Self-Similar Vector Fields

Pouya Dehghani Tafti

Thèse N° 5167 (Novembre 2011)

Thèse présentée à la faculté des sciences et techniques de l'ingénieur
pour l'obtention du grade de docteur ès sciences
et acceptée sur proposition du jury

Prof. Pierre Vandergheynst, *président*
Prof. Michael Unser, *directeur de thèse*
Dr. Patrice Abry, *rapporteur*
Prof. Robert Dalang, *rapporteur*
Dr. Jacques Lévy-Véhel, *rapporteur*

École polytechnique fédérale de Lausanne—2011

Cover design by Annette Unser
Printing and binding by Repro-EPFL
Typeset in Concrete Roman using LuaLaTeX and the memoir package

Available at **http://bigwww.epfl.ch/**

علم رسمی سر به سر قیل است و قال

نه از او کیفیتی حاصل نه حال

علم نبود غیر علم عاشقی

باقی تلبیس ابلیس شقی

هر که را در سر نباشد عشق یار

بهر او پالان و افساری بیار

شیخ بهایی

And each man hears as the twilight nears,
to the beat of his dying heart,
The Devil drum on the darkened pane:
'You did it, but was it Art?'

Rudyard Kipling
The Conundrum of the Workshops

To my Parents

Abstract

We propose statistically self-similar and rotation-invariant models for vector fields, study some of the more significant properties of these models, and suggest algorithms and methods for reconstructing vector fields from numerical observations, using the same notions of self-similarity and invariance that give rise to our stochastic models. We illustrate the efficacy of the proposed schemes by applying them to the problems of denoising synthetic flow phantoms and enhancing flow-sensitive magnetic resonance imaging (MRI) of blood flow in the aorta.

In constructing our models and devising our applied schemes and algorithms, we rely on two fundamental notions. The first of these, referred to as 'innovation modelling' in the thesis, is the principle—applicable both analytically and synthetically—of reducing complex phenomena to combinations of simple independent components or 'innovations'. The second fundamental idea is that of 'invariance', which indicates that in the absence of any distinguishing factor, two equally valid models or solutions should be given equal consideration.

Keywords: invariance, innovation modelling, fractional Brownian vector fields, fractional Brownian motion, fractional stable motion, generalized random fields, variational methods, vector fields

Résumé

Nous proposons des modèles stochastiques autosimilaires et invariant par rotation pour les champs vectoriels, étudions certaines de leurs propriétés les plus importantes, et suggérons des algorithmes et des méthodes pour la reconstruction de champs vectoriels à partir de mesures généralisées. Pour developper ces méthodes, nous utilisons les mêmes notions d'autosimilarité et d'invariance par rotation qui caractérisent nos modèles stochastiques. Nous illustrons les méthodes proposés dans un premier temps en les appliquant au débruitage de champs vectoriels simulés. Nous traitons ensuite des données de flux sanguin dans l'aorte obtenues par IRM (imagerie par résonance magnétique).

La base conceptuelle de cette thèse se resume en deux notions fondamentales. La première, que nous avons nommée « modélisation d'innovations », se fonde sur l'idée de décrire un phénomène complexe par la composition d' « innovations » simples et indépendantes. La deuxième est liée au principe d' « invariance », qui nous amène à établir des relations d'équivalence entre observations possibles d'un phénomène vis à vis de critères naturels définis à apriori. Dans la recherche d'une solution, l'on accorde alors la même priorité aux candidats appartenant à la même classe d'équivalence.

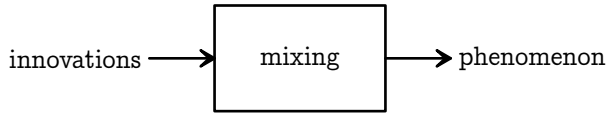
Mots-clés : invariance, modèles d'innovation, champs vectoriels brownien fractionnaire, mouvement brownien fractionnaire, mouvement stable fractionnaire, champs stochastiques généralisés, méthodes variationnelles, champs vectoriels

Preface

This thesis is concerned with developing models and methods of solution for problems involving *vector fields.* Vector fields are mathematical idealizations of physical phenomena that vary inside a spatio-temporal domain (the 'field' bit), and which, in some sense, specify a direction at or about each point inside the domain (the 'vector' bit). An example is the velocity field inside a water pipe (the domain).

Philosophically, this thesis rests on two pillars. One is the idea of 'innovation modelling', and the other is the notion of 'invariance'.

Broadly speaking, *innovation is that which cannot be predicted.* The view put forward by innovation modelling is that the phenomenon being modelled consists of a number of independent innovations that are combined together through a process of 'mixing' in order to create the final object of interest. We can represent this idea schematically as follows:



Going in the other direction, as an analytical principle, innovation modelling suggests that to analyse a complex phenomenon, one might try to identify (or hypothesize) the independent sources of innovation hiding underneath, and the process through which they are compounded. In the role of a model of creation, this view can be applied at different levels, from the representation of natural phenomena such as physical systems of many particles, to the study of human creativity and the works of art.

In this thesis we shall be concerned with a very narrow and specific interpretation and application of this general idea, where we assume, first, that all of the innovation sources are *probabilistically comparable,* meaning roughly that if we were to look at a large number of innovations originating from each

of these sources, they would all appear similar to us. Secondly, we assume that the process of mixing can be modelled by a special kind of mathematical object, called a *continuous linear operator*. 'Operator' is jargon for anything that takes an object and returns something known in return (the box in the above schematic). 'Continuity' of the operator means that if the input to the box is only slightly modified, its output will also change only slightly. Its 'linearity' signifies that its output corresponding to the superposition of some collection of inputs is the superposition of individual outputs associated with each of the inputs. The mathematical framework is presented in Chapter 1.

The second philosophical underpinning of this thesis, which we have codified by the term 'invariance', is a form of the *principle of insufficient reason,* which in our application roughly states that in the absence of any factor based on which to prefer one model or solution for a problem to another, both should be given equal consideration. This quality is expressed by saying that the models or methods of solution are *invariant* under (or insensitive to) transformations that exchange objects uniquely for other equally valid ones. Once again, our application of the general philosophical principle is very specific: we concentrate on particular types of geometric invariances, namely with respect to rotations and changes of scale as they apply to vector fields. Mathematically, this is done by identifying homogeneous and rotation-invariant operators and potential functions. We introduce a family of such operators in Chapter 2, and suggest some suitable potential functions indirectly in Chapters 1 and 4 (although not always by this name).

Chapters 3 and 4 consist of applications of the above principles as developed in the first two chapters. In Chapter 3, we transplant the invariant operators of Chapter 2 in the innovation modelling framework of Chapter 1 in order to define self-similar and rotationally invariant random models for vector fields. In the following chapter, we rely on similar ideas, but this time in a heuristic fashion, in order to devise methods of solution for the concrete problem of refining measurements of a physical vector field.

We bring our unannounced overview of the thesis to an end by informing the reader that Chapter 4 is followed by conclusions and two appendices, the first concerned with some background material and musings about probability theory, and the second consisting of older publications that were not fully incorporated in the text. A more detailed overview of the contents and results is given after this preface. Apart from the above brief sketch and the noted obligatory overview of contributions that will shortly follow, there is no separate introduction. Instead, each chapter begins with introductory re-

marks that are intended to clarify its purpose and connection with the rest of the thesis. The order in which topics are presented follows loosely the principle of introducing mathematical tools and methods before they are applied (references are given when this order is violated). Since the justification for a definition often lies in its application, the purpose of some sections may only become apparent later (this is particularly true of Chapter 2). Other than that, the unbalanced presentation, which puts undue emphasis on trivialities at times while passing lightly over some of the more technically challenging concepts on other occasions, reflects the author's variable mood and the wax and wane of his enthusiasm.

Each chapter is divided into a number of sections and subsections corresponding to the main subjects considered therein, and independently into a relatively large number of 'paragraphs', each of which develops a single idea or two (regarding numbering: subsections of the third section of the second chapter are numbered as **§2.3.1**, **§2.3.2**, *&c.*; paragraphs in the same chapter are numbered independetly as **2.a**, **2.b**, *&c.*, with **2.z** then followed by **2.aa**, **2.ab**, and so on). Particularly fanciful digressions and generalizations that may be skipped without compromise are marked by an asterisk, as in ***2.j**.

***

My debt of gratitude to those from whose support, kindness, and generosity I have benefited is far too heavy, my space here too restricted, my memory too weak, and my literary ability too limited, to thank them all in the way they deserve. I hope those that I do not mention, or mention only briefly where even elaborate thanks would not have sufficed, will look forgivingly at my feeble attempt.

I thank all past and present members of the Biomedical Imaging Group of the EPFL for their delightful company, and the friendly atmosphere they have all helped create in the lab. I am particularly grateful to my adviser, Professor Michael Unser, for his support, patience, and generosity, and for the excellent conditions in which I was able to work. I thank our secretary, Manuelle, for her kind presence and help, and the organizational skill with which she has kept the life of an absent-minded and bureaucratically-challenged PhD student on track (and also for the delicious homemade sweets with which she regularly spoils us). I am grateful to senior members of our lab, Daniel and Philippe, for their willingness to share their expertise and their readiness to offer their help in all sorts of matters, from the technicalities of computing to the practicalities of life in Switzerland, as well as for their thoroughly enjoyable company throughout these years. For the same reasons, I also thank former senior members of the lab, Dimitri and Thierry. My former and cur-

rent officemates, Hagai, Chandra, and Kai have my thanks for their friendship and for the patience with which they have overlooked my disorganized expansion and ever-increasing entropy. I equally thank past and present post-docs, visitors, and fellow students, Ricard (whose unfailing optimism and friendship I have often relied upon to the point of abuse), Matthieu, Jean-Charles, Raquel, Kunal, François, Stamatis, Simona, Nicolas, Arash, Masih, Cédric, Ildar, Aurélien, Djano, Katarina, Florian, Ayush, Sathish, Emrah, Ulugbek, Chiara, Patrick, Julien, Qiyu, Torsten, İlker, Nicolas, Olivier, and John Paul. I am grateful to Jean-Charles and Julien also for their help in bringing my original pseudo-French translation of the abstract somewhat closer to actual French, and to John Paul for pointing out a few typos in the appendix.

I thank Professors Dalang, Abry, Lévy Véhel, and Vandergheynst for accepting to be part of my thesis committee, and for the helpful feedback and encouragement they provided. I also thank Professor A. Chorin who had kindly agreed to take part in the jury, but from whose participation I could not benefit due to scheduling difficulties on my side.

The financial support for my work came from the Swiss National Science Foundation (under grants number 3252B0-108486, 200020-109415, and 200020-121763), the European Research Council (under grant number 267439 FUN-SP), and the STI faculty. I thank them (and, ultimately, the tax-payers who have indirectly funded my studies) for their generous support. The MRI data and pathline visualizations of aortic blood flow used to demonstrate our algorithm in Chapter 4 were produced by Dr A. F. Stalder, who has kindly permitted me to reuse them here. The papers that appear in Appendix B are reprinted with the permission of their respective publishers.

Outside the circle of my academic collaborators, I am deeply grateful to Mojgan and Farhad, Sousan and Massoud, Keyvan, Alex, and Mahshid and Arash, for their help, friendship, and support. I also thank my other friends at the EPFL (whom I shall not list for the fear of omitting some of them), and Dr W. Froestl whose warm company I had the privilege of sharing from time to time on the bus trip home. My long-time friends Mohammad, Foad, Ehsan, Maryam, and Amir, have my thanks for their lasting friendship.

I wish to close this preface by thanking my family members: my sister Leila, my grand parents, and, above all, my parents, Mina and Javad, to whom I owe, directly or indirectly, everything I have and can hope ever to achieve in this life. Although any amount of thanks would be embarrassingly insufficient in the face of their innumerable sacrifices, invaluable advice, and unconditional love and support, I do thank them here from the depths of my soul. This thesis is humbly dedicated to them.

Technical Overview

After the overview of ideas and format given in the beginning of the preface, I now give an overview of the main results and contributions. In doing so, I also take opportunities, as they present themselves, to re-iterate some of my main intellectual debts and influences, wrapping them now and then in pieces of a quasi-narrative of how this thesis came to be.

With the excuse of its more personal nature, the following overview—like the acknowledgements that precede it—is written mostly in the first person singular. For the remainder of the thesis, I shall revert to the collective 'we'.

***

This thesis is the outgrowth of my work as a Ph.D. student on mathematical models with specific geometric invariances. Our work in this area resulted in a number of journal publications [TVU09, TU10b, UT11] and papers delivered at various conferences [TVU08, TU09, TDGSU10, TU10a, TDGSU11]. To evade the tedium of repetition and exploit the benefit of hindsight, in the present thesis I have often departed, sometimes widely, from our published work, revisiting and reworking the subject of study and our methodology with the intention of reducing them to their essentials and presenting them as coherently and completely as I could, while filling any gaps as I had noticed in our original presentation, and generalizing the methodology and the results where there seemed to be some value in doing so.

On the other hand, I have made no attempt to produce an exhaustive account of our publications, preferring instead to leave out some material for the sake of coherence and conciseness. To counter this omission, some of the mentioned papers have been included as appendices, but the main text does not depend on them. Other than the noted appendices, a few philosophical passages in Appendix A that are quoted directly from Kolmogorov [Kol56] and specifically marked as such, parts of the proof of **2.aw** that also appears in one of our published papers, no part of the text has been copied verbatim from elsewhere, including from our previous publications. In addition, except

where specifically noted, the statements of the results as given in the main text and their proofs are mine.

The first chapter of the thesis introduces what is called here the 'innovation modelling' framework for defining stochastic models. I was introduced to the basic idea of innovation modelling (in its strict interpretation as the application of a linear shift-invariant operator to a 'white noise' process) and the theory of generalized random processes soon after my arrival at the EPFL in 2006, through the work of (and proximity with) my adviser, Professor Michael Unser, and Professor Thierry Blu who collaborated with our group at the time. Without formal training as a mathematician, in the years that followed I tried to educate myself about the subject and its subtleties and technicalities, first through the volume by Gel'fand and Vilenkin [GV64], and later by referring to the works of Schwartz [Sch73, Sch81], among others. I do not pretend to have mastered this theory, but I believe that—at least from the perspective of someone mainly interested in using this formalism to construct stochastic models—a more or less systematic methodology has eventually emerged through these contemplations, which I have tried to distil and present in the first chapter. Some (less directly relevant and/or more philosophically oriented) background material has found a place in Appendix A. The novelty of this chapter, whatever its extent, is to be sought in the presentation of a general framework rather than basic matematical definitions and results, which can all be found in the references cited therein.

Chapter 2, in its quasi-totality, is concerned with identifying linear operators with prescribed invariance and continuity properties, and the derivation of their other properties used in the following chapters. After a somewhat non-standard and general overview of the notion of invariance in §2.1, attention is directed to specific geometric invariances for operators acting on scalars and vectors (the main distinction between the two being in the different ways in which they transform under rotations of the coordinate system). A study of the scalar theory appears in §2.2, followed by its (not always comparable or similar) vector counterpart in the next two sections.

The first subsection of the second section of this chapter, §2.2.1, where the general theory of isotropic homogeneous generalized functions in $\mathbb{R}^d$ and their associated convolution operators is considered, follows closely and draws heavily from Gel'fand and Shilov [GS64] in terms of technique, results, and arguments. I have, nevertheless, adapted these results and arguments to our purpose, and derived some new formulae and results, which are used later to define matrix-valued self-similar distributions. The inclusion of this subsection was necessary also in order to have an umambiguous interpretation

of the singular integrals that abound in that chapter.

Next, in **§2.2.2**, attention is turned to 'L_p-continuous' modifications of the scalar convolution operators introduced previously. The one-dimensional L_2-continuous subset of these operators was considered by Blu and Unser [BU07] without explicit mention of their continuity property. I later generalized this definition to several dimensions ([TVU09, TU10a]), in the first case without direct reference to topological considerations and later only commenting on it in passing (by this time, I had come to appreciate the central importance of continuity in a certain topology in the theory of characteristic functionals). Professor Qiyu Sun, who visited our lab a little later, extended these definitions to the L_p setting, $1 \leq p < \infty$, emphasizing continuity throughout [SU]. His definitions and results, as well as all of the previously cited ones, are formulated in the Fourier domain. While I benefited from discussions with Professor Sun during his visit, the definition of L_p-continuous operators given in **§2.2.2**, which is stated in the spatial domain, and the proof of their continuity, are not directly based on his paper (this is partly due to the fact that I found it more convenient to define and derive them in the way they appear here, but also because for $p \neq 2$, Fourier-domain definitions are, *a priori,* somewhat ambiguous and require careful interpretation).

The previous scalar theory is developed here primarily in order that its results might be used to construct and study homogeneous and rotation-invariant operators acting on vector fields. This is done, first, in **§2.3**, where the vector operators and their associated matrix-valued distributions are introduced and their basic properties derived and, later, in **§2.4**, where their vector calculus is developed.

Our definition of homogeneous and vectorially rotation-invariant operators rests on the introduction of a family of matrix-valued (tensor) distributions belonging to $(\mathscr{D}')^{d\times d}$ and $(\mathscr{S}')^{d\times d}$, which can be seen as the natural tensor generalizations of scalar homogeneous isotropic distributions, and this in more than one sense. Specifically, in addition to being homogeneous and rotation-invariant (the latter in a different sense than for scalars), the family, which is parametrized by the (complex) homogeneity order and two additional coefficients, is closed under Fourier transforms. It also has very interesting properties regarding the Helmholtz docomposition of its elements into curl- and divergence-free parts. I investigate these properties in the last two sections of Chapter 2.

One of the definitions of the noted matrix distributions (the first line of (2.34)) is inspired by the work of Arigovindan [Ari05]. During his doctoral studies in the Biomedical Imaging Laboratory of EPFL, which preceded my

affiliation with the same group, Dr Muthuvel Arigovindan worked on scale- and rotation-invariant reconstruction of vector fields and identified the general form of quadratic potential functionals with the noted invariances in two and three dimensions, obtaining Fourier-domain formulae similar to (2.34) for their associated operators. I have extracted what I thought was fundamental in his definition to introduce the said family of matrix-valued homogeneous and rotation-invariant distributions in §2.3, giving new proofs of their invariance properties in any number of dimensions, as well as finding different expressions for them (they come in a number of essentially equivalent forms and parametrizations, each suitable for some purpose) and deriving many of their properties in relationship to each other. As already noted, these distributions are in essence the matrix generalizations of scalar homogeneous isotropic distributions and retain many of their nice properties, while adding others with no scalar equivalent.

The noted matrix distributions give definition to convolution operators that map vector fields to vector fields. This development is then followed by the identification of L_p-continuous rotation-invariant and homogeneous vector operators in §2.3.2, along similar lines as in the scalar case. Finally, in §2.4, I present the vector calculus of these operators, which brings to the forefront their directional characteristics. With the exception of parts of the proof of Theorem **2.aw** and some of the review material in §2.4.1, these sections overlap very little, both in terms of presentation and results, with our previous published work.

Chapter 3, where self-similar random vector field models are introduced, begins with a few standard observations that motivate and suggest a number of ways to define and characterize scalar fractional Brownian motion (fBm) and its most trivial vector generalization (a vector composed of independent fBm fields for each coordinate). The same scalar fields are then characterized in the innovation modelling framework of Chapter 1, to prepare the scene for introducing their non-trivial but nevertheless quite natural extension as random vector fields that are probabilistically self-similar and rotation-invariant in the vector sense.

The definition of these models—which I have called, interchangeably, 'fractional Brownian vector fields' or 'vector fBm'—in the innovation modelling framework utilizes the L_2-continuous vector invariant operators of Chapter 2. This makes it possible to use the results of Chapter 2 quite easily to obtain the main properties of these new models. Finally, in the last section of the chapter, I use the general L_p-continuous vector operators (with p typically different from 2), and superpositions of innovation models to define more gen-

eral self-similar random vector field models with α-stable statistics, to which I refer as 'fractional stable vector fields' and their 'subspace-independent' extensions. The significance of these models is that they are in a sense the largest family of vector models with the required invariances; they also permit us directly to account for the directional and scaling properties of the phenomenon we wish to model, such as its curl and divergence.

Among the models presented in Chapter 3, the Gaussian ones were discussed in our 2010 publication [TU10b], but their non-Gaussian and subspace-independent extensions have not been published previously.

Having disqualified myself as a mathematician early on in this overview, it therefore seemed necessary to complement the previous theoretical developments with at least a semblance of the mixture of inspiration and science and guesswork otherwise known as 'engineering' (in retrospect, what was sacrificed in mathematical formalism seems to have been partially counterbalanced by successful applications, which generated some level of enthusiasm at a certain IEEE symposium [TDGSU11]).

