^\infty = \mathscr{F}^{-1}\left(\frac{1}{i\omega + g(\xi)} M_X\right) = \mathscr{F}^{-1}\left(\frac{1}{i\omega + g(\xi)} \mathscr{F}(X)\right). \tag{C.70}$$

We study right now the relation between the solution U^∞ and the solution U to the associated Cauchy problem studied in the previous section whose existence and uniqueness is stated in Theorem C.1.2. The next Theorem states that, actually, the solution U^∞ describes how the solution U behaves *spatio-temporally* once the time has flown long enough.

We remark that since for the Cauchy problem (C.49) we work over $\mathbb{R}^d \times \mathbb{R}^+$, the distribution X is interpreted as a restriction $\mathbf{1}_{\mathbb{R}^d \times \mathbb{R}^+} X$ when facing this Cauchy problem. This can be done since X acts as a measure in time.

Theorem C.1.3. *Let $X \in \mathscr{V}'(\mathbb{R}^d \times \mathbb{R}) \cap \mathscr{F}_S^{-1}(\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}))$. Let $g : \mathbb{R}^d \to \mathbb{C}$ be a continuous symbol function such that there exists $\kappa > 0$ such that $g_R \geqslant \kappa$. Let $U_0 \in \mathscr{V}'(\mathbb{R}^d)$. Let $U \in \left(\mathscr{V}(\mathbb{R}^d)\widehat{\boxtimes} C_{FD}(\mathbb{R}^+)\right)'$ be the unique solution to the Cauchy problem* (C.49) *(interpreting X as $\mathbf{1}_{\mathbb{R}^d \times \mathbb{R}^+} X$) satisfying conditions in Theorem C.1.2. Let U^∞ be the unique solution in $\mathscr{V}'(\mathbb{R}^d \times \mathbb{R})$ to the equation* (C.68)*. Then, for every $\epsilon > 0$ and for every $\varphi \in \mathscr{S}(\mathbb{R}^d)$, there exists $t_{\epsilon,\varphi} \in \mathbb{R}^+$ such that*

$$|\langle U^\infty - U, \varphi \boxtimes \theta\rangle| < \epsilon, \quad \forall \theta \in \mathscr{S}(\mathbb{R}) \text{ such that } \operatorname{supp}(\theta) \subset [t_{\epsilon,\varphi}, \infty) \text{ and } \int_{\mathbb{R}^+} |\theta(t)| dt = 1. \tag{C.71}$$

Proof: We note as usual $Y = \mathscr{F}_S(X) \in \mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R})$ and $V_0 = \mathscr{F}_S(U_0) \in \mathscr{M}_{SG}(\mathbb{R}^d)$. We recall that the Cauchy problem (C.49) is analysed with $Y \in \mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R}^+)$ rather than in $\mathscr{M}_{SG}(\mathbb{R}^d \times \mathbb{R})$. Hence, the solution U in this case is expressed through the restriction of V to $\mathbb{R}^d \times \mathbb{R}^+$ as

$$U = \mathscr{F}_S^{-1}\left(e^{-tg}\left((V_0 - \mathcal{D}_g^*(\mathbf{1}_{\mathbb{R}^d \times \mathbb{R}^+} Y)_0) \boxtimes \mathbf{1}_{\mathbb{R}^+}\right) + \mathcal{D}_g^*(\mathbf{1}_{\mathbb{R}^d \times \mathbb{R}^+} Y)\right). \tag{C.72}$$

We consider the difference $U^\infty - U$ acting on a test-function of the form $\psi = \varphi \boxtimes \theta$, with $\varphi \in \mathscr{S}(\mathbb{R}^d)$ and

$\theta \in \mathscr{S}(\mathbb{R})$ with $\text{supp}(\theta) \subset \mathbb{R}^+$. Considering that

$$\begin{aligned} U^\infty - U &= \mathscr{F}^{-1}\left(\frac{1}{i\omega+g}\mathscr{F}(X)\right) - \mathscr{F}_S^{-1}\left(e^{-tg}\left((V_0 - \mathcal{D}_g^*(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}Y)_0)\boxtimes \mathbf{1}_{\mathbb{R}^+}\right) + \mathcal{D}_g^*(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}Y)\right) \\ &= \mathscr{F}_S^{-1}\left(\mathscr{F}_T^{-1}\left(\frac{1}{i\omega+g}\mathscr{F}_T(Y)\right) - \mathcal{D}_g^*(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}Y)\right) - \mathscr{F}_S^{-1}\left(e^{-tg}\left((V_0 - \mathcal{D}_g^*(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}Y)_0)\boxtimes \mathbf{1}_{\mathbb{R}^+}\right)\right), \end{aligned} \tag{C.73}$$

let us compare $\mathscr{F}_T^{-1}\left(\frac{1}{i\omega+g}\mathscr{F}_T(Y)\right)$ and $\mathcal{D}_g^*(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}Y)$ acting on test-functions of the already mentioned form. We consider that

$$\langle \mathscr{F}_T^{-1}\left(\frac{1}{i\omega+g}\mathscr{F}_T(Y)\right), \varphi\boxtimes\theta\rangle = \langle Y, \varphi\mathscr{F}_T\left(\frac{1}{i\omega+g}\mathscr{F}_T^{-1}(\theta)\right)\rangle. \tag{C.74}$$

Using the exchange formula for the temporal Fourier Transform, which holds since the function $\omega \mapsto \frac{1}{i\omega+g(\xi)}$ is in $\mathcal{O}_M(\mathbb{R})$ for all $\xi \in \mathbb{R}^d$ since $g_R \geqslant \kappa > 0$, we have that

$$\mathscr{F}_T\left(\frac{1}{i\omega+g}\mathscr{F}_T^{-1}(\theta)\right) = \frac{1}{\sqrt{2\pi}}\mathscr{F}_T\left(\frac{1}{i\omega+g}\right) \overset{(\mathbb{R})}{*} \theta. \tag{C.75}$$

Here $\overset{(\mathbb{R})}{*}$ denotes a convolution with respect to the temporal component. A known result in Fourier Analysis, which is obtained considering that $g_R \geqslant \kappa > 0$, is that

$$\mathscr{F}_T\left(\frac{1}{i\omega+g}\right)(t) = \sqrt{2\pi}e^{tg}\mathbf{1}_{\mathbb{R}_*^-}(t) \tag{C.76}$$

in distributional sense. Hence,

$$\varphi\mathscr{F}_T\left(\frac{1}{i\omega+g}\mathscr{F}_T^{-1}(\theta)\right) = \varphi\left(e^{tg}\mathbf{1}_{\mathbb{R}_*^-} \overset{(\mathbb{R})}{*} \theta\right). \tag{C.77}$$

And as we have mentioned in Section C.1.3, doing this convolution and applying Duhamel's operator is equivalent over test-functions in $C_{FD}(\mathbb{R}^d \times \mathbb{R}^+)$. Since $\text{supp}(\theta) \subset \mathbb{R}^+$, we are in this case. We conclude that

$$\langle \mathscr{F}_T^{-1}\left(\frac{1}{i\omega+g}\mathscr{F}_T(Y)\right), \varphi\boxtimes\theta\rangle = \langle \mathcal{D}_g^*(Y), \varphi\boxtimes\theta\rangle, \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^d), \theta \in \mathscr{S}(\mathbb{R}) \text{ such that } \text{supp}\,\theta \subset \mathbb{R}^+. \tag{C.78}$$

It follows that

$$\begin{aligned} |\langle U^\infty - U, \varphi\boxtimes\theta\rangle| &= \left|\langle \mathscr{F}_S^{-1}\left(e^{-tg}\left((V_0 - \mathcal{D}_g^*(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}Y)_0)\boxtimes\mathbf{1}_{\mathbb{R}^+}\right)\right), \varphi\boxtimes\theta\rangle\right| \\ &= \left|\langle e^{-tg}\left((V_0 - \mathcal{D}_g^*(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}Y)_0)\boxtimes\mathbf{1}_{\mathbb{R}^+}\right), \mathscr{F}_S^{-1}(\varphi)\boxtimes\theta\rangle\right| \\ &= \left|\int_{\mathbb{R}^d}\int_{\mathbb{R}^+} e^{-tg(\xi)}\mathscr{F}_S^{-1}(\varphi)(\xi)\theta(t)dtd(V_0 - \mathcal{D}_g^*(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}Y)_0)(\xi)\right|. \end{aligned} \tag{C.79}$$

We suppose that $V_0 \neq \mathcal{D}_g^*(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}Y)_0$. Otherwise, the result is obvious. Let us fix $\varphi \in \mathscr{S}(\mathbb{R}^d)$. Since $g_R \geqslant \kappa > 0$, $|e^{-tg}| \leqslant e^{-t\kappa}$. We can hence bound (C.79) by

$$\begin{aligned} |\langle U^\infty - U, \varphi \boxtimes \theta\rangle| &\leqslant \int_{\mathbb{R}^d}\int_{\mathbb{R}^+} \left|e^{-tg(\xi)}\mathscr{F}_S^{-1}(\varphi)(\xi)\theta(t)\right| dt d \left|V_0 - \mathcal{D}_g^*(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}Y)_0\right|(\xi) \\ &\leqslant \int_{\mathbb{R}^d} \left|\mathscr{F}_S^{-1}(\varphi)(\xi)\right| d \left|V_0 - \mathcal{D}_g^*(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}Y)_0\right|(\xi) \int_{\mathbb{R}^+} e^{-t\kappa}|\theta(t)|dt \end{aligned} \tag{C.80}$$