For the above reason, in the final chapter of the thesis I put forward a relatively general and systematic method of solution for practical problems involving the reconstruction and refinement/enhancement of numerical representations of vector fields. This chapter diverges from the preceding ones in being more application-oriented, and the principles that were previously applied mathematically to define new stochastic models now play the role of heuristics meant to guide us to schemes whose utility is to be demonstrated and evaluted through experiments. Near the end of the chapter, specific solution methods for the vector field denoising/enhancement problem are presented. Similar approaches have been suggested in the literature (I highlight, again, Arigovindan [Ari05], which was my point of departure in the study of such methods; other relevant works are cited in the chapter itself), their main point of difference being that the ones I know about are based on quadratic optimization, whereas here we mainly emphasize their non-quadratic alternatives, and also generalize them to arbitrary differential orders while maintaining their invariance properties. To demonstrate the relative efficacy of the proposed solution methods a few experiments are conducted, on simulated as well as real-world measurements (obtained using flow-sensitive magnetic resonance imaging), and their (state-of-the-art) results reported along with some discussion/interpretation.

The chapter on applications is followed by the one titled 'Conclusion and Outlook', which, contradictorily, does not conclude the thesis, being in effect succeeded by two appendices, a table of notation, a list of references, and the

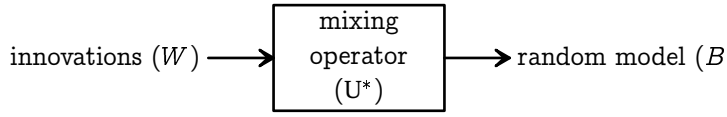
institution-required c.v. The first of the appendices, which was already referred to, is a collection of personal ruminations on the set-theoretical model of probability, sprinkled with standard theorems, definitions, and results. It is followed by a second appendix, also commented on already, where I have collected some of our publications reporting other related work that I did in the course of my doctoral studies, which I have left out of the thesis for the reasons noted previously (these mostly relate to the theory of splines and wavelets—including matrix-valued spline and wavelet bases for vector fields—and their applications in estimating the parameters of the models of Chapter 3). Also included for the purpose of comparison and historical reference are our earlier publications where similar models as in the thesis were originally introduced in a manner that would eventually lead us to—but which was nevertheless less refined than, and not as fully developed as—the formalism and models developed in Chapters 1 to 3.

Contents

— I —

Innovation Modelling

1.a Innovation modelling as presented here is a framework for constructing stochastic models, guided by a separation of the (practical or essential) unpredictability (or unwillingness to predict) ascribed to the model—imagined in the form of many independent contributions or 'atoms [1] of randomness' (innovations)—from factors that 'mix' and combine these independent random contributions. In other words, it is postulated that at some fundamental conceptual level, the randomness that exists in the model depends on the contribution of many independent factors that are then combined in a deterministic fashion. Schematically, we have:



In a variety of situations, the innovation atoms may be thought of as being probabilistically identical. The innovations (collectively denoted as W) may then be modelled by a stationary stochastic field with some form of independence built into it.

We shall not attempt to give a more profound philosophical justification for the innovation framework, although it should not be difficult to identify in it a certain level of intellectual appeal and practical relevance, at least in some physically-motivated models. Instead of philosophizing, we simply specialize and clarify the mathematical framework we shall employ for the purpose of our application, which consists in modelling a random field depending

1. Strictly speaking, we may be able to subdivide the innovations *ad infinitum,* in which case our 'atoms' would be infinitesimal.

on a spatial variable $x \in \mathbb{R}^d$. Because of spatial dependence, it is natural (although by no means exclusive) to take the innovations to be distributed 'spatially'. This idea is captured in the notion of *random fields with independent values at every point* (Gel'fand and Vilenkin [GV64]), which we shall introduce in due course. But, before doing so, let us first put in place our analytical and probabilistic framework, which involves random objects known as *generalized random fields.*

§1.1 Generalized random fields

1.b Both from the point of view of constructing a probability measure for the random model in the above schematic, and also in order to take advantage of the powerful machinery of distribution theory and related techniques (including those we shall employ in Chapter 2), it is fruitful to define our stochastic models as random elements of a space of generalized functions (distributions), and this is the course we shall follow. The underlying theory of *generalized* random fields—which provides a sound probabilistic basis for the construction—is well-developed, with a number of monographs available that are dedicated either entirely, or in part, to its systematic exposition. We highlight, in particular, the volume by Gel'fand and Vilenkin [GV64] (our primary reference) and the book by Schwartz [Sch73]. A condensed summary is given in the next few paragraphs; a slightly more detailed overview can be found in the last section of Appendix A.

Since the reader is likely to be familiar with classical random fields, we shall define generalized random fields in analogy to the former. The 'generalization' from a run-of-the-mill random field to a generalized one is similar, in essence, to the passage from a classical function to a distribution in the sense of Schwartz [Sch66] (like Dobrushin [Dob79] and the Russian school more generally, we prefer the term 'generalized function' to 'distribution' in probabilistic contexts, as the former choice avoids semantic absurdities such as 'the distribution of a distribution').

§1.1.1 Classical functions and random fields

1.c Classical functions. In the classical setting, a (deterministic) function f is a map from some domain $\mathscr{I}$ into some value space such as $\mathbb{R}$. As is standard, we define the space of all functions $\mathscr{I} \to \mathbb{R}$ as the infinite product space $\mathbb{R}^{\mathscr{I}} = \prod_{x \in \mathscr{I}} \mathbb{R}$.

1.d Classical random fields. A classical real-valued random field F, defined on the domain $\mathscr{I}$ (typically $= \mathbb{R}^d$), is identified with a probability law for drawing a function from the space $\mathbb{R}^{\mathscr{I}}$. This probability law is constructed as follows. One begins with a collection of random variables, denoted as $F(x)$, $x \in \mathscr{I}$, corresponding to the point-wise values of the random function F. These variables are characterized by the provision of the joint probabilities of $F(x)$, $x \in I$, for any *finite* subset $I \subset \mathscr{I}$. We shall denote this joint probability as $\mathscr{P}_{F,I}$.

The joint probabilities $\mathscr{P}_{F,I}$, $I \subset \mathscr{I}$ finite, are subject to certain *consistency conditions,* which guarantee the existence of a probability measure on the product space $\mathbb{R}^{\mathscr{I}}$ corresponding to F (Kolmogorov's Extension Theorem; see **§A.5**). The noted probability measure on $\mathbb{R}^{\mathscr{I}}$, which we shall denote as $\mathscr{P}_F$, is sometimes called the *stochastic law* of the random field F. We may thus view F as a *random element* of the function space $\mathbb{R}^{\mathscr{I}}$ drawn in accordance with the probability $\mathscr{P}_F$.

$\mathscr{P}_F$ has the following defining property: the *finite-dimensional marginals* of $\mathscr{P}_F$, defined as its push-forwards through the *canonical projections*

$$\pi_I : \mathbb{R}^{\mathscr{I}} \to \mathbb{R}^I : f \mapsto \{(x, f(x)) : x \in I\}, \quad I \subset \mathscr{I} \text{ finite},$$

correspond to the initial joint probabilities $\mathscr{P}_{F,I}$, $I \subset \mathscr{I}$ finite (see **A.ab** in Appendix A for the definition of the push-forward and **§A.5.2** for an overview of the connection between $\mathscr{P}_F$ and its finite-dimensional projections).

Finally, since any probabilistic computation about F depends only on $\mathscr{P}_F$, F itself becomes redundant (even irrelevant) and we may say that the random field *is* the probability measure $\mathscr{P}_F$ (this point of view is adopted in **§A.5**, where we discuss only probability measures and not random fields as such).

§1.1.2 Generalized functions and random fields

1.e Generalized functions. In intuitive terms, the passage from classical to generalized functions consists in replacing punctual 'evaluations' $f(x)$, $x \in \mathscr{I}$, by linear 'observations' $\langle \phi, f \rangle$, with ϕ belonging to some vector space $\mathscr{X}$ of *test functions.* The space $\mathscr{X}$ is typically assumed to be rich enough so as to distinguish between (or *separate*) any two non-equal generalized functions f_1, f_2. In addition, the observations must be consistent with the vector space structure of $\mathscr{X}$, meaning that

$$\langle \alpha\phi + \beta\psi, f \rangle = \alpha\langle \phi, f \rangle + \beta\langle \psi, f \rangle$$

for all $\phi, \psi \in \mathscr{X}$ and scalars α, β.

To make the theory richer and more useful, $\mathscr{X}$ is given the additional structure of a (complete, locally convex) topological vector space, and a generalized function f is required to be continuous with respect to the topology of $\mathscr{X}$, that is,

$$\phi_\nu \to \phi \quad \Rightarrow \quad \langle \phi_\nu, f\rangle \to \langle \phi, f\rangle,$$

for any convergent net $\{\phi_\nu\}$ in $\mathscr{X}$. We can then identify f with an element of the continuous dual of $\mathscr{X}$, which we denote as $\mathscr{X}'$.

1.f Scalar vs vector fields. Generally speaking, if the space $\mathscr{X}$ consists of 'scalar'-valued (typically real or complex) functions, we refer to elements of $\mathscr{X}'$ as *generalized scalar fields.* If, on the other hand, the space $\mathscr{X}$ corresponds to the dth Cartesian power of some scalar test function space $\mathscr{Y}$, so that the test functions $\phi = (\phi_1, \ldots, \phi_d) \in \mathscr{X} = \mathscr{Y}^d$ are 'vector'-valued with d coordinates, we may refer to the elements of $\mathscr{X}' = (\mathscr{Y}^d)' = (\mathscr{Y}')^d$ as *generalized vector fields.* The scalar product $\langle \phi, f\rangle$ of a generalized vector field and a vector test function is defined in terms of the products of their coordinates:

$$\langle \phi, f\rangle := \sum_{1\le i\le d} \langle \phi_i, f_i\rangle.$$

1.g Generalized random fields. A generalized random field is a mathematical object that allows one to model finite collections of linear observations of some variable (stochastic) phenomenon, and in this sense, it can be viewed as the stochastification of the concept of generalized function, in much the same way that classical random fields correspond to the stochastification of the classical notion of function.

Specifically, let $\mathscr{X}$ be our space of test functions as before. Our objective is to construct a mathematical entity that models random observations associated with test functions $\phi \in \mathscr{X}$. To this end, suppose that we are given a map $F : \phi \to F(\phi) = \langle \phi, F\rangle$ where for each $\phi \in \mathscr{X}$, $F(\phi)$ is a random variable ($\langle \phi, F\rangle$ here is merely a suggestive notation, not a true scalar product). Suppose, further, that for each finite set $E \subset \mathscr{X}$ of test functions we are given a probability measure $\mathscr{P}_{F,E}$ on $\mathbb{R}^E$ (or $\mathbb{C}^E$ in the complex case), corresponding to the joint probability of the random variables $F(\phi)$, $\phi \in E$ (when E consists of a single element ϕ, we shall write $\mathscr{P}_{F,\phi}$ in place of $\mathscr{P}_{F,\{\phi\}}$). To account for the intuitive ideas that (1) our measurements are linear, and (2) we may approximate any observation by limits of observations that are 'close' to it, we impose the following requirements on F (or, more precisely, on the finite joint probabilities $\mathscr{P}_{F,E}$, $E \subset \mathscr{X}$ finite):

(C1") *Consistency:* For any pair of scalars α, β and test functions ϕ, ψ,

$$\langle \alpha\phi + \beta\psi, F\rangle = \alpha\langle \phi, F\rangle + \beta\langle \psi, F\rangle \quad \text{in probability law.}$$

(C2") *Continuity:* For any convergent net $E_\nu \to E$ in the space $\mathscr{X}^E$ where $E \subset \mathscr{X}$ is finite,

$$E_\nu \to E \;\Rightarrow\; \mathscr{P}_{F,E_\nu} \xrightarrow{w} \mathscr{P}_{F,E},$$

where $\xrightarrow{w}$ denotes convergence in probability measure.

F—or the above collection of probability measures that defines it—is called a *generalized random field* (Gel'fand and Vilenkin [GV64]).

We can then associate with F a *cylindrical probability* $\mathscr{P}_F$ on $\mathscr{X}'$, which is a finitely-additive probability measure on the algebra of Borel cylinder sets [2] in $\mathscr{X}'$, with the property that its finite-dimensional marginals, corresponding to the above probabilities $\mathscr{P}_{F,E}$, $E \subset \mathscr{X}$ finite, are countably additive (we give only the bare minimum here; see **§A.6** for more detailed definitions). The noted finite-dimensional marginals are recovered from $\mathscr{P}_F$ as its push-forwards through the canonical projections

$$\pi_E : \mathscr{X}' \ni f \mapsto \{(\phi, \langle \phi, f\rangle) : \phi \in E\}, \quad E \subset \mathscr{X} \text{ finite}.$$

1.h The intuition behind the definition of generalized random fields, and the cylindrical probability measures that characterize them, is that these notions permit us to produce probabilistic models for any finite number of linear observations of a phenomenon without contradiction. Concretely, the probabilistic model for the observations $\langle \phi, F\rangle$, $\phi \in E$, E finite, is the finite-dimensional marginal measure $\mathscr{P}_{F,E}$, which describes the joint probability of the aforementioned observations. On the other hand, one might reasonably enquire as to whether this model permits of a 'sample path' interpretation, meaning whether it is possible to interpret F itself as an element of the space $\mathscr{X}'$ chosen randomly according to some probability measure.

The above enquiry can be formulated more precisely as follows. The cylindrical measure $\mathscr{P}_F$ is identified with a collection of 'finite-dimensional marginals' $\mathscr{P}_{F,E}$ that fulfil properties (C1") and (C2") given above. By construction, these finite-dimensional marginals are countably additive, while $\mathscr{P}_F$ in general is only known *a priori* to be finitely additive. Answering the question posed in the previous paragraph then amounts to identifying sufficient conditions in order for the cylindrical probability $\mathscr{P}_F$ to have a countably-additive extension to (a σ-algebra in) $\mathscr{X}'$.

In the next two paragraphs, we provide sufficient conditions for a unique countably-additive extension of $\mathscr{P}_F$ to exist in specific cases. These condi-

2. In brief, a Borel cylinder set in $\mathscr{X}'$ is the inverse image of some Borel set in $\mathbb{R}^E$, E finite, under a continuous finite-dimensional linear projection $\pi_E : \mathscr{X}' \to \mathbb{R}^E$; see the final section of Appendix A for details.

tions are formulated in terms of the continuity of a certain 'characteristic' functional associated with $\mathscr{P}_F$ (the definition of this functional goes back to Kolmogorov [Kol35]).

1.i Characteristic functionals. The *characteristic functional* of a generalized random field (cylindrical probability measure) $\mathscr{P}_F$ is the map

$$\widehat{\mathscr{P}_F} : \mathscr{X} \to \mathbb{C} : \phi \mapsto \int_{\mathscr{X}'} \mathrm{e}^{\mathrm{i}\langle \phi, f\rangle} \mathscr{P}_F(\mathrm{d}f) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}t} \mathscr{P}_{F,\phi}(\mathrm{d}t) = \mathbb{E}_F\{\mathrm{e}^{\mathrm{i}\langle \phi, F\rangle}\}.$$

The characteristic functional can be seen as the infinite-dimensional counterpart of the classical characteristic function in probability theory. In particular, we can recover from it the characteristic function of any finite-dimensional marginal $\mathscr{P}_{F,E}$, $E = \{\phi_1, \dots, \phi_m\} \subset \mathscr{X}$, as follows (see §A.6.3):

$$\begin{aligned} \widehat{\mathscr{P}_{F,E}}(\xi_1, \dots, \xi_m) &= \int_{\mathbb{R}^m} \mathrm{e}^{\mathrm{i}\sum_k \xi_k t_k} \quad \mathscr{P}_{F,E}(\mathrm{d}t) \\ &= \int_{\mathscr{X}'} \mathrm{e}^{\mathrm{i}\langle \sum_k \xi_k \phi_k, f\rangle} \mathscr{P}_F(\mathrm{d}f) = \widehat{\mathscr{P}_F}\big(\sum_k \xi_k \phi_k\big). \end{aligned} \tag{1.1}$$

The characteristic functional of a generalized random field is continuous in the topology of $\mathscr{X}$, and *non-negative definite* in the sense that

$$\sum_{i,j} \overline{\zeta}_i \zeta_j \widehat{\mathscr{P}}(\phi_i - \phi_j) \geq 0$$

for all finite vectors $(\zeta_1, \dots, \zeta_n) \in \mathbb{C}^n$ and $(\phi_1, \dots, \phi_n) \in \mathscr{X}^n$ of arbitrary length $n \in \mathbb{N}$. Conversely, if a functional has the above two properties and is normalized so that $\widehat{\mathscr{P}_F}(0) = 1$, it defines a cylindrical probability $\mathscr{P}_F$ on $\mathscr{X}'$.

If the space $\mathscr{X}$ of test functions is *nuclear* (***A.bs**), then the above conditions are sufficient for the cylindrical probability to admit a unique countably-additive extension to the σ-algebra generated by Borel cylinder sets, thanks to

1.j Minlos's theorem. A functional $\widehat{\mathscr{P}_F} : \mathscr{X} \to \mathbb{C}$ where the space $\mathscr{X}$ is nuclear is the characteristic functional of a countably-additive probability measure $\mathscr{P}_F$ on (the Borel cylinder σ-algebra in) $\mathscr{X}'$ iff it fulfils the following conditions:

(R1) $\widehat{\mathscr{P}_F}(0) = 1$;

(R2) $\widehat{\mathscr{P}_F}$ is non-negative definite;

(R3) $\widehat{\mathscr{P}_F}$ is continuous in the (nuclear) topology of $\mathscr{X}$.

1.k Important examples of nuclear spaces that we shall encounter in this thesis are the test function spaces $\mathscr{D}(\mathbb{R}^d)$ and $\mathscr{S}(\mathbb{R}^d)$, and their powers, specifically the spaces $\mathscr{D}^d(\mathbb{R}^d)$ and $\mathscr{S}^d(\mathbb{R}^d)$ which are dual to generalized vector field spaces. More generally, products, subspaces, and quotients with respect to closed subspaces of nuclear spaces are again nuclear (Pietsch [Pie72]).

1.l Equivalence and invariance in law. Since, as noted, the stochastic law of a generalized random field is determined by its characteristic functional, two generalized random fields are *equivalent in law* (have the same finite-dimensional marginals) iff their characteristic functionals are equal. On account of this definition, given a transformation (linear homeomorphism) $\mathrm{G} : \mathscr{X} \to \mathscr{X}$ with adjoint $\mathrm{G}^* : \mathscr{X}' \to \mathscr{X}'$, a random field F with characteristic functional $\widehat{\mathscr{P}_F}$ is G^*-*invariant in law* iff

$$\widehat{\mathscr{P}_F} \circ \mathrm{G} = \widehat{\mathscr{P}_F}$$

(we say G^*-invariant and not G-invariant because the probability law of F is defined on the dual space $\mathscr{X}'$, which is the domain of the adjoint operator G^*; cf. **1.n** below).

§1.2 Innovation modelling with generalized random fields

1.m The above theory of generalized functions provides a convenient formalism for the innovation modelling framework presented in **1.a**. The following sequence of steps describes a procedure that sets the latter framework in the former theory:

(IM1) Characterize the innovation W as a generalized random field over some test function space $\mathscr{X}$ by specifying its (normalized, continuous, and non-negative-definite) characteristic functional $\widehat{\mathscr{P}_W}$. The space $\mathscr{X}$ need not be nuclear (and the cylindrical probability $\mathscr{P}_W$ associated with $\widehat{\mathscr{P}_W}$ need not be countably additive).

(IM2) Identify a continuous linear operator $\mathrm{U} : \mathscr{E} \to \mathscr{X}$ with prescribed properties, where the space $\mathscr{E}$ is nuclear.

(IM3) Define the random model B as the generalized random field with characteristic functional

$$\widehat{\mathscr{P}_B} := \widehat{\mathscr{P}_W} \circ \mathrm{U}. \tag{1.2}$$

More explicitly, for any $\phi \in \mathscr{E}$,

$$\widehat{\mathscr{P}_B}(\phi) := \widehat{\mathscr{P}_W}(\mathrm{U}\phi). \tag{1.3}$$

The characteristic functional defined above then identifies a unique countably-additive probability measure $\mathscr{P}_B$ on $\mathscr{E}'$ (see below), which we interpret as the stochastic law of the model B.

1.n Note that in the last step, the non-negative-definiteness of $\widehat{\mathscr{P}_B}$ and its normalization such that $\widehat{\mathscr{P}_B}(0) = 1$ follow from the corresponding properties of $\widehat{\mathscr{P}_W}$ and the linearity of the operator U, while the continuity of $\widehat{\mathscr{P}_B}$ results from the fact that it is the composition of two continuous maps. Consequently, the requirements of Theorem **1.j** are fulfilled, meaning that we can associate with B a unique and well-defined (countably additive) probability measure on $\mathscr{E}'$, which is related to its characteristic functional by the identity

$$\widehat{\mathscr{P}_B}(\phi) = \int_{\mathscr{E}'} \mathrm{e}^{\mathrm{i}\langle \phi, f\rangle}\, \mathscr{P}_B(\mathrm{d}f).$$

Moreover, we have the following equivalence 'in probability' between random variables:

$$\begin{aligned} \forall \phi \in \mathscr{E} \qquad \langle \phi, B\rangle &= \langle \mathrm{U}\phi, W\rangle \qquad \text{by (1.3);} \\ &= \langle \phi, \mathrm{U}^* W\rangle \qquad \text{by the definition of the adjoint.} \end{aligned}$$

The second identity *defines* the application of the adjoint $\mathrm{U}^* : \mathscr{X}' \to \mathscr{E}'$ to the generalized random field W. As, by assumption, $\mathscr{E}$ separates the points in $\mathscr{E}'$, we may then symbolically write

$$B = \mathrm{U}^* W. \tag{1.4}$$

This gives precise meaning to our original innovation schematic given in the beginning of this chapter. We call (1.4) the *innovation equation* for generalized random field B.

1.o In Chapter 2, we shall identify families of operators U distinguished by their invariances and continuity, taking as our innovation space $\mathscr{X}$ some $\mathrm{L}_p(\mathbb{R}^d)$ space in the scalar case and its dth Cartesian power $\mathrm{L}_p^d(\mathbb{R}^d)$ in the vector case, and choosing as our nuclear space $\mathscr{E}$, accordingly, either $\mathscr{D}(\mathbb{R}^d)$ or its dth power $\mathscr{D}^d(\mathbb{R}^d)$, where $\mathscr{D}(\mathbb{R}^d)$ is the space of smooth compactly supported test functions with its standard topology.

To complete the procedure described in (IM1)–(IM3), we also need to identify another component, namely the innovation W. This will be the subject of our next section.

§1.3 Generalized random fields with independent values at every point

1.p Definition. Let $\mathscr{X}$ be a space of functions defined on the domain $\mathbb{R}^d$. Following Gel'fand and Vilenkin [GV64], we say that a generalized random field W on $\mathscr{X}$ with characteristic functional $\widehat{\mathscr{P}_W}$ has *independent values at every point* iff, for any two functions $\phi_1, \phi_2 \in \mathscr{X}$ with disjoint supports in $\mathbb{R}^d$, the random variables $\langle \phi_1, W \rangle$ and $\langle \phi_2, W \rangle$ are independent.

Note that $\langle \phi_1, W \rangle$ and $\langle \phi_2, W \rangle$ are independent iff their joint characteristic function factorizes according to

$$\widehat{\mathscr{P}}_{W,\{\phi_1,\phi_2\}}(\xi_1, \xi_2) = \widehat{\mathscr{P}}_{W,\phi_1}(\xi_1)\widehat{\mathscr{P}}_{W,\phi_2}(\xi_2).$$

Moreover, by (1.1), the above characteristic functions are related to $\widehat{\mathscr{P}_W}$ by

$$\widehat{\mathscr{P}}_{W,\phi_i}(\xi) = \widehat{\mathscr{P}_W}(\xi\phi_i), \quad i = 1, 2,$$

and

$$\widehat{\mathscr{P}}_{W,\{\phi_1,\phi_2\}}(\xi_1, \xi_2) = \widehat{\mathscr{P}_W}(\xi_1\phi_1 + \xi_2\phi_2).$$

We can therefore say that W has independent values at every point iff

$$\widehat{\mathscr{P}_W}(\xi_1\phi_1 + \xi_2\phi_2) = \widehat{\mathscr{P}_W}(\xi_1\phi_1)\widehat{\mathscr{P}_W}(\xi_2\phi_2)$$

for all $\phi_1, \phi_2 \in \mathscr{X}$ with disjoint supports and all $\xi_1, \xi_2 \in \mathbb{R}$. Taking (multivalued) logs, we find the equivalent criterion

$$\log \widehat{\mathscr{P}_W}(\xi_1\phi_1 + \xi_2\phi_2) = \log \widehat{\mathscr{P}_W}(\xi_1\phi_1) + \log \widehat{\mathscr{P}_W}(\xi_2\phi_2), \tag{1.5}$$

for all disjoint $\phi_1, \phi_2 \in \mathscr{X}$ and all $\xi_1, \xi_2 \in \mathbb{R}$.

1.q In particular, in the event that $\widehat{\mathscr{P}_W}(\phi)$, $\phi \in \mathscr{X}$, is of the form

$$\widehat{\mathscr{P}_W}(\phi) = \exp \int_{\mathbb{R}^d} u(\phi(x))\, \mathrm{d}x \tag{1.6}$$

for some continuous u with $u(0) = 0$, (1.5) is verified, since we then have

$$\begin{aligned} \log \widehat{\mathscr{P}_W}(\xi_1\phi_1 + \xi_2\phi_2) &= \int_{\mathbb{R}^d} u(\xi_1\phi_1(x) + \xi_2\phi_2(x))\, \mathrm{d}x \\ &= \int_{\operatorname{Supp}\phi_1} u(\xi_1\phi_1(x))\, \mathrm{d}x + \int_{\operatorname{Supp}\phi_2} u(\xi_2\phi_2(x))\, \mathrm{d}x \\ &= \log \widehat{\mathscr{P}_W}(\xi_1\phi_1) + \log \widehat{\mathscr{P}_W}(\xi_2\phi_2). \end{aligned}$$

The same holds more generally when $\log \widehat{\mathscr{P}_W}(\phi)$ has the form

$$\log \widehat{\mathscr{P}_W}(\phi) = \int_{\mathbb{R}^d} u(\phi(x), \ldots, \phi^{(n)}(x))\, \mathrm{d}x$$

for some $n \in \mathbb{N} \cup \{0\}$ and some function u, since we may again separate the integral over the supports of ϕ_1 and ϕ_2 (in the above formula $\phi^{(n)}$, $n \in \mathbb{N} \cup \{0\}$, denotes the vector of nth-order derivatives of ϕ).

The innovations we shall consider in this thesis are all of the type identified in (1.6), which equation we shall use to define both scalar and vector innovations by choosing suitable functions u.

1.r Scalar vs vector innovations. Since, in order to define a scalar innovation field, our test functions $\phi \in \mathscr{X}$ need to be scalar-valued, for scalar innovations u in (1.6) must be a function of a single variable.

On the other hand, for a vector-valued innovation (a generalized random 'vector' field), the test functions $\phi \in \mathscr{X}$ are vector-valued with d components $\phi_1(x), \dots, \phi_d(x)$ at any point x (cf. **1.f**). Thus, in this case, u needs to be a (scalar-valued) function of d variables.

1.s Naturally, not every choice of the function u in (1.6) leads to a valid (properly normalized, non-negative-definite, and continuous) characteristic functional. However, with the aid of the following lemma we can easily identify many suitable choices (the lemma is inspired by Theorem 2 of Gel'fand and Vilenkin [GV64, §III.4.2], where a similar result is obtained in one dimension for continuous test functions ϕ using the same trick of applying Schur's theorem to a finite sum approximation of the integral).

1.t Lemma. Let the continuous function u, with $u(0) = 0$ and $\operatorname{Re} u(x) \leq 0$ everywhere, be such that the functional $\widehat{\mathscr{P}_W}$ defined in (1.6) is continuous on some space $\mathscr{X}$. Assume, furthermore, that the elements of $\mathscr{X}$ are Lebesgue measurable. Then, in order that $\widehat{\mathscr{P}_W}$ be non-negative definite it is sufficient that the function $\xi \mapsto \mathrm{e}^{s u(\xi)}$ be non-negative definite for all $s > 0$.

Proof. We need to show that for a function u fulfilling the noted conditions,

$$\sum_{i,j} \overline{\zeta}_i \zeta_j \widehat{\mathscr{P}}(\phi_i - \phi_j) \geq 0 \tag{1.7}$$

for all finite vectors $(\zeta_1, \dots, \zeta_n) \in \mathbb{C}^n$ and $(\phi_1, \dots, \phi_n) \in \mathscr{X}^n$ of arbitrary length $n \in \mathbb{N}$. Using (1.6) we have

$$\sum_{i,j} \overline{\zeta}_i \zeta_j \widehat{\mathscr{P}}(\phi_i - \phi_j) = \sum_{i,j} \overline{\zeta}_i \zeta_j \exp\Big[\int_{\mathbb{R}^d} u(\phi_i(x) - \phi_j(x)) \, \mathrm{d}x\Big] \tag{1.8}$$

Since, per assumption, the ϕ_is are Lebesgue measurable, by Lusin's theorem for any $\epsilon > 0$ a set M with $\lambda(M^{\complement}) < \epsilon$ such that any ϕ_i is equal to some continuous function $\tilde{\phi}_i$ on M (λ here denotes the Lebesgue measure).

Fix ϵ, and let M and $\tilde{\phi}_i$ be as above.

For $k = (k_1, \ldots, k_d) \in \mathbb{Z}^d$ and $l \in \mathbb{N}$, let $\alpha_{i,k} = \tilde{\phi}_i(\frac{k_i}{l})$. Then, since u is continuous,

$$\exp\Big[\int_{\mathbb{R}^d} u(\tilde{\phi}_i(x) - \tilde{\phi}_j(x))\mathrm{d}x\Big] = \lim_{l\to\infty} \exp\Big[l^{-d}\sum_{k\in\mathbb{Z}^d} u(\alpha_{i,k} - \alpha_{j,k})\Big]$$
$$= \lim_{K\to\infty}\lim_{l\to\infty} \exp\Big[l^{-d}\sum_{|k|\le K} u(\alpha_{i,k} - \alpha_{j,k})\Big] \quad (1.9)$$

(where, for multi-integer k, $|k| := |k_1| + \ldots + |k_d|$ per the multi-index notation).

Let $\mathrm{A} = [a_{ij}]_{ij}$ be the matrix with entries

$$a_{ij} = \exp\Big(l^{-d}\sum_{|k|\le K} u(\alpha_{i,k} - \alpha_{j,k})\Big) = \prod_{|k|\le K} \exp\Big(l^{-d}u(\alpha_{i,k} - \alpha_{j,k})\Big).$$

Also define, for $|k| \le K$, the matrices $\mathrm{A}_k = [a_{kij}]_{ij}$ with entries

$$a_{kij} = \exp\Big[l^{-d}u(\alpha_{i,k} - \alpha_{j,k})\Big].$$

For each k, the matrix A_k defined above is non-negative definite by assumption (since the function $\mathrm{e}^{su(x)}$ is assumed to be positive definite for any s, and in particular for $s = l^{-d}$). Moreover, the matrix A is equal to the Hadamard (i.e. entry-wise) product of the matrices A_k, $|k| \le K$. Hence, by Schur's product theorem, A itself is non-negative definite. In other words we have

$$\sum_{i,j} \overline{\zeta}_i\zeta_j \exp\Big[l^{-d}\sum_{|k|\le K} u(\alpha_{i,k} - \alpha_{j,k})\Big] \ge 0$$

for any $(\zeta_1, \ldots, \zeta_n) \in \mathbb{C}^n$. Taking the limit in (1.9) we then obtain

$$\sum_{i,j} \overline{\zeta}_i\zeta_j \exp\Big[\int_{\mathbb{R}^d} u(\tilde{\phi}_i(x) - \tilde{\phi}_j(x))\mathrm{d}x\Big] \ge 0. \quad (1.10)$$

Finally, since $\tilde{\phi}_i = \phi_i$ outside a set of Lebesgue measure $< \epsilon$ with ϵ arbitrary, we can obtain the integral in (1.8) as a limit of (1.10) as $\epsilon \to 0$ by forming a sequence of $p\tilde{h}i_i$s for $\epsilon \to 0$, whence we find

$$\sum_{i,j} \overline{\zeta}_i\zeta_j \exp\Big[\int_{\mathbb{R}^d} u(\phi_i(x) - \phi_j(x))\mathrm{d}x\Big] \ge 0.$$

for any $(\zeta_1, \ldots, \zeta_n) \in \mathbb{C}^n$ and $(\phi_1, \ldots, \phi_n) \in \mathscr{X}^n$ with $n \in \mathbb{N}$. In other words, $\widehat{\mathscr{P}_W}(\phi) = \exp\Big[\int_{\mathbb{R}^d} u(\phi(x))\mathrm{d}x\Big]$ is non-negative definite. ∎

Using the above lemma, we can find families of non-negative-definite functionals.[3] Of the remaining conditions for $\widehat{\mathscr{P}_W}$ to be the characteristic functional of a cylindrical measure, the normalization property ($\widehat{\mathscr{P}_W}(0) = 1$) follows from $u(0) = 0$, while continuity can be verified on a case-by-case basis.

1.u Examples: Gaussian and α-stable innovations. We note that for any $s > 0$, $\mathrm{e}^{-\frac{1}{2}s|\xi|^2}$ is the characteristic function of a (jointly) Gaussian random variable (vector) with variance s, hence positive definite. Therefore, for scalar ξ, $u(\xi) = -\frac{1}{2}|\xi|^2$ fulfils the requirement of the previous lemma, and defines the *standard scalar Gaussian innovation* W_G on $\mathrm{L}_2(\mathbb{R}^d)$ with characteristic functional

$$\widehat{\mathscr{P}_{W_G}}(\phi) = \mathrm{e}^{-\frac{1}{2}\|\phi\|_2^2}, \quad \phi \in \mathrm{L}_2(\mathbb{R}^d) \tag{1.11}$$

(the continuity of $\widehat{\mathscr{P}_{W_G}}$ in the topology of L_2 is obvious as it is the composition of two continuous functions, namely the norm[4] and $\mathrm{e}^{-\frac{1}{2}|\cdot|^2}$). It is easy to see that finite-dimensional marginals obtained from the above characteristic functional are all jointly Gaussian distributed (hence the Gaussian name).

We obtain a vector field counterpart of W_G, i.e. a generalized random vector field in $\mathrm{L}_2^d(\mathbb{R}^d)$, by choosing $u(\xi) = -\frac{1}{2}|\xi|^2$ where $\xi = (\xi_1, \ldots, \xi_d)$ is now a vector and $|\xi|$ denotes its magnitude. This choice amounts to replacing the $\mathrm{L}_2(\mathbb{R}^d)$ norm in (1.11) by the vector $\mathrm{L}_2^d(\mathbb{R}^d)$ norm

$$\|\phi\|_2^2 := \int_{\mathbb{R}^d} \phi^{\mathrm{H}}\phi = \int_{\mathbb{R}^d} |\phi|^2 = \sum_{1\le k\le d} \|\phi_k\|_2^2,$$

where ϕ now denotes the column vector $(\phi_1, \ldots, \phi_d)$ and ϕ^{H} its Hermitian transpose. We shall denote this innovation by $\underline{W}_G$ and refer to it as the *standard vector Gaussian innovation.*

More generally, for $\alpha \in (0, 2]$, the function $\mathrm{e}^{-2^{-\frac{\alpha}{2}}s|\xi|^\alpha}$ defines the positive-definite characteristic functional of a symmetric α-stable random variable (vector) centred at 0 (the previous Gaussian example is a special case corresponding to $\alpha = 2$). Thus, for $1 \le \alpha \le 2$ where the corresponding L_α space is

3. It is easy to see that the requirement that $\mathrm{e}^{su(\xi)}$ be non-negative definite for any $s > 0$ is verified if $\mathrm{e}^{u(\xi)}$ is the characteristic function of an infinitely-divisible random variable (r is *infinitely divisible* iff it can be written as the sum of n independent identically distributed random variables for any $n \in \mathbb{N}$).

4. Any norm is continuous in the norm topology by definition, but those readers who prefer a more explicit demonstration may note that by triangle's inequality we have

$$0 \le \Big|\|\phi\| - \|\psi\|\Big| \le \|\phi - \psi\|,$$

which guarantees that if $\|\phi - \psi\|$ goes to zero then $\|\phi\|$ converges to $\|\psi\|$.

locally convex, we may define the scalar *standard scalar α-stable innovation* $W_{S,\alpha}$ as the generalized random field on $\mathrm{L}_{\alpha'} = (\mathrm{L}_\alpha)'$ (with $\frac{1}{\alpha} + \frac{1}{\alpha'} = 1$) whose characteristic functional is defined as

$$\widehat{\mathscr{P}}_{W_{S,\alpha}}(\phi) = \mathrm{e}^{-2^{-\frac{\alpha}{2}}\|\phi\|_\alpha^\alpha} \tag{1.12}$$

(again, L_α continuity is evident).

As in the Gaussian case, to define a vector innovation field (denoted $\underline{W}_{S,\alpha}$) we replace the $\mathrm{L}_\alpha(\mathbb{R}^d)$ norm above by a vector norm on $\mathrm{L}_\alpha^d(\mathbb{R}^d)$. Among all topologically-equivalent norms on L_α^d, we choose the rotation-invariant one given below:

$$\|\phi\|_\alpha^\alpha := \int_{\mathbb{R}^d} (\phi^{\mathrm{H}}\phi)^{\frac{\alpha}{2}} = \int_{\mathbb{R}^d} |\phi|^\alpha$$

with $\phi = (\phi_1, \ldots, \phi_d)$ now denoting a column vector in $\mathrm{L}_\alpha^d(\mathbb{R}^d)$. This choice corresponds to taking $u(\xi) = -2^{-\frac{\alpha}{2}}|\xi|^\alpha$ with $\xi = (\xi_1, \ldots, \xi_d) \in \mathbb{R}^d$ and $|\xi|$ denoting the magnitude of ξ.

Signfinicantly, if (and only if) $\alpha = 2$ (corresponding to a Gaussian innovation) then the coordinates of the above vector innovation become independent, since in this case the characteristic functional of $\underline{W}_{S,\alpha}$ factorizes along its vector coordinates:

$$\widehat{\mathscr{P}}_{\underline{W}_{S,2}}(\phi) = \prod_{1 \le i \le d} \widehat{\mathscr{P}}_{W_{S,2}}(\phi_i).$$

The random field $\underline{W}_G = \underline{W}_{S,2}$ is therefore equivalent to d independent standard Gaussian scalar innovations glued together to form a vector.

1.v Homogeneity and rotation invariance. As far as stochastic modelling is concerned, in this thesis we are interested primarily in those random models that exhibit homogeneity (self-similarity) and rotation-invariance in law (**1.I**). To construct random fields with these properties, in the following chapters we shall first identify some families of invariant mixing operators, and later apply them to innovations that are in turn homogeneous and rotation-invariant. We therefore end the present chapter with a discussion of conditions that force an innovation characterized by (1.6) to be homogeneous and rotation-invariant (the different senses of rotation-invariance in scalar and vector settings are further clarified in **§2.1**; also, we shall always understand the term 'rotation' to include all real orthogonal transformations).

Let W be an innovation whose characteristic functional $\widehat{\mathscr{P}}_W$ is of the form specified in (1.6). We recall that for a scalar innovation, the corresponding u in (1.6) is a function of one variable, whereas in the case of a vector innovation, u is a (scalar-valued) function of d variables. In this context, we have the

1.w Lemma. In the scalar setting, a characteristic functional $\widehat{\mathscr{P}_W}$ of the form given in (1.6) defines an innovation W that is always rotation invariant in law, and homogeneous of degree $\lambda = \frac{d}{\alpha} - d$ as long as $u(\xi)$ is homogeneous of degree α. In the vector field setting, W is again homogeneous in law of degree $\lambda = \frac{d}{\alpha} - d$ if $u(\xi)$ is α-homogeneous, and it is vector rotation-invariant if the d-variate function u is scalarly rotation-invariant.

Proof. Let S_σ and R_ω be the operators corresponding to scaling by $\sigma > 0$ and rotation by the $d \times d$ orthogonal matrix $\omega \in \mathrm{O}(d)$, as defined in **2.n** (with the definition of R_ω changing according as whether its operand is scalar or vector). We wish to find the conditions in order to have

$$\forall \omega \in \mathrm{O}(d) \qquad \mathrm{R}_\omega^* W = W \qquad \text{in law} \qquad \text{(rotation-invariance)}$$

and

$$\forall \sigma > 0 \qquad \mathrm{S}_\sigma^* W = \sigma^{\frac{d}{2}+\lambda} W \qquad \text{in law.} \qquad \text{(degree } \lambda \text{ homogeneity)}$$

The above equivalences in law translate to the following conditions on the characteristic functional:

$$\widehat{\mathscr{P}_{\mathrm{R}_\omega^* W}}(\phi) := \widehat{\mathscr{P}_W}(\mathrm{R}_\omega \phi) = \widehat{\mathscr{P}_W}(\phi) \tag{1.13}$$

and

$$\widehat{\mathscr{P}_{\mathrm{S}_\sigma^* W}}(\phi) := \widehat{\mathscr{P}_W}(\mathrm{S}_\sigma \phi) = \widehat{\mathscr{P}_{\sigma^{\frac{d}{2}+\lambda} W}}(\phi) =: \widehat{\mathscr{P}_W}(\sigma^{\frac{d}{2}+\lambda}\phi), \tag{1.14}$$

for all $\phi \in \mathscr{X}$. Combined together, the two requirements read

$$\widehat{\mathscr{P}_{\mathrm{S}_\sigma^* \mathrm{R}_\omega^* W}}(\phi) = \widehat{\mathscr{P}_W}(\sigma^{\frac{d}{2}+\lambda}\phi).$$

We shall consider the scalar and vector cases separately.