We set

$$C_\varphi := \int_{\mathbb{R}^d} \left|\mathscr{F}_S^{-1}(\varphi)(\xi)\right| d \left|V_0 - \mathcal{D}_g^*(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}Y)_0\right|(\xi)$$

which is a non-null positive real number since the measure $V_0 - \mathcal{D}_g^*(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}Y)_0 \neq 0$ is in $\mathscr{M}_{SG}(\mathbb{R}^d)$. Let $\epsilon > 0$. if we chose $t_{\epsilon,\varphi} \in \mathbb{R}^+$ such that $e^{-\kappa t_{\epsilon,\varphi}} \leqslant \frac{\epsilon}{C_\varphi}$, it follows that if we take $\theta \in \mathscr{S}(\mathbb{R})$ such that $\mathrm{supp}(\theta) \subset [t_{\epsilon,\varphi}, \infty)$ and $\int_{\mathbb{R}^+}|\theta(t)|dt = 1$,

$$\begin{aligned} |\langle U^\infty - U, \varphi \boxtimes \theta\rangle| &\leqslant C_\varphi \int_{\mathbb{R}^+} e^{-t\kappa}|\theta(t)|dt \\ &\leqslant C_\varphi \int_{[t_{\epsilon,\varphi},\infty)} e^{-t_{\epsilon,\varphi}\kappa}|\theta(t)|dt = C_\varphi\frac{\epsilon}{\varphi}\int_{[t_{\epsilon,\varphi},\infty)}|\theta(t)|dt = \epsilon. \end{aligned} \tag{C.81}$$

This proves the desired result. ■

Hence, as the time flows, the solution U is more and more similar to the solution U^∞. We remark that, although U has a càdlàg-in-time representation, it is not clear if the solution U^∞ does. The *asymptotic convergence* described in Theorem C.1.3 is obtained only in a suitable sense of distributions. Let us consider, however, the case where $M_{U^\infty} = \mathscr{F}(U^\infty)$ is temporally integrable, which is the case when

$$\int_{A\times\mathbb{R}} \frac{dM_X(\xi,\omega)}{i\omega + g(\xi)} < \infty, \quad \forall A \in \mathcal{B}_B(\mathbb{R}^d). \tag{C.82}$$

Then, U^∞ is a *continuous-in-time distribution* (see Eq. (5.24), interpreted in a deterministic context). There exists then a continuous-in-time representation $(U_t^\infty)_{t\in\mathbb{R}} \subset \mathscr{V}'(\mathbb{R}^d)$.

Proposition C.1.7. *If $U_0 = U_0^\infty$ in the Cauchy problem* (C.49) *(with X restricted to $\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}$), then its solution U and the distribution U^∞ coincide over $\mathbb{R}^+$.*

Proof: It suffices to show that

$$\mathscr{F}_S^{-1}\left(\mathcal{D}_g^*\left(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}\mathscr{F}_S(X)\right)\right) = \mathscr{F}^{-1}\left(\frac{1}{i\omega+g}\mathscr{F}(X)\right)$$

over $\mathbb{R}^d \times \mathbb{R}^+$. For that, consider $\varphi \in \mathscr{S}(\mathbb{R}^d)$ and $\theta \in \mathscr{S}(\mathbb{R})$ such that $\mathrm{supp}(\theta) \subset \mathbb{R}^+$. We have

$$\begin{aligned}\left\langle \mathscr{F}_S^{-1}\left(\mathcal{D}_g^*\left(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}\mathscr{F}_S(X)\right)\right), \varphi\boxtimes\theta\right\rangle &= \left\langle \mathcal{D}_g^*\left(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}\mathscr{F}_S(X)\right), \mathscr{F}_S^{-1}(\varphi)\boxtimes\theta\right\rangle \\ &= \left\langle \mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}\mathscr{F}_S(X), \mathcal{D}_g\left(\mathscr{F}_S^{-1}(\varphi)\boxtimes\theta\right)\right\rangle \\ &= \left\langle \mathscr{F}_S(X), \mathcal{D}_g\left(\mathscr{F}_S^{-1}(\varphi)\boxtimes\theta\right)\right\rangle \\ &= \left\langle X, \mathscr{F}_S\left(\mathcal{D}_g\left(\mathscr{F}_S^{-1}(\varphi)\boxtimes\theta\right)\right)\right\rangle,\end{aligned} \tag{C.83}$$

where the equality $\left\langle \mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}\mathscr{F}_S(X), \mathcal{D}_g\left(\mathscr{F}_S^{-1}(\varphi)\boxtimes\theta\right)\right\rangle = \left\langle \mathscr{F}_S(X), \mathcal{D}_g\left(\mathscr{F}_S^{-1}(\varphi)\boxtimes\theta\right)\right\rangle$ is justified since $\mathscr{F}_S(X)$ is a measure and $\mathrm{supp}(\theta) \subset \mathbb{R}^+$ and hence we have $\mathrm{supp}(\mathcal{D}_g\left(\mathscr{F}_S^{-1}(\varphi)\boxtimes\theta\right)) \subset \mathbb{R}^d \times \mathbb{R}^+$. We continue our development to obtain

$$\begin{aligned}\left\langle X, \mathscr{F}_S\left(\mathcal{D}_g\left(\mathscr{F}_S^{-1}(\varphi)\boxtimes\theta\right)\right)\right\rangle &= \left\langle \mathcal{D}_g^*\left(\mathscr{F}_S(X)\right), \mathscr{F}_S^{-1}(\varphi)\boxtimes\theta\right\rangle \\ &= \left\langle \mathscr{F}_S^{-1}\left(\mathscr{F}_T^{-1}\left(\frac{1}{i\omega+g}\mathscr{F}_T\left(\mathscr{F}_S(X)\right)\right)\right), \varphi\boxtimes\theta\right\rangle \\ &= \left\langle \mathscr{F}^{-1}\left(\frac{1}{i\omega+g}\mathscr{F}(X)\right), \varphi\boxtimes\theta\right\rangle,\end{aligned} \tag{C.84}$$

where we have used that $\mathcal{D}_g^* = \mathscr{F}_T^{-1}(\frac{1}{i\omega+g}\mathscr{F}_T(\cdot))$ over $\mathbb{R}^d \times \mathbb{R}^+$, as we have already proved in Eq. (C.78). The equality follows, and hence both $\mathscr{F}_S^{-1}\left(\mathcal{D}_g^*\left(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}\mathscr{F}_S(X)\right)\right)$ and $\mathscr{F}^{-1}\left(\frac{1}{i\omega+g}\mathscr{F}(X)\right)$ have the same representation over $\mathbb{R}^+$ which is actually a continuous-in-time representation. In particular,

$$\mathscr{F}_S^{-1}\left(\mathcal{D}_g^*\left(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}\mathscr{F}_S(X)\right)\right)_0 = \mathscr{F}^{-1}\left(\frac{1}{i\omega+g}\mathscr{F}(X)\right)_0 = U_0^\infty.$$

Since the solution U is given by

$$U = \mathscr{F}_S^{-1}\left(e^{-tg}\left(U_0^\infty - \mathcal{D}_g^*(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}\mathscr{F}_S(X))_0\right)\boxtimes\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}\right) + \mathscr{F}_S^{-1}\left(\mathcal{D}_g^*\left(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}\mathscr{F}_S(X)\right)\right),$$

it follows that over $\mathbb{R}^d \times \mathbb{R}^+$:

$$U = 0 + \mathscr{F}_S^{-1}\left(\mathcal{D}_g^*\left(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}\mathscr{F}_S(X)\right)\right) = \mathscr{F}^{-1}\left(\frac{1}{i\omega+g}\mathscr{F}(X)\right) = U^\infty. \quad \blacksquare$$

C.2 *Stochastized* version

Let us now consider the problem (C.49) but with the elements being interpreted as convenient GeRFs with analogue properties to the deterministic case. Say, X is a GeRF over $\mathbb{R}^d \times \mathbb{R}$ such that its spatial Fourier Transform $Y = \mathscr{F}_S(X)$ is a slow-growing Random Measure concentrated on $\mathbb{R}^d \times \mathbb{R}^+$, and U_0 is a GeRF over $\mathbb{R}^d$ such that its (spatial) Fourier Transform is a slow-growing Random Measure. The function g is,

as usual, a continuous spatial symbol function such that $g_R \geqslant 0$. The resolution in this case is done in a *completely analogous way* to the method of the deterministic case. Indeed, Duhamel's operator and the other operations used to construct the solution (C.63) can be applied without technical difficulties to the stochastic case, since they are all defined through their actions over test-functions. The only concept which deserves a little review is the one of càdlàg-in-time representation.

We suppose that all the random objects have zero mean.

C.2.1 Some facts about mean-square càdlàg Random Functions

We first make explicit the definition of a càdlàg in mean-square temporal Random Function. Let $(Z(t))_{t\in\mathbb{R}}$ be a real Random Function over $\mathbb{R}$. We say that Z is **càdlàg in mean-square**, or simply **càdlàg**, if

- for all $t_0 \in \mathbb{R}$, $Z(t) \overset{L^2(\Omega)}{\to} Z(t_0)$ if $t \to t_0^+$,
- for all $t_0 \in \mathbb{R}$, there exists a random variable $Z(t_0^-) \in L^2(\Omega, \mathcal{A}, \mathbb{P})$ such that $Z(t) \overset{L^2(\Omega)}{\to} Z(t_0^-)$ if $t \to t_0^-$.

Càdlàg temporal Random Functions have many similar properties to deterministic ones. In particular, they are continuous outside an at most countable set. To see this, we can apply the same arguments as used to prove such a property for deterministic cases. See for instance the Theorems in Swanson (2011), which can be applied to our case since a temporal Random Function is a function from $\mathbb{R}$ to the metric space $L^2(\Omega, \mathcal{A}, \mathbb{P})$. The main consequence of this fact is that we can define the integral with respect to the Lebesgue measure of a càdlàg Random Function without technical difficulties, just by considering the sum of the integrals outside the discontinuity points, where the procedure presented in Section 3.2.2 can be used. Hence, for every compactly supported and bounded measurable function $f : \mathbb{R} \to \mathbb{C}$, one can define

$$\int_{\mathbb{R}} f(t)Z(t)dt. \tag{C.85}$$

The following Theorem is a simplified stochastic version of Fubini's Theorem.

Theorem C.2.1 (Stochastic Fubini's Theorem). *Let M be a Random Measure over $\mathbb{R}^d$ and let $\mu \in \mathscr{M}(\mathbb{R}^m)$. Let $f : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{C}$ be a measurable and bounded function with compact support such that the Random Function over $\mathbb{R}^m$, $y \mapsto \int_{\mathbb{R}^d} f(x,y)dM(x)$, is continuous outside a $\mu-$null set. Then,*

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^m} f(x,y)d\mu(y)dM(x) = \int_{\mathbb{R}^m}\int_{\mathbb{R}^d} f(x,y)dM(x)d\mu(y). \tag{C.86}$$

The condition on f in Theorem C.2.1 is required for the simple reason that in this dissertation we have not defined the stochastic integral of a Random Function that is not continuous outside a non-null-measure Borel

set (we have only done it for continuous Random Functions or for càdlàg Random Functions in dimension $d = 1$). In order to prove this Theorem, we need the following Lemma.

Lemma C.2.1. *Let $\mu \in \mathscr{M}(\mathbb{R}^d)$ and let Z be a Random Function over $\mathbb{R}^d$ which is continuous outside a $\mu-$null set and such that it is null outside a compact set. Then,*

$$\mathbb{C}ov\left(\int_{\mathbb{R}^d} Z(x)d\mu(x), X\right) = \int_{\mathbb{R}^d} \mathbb{C}ov\left(Z(x), X\right) d\mu(x), \quad \forall X \in L^2(\Omega, \mathcal{A}, \mathbb{P}). \tag{C.87}$$

Proof: Let $X \in L^2(\Omega, \mathcal{A}, \mathbb{P})$. Let $B \in \mathcal{B}_B(\mathbb{R}^d)$ such that Z is null outside B and let $D \in \mathcal{B}(\mathbb{R}^d)$ such that $|\mu|(D) = 0$ and such that Z is continuous outside D. Set $K = B \cap D^c$. Since Z is continuous over K, the integral of Z with respect to μ is well-defined through the use of a Riemann sequence of partitions of K, $(V_j^N)_{j\in\{1,\ldots,N\},N\in\mathbb{N}_*} \subset \mathcal{B}_B(\mathbb{R}^d)$ together with associated tag-points $(x_j^N)_{j\in\{1,\ldots,N\},N\in\mathbb{N}_*}$, having

$$\int_{\mathbb{R}^d} Z(x)d\mu(x) = \int_K Z(x)d\mu(x) = \lim_{N\to\infty} \sum_{j=1}^N Z(x_j^N)\mu(K \cap V_j^N), \tag{C.88}$$

where the limit is taken in the sense of $L^2(\Omega, \mathcal{A}, \mathbb{P})$ (Definition 3.2.1). We remark that the deterministic function $x \mapsto \mathbb{C}ov(Z(x), X)$ is continuous over K, which can be verified using the Cauchy-Schwarz inequality and using the continuity of Z over K. Its integral over K can also then be obtained as the limit of its corresponding evaluations in the tag points of the sequence $(V_j^N)_{j\in\{1,\ldots,N\},N\in\mathbb{N}_*}$. One has then,

$$\begin{aligned}
\mathbb{C}ov\left(\int_{\mathbb{R}^d} Z(x)d\mu(x), X\right) &= \lim_{N\to\infty} \sum_{j=1}^N \mathbb{C}ov(Z(x_j^N), X)\mu(V_j^N \cap K) \\
&= \lim_{N\to\infty} \mathbb{C}ov\left(\sum_{j=1}^N Z(x_j^N)\mu(V_j^N \cap K), X\right) \\
&= \lim_{N\to\infty} \sum_{j=1}^N \mathbb{C}ov\left(Z(x_j^N), X\right)\mu(V_j^N \cap K) \\
&= \int_K \mathbb{C}ov(Z(x), X)d\mu(x) = \int_{\mathbb{R}^d} \mathbb{C}ov(Z(x), X)d\mu(x). \quad \blacksquare
\end{aligned} \tag{C.89}$$

Proof of Stochastic Fubini's Theorem: Since the deterministic function $x \in \mathbb{R}^d \mapsto \int_{\mathbb{R}^m} f(x, y)d\mu(y)$ is a measurable function with compact support (typical result from Measure Theory) which in addition is bounded, it is integrable with respect to M (Proposition 3.3.1). Both integrals in (C.86) are then well-defined as random variables in $L^2(\Omega, \mathcal{A}, \mathbb{P})$. We consider then the value

$$\mathbb{V}ar\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^m} f(x, y)d\mu(y)dM(x) - \int_{\mathbb{R}^m}\int_{\mathbb{R}^d} f(x, y)dM(x)d\mu(y)\right), \tag{C.90}$$

which equals

$$\begin{aligned}
&\mathbb{C}ov\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^m} f(x,y)d\mu(y)dM(x), \int_{\mathbb{R}^d}\int_{\mathbb{R}^m} f(u,v)d\mu(v)dM(u)\right)\\
&+\mathbb{C}ov\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^m} f(x,y)d\mu(y)dM(x), \int_{\mathbb{R}^m}\int_{\mathbb{R}^d} f(u,v)dM(u)d\mu(v)\right)\\
&-\mathbb{C}ov\left(\int_{\mathbb{R}^m}\int_{\mathbb{R}^d} f(x,y)dM(x)d\mu(y), \int_{\mathbb{R}^d}\int_{\mathbb{R}^m} f(x,y)d\mu(y)dM(x)\right)\\
&-\mathbb{C}ov\left(\int_{\mathbb{R}^m}\int_{\mathbb{R}^d} f(x,y)dM(x)d\mu(y), \int_{\mathbb{R}^m}\int_{\mathbb{R}^d} f(u,v)dM(u)d\mu(v)\right).
\end{aligned} \tag{C.91}$$

Using conveniently Lemma C.2.1 and Proposition 3.3.1, one concludes that this expression equals

$$\begin{aligned}
&\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_{\mathbb{R}^m} f(x,y)d\mu(y)\overline{\int_{\mathbb{R}^m} f(u,v)d\mu(v)}dC_M(x,u)\\
&+\int_{\mathbb{R}^m}\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_{\mathbb{R}^m} f(x,y)d\mu(y)\overline{f(u,v)}dC_M(x,u)d\overline{\mu}(v)\\
&-\int_{\mathbb{R}^m}\int_{\mathbb{R}^d\times\mathbb{R}^d} f(x,y)\overline{\int_{\mathbb{R}^m} f(u,v)d\mu(v)}dC_M(x,u)d\mu(y)\\
&-\int_{\mathbb{R}^m}\int_{\mathbb{R}^m}\int_{\mathbb{R}^d\times\mathbb{R}^d} f(x,y)\overline{f(u,v)}dC_M(x,u)d\mu(y)d\overline{\mu}(v),
\end{aligned} \tag{C.92}$$

where C_M is the covariance measure of M. The classical deterministic Fubini's Theorem guarantees that all the repeated integrals in (C.92) are equal, hence expression (C.90) equals 0. This proves the Theorem. ■

The next Proposition presents a simple analogue to the case of deterministic functions whose derivatives are measures.

Proposition C.2.1. *Let M be a Random Measure over $\mathbb{R}$. Then, the Random Function defined by*

$$Z(t) = M([0,t])\mathbf{1}_{\mathbb{R}^+}(t) - M\left((t,0)\right)\mathbf{1}_{\mathbb{R}_*^-}(t), \quad t\in\mathbb{R}, \tag{C.93}$$

is a càdlàg Random Function whose derivative in distributional sense is M.

Proof: Let C_M be the covariance measure of M. Let $t_0\in\mathbb{R}^+$, and let $\Delta t>0$. One has that

$$\begin{aligned}
\mathbb{E}\left(|Z(t_0+\Delta t)-Z(t_0)|^2\right) &= \mathbb{E}\left(|M([0,t_0+\Delta t])-M([0,t_0])|^2\right)\\
&= \mathbb{E}\left(|M((t_0,t_0+\Delta t])|^2\right)\\
&= C_M((t_0,t_0+\Delta t]\times(t_0,t_0+\Delta t]).
\end{aligned} \tag{C.94}$$

Since C_M is a measure over $\mathbb{R}\times\mathbb{R}$ and the set $(t_0,t_0+\Delta t]\times(t_0,t_0+\Delta t]$ decreases to $\varnothing$ as $\delta t\to 0$, one

has that expression (C.94) goes to 0 as $\Delta t \to 0$, and hence Z is *mean-square right-continuous*. If $t_0 < 0$ and Δt is small enough, the same argument can be used to prove that Z is also right-continuous in mean-square over $\mathbb{R}_*^-$.

Let now $t_0 < 0$. We consider that

$$\begin{aligned}\mathbb{E}\left(|Z(t_0-\Delta t)-M([t_0,0))|^2\right) &= \mathbb{E}\left(|M((t_0-\Delta t,0))-M([t_0,0))|^2\right)\\ &= \mathbb{E}\left(|M((t_0-\Delta t,t_0))|^2\right)\\ &= C_M\left((t_0-\Delta t,t_0)\times(t_0-\Delta t,t_0)\right),\end{aligned} \tag{C.95}$$

expression which goes to 0 as $\Delta t \to 0$ since $(t_0-\Delta t,t_0)\times(t_0-\Delta t,t_0)$ decreases to $\varnothing$ as $\Delta t \to 0$ and C_M is a measure. Hence Z has *left-limits in mean-square* over every $t_0 \in \mathbb{R}_*^-$, which are given by the random variables of the form $M([t_0,0))$. This same argument is applied for $t_0 = 0$, using as limit a null random variable, and for $t_0 > 0$ using as limit the random variable $M([0,t_0))$. This proves that Z is càdlàg in mean-square.