First, let u be a function of a single variable, corresponding to a generalized random scalar field (cf. **1.r**). In this case, ϕ is rotated 'scalarly' by ω according to the formula (see **2.n**):

$$\mathrm{R}_{\omega,s}\phi = \phi(\omega^{\mathrm{T}} \cdot).$$

We therefore want to verify if

$$\begin{aligned} \log \widehat{\mathscr{P}_W}(\sigma^{-\frac{d}{2}}\phi(\sigma^{-1}\omega^{\mathrm{T}}\cdot)) &= \int_{\mathbb{R}^d} u(\sigma^{-\frac{d}{2}}\phi(\sigma^{-1}\omega^{\mathrm{T}}x))\,\mathrm{d}x \\ &\overset{?}{=} \int_{\mathbb{R}^d} u(\sigma^{\frac{d}{2}+\lambda}\phi(x))\,\mathrm{d}x. \end{aligned}$$

After a simple change of variables from $\sigma^{-1}\omega^{\mathrm{T}}x$ to x in the middle integral, we see that W is already rotation-invariant without any additional requirement on u, since, fixing $\sigma = 1$, (1.13) is always verified. It also becomes clear that if u is α-homogeneous then the above identity holds for $\lambda = \frac{d}{\alpha} - d$.

Next, let $\mathscr{X}$ be a space of vector-valued functions with d coordinates so that $\mathscr{X} \ni \phi = (\phi_1, \dots, \phi_d)$, meaning that u is now a function of d variables, and W is a generalized random vector field (**1.r**). Viewing the test function $\phi = (\phi_1, \dots, \phi_d)$ as a column vector, we have for its 'vector' rotation the formula

$$\mathrm{R}_{\omega,v}\phi = \omega\phi(\omega^{\mathrm{T}}\cdot),$$

which simply indicates that when rotating the coordinate system, the direction of the vector at each point should nevertheless remain constant (see **2.n**). Hence, in this case, we want to find out when

$$\int_{\mathbb{R}^d} u(\sigma^{-\frac{d}{2}}\omega\phi(\sigma^{-1}\omega^{\mathrm{T}}x))\,\mathrm{d}x \stackrel{?}{=} \int_{\mathbb{R}^d} u(\sigma^{\frac{d}{2}+\lambda}\phi(x))\,\mathrm{d}x.$$

Once again, by the same change of variables as before, we find that, as in the scalar case, if u is α-homogeneous then W is homogeneous of degree $\lambda = \frac{d}{\alpha} - d$. On the other hand, for vector rotation-invariance as per (1.13) to hold we now need

$$\int_{\mathbb{R}^d} u(\omega\phi(x))\,\mathrm{d}x = \int_{\mathbb{R}^d} u(\phi(x))\,\mathrm{d}x$$

for all $\phi \in \mathscr{X}$ and $\omega \in \mathrm{O}(d)$, which requirement is clearly satisfied if u is scalarly rotation-invariant in the sense that $u(\omega\cdot) = u(\cdot)$ for all $\omega \in \mathrm{O}(d)$. ∎

1.x Remark. As a consequence of the above lemma, we see that the scalar and vectorial Gaussian and α-stable innovations introduced in **1.u** are rotation-invariant and homogeneous (self-similar) of degree $\lambda = \frac{d}{\alpha} - d$.

§1.4 Motivating examples

§1.4.1 White Gaussian Noise

1.y The first example we give is that of (scalar) white Gaussian noise on $\mathbb{R}^d$. The Gaussian innovation W_G introduced in **1.u** describes a generalized random field in $\mathrm{L}_2(\mathbb{R}^d)$ whose definition is sufficient for the purpose of modelling *finitely many* observations of a point-wise independent field with stationary Gaussian statistics. We first employ a more-or-less standard argument to show that this is no longer true if we wish to consider an *infinite* number of

observations at the same time, which is necessary in order to have a sample path interpretation of white Gaussian noise. Next, we use the methodology of **1.m** to obtain such an interpretation of white Gaussian noise in some larger space (specifically, either $\mathscr{D}'(\mathbb{R}^d)$ or $\mathscr{S}'(\mathbb{R}^d)$).

Let us put down the first statement plainly:

1.z Claim. The Gaussian cylinder probability measure $\mathscr{P}_{W_G}$ on $\mathrm{L}_2(\mathbb{R}^d)$ associated with the characteristic functional given in **1.u** does *not* extend to a countably-additive probability measure on $\mathrm{L}_2(\mathbb{R}^d)$.

Proof. Let $\{\phi_k\}_{k\in\mathbb{N}}$ be an orthonormal basis of $\mathrm{L}_2(\mathbb{R}^d)$. Thanks to orthonormality, the joint probability density of the random variables

$$\langle \phi_k, W_G\rangle, \quad k \le K,$$

is found from their joint characteristic function (cf. (1.1) and (1.11)) to be given by the multi-variate Gaussian function

$$g_K(t_1,\ldots,t_K) := (2\pi)^{-\frac{K}{2}} \mathrm{e}^{-\frac{1}{2}\sum_k t_k^2}.$$

For $K \in \mathbb{N}$, we define the cylinder set

$$C_K = \{f \in \mathrm{L}_2 : \forall k \le K \ \langle \phi_k, f\rangle \in [-1,1]\}.$$

Note that the sets C_K, $K \in \mathbb{N}$, are nested and decreasing.

We can measure C_K by $\mathscr{P}_{W_G}$, the cylinder measure associated with W_G. Indeed,

$$\begin{aligned}\mathscr{P}_{W_G}(C_K) &= \int_{[-1,1]^K} g_K(t_1,\ldots,t_K)\,\mathrm{d}t \\ &= \left(\frac{1}{\sqrt{2\pi}}\int_{-1}^{+1} \mathrm{e}^{-\frac{1}{2}t^2}\,\mathrm{d}t\right)^K \\ &= \mathrm{erf}(\tfrac{\sqrt{2}}{2})^K =: \theta^K,\end{aligned}$$

where for simplicity we have defined $\theta = \mathrm{erf}(\frac{\sqrt{2}}{2}) < 1$.

Now, consider the unit closed ball of $\mathrm{L}_2(\mathbb{R}^d)$, i.e., the set

$$B = \{f \in \mathrm{L}_2 : \|f\|_2 \le 1\}.$$

Clearly, $B \subset C_K$ for any $K \in \mathbb{N}$ (in fact, $B = \bigcap_K C_K$). Thus, the outer measure of B with respect to $\mathscr{P}_{W_G}$ is upper bounded by θ^K for any $K \in \mathbb{N}$, and this upper bound goes to zero as $K \to \infty$. In the light of this fact,

we find B to have zero $\mathscr{P}_{W_G}$ outer measure. Taking $B + \psi$ to denote the translate of B by some $\psi \in \mathrm{L}_2(\mathbb{R}^d)$, one sees similarly that $B + \psi$ has zero outer measure as well.

As $\mathrm{L}_2(\mathbb{R}^d)$ is separable, it has a countable dense subset $\{\psi_l\}_{l\in\mathbb{N}}$. $\mathrm{L}_2(\mathbb{R}^d)$ is then covered by the sets $B + \psi_l$, $l \in \mathbb{N}$. Therefore, the outer measure of $\mathrm{L}_2(\mathbb{R}^d)$ is no more than

$$\sum_{l\in\mathbb{N}} \overline{\mathscr{P}_{W_G}}(B + \psi_l) = \sum_{l\in\mathbb{N}} 0 = 0.$$

In other words, $\mathrm{L}_2(\mathbb{R}^d)$ has zero outer measure, contradicting the possibility that $\mathscr{P}_{W_G}$ could be extended to a probability measure on it. ∎

1.aa We therefore conclude that $\mathrm{L}_2(\mathbb{R}^d)$ cannot *support* a countably-additive Gaussian law with independent values at every point, which is another way of saying that our idea of white Gaussian noise does not have a sample path interpretation in $\mathrm{L}_2(\mathbb{R}^d)$.

We notice, however, that for both of the spaces $\mathscr{D}(\mathbb{R}^d)$ and $\mathscr{S}(\mathbb{R}^d)$—we shall write $\mathscr{E}(\mathbb{R}^d)$ to refer to either—the natural injection $i : \mathscr{E} \to \mathrm{L}_2$ is continuous, and so is therefore also its adjoint $i^* : \mathrm{L}_2 \to \mathscr{E}'$ (the chain $\mathscr{E} \xrightarrow{i} \mathrm{L}_2 \xrightarrow{i^*} \mathscr{E}'$ is sometimes called a Gel'fand triple). Thus, i fulfils the requirements of step (IM2) of **1.m**, meaning we can use it to define a countably-additive probability law on $\mathscr{E}'$ whose characteristic functional is given by step (IM3) of **1.m** as

$$\mathscr{E} \ni \phi \mapsto \widehat{\mathscr{P}_{W_G}}(i\phi) = \widehat{\mathscr{P}_{W_G}}(\phi) = \mathrm{e}^{-\frac{1}{2}\|\phi\|_2^2}, \tag{1.15}$$

which is formulaically the same as the characteristic functional of W_G, but with its domain restricted to $\mathscr{E}$. This shows that, unlike $\mathrm{L}_2(\mathbb{R}^d)$, the larger spaces $\mathscr{D}'(\mathbb{R}^d)$ and $\mathscr{S}'(\mathbb{R}^d)$ can support a countably-additive white Gaussian noise model with the characteristic functional given by (1.15).

§1.4.2 A simple innovation model—

1.ab We end this chapter with another example (divided between this subsection and the next), which is meant to further illustrate the approach described in **§1.2**, and also give a flavour of the forthcoming chapters.

In one dimension, the sign function, $h(x) := \mathbb{1}_{[0,\infty)}(x) - \mathbb{1}_{(-\infty,0]}(x)$, is a homogeneous distribution of order 0 in the (real) space $\mathscr{D}'(\mathbb{R})$ ($h(\sigma^{-1}x) \equiv h(x)$ for all $\sigma > 0$), which, however, is not rotation-invariant (i.e. not symmetric). h gives definition to the convolution operator

$$\mathrm{H} : \mathscr{D}(\mathbb{R}) \to \mathscr{D}'(\mathbb{R}) : \phi(x) \mapsto h * \phi(x) = \int_{-\infty}^{x} \phi(y)\,\mathrm{d}y - \int_{x}^{+\infty} \phi(y)\,\mathrm{d}y.$$

The above operator is continuous $\mathscr{D} \to \mathscr{D}'$ (simple application of the kernel theorem of Schwartz) and homogeneous of order -1 in the sense that

$$\mathrm{S}_\sigma \, \mathrm{H} \, \phi \equiv \sigma^{-1} \, \mathrm{H} \, \mathrm{S}_\sigma \, \phi$$

for the normalized dilation operator $\mathrm{S}_\sigma : \phi \mapsto \sigma^{\frac{1}{2}} \phi(\sigma^{-1} \cdot)$, $\sigma > 0$ (the normalization of S_σ is irrelevant here, but we include it for consistency with the following chapters, where it will be convenient to have $\mathrm{S}_\sigma^* = \mathrm{S}_{\sigma^{-1}}$ for its adjoint on $\mathscr{D}$). In addition, H is a left inverse of the derivative $\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}x}$. This suggests that we can formally relate H to a Fourier multiplier of the form $\widehat{h}(\xi) = 2(\mathrm{i}\xi)^{-1}$, although the said Fourier multiplier is singular at $\xi = 0$, and its corresponding Fourier integral must therefore be interpreted (or 'regularized') properly to avoid inconsistencies. In our formalism, this consistent regularization is incorporated in the definition of the 'distribution' or 'generalized function' $\widehat{h}(\xi)$ itself, and is given by the Cauchy principal value of the Fourier integral (denoted as p.v. $\int$):

$$\mathrm{H} \, \phi(x) = \mathrm{p.v.} \int_{-\infty}^{+\infty} \frac{\mathrm{e}^{\mathrm{i}x\xi} \widehat{\phi}(\xi)}{2\sqrt{2\pi}\mathrm{i}\xi} \, \mathrm{d}\xi = \int_0^{+\infty} \frac{\mathrm{e}^{\mathrm{i}x\xi} \widehat{\phi}(\xi) - \mathrm{e}^{-\mathrm{i}x\xi} \widehat{\phi}(-\xi)}{2\sqrt{2\pi}\mathrm{i}\xi} \, \mathrm{d}\xi.$$

Although presented for the sake of completeness, the above Fourier-domain characterization of H is not needed in this section, hence we shall not burden the reader with the technicalities involved therein (the concerned reader may rest assured that s/he will find more than enough of them in the next chapter). [5]

1.ac Intuitively, we might wish to apply the (adjoint of the) above homogeneous operator to, say, a homogeneous Gaussian innovation W_G with characteristic functional

$$\widehat{\mathscr{P}_{W_G}}(\phi) = \mathrm{e}^{-\frac{1}{2}\|\phi\|_2^2}, \quad \phi \in \mathrm{L}_2(\mathbb{R}),$$

as introduced in **1.u**, in order to obtain a homogeneous (self-similar) random process $B = \mathrm{H}^* W_G$, which could be interpreted as an accumulation of Gaussian innovations. This, however, poses the technical difficulty that for test functions ϕ with non-vanishing integral (0th moment), $\mathrm{H}\phi$ does not even belong to $\mathrm{L}_2(\mathbb{R})$ (it goes asymptotically to $\int_{\mathbb{R}} \phi$ at $+\infty$ and to minus that value at $-\infty$), let alone be *continuously* mapped into it, as required by step (IM2) of **§1.2**.

5. The above regularization of $(\mathrm{i}\xi)^{-1}$ is standard in the study of the Hilbert transform. In the next chapter we shall consider integrals with kernels of the form $|\xi|^\lambda$, $\xi \in \mathbb{R}^d$, which become singular for $\mathrm{Re}\,\lambda \leq -d$; the presentation there is essentially self-contained and the relevant aspects of the theory of singular integrals and distributions will be developed from the ground up.

1.ad In our framework, there exist at least two ways around the above problem. One is to find a nuclear space $\mathscr{N}$ of test functions with vanishing 0th moment, say as a subspace of $\mathscr{D}$, such that the restriction of H to this space will be continuous $\mathscr{N} \to \mathrm{L}_2$. This permits us to employ the scheme outlined in §1.2 and characterize $\mathrm{H}^* W_G$ as a generalized random process in $\mathscr{N}'$ with characteristic functional $\widehat{\mathscr{P}_{W_G}}(\mathrm{H}\phi)$, $\phi \in \mathscr{N}$.

1.ae Alternatively (and this is the primary path we shall follow), we might find a modification $\tilde{\mathrm{H}}$ of H that agrees with H on the smaller subspace of test functions with vanishing 0th moment suggested previously, but which differs from H for other $\phi \in \mathscr{D}$ (where $\mathrm{H}\phi \notin \mathrm{L}_2$), and maps continuously into L_2 also for such ϕ. Concretely, we define

$$\tilde{\mathrm{H}} \;:\; \phi(x) \;\mapsto\; h * \phi(x) - h(x) \int_{-\infty}^{+\infty} \phi(y)\, \mathrm{d}y. \tag{1.16}$$

Note that $\tilde{\mathrm{H}}$ is homogeneous in the same fashion as H, namely,

$$\mathrm{S}_\sigma \tilde{\mathrm{H}} \phi \equiv \sigma^{-1} \tilde{\mathrm{H}} \mathrm{S}_\sigma \phi.$$

From the above definition, it is clear that for any $\phi \in \mathscr{D}$, $\tilde{\mathrm{H}}\phi$ is, first, bounded, and, secondly, compactly-supported. *A fortiori,* it belongs to L_p for any p. Furthermore, by considering the limit of $\tilde{\mathrm{H}}$ for convergent sequences in $\mathscr{D}$, we can convince ourselves without too much difficulty that the map into L_p, $p \geq 1$, is in fact continuous. Thus, we may define

$$\tilde{B} = \tilde{\mathrm{H}}^* W_G$$

with

$$\widehat{\mathscr{P}_{\tilde{B}}}(\phi) := \widehat{\mathscr{P}_{W_G}}(\tilde{\mathrm{H}}\phi) = \mathrm{e}^{-\frac{1}{2}\|\tilde{\mathrm{H}}\phi\|_2^2}, \quad \phi \in \mathscr{D}(\mathbb{R}). \tag{1.17}$$

1.af We can relate the processes B and $\tilde{B}$ described in paragraphs **1.ad** and **1.ae** by observing that, since the restriction of $\tilde{\mathrm{H}}$ to the space $\mathscr{N}$ identified in the first approach is identical to the corresponding restriction of H, the generalized random process B, which was characterized in $\mathscr{N}'$, can be interpreted (at least formally) as an equivalence class of random processes in $\mathscr{D}'$, of which $\tilde{B}$ is a representative.

§1.4.3 —that describes Brownian motion

1.ag Since, unlike H, its modification $\tilde{\mathrm{H}}$ defined in (1.16) is not shift-invariant, the generalized random process $\tilde{B}$ introduced above is non-stationary. It is,

however, stationarized by taking first-order finite differences. This follows from the fact that first-order differences of $\tilde{\mathrm{H}}$ are the same as those of H (the non-stationary term $h(x)\int_{\mathbb{R}}\phi$ in (1.16) is annihilated by finite differences). Indeed, for $\psi = \phi(\cdot+u) - \phi(\cdot+v)$ we have:

$$\tilde{\mathrm{H}}\psi \;=\; h*\phi(\cdot+u) \;-\; h(\cdot)\int_{\mathbb{R}}\phi \;-\; h*\phi(\cdot+v) \;+\; h(\cdot)\int_{\mathbb{R}}\phi \;=\; \mathrm{H}\psi. \quad (1.18)$$

Consequently, the variance of a finite increment of $\tilde{B}$ should depend only on the step size and not on its location. Let us compute this variance. We first define an increment process $I = \tilde{B}(\cdot-u) - \tilde{B}(\cdot-v)$. Now, for a test function $\phi \in \mathscr{D}$, we have

$$\begin{aligned}\mathbb{E}\{|\langle\phi, I\rangle|^2\} &= \mathbb{E}\{|\langle\phi, \tilde{B}(\cdot-u) - \tilde{B}(\cdot-v)\rangle|^2\}\\ &= \mathbb{E}\{|\langle\phi(\cdot+u) - \phi(\cdot+v), \tilde{B}\rangle|^2\} = \mathbb{E}\{|\langle\psi, \tilde{B}\rangle|^2\},\end{aligned}$$

using ψ once again to denote $\phi(\cdot+u) - \phi(\cdot+v)$.

Next, by noting (as we did in (1.1)) that the characteristic function of the random variable $\langle\psi, \tilde{B}\rangle$ is nothing but

$$\widehat{\mathscr{P}}_{\langle\psi,\tilde{B}\rangle}(\xi) = \widehat{\mathscr{P}}_{\tilde{B}}(\xi\psi) = \widehat{\mathscr{P}}_{W_G}(\xi\tilde{\mathrm{H}}\phi),$$

we find:

$$\begin{aligned}\mathbb{E}\{|\langle\psi, \tilde{B}\rangle|^2\} &= -\tfrac{\mathrm{d}^2}{\mathrm{d}\xi^2}\,\widehat{\mathscr{P}}_{\langle\psi,\tilde{B}\rangle}(\xi)\Big|_{\xi=0}\\ &= -\tfrac{\mathrm{d}^2}{\mathrm{d}\xi^2}\mathrm{e}^{-\frac{1}{2}\xi^2\|\tilde{\mathrm{H}}\phi\|_2^2}\Big|_{\xi=0} && \text{by (1.17)},\\ &= \|\tilde{\mathrm{H}}\psi\|_2^2\\ &= \|\mathrm{H}\psi\|_2^2 && \text{by (1.18)},\\ &= \|h*\phi(\cdot+u) - h*\phi(\cdot+v)\|_2^2\\ &= 4\|\mathbb{1}_{[0,a]}*\phi\|_2^2 && \text{with } a := |u-v|, \quad (1.19)\\ &= 4\langle\mathbb{1}_{[-a,0]}*\mathbb{1}_{[0,a]}*\phi, \phi\rangle, && (1.20)\end{aligned}$$

where in the penultimate equality we have used the shift-invariance of the L_2 norm together with the fact that $h(\cdot) - h(\cdot-a)$ is equal to twice the indicator of the interval $[0, a]$.