Let us now prove that $\frac{dZ}{dt} = M$ in distributional sense. For that, we consider $\theta \in \mathcal{D}(\mathbb{R})$ and we see that,

$$\begin{aligned}\langle M,\theta\rangle &= \int_{\mathbb{R}}\theta(s)dM(s)\\ &= \int_{\mathbb{R}^+}\theta(s)dM(s)+\int_{\mathbb{R}_*^-}\theta(s)dM(s)\\ &= \int_{\mathbb{R}^+}-\int_s^\infty\frac{d\theta}{dt}(t)dtdM(s)+\int_{\mathbb{R}_*^-}\int_{-\infty}^s\frac{d\theta}{dt}(t)dtdM(s)\\ &= -\int_{\mathbb{R}^+}\int_{\mathbb{R}^+}\frac{d\theta}{dt}(t)\mathbf{1}_{[s,\infty)}(t)dtdM(s)+\int_{\mathbb{R}_*^-}\int_{\mathbb{R}_*^-}\frac{d\theta}{dt}(t)\mathbf{1}_{(-\infty,s)}(t)dtdM(s)\\ &= -\int_{\mathbb{R}^+}\int_{\mathbb{R}^+}\frac{d\theta}{dt}(t)\mathbf{1}_{[0,t]}(s)dtdM(s)+\int_{\mathbb{R}_*^-}\int_{\mathbb{R}_*^-}\frac{d\theta}{dt}(t)\mathbf{1}_{(t,0)}(s)dtdM(s),\end{aligned} \tag{C.96}$$

we use Stochastic Fubini's Theorem C.2.1 to obtain

$$\begin{aligned}\langle M,\theta\rangle &= -\int_{\mathbb{R}^+}\int_{\mathbb{R}^+}\frac{d\theta}{dt}(t)\mathbf{1}_{[0,t]}(s)dM(s)dt+\int_{\mathbb{R}_*^-}\int_{\mathbb{R}_*^-}\frac{d\theta}{dt}(t)\mathbf{1}_{(t,0)}(s)dM(s)dt\\ &= -\int_{\mathbb{R}^+}\frac{d\theta}{dt}(t)M\left([0,t]\right)dt+\int_{\mathbb{R}_*^-}\frac{d\theta}{dt}(t)M\left((t,0)\right)dt\\ &= -\int_{\mathbb{R}}\frac{d\theta}{dt}(t)Z(t)dt\\ &= -\langle Z,\frac{d\theta}{dt}\rangle = \langle\frac{dZ}{dt},\theta\rangle. \quad \blacksquare\end{aligned} \tag{C.97}$$

We finish by describing the covariance of a càdlàg Random Function.

Proposition C.2.2. *Let Z be a càdlàg real Random Function over $\mathbb{R}$. Then, its covariance function $C_Z : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ satisfies:*

- *for all $t \in \mathbb{R}$, $C_Z(t, \cdot)$ is a càdlàg function, and*
- *the function $t \mapsto C_Z(t, t)$ is càdlàg.*

The first condition in Proposition C.2.2 is equivalent to say that C_Z is *separately càdlàg*, that is, that the one-variable function obtained when fixing any of the two components is a càdlàg function. The equivalence is obtained immediately from the symmetry of a covariance function.

Proof: Let us prove the necessity. Let $t_0 \in \mathbb{R}$ be fixed. Then, by Cauchy-Schwarz inequality one has for every $s, s_0 \in \mathbb{R}$,

$$|C_Z(t_0, s) - C_Z(t_0, s_0)| = |\mathbb{E}\left((Z(s) - Z(s_0))Z(t_0)\right)| \leqslant \sqrt{\mathbb{V}ar\left(Z(t_0)\right)}\sqrt{\mathbb{V}ar\left(Z(s) - Z(s_0)\right)}. \quad \text{(C.98)}$$

Since Z is càdlàg, then $\mathbb{V}ar\left(Z(s) - Z(s_0)\right) \to 0$ as $s \to s_0^+$. This proves that the function $C_Z(t_0, \cdot)$ is right-continuous. Now, let $(t_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ be a sequence of real numbers such that $s_n \to s_0^-$. Let us consider the sequence of real numbers $(C_Z(t_0, s_n))_{n\in\mathbb{N}}$. Let $n, m \in \mathbb{N}$. Using again the Cauchy-Schwarz inequality, one obtains,

$$|C_Z(t_0, s_n) - C_Z(t_0, s_m)| \leqslant \sqrt{\mathbb{V}ar\left(Z(t_0)\right)}\sqrt{\mathbb{V}ar\left(Z(s_n) - Z(s_m)\right)}. \quad \text{(C.99)}$$

Since Z is càdlàg, the sequence $(Z(s_n))_{n\in\mathbb{N}}$ converges in $L^2(\Omega, \mathcal{A}, \mathbb{P})$, and hence it is a Cauchy sequence. It follows that $\mathbb{V}ar\left(Z(s_n) - Z(s_m)\right) \to 0$ as $n, m \to \infty$. It follows that the sequence $(C_Z(t_0, s_n))_{n\in\mathbb{N}}$ is Cauchy and hence it converges to a limit as $s_n \to s_0^-$. This proves that $C_Z(t_0, \cdot)$ has left-limits, and hence it is càdlàg. In order to prove the second condition, we consider that if $t \in \mathbb{R}$, then

$$\begin{aligned} |C_Z(t, t) - C_Z(t_0, t_0)| &= \left|\mathbb{E}\left(Z^2(t) - Z^2(t_0)\right)\right| \\ &= \left|\mathbb{E}\left(\left(Z(t) - Z(t_0)\right)^2\right) + 2\mathbb{E}\left(Z(t_0)\left(Z(t) - Z(t_0)\right)\right)\right| \\ &\leqslant \underbrace{\mathbb{E}\left(|Z(t) - Z(t_0)|^2\right)}_{\to 0 \text{ as } t\to t_0^+} + 2\sqrt{\mathbb{V}ar\left(Z(t_0)\right)}\underbrace{\sqrt{\mathbb{V}ar\left(Z(t) - Z(t_0)\right)}}_{\to 0 \text{ as } t\to t_0^+}, \end{aligned} \quad \text{(C.100)}$$

where we have used the Cauchy-Schwarz inequality and that Z is càdlàg. This proves that $t \mapsto C_Z(t, t)$ is right-continuous. Finally, consider a sequence $(t_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ such that $t_n \to t_0^-$. We consider then the sequence

$$C_Z(t_n, t_n) = \mathbb{E}\left(Z(t_n)^2\right), \quad n \in \mathbb{N}. \quad \text{(C.101)}$$

Since Z is càdlàg, and $t_n \to t_0^-$, $Z(t_n)$ is Cauchy on $L^2(\Omega, \mathcal{A}, \mathbb{P})$, and hence it converges. It follows that

its squared-norm also does it, and hence $C_Z(t_n, t_n)$ has a limit as $n \to \infty$. This proves that the function $t \mapsto C_Z(t,t)$ has left-limits, and hence it is càdlàg. ■

C.2.2 Solving the stochastic Cauchy problem

Let Z be a GeRF over $\mathbb{R}^d \times \mathbb{R}$. We say that Z has a *càdlàg-in-time representation* if there exists a family of spatial GeRFs such that

- for all $\varphi \in \mathscr{S}(\mathbb{R}^d)$, the Random Function $t \mapsto \langle Z_t, \varphi \rangle$ is càdlàg in mean-square,
- for all $\varphi \in \mathscr{S}(\mathbb{R}^d)$ and for all $\theta \in \mathscr{S}(\mathbb{R})$, it holds that

$$\langle Z, \varphi \boxtimes \theta \rangle = \int_{\mathbb{R}} \langle Z_t, \varphi \rangle \theta(t) dt. \tag{C.102}$$

If Z has a càdlàg-in-time representation, then its covariance distribution $C_Z \in \mathscr{S}'((\mathbb{R}^d \times \mathbb{R}) \times (\mathbb{R}^d \times \mathbb{R}))$ has a separately-càdlàg-in-time representation, that is, there exists a family of tempered distributions in $\mathscr{S}'(\mathbb{R}^d \times \mathbb{R}^d)$, $(C_Z^{t,s})_{(t,s)} \in \mathbb{R} \times \mathbb{R}$, such that

- for all $\varphi, \phi \in \mathscr{S}(\mathbb{R}^d)$ and for all $t \in \mathbb{R}$, the function $s \mapsto \langle C_Z^{t,s}, \varphi \otimes \phi \rangle$ is càdlàg,
- for all $\varphi, \phi \in \mathscr{S}(\mathbb{R}^d)$ and for all $s \in \mathbb{R}$, the function $t \mapsto \langle C_Z^{t,s}, \varphi \otimes \phi \rangle$ is càdlàg,
- for all $\varphi, \phi \in \mathscr{S}(\mathbb{R}^d)$, and for all $\theta_1, \theta_2 \in \mathscr{S}(\mathbb{R}^d)$, it holds that

$$\langle C_Z, (\varphi \boxtimes \theta_1) \otimes (\phi \boxtimes \theta_2) \rangle = \int_{\mathbb{R} \times \mathbb{R}} \langle C_Z^{(t,s)}, \varphi \otimes \phi \rangle \theta_1(t) \theta_2(s) d(t,s). \tag{C.103}$$

This claim can be easily proven using Proposition C.2.2.