Putting everything together and observing that

$$\mathbb{1}_{[-a,0]} * \mathbb{1}_{[0,a]}(x) = \tfrac{1}{2}|a-x| - |x| + \tfrac{1}{2}|a+x|,$$

we then arrive at the identity

$$\mathbb{E}\{|\langle\phi, I\rangle|^2\} = \int_{\mathbb{R}^2}\phi(s)\big(\tfrac{1}{2}|a-s+t| - |s-t| + \tfrac{1}{2}|a+s-t|\big)\phi(t)\,\mathrm{d}s\mathrm{d}t,$$

which (using the kernel theorem [6]) indicates that the correlation function of the increment random process I is

$$C_I(s,t) = \tfrac{1}{2}|a-s+t| - |s-t| + \tfrac{1}{2}|a+s-t|.$$

In particular, we have

$$\mathbb{E}\{|\tilde{B}(u) - \tilde{B}(v)|^2\} = \mathbb{E}\{|I(0)|^2\} = C_I(0,0) = |a| = |u-v|.$$

Using the above equation and the parallelogram identity, we further find that for any sequence of points $u < v < w$,

$$\mathbb{E}\{|\tilde{B}(u) - \tilde{B}(w)|^2\} = \mathbb{E}\{|\tilde{B}(u) - \tilde{B}(v)|^2\} + \mathbb{E}\{|\tilde{B}(v) - \tilde{B}(w)|^2\}.$$

As we know that $\tilde{B}$ is Gaussian with zero mean, the above identity effectively proves that its disjoint increments are independent. Since Gaussianity, zero mean, and independent increments are the main properties that give definition to Brownian motion, we conclude that our innovation model $\tilde{B}$ describes such a process.

The novel aspect of the approach to stochastic modelling discussed here is that it reduces the task of characterizing the random process to that of choosing an innovation model (in this example, the Gaussian innovation) and defining a continuous operator (here $\tilde{\mathrm{H}}$).

6. The kernel theorem (cf. Gel'fand and Vilenkin [GV64, Ch. I, §1.3]) states that for $\mathscr{E} = \mathscr{D}$ or $\mathscr{S}$, any bilinear functional l on $\mathscr{E}(\mathbb{R}^d) \times \mathscr{E}(\mathbb{R}^d)$, continuous in each argument, can be identified with a generalized function f on $\mathscr{E}(\mathbb{R}^d \times \mathbb{R}^d)$ by means of an identity of the form

$$l(\phi,\psi) = \langle \phi(x)\psi(y), f\rangle.$$

The above theorem has an abstract generalization for arbitrary nuclear spaces [GV64, Ch. I, §3.5], which we shall not need.

— 2 —

Some Similarity-Preserving Operators

In this chapter we identify certain families of operators that commute with geometrical transformations. The idea is that these operators can then be used to define random fields that are statistically invariant under the same transformations (Chapter 3), as well as algorithms that are insensitive to transformations of the coordinate system (Chapter 4). This chapter, almost in its entirety, serves as a collection of definitions, techniques, and results. Motivation for the particular definitions and results should therefore be sought in their applications in the following chapters.

Briefly, we review the contents of the chapter. In the following section, the mathematical concept of invariance—as insensitivity to the exchange of objects that are transformations of one another—is developed, and special types of geometric transformations, in particular translation, scaling, and scalar and vector rotations, are introduced. This is followed in §2.2 by a review of of translation-, scaling-, and rotation-invariant operators acting on scalars (§2.2.1), and modifications of them that acquire continuity $\mathscr{D}(\mathbb{R}^d) \to \mathrm{L}_p(\mathbb{R}^d)$ for some p, usually at the expense of translation-invariance (§2.2.2). The introduction of these new continuous modifications is necessary in order to use the framework of the previous chapter to construct and characterize probabilistically homogeneous and rotation-invariant random scalar fields with α-stable statistics, as we shall do in the next chapter.

Next, in §2.3.1, we present a new family of matrix-valued distributions that serve as kernels of homogeneous translation-invariant operators acting on vector fields, with the additional property of being rotation-invariant *in the vector sense.* This is again followed by introducing their L_p-continuous variants (§2.3.2), which we shall use in Chapter 3 to define homogeneous rotation-invariant random vector fields. Finally, in §2.4, we study the vector calculus

of the operators introduced in §2.3, particularly in connection with their Helmholtz decomposition into curl- and divergence-free components.

In addition to providing one of the two main ingredients (the 'continuous operator') for defining self-similar random models in the framework of Chapter 1 (the other ingredient, the 'innovation', having been introduced already in the previous chapter), the results obtained in this chapter make it possible to have straightforward derivations of the main properties of the said random models in Chapter 3. Later, in Chapter 4, we shall see another application of these operators, this time in the context of devising solution methods for engineering problems involving vector fields.

§2.1 Domains and transformations

Intuitively, geometrical transformations are understood as transformations of the domain of functions. This idea is made precise below.

Let us begin by considering *geometric transformations* of the Euclidean domain $\mathbb{R}^d$. Throughout this chapter x, y will denote column vectors in $\mathbb{R}^d$, with, for instance, $x = (x_1, \ldots, x_d)$. $\xi = (\xi_1, \ldots, \xi_d)$ will be used to denote a column vector in the dual of the translation group $\mathbb{R}^d$, which we identify with $\mathbb{R}^d$ itself (in simpler terms, ξ is the 'Fourier' variable for the 'spatial' variable x). In addition, we use the notation $|x| := \sqrt{x_1^2 + \ldots + x_d^2}$.

2.a We define a geometric transformation of $\mathbb{R}^d$ to be a bijection $\mathbb{R}^d \leftrightarrow \mathbb{R}^d$. Examples include any
- rotation (proper or improper) [1] by some orthogonal matrix $\omega \in \mathrm{O}(d)$: $x \mapsto \omega^{\mathrm{T}} x$;
- translation by some vector $\tau \in \mathbb{R}^d$: $x \mapsto x - \tau$; and,
- scaling by some real $\sigma > 0$: $x \mapsto \sigma^{-1} x$.

In the remainder of the chapter, the symbols ω, τ, and σ will be used exclusively to denote, respectively, orthogonal matrices, translation vectors, and scaling factors.

2.b We also consider the punctured Euclidean domain $\dot{\mathbb{R}}^d := \mathbb{R}^d \backslash \{0\}$. Geometric transformations of $\dot{\mathbb{R}}^d$ include rotations by $\omega \in \mathrm{O}(d)$ and scalings by $\sigma > 0$ as defined previously, but *not* translations. On the other hand, we can define a new geometric transformation of $\dot{\mathbb{R}}^d$ corresponding to
- inversion with respect to the unit sphere: $x \mapsto \dfrac{x}{|x|^2}$.

In more generality, we put forward the

1. A rotation is *improper* iff its corresponding matrix has determinant -1.

2.c Definition. Given a set D, a *geometric transformation* of D is any permutation of D or, what is the same, any element of the *symmetric group* $\mathrm{Sym}(D)$ consisting of all bijections on D under composition. Any subgroup of $\mathrm{Sym}(D)$ is a *transformation group* of D.

2.d We now consider geometric transformations of functions with domain D. Given a geometeric transformation $g : D \to D$ of the domain, the transformation of a function $f(x)$, $x \in D$, by g is defined thus:

$$f \mapsto f \circ g. \tag{2.1}$$

A space $\mathscr{X}$ of functions on D is said to be invariant with respect to the transformation group $G \subset \mathrm{Sym}(D)$, in short *$G$-invariant,* iff the transformation of any $f \in \mathscr{X}$ by any $g \in G$ belongs back in $\mathscr{X}$; in symbols,

$$f \circ g \in \mathscr{X} \quad \text{for all } f \in \mathscr{X}, g \in G.$$

2.e We recall some familiar examples of function spaces, namely the Lebesgue spaces $\mathrm{L}_p(\mathbb{R}^d)$, $1 \leq p \leq \infty$, with the norm topology, the space $\mathscr{D}(\mathbb{R}^d)$ of compactly-supported test functions with its standard topology, the space $\mathscr{D}'(\mathbb{R}^d)$ of distributions that is dual to $\mathscr{D}(\mathbb{R}^d)$, the space $\mathscr{S}(\mathbb{R}^d)$ of Schwartz test functions with its usual topology, and its dual $\mathscr{S}'(\mathbb{R}^d)$, the space of tempered distributions. Notational convention: in this chapter the symbol $\mathscr{E}$, wherever it appears, can be replaced by either $\mathscr{D}$ or $\mathscr{S}$.

Note that, with the exception of $\mathscr{D}(\mathbb{R}^d)$ and $\mathscr{S}(\mathbb{R}^d)$, in the above examples of 'function' spaces, an element $f(x)$ of the space in question is *not* a function of $x \in \mathbb{R}^d$ in the familiar sense (i.e. point-wise). Rather, the *domain* $\mathbb{R}^d$ and the argument x serve to remind us, among other things, that (some of) the geometric transformations of $\mathbb{R}^d$ can be applied to elements of the space in a manner reminiscent of (2.1). In particular, we can translate, rotate, and scale elements of each of the above spaces. Moreover, the said transformations preserve the *algebraic* (vector space) and *set-theoretical* (topological) structure of the noted spaces. More precisely, they are *linear* (vector space morphisms) and *bi-continuous* (topology-preserving).

These observations give rise to the following generalizations.

2.f Definition. We define the *domain* $\mathrm{Dm}\,\mathscr{X}$ of a space $\mathscr{X}$ to be the group of all *automorphisms* of $\mathscr{X}$. An automorphism is a *structure-preserving* bijection. *Structures* that are of interest to us are either *algebraic* (e.g. in vector spaces) or *set-theoretical* (e.g. in topological or measurable spaces), or a combination thereof (in topological vector spaces); the various types of structure are defined below.

We re-iterate that structure-preserving bijections on a vector space are nothing but linear bijections, while on a topological space they are precisely the *homeomorphisms* or bi-continuous maps (continuous maps with continuous inverse). The generalizations of the different notions of structure given below will not be used afterwards, and may safely be skipped (go to **2.m**).

***2.g Definition.** The *algebraic* structure of a space $\mathscr{X}$ is described, in the language of universal algebra, by a (possibly infinite) set M of *finitary operations,* each of which is a map

$$m : \mathscr{X}^{r(m)} \to \mathscr{X}$$

where $r(m)$ is a non-negative integer called the *arity* of m (by convention, $\mathscr{X}^0$ is a singleton). A map $h : \mathscr{X} \to \mathscr{X}$ *preserves* the algebraic structure of $\mathscr{X}$ iff

$$\forall m \in M \quad m \circ h^{r(m)} = h \circ m,$$

where $h^{r(m)}$ is the $r(m)$-fold Cartesian product of h.

***2.h Example.** For a vector space on the field $\mathbb{F}$ we have the binary addition map $\mathscr{X}^2 \to \mathscr{X} : (f, g) \mapsto f + g$ and a (possibly infinite) number of unary scalar multiplication maps $\mathscr{X} \to \mathscr{X} : f \mapsto \alpha f$, $\alpha \in \mathbb{F}$. Maps that preserve these operations are precisely those that are *linear.*

***2.i** We take a convenient if not particularly elegant approach to introducing *set-theoretical* structure, aimed at bringing together the notions of topological and measurable spaces. Given a space (set) $\mathscr{X}$, the power set of $\mathscr{X}$, denoted $\mathfrak{P}(\mathscr{X})$, is a Boolean algebra. A bijection h on $\mathscr{X}$ defines a set function $h : \mathfrak{P}(\mathscr{X}) \to \mathfrak{P}(\mathscr{X})$ taking each set to its image under h, as well as an inverse set map h^{-1} taking each set to its pre-image, as discussed in **§A.1**. Every set function h obtained from a bijection is a bijection on $\mathfrak{P}(\mathscr{X})$, which automatically preserves the Boolean algebraic structure of $\mathfrak{P}(\mathscr{X})$ (it preserves joins and meets, i.e. unions and intersections, as well as complements and the partial order defined by inclusion).

***2.j Definition.** The *set-theoretical structure* of $\mathscr{X}$ is specified by a number of set systems $\mathfrak{T}_\nu \subset \mathfrak{P}(\mathscr{X})$ (ν is an index). These set systems are in general not independent of one another (think of the systems of open, closed, compact, or Borel sets), and they are not arbitrary in the sense that they have some algebraic 'closedness' properties (topology, σ-algebra, *&c.*), but these will be of no consequence to us. We define structure-preserving automorphisms in $\mathscr{X}$ as those bijection h which preserve all $\mathfrak{T}_\nu$s in the sense of **§A.1**, meaning that we have,

$$\forall \nu \quad T \in \mathfrak{T}_\nu \;\Leftrightarrow\; h(T) \in \mathfrak{T}_\nu. \tag{2.2}$$

*2.k **Examples.** Thus the automorphisms of a Borel space are Borel-measurable bijections with Borel-measurable inverse. For a topological space $\mathscr{X}$ with set systems $\mathrm{Open}(\mathscr{X})$, $\mathrm{Closed}(\mathscr{X})$, and $\mathrm{Comp}(\mathscr{X})$, automorphisms are precisely the *homeomorphisms* of $\mathscr{X}$. In this case, each of the two requirements

$$T \in \mathrm{Open}(\mathscr{X}) \;\Leftrightarrow\; h(T) \in \mathrm{Open}(\mathscr{X})$$

or

$$T \in \mathrm{Closed}(\mathscr{X}) \;\Leftrightarrow\; h(T) \in \mathrm{Closed}(\mathscr{X})$$

implies the other, and they both imply

$$T \in \mathrm{Comp}(\mathscr{X}) \;\Leftrightarrow\; h(T) \in \mathrm{Comp}(\mathscr{X}),$$

so there is some redundancy in (2.2). Moreover, if $\mathscr{X}$ is a Hausdorff k-space, the last condition implies the former two; see **A.d** and Proposition ***A.e**.

*2.l **Definition.** Finally, an automorphism on a space with both algebraic and set theoretical structure is a bijection that preserves both algebraic and set theoretical structures (in a topological vector space this is precisely a linear homeomorphism).

2.m **Definition.** A *transformation group* of $\mathscr{X}$ is a group G that is isomorphic to a subgroup of $\mathrm{Dm}\,\mathscr{X}$. The isomorphism $G \to \subset \mathrm{Dm}\,\mathscr{X}$ must be given, and is referred to as the *realization* of G in $\mathscr{X}$. Once a transformation group G and its realization in $\mathscr{X}$ are given, the space $\mathscr{X}$ is labelled *G-invariant.*

In this chapter we shall be concerned only with certain transformation groups that have finite-dimensional representations as groups of operators on $\mathbb{R}^d$ or $\dot{\mathbb{R}}^d$; these representations were considered in **2.a** and **2.b**.

2.n **Notational convention.** The realization of $g \in G$ in $\mathscr{X}$ will be denoted as $\blacksquare_g$ with $\blacksquare$ replaced by an appropriate symbol indicating the type of the transformation, namely $\blacksquare = \mathrm{T}$ for translations, $\blacksquare = \mathrm{R}$ for rotations, $\blacksquare = \mathrm{S}$ for scalings, and $\blacksquare = \mathrm{O}$ for inversion with respect to the unit sphere. When dealing with different G-invariant spaces $\mathscr{X}$ and $\mathscr{Y}$, we understand that $\blacksquare_g$ denotes *different operators* according as whether it is acting on $\mathscr{X}$ or $\mathscr{Y}$. The distinction is typically irrelevant for spaces of the same type (scalar, vector, tensor, *&c.*), and even for spaces of different types, it is only in the case of rotations that we really need to distinguish between scalar and vector rotations, which we shall do by indexing the two respectively as $\mathrm{R}_{\omega,s}$ (s for scalar) and $\mathrm{R}_{\omega,v}$ (v for vector).

We give the realizations of the mentioned transformation groups below:

(G1) Translation: $T = \{\tau\} = \mathbb{R}^d$ under addition;
realization: $\mathrm{T}_\tau : f \mapsto f(\cdot - \tau)$;
adjoint:[2] $\mathrm{T}^*_\tau : f \mapsto f(\cdot + \tau)$.

(G2) Scaling: $\Sigma = \{\sigma\} = \mathbb{R}_+$ under multiplication;
realization: $\mathrm{S}_\sigma : f \mapsto \sigma^{-\frac{d}{2}} f(\sigma^{-1}\cdot)$;
adjoint: $\mathrm{S}^*_\sigma : f \mapsto \sigma^{\frac{d}{2}} f(\sigma\cdot)$.

(G3) Rotation: $\Omega = \{\omega\} = \mathrm{O}(d)$ (the real orthogonal group);
realization (scalar): $\mathrm{R}_{\omega,s} : f \mapsto f(\omega^{\mathrm{T}}\cdot)$,
realization (vector):[3] $\mathrm{R}_{\omega,v} : f \mapsto \omega f(\omega^{\mathrm{T}}\cdot)$;
adjoints: $\mathrm{R}^*_{\omega,s} : f \mapsto f(\omega\cdot)$, $\mathrm{R}^*_{\omega,v} : f \mapsto \omega^{\mathrm{T}} f(\omega\cdot)$.

(G4) Spherical inversion: $O =$ the trivial group;
realization: $\mathrm{O} : f(x) \mapsto |x|^{-d} f(\frac{x}{|x|^2})$;
adjoint: $\mathrm{O}^* : f(x) \mapsto |x|^{-d} f(\frac{x}{|x|^2})$.

Blurring further the distinction between realizations in different function spaces of the same type, with a harmless abuse of notation we may write $\mathrm{T}^*_\tau = \mathrm{T}_{-\tau}$, $\mathrm{S}^*_\sigma = \mathrm{S}_{\sigma^{-1}}$, $\mathrm{R}^*_{\omega,\bullet} = \mathrm{R}_{\omega^{\mathrm{T}},\bullet}$, $\bullet = s, v$, and $\mathrm{O}^* = \mathrm{O}$, thus identifying adjoints with group inverses in these particular examples.

2.o Some invariant function spaces. The familiar spaces $\mathscr{D}(\mathbb{R}^d)$, $\mathscr{D}'(\mathbb{R}^d)$, $\mathscr{S}(\mathbb{R}^d)$, $\mathscr{S}'(\mathbb{R}^d)$, and $\mathrm{L}_p(\mathbb{R}^d)$, $0 \le p \le \infty$, are invariant with respect to the transformation groups T, Σ, and Ω, but not with respect to O.

We define their vector counterparts as their dth Cartesian power, namely the spaces $\mathscr{D}^d(\mathbb{R}^d)$, $(\mathscr{D}')^d(\mathbb{R}^d)$, $\mathscr{S}^d(\mathbb{R}^d)$, $(\mathscr{S}')^d(\mathbb{R}^d)$, $\mathrm{L}^d_p(\mathbb{R}^d)$, and the like. These spaces are also T, Σ, Ω-invariant.

Later on, in the study of homogeneous operators, we shall find it convenient to consider certain spaces of test functions and distributions that are O-invariant (or whose Fourier transforms are so). For this reason, we introduce the space $\mathscr{D}(\dot{\mathbb{R}}^d)$ as a vector subspace of $\mathscr{D}(\mathbb{R}^d)$ obtained thus:

$$\mathscr{D}(\dot{\mathbb{R}}^d) = \{\phi \in \mathscr{D} : \mathrm{O}\phi \in \mathscr{D}\},$$

and give this new space a stronger topology than $\mathscr{D}(\mathbb{R}^d)$ by defining the convergent nets (ϕ_ν) of $\mathscr{D}(\dot{\mathbb{R}}^d)$ as those nets for which both (ϕ_ν) and $(\mathrm{O}\phi_\nu)$ converge in the topology of $\mathscr{D}(\mathbb{R}^d)$ (recall that the set $\dot{\mathbb{R}}^d = \mathbb{R}^d \backslash \{0\}$ is

2. With respect to the bilinear map $(f, g) \mapsto \langle f, g\rangle = \int_{\mathbb{R}^d} f^{\mathrm{T}} g$.

3. The different realization of Ω for vector fields can be justified by the physical requirement that the direction of the vector should remain the same in the rotated coordinate system. Mathematically, it can be seen as a consequence of the chain rule of differentiation following the identification of vector fields with differential 1-forms; cf. Rudin [Rud76, §§10.21, 10.42].

O-invariant). The space $\mathscr{S}(\dot{\mathbb{R}}^d)$ is defined similarly, replacing $\mathscr{D}(\mathbb{R}^d)$ everywhere by $\mathscr{S}(\mathbb{R}^d)$ in the above definition. The spaces $\mathscr{D}'(\dot{\mathbb{R}}^d)$ and $\mathscr{S}'(\dot{\mathbb{R}}^d)$ are defined as the corresponding topological duals. The newly introduced spaces $\mathscr{D}(\dot{\mathbb{R}}^d)$, $\mathscr{D}'(\dot{\mathbb{R}}^d)$, *&c.*, and their vector counterparts, are O, Σ, Ω-invariant but not T-invariant.

We shall denote by $\widehat{\mathscr{S}}(\dot{\mathbb{R}}^d)$ the homeomorphic image of $\mathscr{S}(\dot{\mathbb{R}}^d)$ under the Fourier transform, and use similar notations for the Fourier images of the other spaces defined above.

***2.p** The inclusion map i injects $\mathscr{D}(\dot{\mathbb{R}}^d)$ continuously into $\mathscr{D}(\mathbb{R}^d)$ (same for $\mathscr{S}(\dot{\mathbb{R}}^d)$ and $\mathscr{S}(\mathbb{R}^d)$), while its adjoint i' maps $\mathscr{D}'(\mathbb{R}^d)$ continuously into $\mathscr{D}'(\dot{\mathbb{R}}^d)$ (similarly for $\mathscr{S}'(\mathbb{R}^d)$ and $\mathscr{S}'(\dot{\mathbb{R}}^d)$). Putting topology aside, one then obtains the following direct sum decompositions of $\mathscr{D}'(\mathbb{R}^d)$ and $\mathscr{D}'(\dot{\mathbb{R}}^d)$ as vector spaces:

$$\mathscr{D}'(\mathbb{R}^d) = \mathscr{D}'(\mathbb{R}^d)/i' \oplus \operatorname{Ker} i' \quad \text{and} \quad \mathscr{D}'(\dot{\mathbb{R}}^d) = \mathscr{D}'(\mathbb{R}^d)/i' \oplus \operatorname{coKer} i'.$$

§2.1.1 Invariant operators

2.q Intuitively, an operator $\mathrm{U} : \mathscr{X} \to \mathscr{Y}$ is *invariant* under some transformation g iff the order in which g and U are applied is inconsequential. This assumes that, in some sense, the *same* transformation g is defined on $\mathscr{X}$ and $\mathscr{Y}$. We extract this notion of sameness from the language of transformation groups and realizations:

2.r **Definition.** Given G-invariant spaces $\mathscr{X}, \mathscr{Y}$, an operator $\mathrm{U} : \mathscr{X} \to \mathscr{Y}$ is said to be *(strictly) G-invariant* iff

$$\mathrm{U}\,\blacksquare_g = \blacksquare_g\,\mathrm{U} \quad \text{for all } g \in G, \tag{2.3}$$

where $\blacksquare_g$ denotes the realization of $g \in G$ in $\mathscr{X}$ on the left-hand side, and its realization in $\mathscr{Y}$ on the right.

2.s We shall adopt the above 'strict' definition of invariance with respect to rotations and translations. On the other hand, it is observed that in practice homogeneity often happens to be a more interesting property than strict scale-invariance (in fact, we shall see that the only continuous linear operators $\mathscr{D}(\mathbb{R}^d) \to \mathscr{D}'(\mathbb{R}^d)$ that are simultaneously scale-, translation-, and rotation-invariant are multiples of the identity map). For this reason, for the transformation group Σ of scalings we relax (2.3) as

$$\mathrm{U}\mathrm{S}_\sigma = \sigma^{-\lambda}\mathrm{S}_\sigma\mathrm{U} \quad \text{for all } \sigma \in \mathbb{R}_+, \tag{2.4}$$

to which property we refer as *homogeneity* or *self-similarity* of order λ (with λ a real or eventually complex scalar). Notationally, we shall henceforth indicate the homogeneity order of U as a superscript thus: U^λ.

Let us point out that the mathematical form of the gain $\sigma^{-\lambda}$ in (2.4) is not arbitrary, and is in fact the most general kind of gain-relaxation possible for scale-invariance. Moreover, a similar relaxation for rotations would not yield any further generalization (at least when limiting oneself to real-valued gains). These assertions are captured in the next two lemmata.

2.t Lemma. Suppose

$$\mathrm{US}_\sigma = \alpha_\sigma \mathrm{S}_\sigma \mathrm{U} \quad \text{for all } \sigma \in \Sigma$$

with $\alpha_\sigma \in \mathbb{R}_+$ and $\sigma \mapsto \alpha_\sigma$ continuous. Then there exists λ such that $\alpha_\sigma \equiv \sigma^{-\lambda}$.

Proof. $\sigma \mapsto \alpha_\sigma$ is a homomorphism $(\mathbb{R}_+, \cdot) \to (\mathbb{R}_+, \cdot)$. Let $\Phi(x) := \log \alpha_{\exp x}$. Φ satisfies Cauchy's additive functional equation

$$\Phi(x+y) = \Phi(x) + \Phi(y)$$

since $\Phi(x+y) = \log \alpha_{\exp x \exp y} = \log(\alpha_{\exp x}\alpha_{\exp y}) = \log \alpha_{\exp x} + \log \alpha_{\exp y} = \Phi(x) + \Phi(y)$. By the continuity of $\sigma \mapsto \alpha_\sigma$, Φ is continuous at at least one point. It is known that any solution of Cauchy's equation that is continuous at at least one point is linear: $\Phi(x) \equiv -\lambda x$ for some λ. From there,

$$\alpha_\sigma = \mathrm{e}^{\Phi(\log \sigma)} = \sigma^{-\lambda}.$$

∎

2.u Lemma. Suppose

$$\mathrm{UR}_\omega = \alpha_\omega \mathrm{R}_\omega \mathrm{U} \quad \text{for all } \omega \in \Omega$$

with $\alpha_\omega \in \mathbb{R}_+$. Then $\alpha_\omega \equiv 1$.

Proof. First of all, for those elements of $\Omega = \mathrm{O}(d)$ that are of some finite order m, i.e. for any real orthogonal matrix ω such that $\omega^m = \mathrm{I}$, by the homomorphism $\omega \mapsto \alpha_\omega$ we have $1 = \alpha_\mathrm{I} = \alpha_{\omega^m} = \alpha_\omega^m \Rightarrow \alpha_\omega = 1$, using $\alpha_\omega \in \mathbb{R}_+$.

Now, any element ω of Ω, including those of infinite order, can be written as a product of at most d simple reflections ω_i (this is the Cartan-Dieudonné theorem). Simple reflections are of order 2 and therefore, by the previous paragraph, $\alpha_{\omega_i} = 1$ for all of them. Thus, for arbitrary $\omega \in \Omega$, $\alpha_\omega = \prod_i \alpha_{\omega_i} = 1$. ∎

In the remainder of this chapter, we shall identify continuous linear operators $\mathscr{D}(\mathbb{R}^d) \to \mathscr{D}'(\mathbb{R}^d)$ (*scalar* operators) and $\mathscr{D}^d(\mathbb{R}^d) \to (\mathscr{D}')^d(\mathbb{R}^d)$ (*vector*

operators) that are self-similar with respect to scaling and invariant to translations and rotations (the key distinction between scalar and vector cases being in the different definition of rotation for vector fields vs scalars; cf. (G_3) above). In each case, we also introduce modifications of these families of operators, which map $\mathscr{D}$ continuously into a given L_p space, while maintaining homogeneity and rotation-invariance. This continuity property is needed in order to use the framework of Chapter 1 to define homogeneous and rotation-invariant random fields, but we shall see that it frequently comes at the cost of translation-invariance. We derive some of the main properties of the families of operators that we introduce, most importantly in connection with the interaction of the vector operators with physically-significant operators such as curl and divergence. These properties, as we shall later see, have interesting consequences for our random models.

§2.2 Rotation-invariant homogeneous operators: Scalar case

By the Schwartz kernel theorem, any continuous operator $U : \mathscr{D}(\mathbb{R}^d) \to \mathscr{D}'(\mathbb{R}^d)$ is associated with a unique distribution $K_U \in \mathscr{D}'(\mathbb{R}^d \times \mathbb{R}^d)$ by the identity

$$U\phi = \int_{\mathbb{R}^d} K_U(x, \cdot)\phi(x)\, dx. \tag{2.5}$$

If U is translation-invariant, the above formula takes the form of a convolution:

$$U\phi = \int_{\mathbb{R}^d} K_U(x - \cdot)\phi(x)\, dx = \phi * K_U \tag{2.6}$$

(redefining K_U as a distribution in $\mathscr{D}'(\mathbb{R}^d)$). Moreover, U is rotation-invariant and homogeneous in the sense of (2.4) iff K_U is isotropic and homogeneous of order $-\lambda - d$ in the sense that $S_\sigma K_U = \sigma^{\frac{d}{2}+\lambda} K_U$ (cf. (2.7) below).

The question of identifying homogeneous translation- and rotation-invariant operators is therefore seen to be the same as that of identifying homogeneous rotation-invariant distributions (convolution kernels) in $\mathscr{D}'(\mathbb{R}^d)$. If, in addition, the distributions thus obtained happen to belong to $\mathscr{S}' \subset \mathscr{D}'$ we shall have found all operators $U : \mathscr{S} \to \mathscr{S}'$ with the desired invariances.

§2.2.1 Homogeneous and rotation-invariant scalar distributions

The story of scalar homogeneous distributions on $\mathbb{R}^d$ is relatively well-understood, and we do not claim originality in this subsection, which is meant primarily as a compendium of known results and some of the techniques used

to obtain them; most of the material of this section is essentially covered in Gel'fand and Shilov [GS64] and also in Hörmander [Hör90].

2.v We say that a distribution f is homogeneous of order $\lambda \in \mathbb{R}$ iff

$$S_\sigma f = \sigma^{-\lambda-\frac{d}{2}} f \quad \text{for all } \sigma > 0, \tag{2.7}$$

or, equivalently, iff

$$\langle \phi, S_\sigma f \rangle = \langle S_\sigma^* \phi, f \rangle = \sigma^{-\lambda-\frac{d}{2}} \langle \phi, f \rangle$$

for all test functions ϕ. If a λ-homogeneous distribution is, in addition, rotation-invariant, then on the punctured Euclidean domain $\dot{\mathbb{R}}^d$ (i.e. as a distribution in $\mathscr{D}'(\dot{\mathbb{R}}^d)$) it can be written as a multiple of the isotropic homogeneous function

$$|x|^\lambda.$$

The question we shall consider in this subsection is if and how the above distribution, initially defined on $\mathscr{D}(\dot{\mathbb{R}}^d)$, extends to a distribution on $\mathscr{D}(\mathbb{R}^d)$.

2.w We can immediately say that the required extension of $f(x) = |x|^\lambda$ to $\mathscr{D}(\mathbb{R}^d)$, if it exists for some λ, is determined up to a homogeneous distribution localized at 0, i.e. up to a finite sum of homogeneous derivatives of the δ distribution (which sum we shall call the *local part* of f). Furthermore, since any such sum can at best be homogeneous of some *integer* order $\leq -d$, the extension for $\lambda + d \neq 0, -1, -2, \ldots$ is (again, if it exists) unique.

For $\lambda + d = 0, -1, -2, \ldots$, we can write the local part of f as

$$\sum_{|k|=-\lambda-d} a_k \delta^{(k)} = \Big(\sum_{|k|=-\lambda-d} a_k \partial_k \Big) \delta =: Q(\partial)\delta$$

where k is a multi-index in $\in \mathbb{Z}^d_{\geq 0}$, $|k| = k_1 + \ldots + k_d$, and $\partial_k = \partial_{k_1} \ldots \partial_{k_d}$. Since δ itself is rotation-invariant, the partial differential operator $Q(\partial)$ must be rotation-invariant as well; in particular it cannot include any odd partial derivatives (i.e. all k_i's are even); more specifically, it can only be a multiple of an integer iterate of the Laplacian. Consequently, the local part of f is proportional to

$$\Delta^m \delta \quad \text{for } \lambda + d = -2m = 0, -2, -4, -6, \ldots, \tag{2.8}$$

and is zero for other values of λ.

2.x Let us now turn our attention to the non-local part of f. For real $\lambda > -d$, the regular function $f(x) = |x|^\lambda$ is locally summable in $\mathbb{R}^d$, and hence defines a distribution in $\mathscr{D}'(\mathbb{R}^d)$. In view of the preceding paragraph, this uniquely

identifies f (up to normalization) for $\operatorname{Re}\lambda > -d$ (we shall henceforth allow λ to be complex). Moreover, as a function of λ, the product

$$\langle \phi, |x|^\lambda \rangle = \int_{\mathbb{R}^d} |x|^\lambda \phi(x)\,\mathrm{d}x$$

is analytic, as can be seen by complex differentiation under the integral sign with respect to λ (which is permitted since ϕ is compactly supported). Re-writing the integral in spherical coordinates $r := |x|, \theta := x/|x|$ we obtain

$$\langle \phi, |x|^\lambda \rangle = \int_0^\infty r^{\lambda+d-1} S_\phi(r)\,\mathrm{d}r \tag{2.9}$$

with

$$S_\phi(r) := \int_S \phi(r\theta)\,\mathrm{d}\theta$$

(S as the domain of integration denotes the unit sphere in $\mathbb{R}^d$). S_ϕ is easily seen to be compactly supported (as was ϕ) and infinitely differentiable at any $r > 0$. Also, at $r = 0$, by looking at finite Taylor expansions of ϕ we find that S_ϕ has derivatives of all finite orders and that its odd derivatives vanish at 0 due to anti-symmetry (we shall calculate them explicitly later on).

We noted that $\langle \phi, |x|^\lambda \rangle$ is analytic in λ for $\operatorname{Re}\lambda > -d$. It can be analytically continued to the entire complex plane with the exception of the discrete set of points $\lambda = -d, -d-2, -d-4, -d-6, \ldots$. Within strips of the form $-d-2m-2 < \operatorname{Re}\lambda < -d-2m$, $m \in \mathbb{N}$, this continuation is obtained by the formula

$$\begin{aligned}\langle \phi, |x|^\lambda \rangle &= \int_0^\infty \left[S_\phi(r) - \sum_{n \le \lfloor -\operatorname{Re}\lambda - d \rfloor} \frac{r^n}{n!} S_\phi^{(n)}(0) \right] r^{\lambda+d-1}\mathrm{d}r \\ &= \int_{\mathbb{R}^d} \left[\phi(x) - \sum_{|k| \le \lfloor -\operatorname{Re}\lambda - d \rfloor} \frac{x^k}{k!} \phi^{(k)}(0) \right] |x|^\lambda \mathrm{d}x\end{aligned} \tag{2.10}$$

(using multi-index notation in the last expression), which coincides with the previous definition of $\langle \phi, |x|^\lambda \rangle$ if $\operatorname{Re}\lambda > -d$ (the analytic continuation is of course unique).[4] We can thus define $|x|^\lambda$ as the generalized function whose value at any test function ϕ is given by the above analytic continuation. This completely determines the *non-local* part of f for $\lambda + d \neq 0, -2, -4, -6, \ldots$.[5]

4. We have used the fact that odd derivatives of S_ϕ vanish (i.e. odd terms in its Taylor expansion are zero) to deduce that the integral is well-defined for $\lambda + d = -1, -3, -5, \ldots$. It can also be seen from the above expression that $\langle \phi, |x|^\lambda \rangle$ in general has simple poles at $\lambda + d = 0, -2, -4, -6, \ldots$. We shall later verify this fact in a different manner.

5. The reader may have noted that (2.9) defines the Mellin transform of S_ϕ at $\lambda + d$; the procedure described here then amounts to analytically continuing the Mellin transform to the left half-plane (excepting even negative integers).

The equality of the two integral expressions in (2.10) is a consequence of the fact that each term in the Taylor expansion of S_ϕ about 0 is the spherical integral of the corresponding homogeneous term of the expansion of ϕ. More explicitly,

$$\frac{r^n}{n!} S_\phi^{(n)}(0) = \sum_{|k|=n} \frac{\int_S x^k \, \mathrm{d}\theta}{k!} \phi^{(k)}(0). \tag{2.11}$$

Let us now compute these terms.

2.y The spherical integral of x^k can be evaluated using a known trick described in Folland [Fol01]: we wish to calculate $\int_S x^k \, \mathrm{d}\theta$ for any monomial $x^k = x_1^{k_1} \cdots x_d^{k_d} = (r\theta)^k = r^{|k|}\theta^k$ (where, we recall, $r = |x|, \theta = x/|x|$). Due to anti-symmetry we have

$$\int_S x^k \, \mathrm{d}\theta = 0 \qquad \text{if any of the } k_i\text{'s is odd.}$$

Otherwise,

$$\begin{aligned}
\int_S x^k \, \mathrm{d}\theta &= r^{|k|} \int_S \theta^k \, \mathrm{d}\theta \\
&= r^{|k|} \frac{\int_0^\infty r^{|k|+d-1} \mathrm{e}^{-r^2} \, \mathrm{d}r \int_S \theta^k \, \mathrm{d}\theta}{\int_0^\infty r^{|k|+d-1} \mathrm{e}^{-r^2} \, \mathrm{d}r} \\
&= r^{|k|} \frac{\int_0^\infty \int_S (r\theta)^k \mathrm{e}^{-r^2} r^{d-1} \, \mathrm{d}\theta \mathrm{d}r}{\int_0^\infty r^{|k|+d-1} \mathrm{e}^{-r^2} \, \mathrm{d}r} \\
&= r^{|k|} \frac{\int_{\mathbb{R}^d} x^k \mathrm{e}^{-\sum x_i^2} \, \mathrm{d}x}{\int_0^\infty r^{|k|+d-1} \mathrm{e}^{-r^2} \, \mathrm{d}r} \\
&= r^{|k|} \frac{\prod_{1 \le i \le d} \int_{\mathbb{R}} x_i^{k_i} \mathrm{e}^{-x_i^2} \, \mathrm{d}x_i}{\int_0^\infty r^{|k|+d-1} \mathrm{e}^{-r^2} \, \mathrm{d}r} \\
&= r^{|k|} \frac{\prod_{1 \le i \le d} \Gamma\left(\frac{k_i+1}{2}\right)}{\frac{1}{2}\Gamma\left(\frac{|k|+d}{2}\right)} \qquad \text{if all } k_i\text{'s are even,}
\end{aligned}$$

where we have used the formula

$$\int_0^\infty r^{s-1} \mathrm{e}^{-r^2} \mathrm{d}r = \tfrac{1}{2}\Gamma\left(\tfrac{s}{2}\right) \tag{2.12}$$

for the gamma function. Using the properties of the gamma function, we can then write

$$\frac{\int_S x^k \, \mathrm{d}\theta}{k!} = \begin{cases} r^{|k|} \dfrac{2\pi^{\frac{d}{2}} \left(\frac{|k|}{2}\right)}{2^{|k|}\Gamma\left(\frac{d}{2}\right)\Gamma\left(\frac{|k|}{2}+1\right)\left(\frac{d}{2}\right)\left(\frac{d}{2}+1\right)\cdots\left(\frac{d}{2}+\frac{|k|}{2}-1\right)} & \text{all } k_i\text{'s even,} \\ 0 & \text{otherwise.} \end{cases}$$

Plugging this into (2.11) and comparing with the formula for the iterated Laplacian, with n even,

$$\Delta^{\frac{n}{2}}\phi(0) = \sum_{\substack{|k|=n \\ \text{all } k_i\text{'s even}}} \binom{\frac{|k|}{2}}{\frac{k_1}{2},\dots,\frac{k_d}{2}} \phi^{(k)}(0),$$

we finally get:

$$\frac{1}{n!}S_\phi^{(n)}(0) = \begin{cases} \frac{2\pi^{\frac{d}{2}}}{2^n\Gamma(\frac{d}{2})\Gamma(\frac{n}{2}+1)(\frac{d}{2})(\frac{d}{2}+1)\cdots(\frac{d}{2}+\frac{n}{2}-1)}\,\Delta^{\frac{n}{2}}\phi(0) & n \text{ even}, \\ 0 & n \text{ odd}. \end{cases} \tag{2.13}$$

The last result is a version of Pizetti's formula.

2.7 It remains to identify the non-local part of f for $\lambda + d = 0, -2, -4, -6, \dots$ (the local part having been identified in (2.8)). In other words, we want to find an isotropic homogeneous extension of $|x|^{-d-2m}$, $m = 0, 1, 2, 3, \dots$ from $\mathscr{D}(\dot{\mathbb{R}}^d)$ to $\mathscr{D}'(\mathbb{R}^d)$. We shall now show that such a homogeneous extension does not exist. Indeed, let $\tilde{f}(x)$ be an extension of $|x|^{-d-2m}$ to $\mathscr{D}'(\mathbb{R}^d)$. The distribution $\tilde{f}_0(x) := \tilde{f}(x)|x|^{2m}$ then also belongs to $\mathscr{D}'(\mathbb{R}^d)$. Furthermore, $\tilde{f}_0$ is isotropic like $\tilde{f}$, and if it is true that $\tilde{f}$ is homogeneous of order $-d-2m$, then $\tilde{f}_0$ must be homogeneous of order $-d$. Therefore, $\tilde{f}_0$ corresponds to an extension to $\mathbb{R}^d$ of $|x|^{-d}$. We have already seen that such an extension is determined uniquely up to a term localized at 0. We may therefore write $\tilde{f}_0$ as the sum of the particular extension given by the first two terms of the following expression plus some finite sum of derivatives of δ (corresponding to the third term):

$$\langle \phi, \tilde{f}_0 \rangle = \int_{|x|\le 1} \frac{\phi(x) - \phi(0)}{|x|^d}\,\mathrm{d}x + \int_{|x|>1} \frac{\phi(x)}{|x|^d}\,\mathrm{d}x + \sum_{|k|\le N} a_k \phi^{(k)}(0)$$

for some finite N and a_k's (the reason the above expression is a valid extension is that it agrees with $|x|^{-d}$ for test functions in $\mathscr{D}(\dot{\mathbb{R}}^d)$—which vanish at 0—and is continuous for all $\phi \in \mathscr{D}(\mathbb{R}^d)$).