This definition of a càdlàg-in-time representation can be defined analogously in the specific cases when Z is a slow-growing Random Measure (using $\varphi \in C_{FD}(\mathbb{R}^d)$ and $\theta \in C_{FD}(\mathbb{R})$, the family $(Z_t)_{t \in \mathbb{R}}$ being a family of spatial slow-growing Random Measures), a slow-growing Random Measure over $\mathbb{R}^d \times \mathbb{R}^+$ (that is, $C_Z \in \mathscr{M}_{SG}((\mathbb{R}^d \times \mathbb{R}^+) \times (\mathbb{R}^d \times \mathbb{R}^+))$, using $\varphi \in C_{FD}(\mathbb{R}^d)$ and $\theta \in C_{FD}(\mathbb{R}^+)$, the family $(Z_t)_{t \in \mathbb{R}^+}$ being a family of spatial slow-growing Random Measures), or if Z is such that $\mathscr{F}_S(Z)$ is a slow-growing Random Measure over $\mathbb{R}^d \times \mathbb{R}^+$, that is, it acts analogously as a member of $\left(\mathscr{V}(\mathbb{R}^d) \widehat{\boxtimes} C_{FD}(\mathbb{R}^+)\right)'$ (using $\varphi \in \mathscr{V}(\mathbb{R}^d)$ and $\theta \in C_{FD}(\mathbb{R}^+)$, the family $(Z_t)_{t \in \mathbb{R}^+}$ being a family of spatial GeRF for which the action to test-functions in $\mathscr{V}(\mathbb{R}^d)$ is well defined). For the latter two cases, the mean-square càdlàg condition is only required over $\mathbb{R}^+$. Let us describe the latter case more in detail. Let Z be a real GeRF over $\mathbb{R}^d \times \mathbb{R}$ such that its spatial Fourier Transform is a slow-growing Random Measure concentrated on $\mathbb{R}^d \times \mathbb{R}^+$. Hence, the covariance distribution of $\mathscr{F}_S(Z)$, $C_{\mathscr{F}_S(Z)}$, is in $\mathscr{M}_{SG}\left((\mathbb{R}^d \times \mathbb{R}^+) \times (\mathbb{R}^d \times \mathbb{R}^+)\right)$. Using the density of $\mathscr{S}(\mathbb{R}^d)$ in

$\mathscr{V}(\mathbb{R}^d)$, it is clear that the random variables of the form $\langle Z, \varphi \boxtimes \theta \rangle$ for $\varphi \in \mathscr{V}(\mathbb{R}^d)$ and $\theta \in C_{FD}(\mathbb{R}^+)$ can be defined by a limiting argument in a mean-square sense[4]. It follows that if $\theta \in C_{FD}(\mathbb{R}^+)$, then the application $\varphi \mapsto \langle Z, \varphi \boxtimes \theta \rangle$ defines a GeRF which acts similarly to a member of $\mathscr{V}'(\mathbb{R}^d)$, in the sense that its Fourier Transform is a Slow-growing Random Measure over $\mathbb{R}^d$. Similarly, for every $\varphi \in \mathscr{V}(\mathbb{R}^d)$, the application $\theta \in C_{FD}(\mathbb{R}^+) \mapsto \langle Z, \varphi \boxtimes \theta \rangle$ defines a slow-growing Random Measure over $\mathbb{R}^+$.

The next Proposition is the stochastic analogue of Proposition C.1.3.

Proposition C.2.3. *Let Z be a spatio-temporal GeRF such that its spatial Fourier Transform is a slow-growing Random Measure over $\mathbb{R}^d \times \mathbb{R}^+$. Suppose in addition that $\frac{\partial Z}{\partial t}$ is also such that its spatial Fourier Transform is a slow-growing Random Measure. Then, Z has a càdlàg-in-time representation.*

Proof: Let $U = \mathscr{F}_S(Z)$. U is a Random Measure over $\mathbb{R}^d \times \mathbb{R}^+$, and so does $\frac{\partial U}{\partial t}$. We consider the family of GeRFs over $\mathbb{R}^d$ defined through:

$$\langle U_t, \varphi \rangle := \langle \frac{\partial U}{\partial t}, \varphi \boxtimes \mathbf{1}_{[0,t]} \rangle = \int_{\mathbb{R}^d \times \mathbb{R}^+} \varphi(x) \mathbf{1}_{[0,t]}(s) d\frac{\partial U}{\partial t}(x,s), \quad t \in \mathbb{R}^+, \tag{C.104}$$

for $\varphi \in C_{FD}(\mathbb{R}^d)$. It is immediate that $\varphi \in C_{FD}(\mathbb{R}^d) \mapsto \langle U_t, \varphi \rangle$ defines a spatial slow-growing Random Measure.

If $\varphi \in C_{FD}(\mathbb{R}^d)$ is fixed, the application $A \in \mathcal{B}_B(\mathbb{R}) \mapsto \int_{\mathbb{R}^d \times \mathbb{R}^+} \varphi(x) \mathbf{1}_A(t) d\frac{\partial U}{\partial t}(x,t)$ defines a temporal Random Measure, and hence it is immediate from Proposition C.2.1 that the random function $t \mapsto \langle U_t, \varphi \rangle$ defines a mean-square càdlàg Random Function, for which in addition its covariance function is polynomially bounded since $\frac{\partial U}{\partial t}$ is slow-growing. The stochastic integrals of the form $\int_{\mathbb{R}^+} \langle U_t, \varphi \rangle \theta(t) dt$ are then well-defined for every $\varphi \in C_{FD}(\mathbb{R}^d)$ and $\theta \in C_{FD}(\mathbb{R}^+)$.

Let us now prove that the family (C.104) *represents* U in the sense of Eq. (C.102). As we did in the proof of Proposition C.1.3, we approach the function $\mathbf{1}_{[0,t]}$ by a sequence of positive functions $(\theta_n^{(t)})_{n \in \mathbb{N}} \subset C_c(\mathbb{R}^+)$, continuously differentiable over $\mathbb{R}_*^+$ and such that $\theta_n^{(t)} = 1$ over $[0,t]$ and $\theta_n^{(t)} = 0$ over $\left[t + \frac{1}{n+1}, \infty\right)$, having a decreasing behaviour over the open interval $(t, t + \frac{1}{n+1})$. Then, the functions $(\theta_n^{(t)})_{n \in \mathbb{N}}$ converge point-wise to the function $\mathbf{1}_{[0,t]}$, and they are all bounded by $\mathbf{1}_{[0,t+1]}$. In addition, $(\theta_n^{(t)})_{n \in \mathbb{N}}$ is such that the sequence of derivatives $(\frac{d\theta_n^{(t)}}{dt})_{n \in \mathbb{N}}$ satisfies that $\mathrm{supp}(\frac{d\theta_n^{(t)}}{dt}) \subset \left[t, t + \frac{1}{n+1}\right]$, $\frac{d\theta_n^{(t)}}{dt} \leqslant 0$ and $-\int_{\mathbb{R}^+} \frac{d\theta_n^{(t)}}{dt}(s) ds = 1$ for all $n \in \mathbb{N}$, hence, they *approach* $-\delta_t$ *from the right*. Let us denote by $M^{U,\varphi}$ the temporal Random Measure defined by $M^{U,\varphi}(A) = \langle U, \varphi \boxtimes \mathbf{1}_A \rangle$ for every $A \in \mathcal{B}_B(\mathbb{R})$. It follows from the stochastic versions of the Dominated Convergence Theorem (see Section 3.3.3) and Fubini's Theorem C.2.1

[4]It is a little bit more technical, but also possible to prove that the random variables of the form $\langle Z, \psi \rangle$ for $\psi \in \mathscr{V}(\mathbb{R}^d) \hat{\boxtimes} C_{FD}(\mathbb{R}^+)$ can be defined and hence we can treat Z as a continuous linear functional from $\mathscr{V}(\mathbb{R}^d) \hat{\boxtimes} C_{FD}(\mathbb{R}^+)$ to $L^2(\Omega, \mathcal{A}, \mathbb{P})$. To prove this, it is necessary to prove that the restrictions of functions of the Schwartz space $\mathscr{S}(\mathbb{R}^d \times \mathbb{R})$ to $\mathbb{R}^d \times \mathbb{R}^+$ form a dense subspace of $\mathscr{V}(\mathbb{R}^d) \hat{\boxtimes} C_{FD}(\mathbb{R}^+)$ and of $C_{FD}(\mathbb{R}^d \times \mathbb{R}^+)$. This can be proven following similar arguments as in Proposition C.1.1.

that

$$\begin{aligned}\int_{\mathbb{R}^+}\langle\frac{\partial U}{\partial t},\varphi\boxtimes\mathbf{1}_{[0,t]}\rangle\theta(t)dt &= \lim_{n\to\infty}\int_{\mathbb{R}^+}\langle\frac{\partial U}{\partial t},\varphi\boxtimes\theta_n^{(t)}\rangle\theta(t)dt \qquad \text{(since } \theta_n^{(t)}\to\mathbf{1}_{[0,t]} \text{ point-wise dominated)}\\ &= \lim_{n\to\infty}-\int_{\mathbb{R}^+}\langle U,\varphi\boxtimes\frac{d\theta_n^{(t)}}{dt}\rangle\theta(t)dt\\ &= \lim_{n\to\infty}\int_{\mathbb{R}^+}\int_{\mathbb{R}^+}-\frac{d\theta_n^{(t)}}{dt}(s)dM^{U,\varphi}(s)\theta(t)dt\\ &= \lim_{n\to\infty}\int_{\mathbb{R}^+}\int_{\mathbb{R}^+}-\frac{d\theta_n^{(t)}}{dt}(s)\theta(t)dtdM^{U,\varphi}(s) \quad \text{(Stochastic Fubini's Theorem)}\\ &= \int_{\mathbb{R}^+}\theta(s)dM^{U,\varphi}(s)=\langle U,\varphi\boxtimes\phi\rangle. \qquad \text{(Lemma C.1.1)}\end{aligned} \tag{C.105}$$

and thus the family of GeRFs (C.104) satisfies all the required conditions. The passage to Z is simply obtained by applying a spatial Inverse Fourier Transform to each member of the family U_t, $t\in\mathbb{R}^+$. ■

Now, in order to construct a solution to the stochastic version of the Cauchy problem (C.49), we see that we have all the necessary tools to justify the use of the solution

$$U=\mathscr{F}_S^{-1}\left(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}V\right)=\mathscr{F}_S^{-1}\left(e^{-tg}\left((V_0-\mathcal{D}_g^*(Y)_0)\boxtimes\mathbf{1}_{\mathbb{R}^+}\right)+\mathcal{D}_g^*(Y)\right), \tag{C.106}$$

where $Y=\mathscr{F}_S(X)$ is a slow-growing Random Measure over $\mathbb{R}^d\times\mathbb{R}^+$ and $V_0=\mathscr{F}_S(U_0)$ is a Random Measure over $\mathbb{R}^d$. Indeed, Duhamel's operator $\mathcal{D}_g^*$ can be applied without problem to a slow-growing Random Measure over $\mathbb{R}^d\times\mathbb{R}^+$ since it is defined through an adjoint, the result still being a slow-growing Random Measure over $\mathbb{R}^d\times\mathbb{R}^+$ which has a càdlàg-in-time representation thanks to Proposition C.2.3 . The rest of the operations such as the restrictions to $\mathbb{R}^+$ are well-defined for slow-growing Random Measures. The uniqueness of the solution is guaranteed using the same arguments which prove the uniqueness in the case of Theorem C.1.2. The analogue of Theorem C.1.2 can be then stated.