Now, if $\tilde{f}_0$ is indeed homogeneous of order $-d$, we must have

$$\langle \phi(\tfrac{\cdot}{\sigma}), \tilde{f}_0 \rangle \equiv \langle \phi, \tilde{f}_0 \rangle$$

for any $\sigma > 0$, and *a fortiori* for $\sigma > 1$, which translates to the identity

$$\int_{|x|\le\frac{1}{\sigma}} \frac{\phi(x) - \phi(0)}{|x|^d}\,\mathrm{d}x + \int_{|x|>\frac{1}{\sigma}} \frac{\phi(x)}{|x|^d}\,\mathrm{d}x + \sum_{|k|\le N} \frac{a_k}{\sigma^{|k|}} \phi^{(k)}(0)$$
$$= \int_{|x|\le 1} \frac{\phi(x) - \phi(0)}{|x|^d}\,\mathrm{d}x + \int_{|x|>1} \frac{\phi(x)}{|x|^d}\,\mathrm{d}x + \sum_{|k|\le N} a_k \phi^{(k)}(0)$$

or, what is the same,

$$\sum_{|k|\le N} a_k\Big(1-\frac{1}{\sigma^{|k|}}\Big)\phi^{(k)}(0)=\phi(0)\int_{\frac{1}{\sigma}<|x|\le 1}|x|^{-d}\,\mathrm{d}x=\phi(0)\log\sigma, \qquad (2.14)$$

for all ϕ and all $\sigma>1$. But it is clear that, for any given N and a_k's (i.e. for any extension), we can always choose $\phi\in\mathscr{D}(\mathbb{R}^d)$ and $\sigma>1$ such that the above identity is violated. Therefore $|x|^{-d}$ and its relatives of the form $|x|^{-d-2m}$ do not possess homogeneous extensions to $\mathscr{D}'(\mathbb{R}^d)$.

2.aa The family ρ_c^λ, $\lambda, c\in\mathbb{C}$. We therefore see that the local and non-local solutions of the problem are in fact mutually exclusive, in the sense that for $\lambda+d\neq 0,-2,-4,-6,\ldots$, λ-homogeneous radial distributions are non-local and are fully characterized by (2.10), while for $\lambda+d=0,-2,-4,-6,\ldots$ they are localized at 0 and have the form specified in (2.8). Thus, for any $\lambda\in\mathbb{C}$, the solution is unique up to normalization. In fact, by a suitable choice of normalization, we can include both local and non-local solutions in the same parametric family.

To this end, first note that, by analytic continuation of (2.12) we have

$$\langle \mathrm{e}^{-|x|^2},|x|^\lambda\rangle=\langle \mathrm{e}^{-r^2},r^{\lambda+d-1}\rangle=\tfrac{1}{2}\Gamma\big(\tfrac{\lambda+d}{2}\big)$$

valid for all λ different from $-d,-d-2,-d-4,\ldots$ where the right-hand side has simple poles. Accordingly, we define a new distribution ρ_c^λ, $c\in\mathbb{C}$, by normalizing $|x|^\lambda$:

$$\rho_c^\lambda(x):=c\frac{|x|^\lambda}{2^{\frac{\lambda}{2}}\Gamma\big(\frac{\lambda+d}{2}\big)} \qquad (2.15)$$

where $|x|^\lambda$ is the distribution defined in (2.10). This normalization, as we now show, will lead to the cancellation of the singularities of $|x|^\lambda$ at negative even integers.

For $\operatorname{Re}\lambda>-d$, and, *a fortiori,* for $\operatorname{Re}\lambda>2-d$, ρ_c^λ, $c\in\mathbb{C}$, is an ordinary function with at most an integrable singularity at 0. By differentiating it we obtain the simple identity

$$\Delta\,\rho_c^\lambda=\lambda\rho_c^{\lambda-2} \qquad (2.16)$$

which, in view of the uniqueness of the analytic continuation of both sides, extends, *a priori,* to all $\lambda\neq -d,-d-2,-d-4,\ldots$. Furthermore, taking the

limit as $\lambda \to -d$ we obtain:

$$\begin{aligned}
\langle \phi, \rho_c^{-d} \rangle &= \frac{\langle \phi, \Delta\, \rho_c^{2-d} \rangle}{2-d} = \frac{\langle \Delta\, \phi, \rho_c^{2-d} \rangle}{2-d} \\
&= \frac{c}{2^{1-\frac{d}{2}}(2-d)\Gamma(1)} \int_{\mathbb{R}^d} |x|^{2-d}\, \Delta\, \phi(x)\, \mathrm{d}x \\
&= \frac{c}{2^{1-\frac{d}{2}}(2-d)} \int_0^\infty r \frac{\mathrm{d}^2}{\mathrm{d}r^2} S_\phi(r) + (d-1) \frac{\mathrm{d}}{\mathrm{d}r} S_\phi(r)\, \mathrm{d}r \\
&= \frac{c}{2^{1-\frac{d}{2}}(2-d)} \int_0^\infty (d-2) \frac{\mathrm{d}}{\mathrm{d}r} S_\phi(r)\, \mathrm{d}r \quad \text{by integration by parts;} \\
&= \frac{c}{2^{1-\frac{d}{2}}} S_\phi(0) \\
&= c \frac{|S_d|}{2^{1-\frac{d}{2}}} \phi(0) \qquad \text{by (2.11),}
\end{aligned}$$

where $|S_d| = \pi^{\frac{d}{2}}/\Gamma(\frac{d}{2})$ is the surface area of the unit sphere in d dimensions. We have thus shown that the continuation of ρ_c^λ at $\lambda = -d$ is the distribution

$$\rho_c^{-d} = c \frac{(2\pi)^{\frac{d}{2}}}{2\Gamma(\frac{d}{2})} \delta.$$

From there, by iterating (2.16), for $\lambda = -d - 2m$, $m = 0, -1, -2, \ldots$ we find:

$$\rho_c^{-d-2m} = c \frac{(2\pi)^{\frac{d}{2}}}{2\Gamma(\frac{d}{2}) \prod_{0 \le k < m} (d+2k)} (-\Delta)^m \delta = c \frac{(2\pi)^{\frac{d}{2}}}{2^{m+1}\Gamma(m+\frac{d}{2})} (-\Delta)^m \delta. \tag{2.17}$$

In fact, in addition to the above, for other values of λ we could have found the regularization formula (2.10) and the coefficients given in (2.13) directly by iterating and analytically continuing (2.16).

The expression of ρ_c^{-d-2m} given in (2.17) coincides with our previous complete characterization of radial homogeneous distributions of order $\lambda = -d$, $-d-2, -d-4, \ldots$ given in (2.8). Therefore, in addition to the non-local isotropic homogeneous distributions, the family ρ_c^λ, $\lambda, c \in \mathbb{C}$, also includes the local ones. We conclude that up to the choice of $c \in \mathbb{C}$, ρ_c^λ is *the only* radial homogeneous distribution of order λ for any given $\lambda \in \mathbb{C}$. Furthermore, it is obvious that ρ_c^λ is *tempered*. We summarize these observations as a

2.ab Theorem. The family of isotropic homogeneous distributions $\in \mathscr{D}'(\mathbb{R}^d)$ of order λ, $\lambda \in \mathbb{C}$, consists precisely of the distributions

$$\rho_c^\lambda, \quad c \in \mathbb{C}.$$

Moreover, all such distributions are tempered, i.e., they all belong to $\mathscr{S}'(\mathbb{R}^d)$.

2.ac Corollary. In order that a continuous linear operator $\mathscr{D}(\mathbb{R}^d) \to \mathscr{D}'(\mathbb{R}^d)$ be simultaneously translation- and rotation-invariant and homogeneous of order $\lambda \in \mathbb{C}$ in the sense of (2.4), it is necessary and sufficient that it can be expressed as

$$\mathrm{U}_c^\lambda : \phi \;\mapsto\; \phi * \widehat{\rho_c^\lambda} = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\langle\cdot,\xi\rangle}\, \rho_c^\lambda(\xi)\, \widehat{\phi}(\xi)\, \mathrm{d}\xi$$

for some $c \in \mathbb{C}$, where $\widehat{\lambda} := -d - \lambda$.

2.ad Remark. For $\lambda = -d - 2m$, $d \in \mathbb{N}_0$, the image of U_c^λ is finite-dimensional (since the Fourier symbol of U_c^λ is then concentrated at $\xi = 0$).

The Fourier-domain characterization of U_c^λ in the above corollary (i.e. the integral) additionally makes use of

2.ae The Fourier transform of ρ_c^λ. We define the (unitary) Fourier transform $\phi^\wedge = \widehat{\phi}$ of a test function $\phi \in \mathscr{S}(\mathbb{R}^d)$ by the integral

$$\phi^\wedge(\xi) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}\langle x,\xi\rangle} \phi(x)\, \mathrm{d}x = (2\pi)^{-\frac{d}{2}} \langle \phi, \mathrm{e}^{-\mathrm{i}\langle\cdot,\xi\rangle}\rangle.$$

The Fourier transform is inverted by the integral

$$\phi^\vee(x) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathrm{e}^{+\mathrm{i}\langle x,\xi\rangle} \phi(\xi)\, \mathrm{d}\xi = (2\pi)^{-\frac{d}{2}} \langle \phi, \mathrm{e}^{+\mathrm{i}\langle x,\cdot\rangle}\rangle$$

which also shows that

$$\bar{\phi}^\vee = \overline{\phi^\wedge}$$

($\bar{\phi}$ is the complex conjugate of ϕ). The same relations hold for $\phi \in \mathscr{D}(\mathbb{R}^d)$ with the additional property that $\widehat{\phi}$ is then complex analytic (since the support of ϕ is compact, to find the complex derivative of $\widehat{\phi}$ we can differentiate under the first integral with respect to complex ξ). We denote the space of $\widehat{\phi}$ with $\phi \in \mathscr{D}$ by $\widehat{\mathscr{D}}$.

The Fourier transform is a continuous operator (in fact a homeomorphism) on $\mathscr{S}(\mathbb{R}^d)$. It also becomes a homeomorphism $\mathscr{D}(\mathbb{R}^d) \to \widehat{\mathscr{D}}(\mathbb{R}^d)$ by identifying the open sets of $\widehat{\mathscr{D}}$ with the Fourier image of open sets in $\mathscr{D}$. In either case ($\mathscr{S}$ or $\mathscr{D}$), one can show that for test functions ϕ, ψ we have the identity $\langle \widehat{\phi}, \psi\rangle = \langle \phi, \widehat{\psi}\rangle$. Using these facts, we define the Fourier transform $\widehat{f}$ of a distribution $f \in \mathscr{S}'(\mathbb{R}^d)$ or $\mathscr{D}'(\mathbb{R}^d)$ by the identity

$$\langle \phi, \widehat{f}\rangle := \langle \widehat{\phi}, f\rangle$$

(in the case of $f \in \mathscr{D}'$, ϕ runs through $\widehat{\mathscr{D}}$).

It follows from the basic properties of the Fourier transform that the transform of an isotropic and λ-homogeneous distribution is also isotropic, and homogeneous of order $-d-\lambda$. Since we now have a complete characterization of such distributions in $\mathscr{S}'(\mathbb{R}^d)$, we determine that the Fourier transform of ρ_c^λ must be of the form $\rho_{\widehat{c}}^{-d-\lambda}$ for some $\widehat{c} \in \mathbb{C}$ proportional to c. It thus remains only to find the proportionality factor. We can compute it using the technique suggested in Gel'fand and Shilov [GS64, Ch. II, §3.3]. By plugging $\phi(x) = \mathrm{e}^{-|x|^2/2}$ with Fourier transform $\widehat{\phi}(\xi) = \phi(\xi)$ in the identity

$$\langle \widehat{\phi}, \rho_c^\lambda \rangle = \langle \phi, (\rho_c^\lambda)^\wedge \rangle = \langle \phi, \rho_{\widehat{c}}^{-d-\lambda} \rangle \tag{2.18}$$

we find

$$\widehat{c} = \frac{2^{-\frac{d+\lambda}{2}}\Gamma(-\frac{\lambda}{2})}{2^{\frac{\lambda}{2}}\Gamma(\frac{d+\lambda}{2})} \frac{\int_0^\infty \mathrm{e}^{-\frac{r^2}{2}} r^{\lambda+d-1}\,\mathrm{d}r}{\int_0^\infty \mathrm{e}^{-\frac{r^2}{2}} r^{-\lambda-1}\,\mathrm{d}r} c = c$$

where in the last equality we have used (2.12). Thus,

$$(\rho_c^\lambda)^\wedge = \rho_c^{-d-\lambda}$$

i.e. $\widehat{c} = c$.

We therefore see that for $\operatorname{Re}\lambda < -d$ where the distribution ρ_c^λ does not correspond to an ordinary function, its Fourier transform $\rho_c^{-d-\lambda}$ does. This suggests the use of (2.18) as an alternative to (2.10) for evaluating $\langle \phi, \rho_c^\lambda \rangle$ for $\operatorname{Re}\lambda < -d$.

As the map $\lambda \mapsto -d-\lambda$ will appear often in the sequel, we introduce a special notation for it (which we already used in Corollary **2.ac**):

$$\widehat{\lambda} := -d-\lambda.$$

Note that

$$\widehat{\widehat{\lambda}} = \lambda \quad \text{and} \quad \widehat{\lambda - a} = \widehat{\lambda} + a.$$

With this notation we have

$$(\rho_c^\lambda)^\wedge = \rho_c^{\widehat{\lambda}}.$$

2.af **Derivatives of ρ_c^λ.** We already noted the Laplacian formula (2.16), reproduced below:

$$\Delta\, \rho_c^\lambda(x) = \lambda \rho_c^{\lambda-2}(x).$$

Taking Fourier transforms we find another formula:

$$|x|^2\, \rho_c^\lambda(x) = -\widehat{\lambda}\rho_c^{\lambda+2}(x). \tag{2.19}$$

We can also symbolically write

$$\Delta^{-1}\,\rho_c^{\lambda-2}(x) = \frac{1}{\lambda}\rho_c^{\lambda}(x)$$

and

$$|x|^{-2}\rho_c^{\lambda+2}(x) = -\frac{1}{\widehat{\lambda}}\rho_c^{\lambda}(x) \tag{2.20}$$

but, to avoid inconsistencies, the symbols Δ^{-1} and $|x|^{-2}$ in the last two formulae must be interpreted simply as shorthand for the maps $\rho_c^{\lambda} \mapsto \rho_c^{\lambda+2}/\lambda$ and $\rho_c^{\lambda} \mapsto -\rho^{\lambda-2}/\widehat{\lambda}$ respectively. In particular, they often *fail to commute* with operations such as taking derivatives and multiplication by polynomials (as seen below).

Similarly, by differentiating ρ_c^{λ} where (in λ) it is a regular function and analytically continuing the resulting formula we obtain the first- and second-order differential identities

$$(\widehat{\lambda}+2)\partial_i\rho_c^{\lambda}(x) = -\lambda x_i \rho_c^{\lambda-2}(x)$$

and

$$\begin{aligned}\delta_{ij}\rho_c^{\lambda} &= -\frac{\widehat{\lambda}}{\lambda+2}\partial_{ij}\rho_c^{\lambda+2} + \frac{\lambda}{\widehat{\lambda}+2}x_i x_j \rho_c^{\lambda-2} \\ &= -\frac{\widehat{\lambda}}{\lambda+2}\partial_{ij}\rho_c^{\lambda+2} - \frac{\lambda}{\widehat{\lambda}+2}\left(\partial_{ij}\rho_c^{\widehat{\lambda}+2}\right)^{\wedge} \\ &= -\big[\widehat{\lambda}\partial_{ij}\,\Delta^{-1} + \lambda\frac{x_i x_j}{|x|^2}\big]\rho_c^{\lambda}\end{aligned} \tag{2.21}$$

(δ_{ij} is the Kronecker delta).

The passage to the last line of the last display equation takes place in accordance with the previous discussion about Δ^{-1} and $|x|^{-2}$ (in particular, although this is not clear from the notation, $|x|^{-2}$ must be applied *before* $x_i x_j$).

2.ag Products and convolutions. We say that the product $f_1 f_2$ of two generalized functions f_1, f_2 equals some generalized function f iff $\langle f_1\phi, f_2\rangle = \langle \phi, f\rangle$ for all test functions ϕ taken from the space over which f is defined. Similarly, we may define the convolution $f_1 * f_2$ to be equal to f iff $\langle f_1 * \phi, f_2\rangle = \langle \phi, f\rangle$ for all test functions ϕ. Note that with these definitions, the product and/or convolution of two arbitrary generalized functions may not always be defined, and even when they are, *a priori,* they need not be associative or commutative. (It is instructive, in considering such questions, to consider the maps $\phi \mapsto f_1\phi$ and $\phi \mapsto f_1 * \phi$ as operators on test functions. The most

convenient situation is of course when these operators map back into the same test function space.)

In the case of the distributions we have just introduced, considered as tempered distributions in $\mathscr{S}'(\dot{\mathbb{R}}^d)$ (and, *a fortiori,* as distributions in $\mathscr{D}'(\dot{\mathbb{R}}^d)$), ρ_c^λ and $\rho_{c'}^{\lambda'}$ can be multiplied and the result is always a distribution in $\mathscr{S}'(\dot{\mathbb{R}}^d)$ that belongs to the same family (this is because $\phi \in \mathscr{S}(\dot{\mathbb{R}}^d) \;\Rightarrow\; \rho_c^\lambda \phi \in \mathscr{S}(\dot{\mathbb{R}}^d)$). Specifically,

$$\rho_c^\lambda \rho_{c'}^{\lambda'} = \rho_{c''}^{\lambda+\lambda'} \qquad \text{with} \quad c'' = cc' \frac{\Gamma(-\widehat{\frac{\lambda+\lambda'}{2}})}{\Gamma(-\widehat{\frac{\lambda}{2}})\Gamma(-\widehat{\frac{\lambda'}{2}})} \tag{2.22}$$

Similarly, from the convolution-multiplication exchange formula we see that as distributions on $\widehat{\mathscr{S}}(\dot{\mathbb{R}}^d)$, $\rho_c^{\widehat{\lambda}}$ and $\rho_{c'}^{\widehat{\lambda'}}$ can be convolved according to

$$\rho_c^{\widehat{\lambda}} * \rho_{c'}^{\widehat{\lambda'}} = \rho_{c''}^{\widehat{\lambda+\lambda'}} \qquad \text{with} \quad c'' = cc' \frac{\Gamma(-\widehat{\frac{\lambda+\lambda'}{2}})}{\Gamma(-\widehat{\frac{\lambda}{2}})\Gamma(-\widehat{\frac{\lambda'}{2}})} \tag{2.23}$$

(in this case, we rely on the implication $\phi \in \widehat{\mathscr{S}}(\dot{\mathbb{R}}^d) \;\Rightarrow\; \rho_c^\lambda * \phi \in \widehat{\mathscr{S}}(\dot{\mathbb{R}}^d)$).

Moreover, convolution and multiplication are associative and commutative in the above cases.

If, on the other hand, we take ϕ from the larger test function space $\mathscr{S}(\mathbb{R}^d)$, the product $\rho_c^\lambda \phi$ and convolution $\rho_c^\lambda * \phi$ will be distributions in $\mathscr{S}'(\mathbb{R}^d)$ that, generally speaking, are no longer confined to $\mathscr{S}(\mathbb{R}^d)$. For this reason, the above formulae for products and convolutions may not always hold, and where they do hold, they may fail to be associative. Yet, at least for $\operatorname{Re}\lambda \geq 0$ and $\operatorname{Re}\lambda' \geq -\operatorname{Re}\lambda$, (2.22) and (2.23) are still valid (operating from left to right). We have already seen a special case of the former result in (2.19).