Theorem C.2.2. *Let X be a GeRF over $\mathbb{R}^d\times\mathbb{R}$ such that its spatial Fourier Transform is a slow-growing Random Measure concentrated on $\mathbb{R}^d\times\mathbb{R}^+$. Let U_0 be a GeRF over $\mathbb{R}^d$ such that its (spatial) Fourier Transform is a slow-growing Random Measure over $\mathbb{R}^d$. Then, there exists a unique-up-to-a-modification GeRF over $\mathbb{R}^d\times\mathbb{R}$, U such that its spatial Fourier Transform is a slow-growing Random Measure and such that*

- *It has a càdlàg-in-time representation whose evaluation at $t=0$ equals U_0 almost surely.*
- *It satisfies* (C.62).

The resolution of the stochastized transformed problem (C.61) is done immediately by using a spatial Fourier Transform.

C.3 Convergence to the stationary solution

Let us now consider the case where X is a stationary GeRF over $\mathbb{R}^d \times \mathbb{R}$ such that its spatial Fourier Transform $Y = \mathscr{F}_S(X)$ is a slow-growing Random measure. An example of such a GeRF is a separable GeRF with any arbitrary spatial stationary structure and a temporal stationary structure which represents a continuous random function or a random measure: $X = X_S \boxtimes X_T$, with X_T representing a stationary temporal stationary random function or, for instance, a White Noise.

In this case, if we suppose in addition that there is $\kappa > 0$ such that $g_R \geqslant \kappa$, we can consider the unique stationary solution to (C.68), which is given by

$$U^{stat} := \mathscr{F}^{-1}\left(\frac{1}{i\omega + g}\mathscr{F}(X)\right). \tag{C.107}$$

We consider hence both GeRFs U^{stat} and the solution over $\mathbb{R}^d \times \mathbb{R}^+$ to the stochastic Cauchy problem U, with U_0 being any spatial real stationary GeRF which we will suppose is independent of X. Consider U to be the solution to the associated Cauchy problem, using the restriction of X to $\mathbb{R}^d \times \mathbb{R}^+$ as source term, which has a càdlàg-in-time representation given by (C.67). The following analogue to Theorem C.1.3 is then obtained.

Theorem C.3.1. *For every $\epsilon > 0$ and for every $\varphi \in \mathscr{S}(\mathbb{R}^d)$, there exists $t_{\epsilon,\varphi} \in \mathbb{R}^+$ such that*

$$\mathbb{E}\left(\left|\langle U - U^{stat}, \varphi \boxtimes \theta\rangle\right|^2\right) < \epsilon, \quad \forall \theta \in \mathscr{S}(\mathbb{R}^d) \text{ such that } \operatorname{supp}(\theta) \subset [t_{\epsilon,\varphi}, \infty) \text{ and } \int_{\mathbb{R}^+} |\theta|(t)dt = 1. \tag{C.108}$$

Proof: The proof is completely analogue to the proof of Theorem C.1.3, considering the fact that both Duhamel's operator and the operator $\mathscr{F}_T^{-1}\left(\frac{1}{i\omega+g}\mathscr{F}_T(Y)\right)$ coincide over $\mathbb{R}^d \times \mathbb{R}^+$ in the sense of Eq. (C.78), and that this also holds in our stochastic case, since every argument for such a claim is applied to the test-functions. It follows that (see Eq. (C.79)),

$$\begin{aligned}\mathbb{E}\left(\left|\langle U^{stat} - U, \varphi \boxtimes \theta\rangle\right|^2\right) &= \mathbb{E}\left(\left|\int_{\mathbb{R}^d}\int_{\mathbb{R}^+} e^{-tg(\xi)}\mathscr{F}_S^{-1}(\varphi)(\xi)\theta(t)dtd(V_0 - \mathcal{D}_g^*(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}Y)_0)(\xi)\right|^2\right)\\ &= \int_{\mathbb{R}^d\times\mathbb{R}^d}\int_{\mathbb{R}^+\times\mathbb{R}^+} e^{-tg(\xi)-s\overline{g(\eta)}}\mathscr{F}_S^{-1}(\varphi)(\xi)\overline{\mathscr{F}_S^{-1}(\varphi)(\eta)}\theta(t)\overline{\theta(s)}d(t,s)dC_{V_0-\mathcal{D}_g^*(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}Y)_0}(\xi,\eta).\end{aligned} \tag{C.109}$$

Since $g_R \geqslant \kappa > 0$ and $C_{V_0-\mathcal{D}_g^*(\mathbf{1}_{\mathbb{R}^d\times\mathbb{R}^+}Y)_0}$ is a slow-growing measure over $\mathbb{R}^d \times \mathbb{R}^d$, we can argue similarly to the case of Theorem C.1.3 to prove the convergence to 0 as the time flows, obtaining and arbitrarily small value for fixed $\varphi \in \mathscr{V}(\mathbb{R}^d)$ and θ such that $\int_{\mathbb{R}^+} |\theta|(t)dt = 1$ with $\operatorname{supp}(\theta)$ being contained in an interval

sufficiently far away from 0. ■

We conclude that as the time flows, the solutions gets closer *spatio-temporally* to the stationary solution U^{stat}.

Appendix D

Some notions on Topological Vector Spaces.

A (complex) Hausdorff locally convex topological vector space E is a vector space endowed with a topology determined by a family of semi-norms indexed by an arbitrary index set I, $(p_i)_{i\in I}$, and satisfying the axiom

$$p_i(x) = 0 \text{ for all } i \in I \;\Rightarrow\; x = 0. \tag{D.1}$$

Precisely, the topology with which E is endowed is the weakest topology in which the addition, the multiplication by scalar and all the semi-norms $(p_i)_{i\in I}$ are continuous.

A family of semi-norms $(p_i)_{i\in I}$ over E is called directed if for all $i, j \in I$ there exists $k \in I$ and $C > 0$ such that

$$p_i(x) + p_j(x) \leqslant C p_k(x), \quad \forall x \in E. \tag{D.2}$$

One can prove that for every Hausdorff locally convex topological vector space endowed with an arbitrary family of semi-norms, we can construct a directed family of semi-norms which is equivalent to the initial one, that is, such that the topologies generated by the two families are the same. If the family of directed semi-norms consists in just one semi-norm (which is then automatically a norm), the space is a normed space, and it is called a Banach space if it is complete. If the family of directed semi-norms consists in a countable family of semi-norms, the space is a metric space, and it is called a Fréchet space if it is complete.

The next Theorem, which can be found in (Reed & Simon, 1980, Theorem V.2) has been widely used in this dissertation:

Theorem D.0.1. *Let E and F two complex Hausdorff locally convex topological vector spaces with families of semi-norms $(p_i)_{i\in I}$ and $(d_j)_{j\in J}$ respectively. Then, a linear map $T : E \to F$ is continuous if and only if for all $j \in J$ there exist $i_1, ..., i_n \in I$ and $C > 0$ such that*

$$d_j\left(\, T(x) \,\right) \leqslant C\left(p_{i_1}(x) + ... + p_{i_n}(x)\right), \quad \forall x \in E. \tag{D.3}$$

In particular, if the family $(p_i)_{i\in I}$ *is directed, then* T *is continuous if and only if for all* $j \in J$ *there exists* $i \in I$ *and* $D > 0$ *such that*

$$d_j\left(\ T\left(x\right)\ \right) \leqslant Dp_i(x), \quad \forall x \in E. \tag{D.4}$$

A particular case of Theorem D.0.1 is when the space F is $\mathbb{C}$. In such a case, $T : E \to \mathbb{C}$ is a continuous linear functional if and only if there exists $C > 0$ and $i_i, ..., i_n \in I$ such that

$$\left|\langle T, x\rangle\right| \leqslant C\left(p_{i_1}(x) + ... + p_{i_n}(x)\right), \quad \forall x \in E, \tag{D.5}$$

or equivalently, if and only if there exists $i \in I$ and $D > 0$ such that

$$\left|\langle T, x\rangle\right| \leqslant Dp_i(x), \quad \forall x \in E, \tag{D.6}$$

if the family $(p_i)_{i\in I}$ is directed.

For the proofs of these claims and for more details about locally convex topological vector spaces, we refer to (Reed & Simon, 1980, Chapter V).

References

Ailliot, P., Baxevani, A., Cuzol, A., Monbet, V., & Raillard, N. (2011). Space-time models for moving fields with an application to significant wave height fields. Environmetrics, 22(3), 354–369.

Angulo, J., Ruiz-Medina, M., & Anh, V. (2000). Estimation and filtering of fractional generalised random fields. Journal of the Australian Mathematical Society, 69(3), 336–361.

Anh, V., Angulo, J., & Ruiz-Medina, M. (1999). Possible long-range dependence in fractional random fields. Journal of Statistical Planning and Inference, 80(1-2), 95–110.

Armstrong, S., Kuusi, T., & Mourrat, J.-C. (2017). Quantitative stochastic homogenization and large-scale regularity. arXiv preprint arXiv:1705.05300.

Barth, A., & Lang, A. (2012). Simulation of stochastic partial differential equations using finite element methods. Stochastics An International Journal of Probability and Stochastic Processes, 84(2-3), 217–231.

Bolin, D., & Kirchner, K. (2017). The SPDE approach for Gaussian random fields with general smoothness. arXiv preprint arXiv:1711.04333.

Bolin, D., & Lindgren, F. (2011). Spatial models generated by nested stochastic partial differential equations, with an application to global ozone mapping. The Annals of Applied Statistics, 5(1), 523–550.

Bourbaki, N. (1965). Intégration. Hermann.

Braess, D. (2007). Finite elements: Theory, fast solvers, and applications in solid mechanics. Cambridge University Press.

Bréhier, C.-E., Hairer, M., & Stuart, A. M. (2016). Weak error estimates for trajectories of SPDEs for Spectral Galerkin discretization. arXiv preprint arXiv:1602.04057.

Cameletti, M., Lindgren, F., Simpson, D., & Rue, H. (2013). Spatio-temporal modeling of particulate matter concentration through the SPDE approach. AStA Advances in Statistical Analysis, 97(2), 109–131.

Canuto, C., Hussaini, M. Y., Quarteroni, A., & Zang, T. A. (2006). Spectral Methods. Springer.

Cartier, P. (1963). Processus aléatoires généralisés. Séminaire Bourbaki, 8, 425–434.

Chilès, J., & Delfiner, P. (1999). Geostatistics: Modeling Spatial Uncertainty. John Wiley & Sons.

Clément, F., & Martin, V. (2016). The Lax-Milgram Theorem. A detailed proof to be formalized in Coq. arXiv preprint arXiv:1607.03618.

Cooley, J. W., & Tukey, J. W. (1965). An algorithm for the machine calculation of complex Fourier series. Mathematics of computation, 19(90), 297–301.

Dai, F., & Xu, Y. (2013). Approximation theory and harmonic analysis on spheres and balls. Springer.

De Iaco, S., Myers, D. E., & Posa, D. (2001). Space-time analysis using a general product-sum model. Statistics & Probability Letters, 52(1), 21–28.

Deitmar, A. (2005). A First Course in Harmonic Analysis (2nd ed.). Springer, New York.

Dellacherie, C., & Meyer, P. (1978). Probabilities and Potential. Hermann.

Demengel, F., & Demengel, G. (2000). Mesures et Distributions. Théorie et illustration par les exemples. Ellipses.

Dierolf, P., & Voigt, J. (1978). Convolution and S'-Convolution of Distributions. Collectanea Mathematica, 29(3), 185–196.

Dong, A. (1990). Estimation géostatistique des phénomènes régis par des équations aux dérivées partielles (Unpublished doctoral dissertation). Paris, ENMP.

Donoghue, W. F. (1969). Distributions and Fourier Transform. Academic Press.

Doob, J. L. (1953). Stochastic processes (Vol. 7) (No. 2). Wiley New York.

Dunford, N., & Schwartz, J. T. (1958). Linear operators part I: general theory (Vol. 7). Interscience publishers New York.

Emery, X., Arroyo, D., & Porcu, E. (2016). An improved spectral turning-bands algorithm for simulating stationary vector Gaussian random fields. Stochastic environmental research and risk assessment, 30(7), 1863–1873.

Fouedjio, F. (2014). Contributions à la modélisation et à l'inférence des fonctions aléatoires non-stationnaires de second ordr (Unpublished doctoral dissertation). Paris, ENMP.

Fuglstad, G.-A., Simpson, D., Lindgren, F., & Rue, H. (2013). Non-stationary spatial modelling with applications to spatial prediction of precipitation. arXiv preprint arXiv:1306.0408.

Gabriel, E. (2014). Estimating second-order characteristics of inhomogeneous spatio-temporal point processes. Methodology and Computing in Applied Probability, 16(2), 411–431.

Gabriel, E., Coville, J., & Chadœuf, J. (2017). Estimating the intensity function of spatial point processes outside the observation window. Spatial Statistics, 22, 225–239.

Gay, R., & Heyde, C. (1990). On a class of random field models which allows long range dependence. Biometrika, 401–403.

Gelfand, I. M. (1955). Generalized random processes. Dokl. Akad. Nauk SSSR(N.S.)(100), 853-856.

Gelfand, I. M., & Vilenkin, N. I. (1964). Generalized functions, Vol. 4: applications of harmonic analysis. Academic Press.

Gneiting, T. (2002). Nonseparable, stationary covariance functions for space–time data. Journal of the American Statistical Association, 97(458), 590–600.

Gneiting, T., Genton, M. G., & Guttorp, P. (2006). Geostatistical space-time models, stationarity, separability, and full symmetry. Monographs On Statistics and Applied Probability, 107, 151.

Gottlieb, D., & Orszag, S. A. (1977). Numerical analysis of spectral methods: theory and applications (Vol. 26). Siam.

Gradshteyn, I. S., & Ryzhik, I. M. (2014). Table of integrals, series, and products. Academic press.

Grothendieck, A. (1955). Produits tensoriels topologiques et espaces nucléaires (Vol. 16). American Mathematical Soc.

Habil, E. D. (2016). Double sequences and double series. IUG Journal of Natural Studies, 14(1).

Hairer, M. (2013). Solving the KPZ equation. Annals of Mathematics, 559–664.

Hairer, M. (2014). A theory of regularity structures. Inventiones mathematicae, 198(2), 269–504.

Heine, V. (1955). Models for two-dimensional stationary stochastic processes. Biometrika, 42(1-2), 170–178.

Holden, H., Øksendal, B., Ubøe, J., & Zhang, T. (2009). Stochastic Partial Differential Equations: a modeling, White Noise functional approach. Springer Science & Business Media.

Horowitz, J. (1977). Une remarque sur les bimesures. Séminaire de probabilités de Strasbourg, 11, 59–64.

Horowitz, J. (1986). Gaussian random measures. Stochastic processes and their applications, 22(1), 129–133.

Hristopulos, D. T., & Tsantili, I. C. (2016). Space-time models based on random fields with local interactions. International Journal of Modern Physics B, 30(15), 1541007.

Huang, J., Malone, B. P., Minasny, B., McBratney, A. B., & Triantafilis, J. (2017). Evaluating a Bayesian modelling approach (INLA-SPDE) for environmental mapping. Science of the Total Environment, 609, 621–632.

Itô, K. (1954). Stationary random distributions. Memoirs of the College of Science, University of Kyoto. Series A: Mathematics, 28(3), 209–223.

Jones, R. H. (1989). Fitting a stochastic partial differential equation to aquifer data. Stochastic Hydrology and Hydraulics, 3(2), 85–96.

Jones, R. H., & Zhang, Y. (1997). Models for continuous stationary space-time processes. In Modelling longitudinal and spatially correlated data (pp. 289–298). Springer.

Kærgaard, E. B. (2013). Spectral methods for uncertainty quantification (Unpublished master's thesis). Master's thesis, Technical University of Denmark.

Kallenberg, O. (2017). Random measures, theory and applications. Springer.

Kardar, M., Parisi, G., & Zhang, Y.-C. (1986). Dynamic scaling of growing interfaces. Physical Review Letters, 56(9), 889.

Kelbert, M. Y., Leonenko, N. N., & Ruiz-Medina, M. (2005). Fractional random fields associated with stochastic fractional heat equations. Advances in Applied Probability, 37(1), 108–133.

Khintchine, A. (1934). Korrelationstheorie der stationären stochastischen Prozesse. Mathematische Annalen, 109(1), 604–615.

Kingman, J. (1967). Completely random measures. Pacific Journal of Mathematics, 21(1), 59–78.

Knapp, A. W. (2005). Basic real analysis: along with companion volume Advanced real analysis. Birkhåuser.

Kolmogorov, A. N. (1956). Foundations of the Theory of Probability: Second English Edition (N. Morrison, Trans.). Chelsea Publishing Company.

Kolmogorov, A. N., & Prokhorov, Y. V. (1992). Random Functions and Limit Theorems. In A. Shiryayev (Ed.), Selected works of A.N. Kolmogorov: Vol. 2, Probability theory and mathematical statistics (chap. 44). Springer Science and Business Media Dordrecht.

Kupka, J. (1978). The Carathéodory extension theorem for vector valued measures. Proceedings of the American Mathematical Society, 72(1), 57–61.

Lang, A. (2007). Simulation of stochastic partial differential equations and stochastic active contours (Unpublished doctoral dissertation). Universität Mannheim.

Lang, A., & Potthoff, J. (2011). Fast simulation of Gaussian random fields. Monte Carlo Methods and Applications, 17(3), 195–214.

Lang, A., Schwab, C., et al. (2015). Isotropic Gaussian random fields on the sphere: Regularity, fast simulation and stochastic partial differential equations. The Annals of Applied Probability, 25(6), 3047–3094.

Lantuéjoul, C. (2013). Geostatistical Simulation: Models and Algorithms. Springer Science & Business Media.

Lim, S., & Teo, L. (2009). Generalized Whittle–Matérn random field as a model of correlated fluctuations. Journal of Physics A: Mathematical and Theoretical, 42(10), 105202.

Lindgren, F., Rue, H., & Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 73(4), 423–498.

Liu, X., Guillas, S., & Lai, M.-J. (2016). Efficient spatial modeling using the SPDE approach with bivariate splines. Journal of Computational and Graphical Statistics, 25(4), 1176–1194.

Loève, M. (1978). Probability theory, Vol. II. Graduate texts in mathematics, 46, 0–387.

Ma, C. (2009). Construction of non-Gaussian random fields with any given correlation structure. Journal of Statistical Planning and Inference, 139(3), 780–787.

Mainardi, F., Luchko, Y., & Pagnini, G. (2007). The fundamental solution of the space-time fractional diffusion equation. arXiv preprint cond-mat/0702419.

Matheron, G. (1965). Les variables régionalisées et leur estimation: une application de la théorie des fonctions aléatoires aux
Masson et CIE.

Mena, H., & Pfurtscheller, L. (2017). An efficient SPDE approach for El Ni\˜ no. arXiv preprint arXiv:1708.04144.

Minasny, B., & McBratney, A. B. (2005). The Matérn function as a general model for soil variograms. Geoderma, 128(3-4), 192–207.

Morando, P. (1969). Mesures aléatoires. Séminaire de probabilités de Strasbourg, 3, 190–229.

Morse, M. (1955). Bimeasures and their integral extensions. Annali di Matematica Pura ed Applicata, 39(1), 345–356.

Neveu, J. (1970). Bases mathématiques du calcul des probabilités (2nd ed.). Masson et Cie.

Øksendal, B. (2003). Stochastic differential equations: an introduction with applications (5th ed.). Springer-Verlag.

Pardo-Iguzquiza, E., & Chica-Olmo, M. (1993). The Fourier integral method: an efficient spectral method for simulation of random fields. Mathematical Geology, 25(2), 177–217.

Pereira, M., & Desassis, N. (2018). Efficient simulation of Gaussian Markov random fields by Chebyshev polynomial approximation. arXiv preprint arXiv:1805.07423.

Porcu, E., Bevilacqua, M., & Genton, M. G. (2016). Spatio-temporal covariance and cross-covariance functions of the great circle distance on a sphere. Journal of the American Statistical Association, 111(514), 888–898.

Rao, M. (2012). Random and Vector Measures. World Scientific Publishing.

Reed, M., & Simon, B. (1980). Methods of modern mathematical analysis: Functional Analysis. Academic Press.

Richards, I., & Youn, H. K. (1995). The theory of distributions: a nontechnical introduction. Cambridge University Press.

Richards, I., & Youn, H.-K. K. (2000). Localization and multiplication of distributions. Journal of the Korean Mathematical Society, 37(3), 371–389.

Rogers, C. (1970). Hausdorff Measures. Cambridge University Press.

Rogers, L., & Williams, D. (2000). Diffusions, Markov Processes and Martingales. Volume 1: Foundations (2nd ed.). Cambridge University Press.

Rozanov, J. A. (1977). Markov random fields and Stochastic Partial Differential Equations. Mathematics of the USSR-Sbornik, 32(4), 515.

Rozanov, Y. A. (1982). Markov Random Fields. Springer.

Rudin, W. (1987). Real and complex Analysis (3rd ed.). McGraw-Hill Book Company.

Ruiz-Medina, M., Angulo, J., & Anh, V. (2003). Fractional generalized random fields on bounded domains.

Ruiz-Medina, M. D., Angulo, J. M., Christakos, G., & Fernández-Pascual, R. (2016). New compactly supported spatiotemporal covariance functions from SPDEs. Statistical Methods & Applications, 25(1), 125–141.

Schwartz, L. (1954). Sur l'impossibilité de la multiplication des distributions. Comptes Rendus Hebdomadaires des Séances de L'Academie des Sciences, 239(15), 847–848. http://sites.mathdoc.fr/OCLS/pdf/OCLS_1954__21__1_0.pdf.

Schwartz, L. (1966). Théorie des Distributions. Hermann.

Shiraishi, R., & Itano, M. (1964). On the Multiplicative Products of Distributions. Journal of Science of the Hiroshima University, Series AI (Mathematics), 28(2), 223–235.

Sigrist, F., Künsch, H. R., & Stahel, W. A. (2015). Stochastic partial differential equation based modelling of large space-time data sets. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 77(1), 3–33.

Simpson, D., Lindgren, F., & Rue, H. (2012). In order to make spatial statistics computationally feasible, we need to forget about the covariance function. Environmetrics, 23(1), 65–74.

Sobczyk, K. (1991). Stochastic Differential Equations: With Applications to Physics and Engineering. Springer Science & Business Media.

Stefanou, G. (2009). The stochastic finite element method: past, present and future. Computer Methods in Applied Mechanics and Engineering, 198(9-12), 1031–1051.

Stein, M. L. (1999). Interpolation of spatial data: some theory for kriging. Springer Science & Business Media.

Stein, M. L. (2005). Space–time covariance functions. Journal of the American Statistical Association, 100(469), 310–321.

Swanson, J. (2011, July). Elementary properties of functions with one-sided limits. http://math.swansonsite.com/instructional/cadlag.pdf.

Takahashi, M. (1966). On topological-additive-group-valued measures. Proceedings of the Japan Academy, 42(4), 330–334.

Teukolsky, S. A., Flannery, B. P., Press, W., & Vetterling, W. (1992). Numerical recipes in C. SMR, 693(1).

Trèves, F. (1967). Topological Vector Spaces, Distributions and Kernels. Academic Press.

Vecchia, A. (1985). A general class of models for stationary two-dimensional random processes. Biometrika, 72(2), 281–291.

Wackernagel, H. (2003). Multivariate Geostatistics: an introduction with applications (3rd ed.). Springer.

White, P., & Porcu, E. (2018). Towards a Complete Picture of Covariance Functions on Spheres Cross Time. arXiv preprint arXiv:1807.04272.

Whittle, P. (1954). On stationary processes in the plane. Biometrika, 434–449.

Whittle, P. (1963). Stochastic processes in several dimensions. Bulletin of the International Statistical Institute, 40(2), 974–994.

Williams, D. (1990). Probability with Martingales. Cambridge University Press.

Yaglom, A. M. (1987). Correlation theory of stationary and related random functions I: Basic Results. Springer-Verlag, New York.

Zhang, T., & Zhang, H. (n.d.). Non-Fully Symmetric Space-Time Matérn Covariance Functions.

Zienkiewicz, O. C., Taylor, R. L., & Zhu, J. Z. (2013). The Finite Element Method: Its Basis and Fundamentals: Its Basis and Fundamentals.

RÉSUMÉ

Ces travaux présentent des avancées théoriques pour l'application de l'approche EDPS (Équation aux Dérivées Partielles Stochastique) en Géostatistique. On considère dans cette approche récente que les données régionalisées proviennent de la réalisation d'un Champ Aléatoire satisfaisant une EDPS. Dans le cadre théorique des Champs Aléatoires Généralisés, l'influence d'une EDPS linéaire sur la structure de covariance de ses éventuelles solutions a été étudiée avec une grande généralité. Un critère d'existence et d'unicité des solutions stationnaires pour une classe assez large d'EDPSs linéaires a été obtenu, ainsi que des expressions pour les mesures spectrales associées. Ces résultats permettent de développer des modèles spatio-temporels présentant des propriétés non-triviales grâce à l'analyse d'équations d'évolution présentant un ordre de dérivation temporel fractionnaire. Des paramétrisations adaptées de ces modèles permettent de contrôler leur séparabilité et leur symétrie ainsi que leur régularité spatiale et temporelle séparément. Des résultats concernant des solutions stationnaires pour des EDPSs issues de la physique telles que l'équation de la Chaleur et l'équation d'Onde sont présentés. Puis, une méthode de simulation non-conditionnelle adaptée à ces modèles est étudiée. Cette méthode est basée sur le calcul d'une approximation de la Transformée de Fourier du champ, et elle peut être implémentée de façon efficace grâce à la Transformée de Fourier Rapide. La convergence de cette méthode a été montrée théoriquement dans un sens faible et dans un sens fort. Cette méthode est appliquée à la résolution numérique des EDPSs présentées dans ces travaux. Des illustrations de modèles présentant des propriétés non-triviales et reliés à des équations de la physique sont alors présentées.

MOTS CLÉS

Modèles géostatistiques, Champs aléatoires généralisés, Équations aux Dérivées Partielles Stochastiques, Approche EDPS, Géostatistique spatio-temporelle, Simulation.

ABSTRACT

This dissertation presents theoretical advances in the application of the Stochastic Partial Differential Equation (SPDE) approach in Geostatistics. This recently developed approach consists in interpreting a regionalised data-set as a realisation of a Random Field satisfying a SPDE. Within the theoretical framework of Generalized Random Fields, the influence of a linear SPDE over the covariance structure of its potential solutions can be studied with a great generality. A criterion of existence and uniqueness of stationary solutions for a wide-class of linear SPDEs has been obtained, together with an expression for the related spectral measures. These results allow to develop spatio-temporal covariance models presenting non-trivial properties through the analysis of evolution equations presenting a fractional temporal derivative order. Suitable parametrizations of such models allow to control their separability, symmetry and separated space-time regularities. Results concerning stationary solutions for physically inspired SPDEs such as the Heat equation and the Wave equation are also presented. A method of non-conditional simulation adapted to these models is then studied. This method is based on the computation of an approximation of the Fourier Transform of the field, and it can be implemented efficiently thanks to the Fast Fourier Transform algorithm. The convergence of this method has been theoretically proven in suitable weak and strong senses. This method is applied to numerically solve the SPDEs studied in this work. Illustrations of models presenting non-trivial properties and related to physically driven equations are then given.

KEYWORDS

Geostatistical models, Generalized random fields, Stochastic Partial Differential Equations, SPDE Approach, Space-time Geostatistics, Simulation.