In particular, for $\lambda' = -\lambda$ with $\operatorname{Re}\lambda > 0$ we have

$$\rho_c^\lambda \; \rho_{c'}^{-\lambda} = \frac{cc'}{\Gamma(-\widehat{\frac{\lambda}{2}})\Gamma(-\widehat{\frac{-\lambda}{2}})} = \frac{c_1 c_2}{\Gamma(\frac{d+\lambda}{2})\Gamma(\frac{d-\lambda}{2})} =: cc' H_\lambda \quad \text{(a constant)} \tag{2.24}$$

and also,

$$\rho_c^{\widehat{\lambda}} * \rho_{c'}^{\widehat{-\lambda}} = cc' H_\lambda \; \delta. \tag{2.25}$$

The last product and convolution are zero iff, at the same time, $\lambda \in \mathbb{R}$ and $|\lambda| - d = 0, 2, 4, \ldots$ (cf. Figure 2.1).

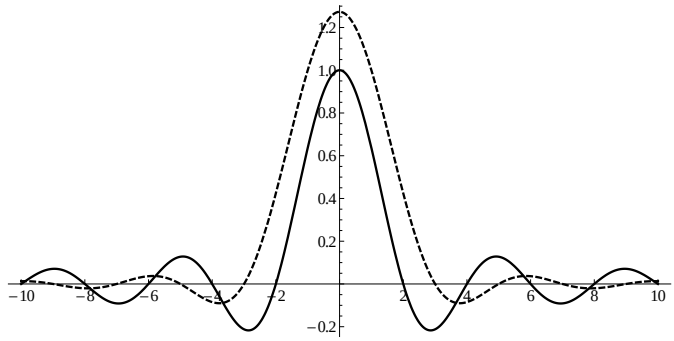


Figure 2.1: Plot of $H_\lambda = \frac{1}{\Gamma(\frac{d+\lambda}{2})\Gamma(\frac{d-\lambda}{2})}$ along the real λ line for $d = 2$ (solid) and 3 (dashed). For $d = 1$, H_λ is equal to the sinc function $\mathrm{sinc}(\lambda) := \frac{\sin(\pi\lambda)}{\pi\lambda}$ (cf. Euler's reflection formula).

§2.2.2 L_p-continuous homogeneous and rotation-invariant operators

2.ah In the previous subsection we considered translation- and rotation-invariant homogeneous continuous operators $\mathscr{E}(\mathbb{R}^d) \to \mathscr{E}'(\mathbb{R}^d)$, $\mathscr{E} = \mathscr{D}, \mathscr{S}$. Looking back at the Fourier-domain characterization of these operators given in Corollary **2.ac** we see that the restriction to $\widehat{\mathscr{E}}(\dot{\mathbb{R}}^d)$ of these operators is continuous $\widehat{\mathscr{E}}(\dot{\mathbb{R}}^d) \to \widehat{\mathscr{E}}(\dot{\mathbb{R}}^d)$ and, *a fortiori,* $\widehat{\mathscr{E}}(\dot{\mathbb{R}}^d) \to L_p(\mathbb{R}^d)$ for all p, $1 \le p \le \infty$.

Moreover, from (2.25), on $\widehat{\mathscr{E}}(\dot{\mathbb{R}}^d)$ we have the following *inversion formula:*

$$U_c^{-\lambda} U_{1/c}^{\lambda} \phi = H_\lambda \, \phi = \frac{1}{\Gamma(\frac{d+\lambda}{2})\Gamma(\frac{d-\lambda}{2})} \, \phi \tag{2.26}$$

By Paragraph **2.ag**, the latter formula extends to all $\phi \in \mathscr{E}(\mathbb{R}^d)$, provided that $\mathrm{Re}\,\lambda \ge 0$ (we are implicitly extending the domain of the leftmost operator). Note, however, that for $\lambda - d = 0, 2, 4, \ldots$, where $H_\lambda = 0$, we cannot use (2.26) to invert $U_{1/c}^{\lambda}$ from the left. But this is not a limitation in our theory since in fact, in this case $U_{1/c}^{\lambda}$ does not have *any* left inverse on either of $\mathscr{E}(\mathbb{R}^d)$ or $\widehat{\mathscr{E}}(\dot{\mathbb{R}}^d)$; this is because the image of $U_{1/c}^{\lambda}$ is finite-dimensional for such λ while its domain isn't; see Remark **2.ad**. Symmetrically, we shall exclude $-\lambda - d = 0, 2, 4, \ldots$ from consideration since there again $H_\lambda = 0$.

The translation-invariant extension of $U_c^{-\lambda}$ from $\widehat{\mathscr{E}}(\dot{\mathbb{R}}^d)$ to $\mathscr{E}(\mathbb{R}^d)$ characterized by the convolution kernel $\rho_c^{\lambda-d}$ (which, assuming the additional invariances and continuity, is unique as we showed in Corollary **2.ac**), in general fails to map into $L_p(\mathbb{R}^d)$. Since the latter property is necessary for our

characterization of generalized random fields in the next chapter, in this subsection we introduce a modified version of the operator $\mathrm{U}_c^{-\lambda}$ that achieves $\mathscr{E}(\mathbb{R}^d) \to \mathrm{L}_p(\mathbb{R}^d)$ continuity at the cost of translation-invariance (which we can afford to relax, since we are comfortable with non-stationary solutions).

2.ai Definition. The operator $\mathrm{Reg}_{c,k}^{\lambda} : \mathscr{E}(\mathbb{R}^d) \to \mathscr{E}'(\mathbb{R}^d)$, $\mathscr{E} = \mathscr{D}, \mathscr{S}$, is defined as

$$\mathrm{Reg}_{c,n}^{\lambda} \;:\; \phi \;\mapsto\; \sum_{|k|\le n} \frac{\partial_k \widehat{\rho_c^{\lambda}}}{k!} \int_{\mathbb{R}^d} (-y)^k \phi(y)\, \mathrm{d}y,$$

where k is a multi-index in $\mathbb{N}_0^d$ with $|k| = \sum_i k_i$, $(-y)^k = \prod_i (-y_i)^{k_i}$, and $k! = \prod_i k_i!$; we further understand the sum to be zero when empty, i.e. for $n < 0$.

The following lemmata are not immediately needed but will prove useful later on.

2.aj Lemma. The adjoint $\mathscr{E}(\mathbb{R}^d) \to \mathscr{E}'(\mathbb{R}^d)$ of $\mathrm{Reg}_{c,n}^{\lambda}$ is the operator

$$\mathrm{Reg}_{c,n}^{\lambda *} \;:\; \phi \;\mapsto\; \sum_{|k|\le n} y^k \frac{\partial_k \mathrm{U}_c^{\lambda} \phi(0)}{k!},$$

which maps ϕ precisely to the nth degree Taylor expansion of $\mathrm{U}_c^{\lambda}\phi$ at 0.

Proof. One needs to verify that $\langle \phi, \mathrm{Reg}_{c,n}^{\lambda *} \psi \rangle = \langle \psi, \mathrm{Reg}_{c,n}^{\lambda *} \phi \rangle$ for $\phi, \psi \in \mathscr{E}$; this can be done simply by replacing the definitions of the two operators in the identity. ∎

2.ak Lemma. $\mathrm{Reg}_{c,n}^{\lambda}\, \partial_i \phi = \partial_i\, \mathrm{Reg}_{c,n-1}^{\lambda}\, \phi$.

Proof. For $n \le 0$ both sides are null, thus equal. For $n > 0$, we re-express $\mathrm{Reg}_{c,n}^{\lambda}$ in the Fourier domain by writing moments in terms of Fourier derivatives and then taking a Fourier transform with respect to x:

$$\begin{aligned}(\mathrm{Reg}_{c,n}^{\lambda}\, \phi)^{\wedge}(\xi) &= \sum_{|k|\le n} \frac{(\mathrm{i}\xi)^k \rho_c^{\lambda}(\xi)}{k!} (-\mathrm{i})^{|k|} \partial_k \widehat{\phi}(0) \\ &= \rho_c^{\lambda}(\xi) \sum_{|k|\le n} \frac{\xi^k}{k!} \partial_k \widehat{\phi}(0) = \rho_c^{\lambda}(\xi) T_{\widehat{\phi},n}(\xi) \qquad (2.27)\end{aligned}$$

where $T_{\widehat{\phi},n}(\xi)$ denotes the nth-degree Taylor series expansion of $\widehat{\phi}$ at $\xi = 0$.

Replace ϕ by $\partial_i \phi$, hence $\widehat{\phi}$ by $\mathrm{i}\xi_i \phi$, to have

$$(\mathrm{Reg}_{c,n}^{\lambda}\, \partial_i \phi)^{\wedge}(\xi) = \mathrm{i} \rho_c^{\lambda}(\xi) T_{\xi_i \widehat{\phi},n}(\xi).$$

Next, note that $T_{\xi_i \widehat{\phi},n} = \xi_i T_{\widehat{\phi},n-1}$. Thus,

$$(\mathrm{Reg}_{c,n}^{\lambda}\, \partial_i \phi)^{\wedge}(\xi) = \mathrm{i}\xi_i \rho_c^{\lambda}(\xi) T_{\widehat{\phi},n-1}(\xi) = (\partial_i \,\mathrm{Reg}_{c,n-1}^{\lambda}\, \phi)^{\wedge}(\xi)$$

by (2.27). ∎

The next lemma concerns homogeneity and rotation-invariance.

2.al Lemma. $\mathrm{Reg}_{c,n}^{\lambda}$ is rotation-invariant and homogeneous of order λ in the sense of (2.4).

Proof. Homogeneity is easily established by putting the following facts together: the distribution $\rho_c^{\widehat{\lambda}}$ is homogeneous of order $\widehat{\lambda}$; $\partial_k \rho_c^{\widehat{\lambda}}$ is therefore homogeneous of order $\widehat{\lambda} - |k| = -d - \lambda - |k|$; the map $\phi \mapsto \int_{\mathbb{R}^d} (-y)^k \phi(y)\, \mathrm{d}y$ is homogeneous of order $-d - |k|$ in the sense of (2.4).

To prove rotation invariance is to show that $\mathrm{Reg}_{c,n}^{\lambda}\, \mathrm{R}_{\omega,s}\phi = \mathrm{R}_{\omega,s}\, \mathrm{Reg}_{c,n}^{\lambda}\, \phi$ for all orthogonal matrices $\omega \in \Omega = O(d)$ and all $\phi \in \mathscr{E}(\mathbb{R}^d)$. We use the Fourier-domain expression of $\mathrm{Reg}_{c,n}^{\lambda}\, \phi$ given in (2.27):

$$(\mathrm{Reg}_{c,n}^{\lambda}\, \phi)^{\wedge}(\xi) = \rho_c^{\lambda}(\xi) T_{\widehat{\phi},n}(\xi).$$

By the uniqueness of the Taylor series expansion, the Taylor expansion of any rotation of ϕ is equal to the corresponding rotation of its Taylor expansion; that is, we have

$$T_{\mathrm{R}_\omega \widehat{\phi},n}(\xi) = \mathrm{R}_\omega T_{\widehat{\phi},n}(\xi).$$

Combining the last identity with the rotation-invariance of ρ_c^{λ} and the rotation property of the Fourier transform (namely, that $(\mathrm{R}_\omega \phi)^{\wedge} = \mathrm{R}_\omega \widehat{\phi}$) yields the lemma. ∎

2.am Corollary. The *corrected* operator

$$\mathrm{U}_{c,n}^{-\lambda} := \mathrm{U}_c^{-\lambda} - \mathrm{Reg}_{c,n}^{-\lambda}$$

is $(-\lambda)$-homogeneous and scalar-rotation-invariant.

2.an Remark. Note that, in general, $\mathrm{U}_{c,n}^{-\lambda}$ is *not* translation-invariant except when $\mathrm{Reg}_c^{-\lambda}$ is nullified either by choosing $n < 0$ or by restricting the domain of $\mathrm{U}_{c,n}^{-\lambda}$ to (a subspace of) the kernel of $\mathrm{Reg}_c^{-\lambda}$. However, a finite difference of sufficiently high order of $\mathrm{Reg}_{c,n}^{-\lambda}$ (and hence that of $\mathrm{U}_{c,n}^{-\lambda}$) becomes translation-invariant:

2.ao Lemma. Let D_h, $h \in \mathbb{R}^d$, be the first-order symmetric finite difference operator

$$f \mapsto f(\cdot + \tfrac{1}{2}h) - f(\cdot - \tfrac{1}{2}h)$$

and let

$$\mathrm{D}_H := \mathrm{D}_{h_n} \mathrm{D}_{H \setminus \{h_n\}}$$

be the recursively-defined finite difference operator of order $n+1$ corresponding to the finite set of vectors $H = \{h_0, \dots, h_n \in \mathbb{R}^d\}$ (for $H = \varnothing$ we define D_H to be the identity operator). Then,

$$\mathrm{Reg}_{c,n}^{-\lambda}\ \mathrm{D}_H = \mathrm{D}_H\ \mathrm{Reg}_{c,n}^{-\lambda} = 0.$$

Consequently, $\mathrm{U}_{c,n}^{-\lambda}\mathrm{D}_H = \mathrm{D}_H\mathrm{U}_{c,n}^{-\lambda} = \mathrm{D}_H\mathrm{U}_c^{-\lambda}$, which is translation-invariant.

Proof. For $H = \varnothing$ we have $n = -1$ and $\mathrm{Reg}_{c,n}^{-\lambda} = 0$ is trivially translation-invariant. For $n \geq 0$, it suffices to note that the moments of $\mathrm{D}_H\phi$ vanish up to order n for any test function ϕ. The latter property is easily proved by induction on n. ∎

Similar to $\mathrm{U}_c^{-\lambda}$, the corrected operator $\mathrm{U}_{c,n}^{-\lambda}$ is also a left inverse of $\mathrm{U}_{1/c}^{\lambda}$. To wit, $\widehat{\mathscr{E}}(\dot{\mathbb{R}}^d)$ lies in the kernel of $\mathrm{Reg}_{c,n}^{-\lambda}$, which, together with (2.26), proves the first part of the following

2.ap Lemma. $\mathrm{U}_{c,n}^{-\lambda}\mathrm{U}_{1/c}^{\lambda}\,\phi = H_\lambda\phi$ for all $\phi \in \widehat{\mathscr{E}}(\dot{\mathbb{R}}^d)$. The same relation holds more generally for $\phi \in \mathscr{E}(\mathbb{R}^d)$ and the corresponding extension of the leftmost operators, provided that $n < \mathrm{Re}\,\lambda$.

The second half of the lemma is proved by observing that, according to the Fourier-domain characterization of $\mathrm{U}_{1/c}^{\lambda}$, given $n < \mathrm{Re}\,\lambda$ the moments of $\mathrm{U}_{1/c}^{\lambda}\phi$ vanish up to degree n at least; hence $\mathrm{Reg}_{c,n}^{-\lambda}\,\mathrm{U}_{1/c}^{\lambda} = 0$ for $n < \mathrm{Re}\,\lambda$, and we are back to (2.26).

Finally, we have the following key result about L_p continuity, which will be used in the next chapter to characterize homogeneous rotation-invariant random fields.

2.aq Theorem. The operator

$$\mathrm{U}_{c,n}^{-\lambda} = \mathrm{U}_c^{-\lambda} - \mathrm{Reg}_{c,n}^{-\lambda}$$

with $n = \lfloor \mathrm{Re}\,\lambda + \frac{d}{p} \rfloor - d$ and $\mathrm{Re}\,\lambda + \frac{d}{p} \notin \mathbb{N}$ is continuous $\mathscr{D}(\mathbb{R}^d) \to \mathrm{L}_p(\mathbb{R}^d)$, $1 \leq p < \infty$.

Proof. For simplicity, we shall assume $\lambda \in \mathbb{R}$. In the general case we may replace λ by $\mathrm{Re}\,\lambda$ and carry on with the same proofs.

(a) The image of $\mathscr{D}(\mathbb{R}^d)$ lies in $\mathrm{L}_p(\mathbb{R}^d)$. We have,

$$\mathrm{U}_{c,n}^{-\lambda}\phi(x) = \rho_c^{\lambda-d} * \phi(x) - \sum_{|k|\le n} \frac{\partial_k \rho_c^{\lambda-d}(x)}{k!} \int_{\mathbb{R}^d} (-y)^k \phi(y)\, \mathrm{d}y, \qquad (2.28)$$

which is infinitely smooth (and therefore locally bounded) in $\dot{\mathbb{R}}^d = \mathbb{R}^d\backslash\{0\}$ (in all of $\mathbb{R}^d$ if $n < 0$), and possibly singular at the origin only if $n \ge 0$, but even then dominated by a multiple of $|x|^{\lambda-d-n}$ in a punctured neighbourhood of 0. Note that

$$\lambda - d - n = \lambda + \tfrac{d}{p} - \lfloor \lambda + \tfrac{d}{p} \rfloor - \tfrac{d}{p} = \{\lambda + \tfrac{d}{p}\} - \tfrac{d}{p}$$
$$\Rightarrow \quad -\tfrac{d}{p} < \lambda - d - n < 1 - \tfrac{d}{p}$$

where $\{\lambda + \frac{d}{p}\} \in (0,1)$ denotes the fractional part of $\lambda + \frac{d}{p}$. This shows that for $n = \lfloor \lambda + \frac{d}{p} \rfloor - d$, the derivatives of $\rho_c^{\lambda-d}$ appearing in the sum—and consequently $\mathrm{U}_{c,n}^{-\lambda}\phi$ itself—are ordinary functions.

Therefore, $\mathrm{U}_{c,n}^{-\lambda}\phi \in \mathrm{L}_{p,\mathrm{loc}}$. We next prove that $\mathrm{U}_{c,n}^{-\lambda}\phi$ decays faster than $|x|^{-\frac{d}{p}}$ as $|x| \to \infty$. Together, these facts will establish that $\mathrm{U}_{c,n}^{-\lambda}\phi \in \mathrm{L}_p$.

To prove a minimal rate of decay for $\mathrm{U}_{c,n}^{-\lambda}\phi$ at infinity, take the function $A(x) := |x|^\alpha |\mathrm{U}_{c,n}^{-\lambda}\phi(x)|$. We shall show that it vanishes as $|x| \to \infty$ for some $\alpha > \frac{d}{p}$. We have,

$$A(x) = \left| \int_{\mathbb{R}^d} \phi(y) \left[\rho_c^{\lambda-d}(x-y) - \sum_{|k|\le n} (-y)^k \frac{\partial_k \rho_c^{\lambda-d}(x)}{k!} \right] \mathrm{d}y \right| |x|^\alpha \qquad (2.29)$$

We consider two cases, depending on the sign of n. For $n < 0$, where the sum over n is null, we have $n = \lfloor \lambda - d + \frac{d}{p} \rfloor < 0$ and therefore $\lambda - d < -\frac{d}{p}$. Thus, in this case, the integral inside $A(x)$ corresponds to the convolution of a rapidly decaying function with an ordinary function with faster than $|x|^{-\frac{d}{p}}$ decay. $A(x)$ as a whole therefore vanishes at ∞ for some $\alpha > \frac{d}{p}$.

In the case of $n \ge 0$, assume $x \ne 0$ and let $h := |x|^{-1}$, $\theta := hx = x/|x|$. Also take $K > 0$ sufficiently large such that $|y| > K \;\Rightarrow\; \phi(y) = 0$ ($\phi \in \mathscr{D}(\mathbb{R}^d)$ is compactly supported). Then,

$$A(x) = A(h^{-1}\theta)$$
$$= \left| \int_{|y|\le K} \phi(y) \left[\rho_c^{\lambda-d}(\theta - hy) - \sum_{|k|\le n} (-hy)^k \frac{\partial_k \rho_c^{\lambda-d}(\theta)}{k!} \right] \mathrm{d}y \right| h^{-\alpha-\lambda+d},$$

using the homogeneity of $\rho_c^{\lambda-d}$ and its derivatives.

The sum inside the brackets above corresponds to the degree n Taylor series expansion of $\rho_c^{\lambda-d}$ about θ. For $h < \frac{1}{2}K^{-1}$ (equiv. for $|x| > 2K$) the difference inside the brackets may therefore be bounded by a term like

$$\sum_{|k|=n+1} R_k \left|(-hy)^k\right| = h^{n+1} \sum_{|k|=n+1} R_k \left|y^k\right|$$

with $R_k = \sup_{|z|\le\frac{1}{2}} |\frac{\partial_k \rho_c^{\lambda-d}(\theta-z)}{k!}| < \infty$. Consequently, $A(x)$ may be bounded as

$$0 \le A(x) \le h^{-\alpha-\lambda+d+n+1} \int_{|y|\le K} \sum_{|k|=n+1} \left|\phi(y) y^k\right| \, \mathrm{d}y =: A_\phi h^{-\alpha-\lambda+d+n+1}$$

(2.30)

with the constant A_ϕ defined as the integral in the middle expression.

This proves that as $h^{-1} = |x| \to \infty$,

$$|x|^\alpha |\mathrm{U}_{c,n}^{-\lambda}\phi(x)| \to 0$$

for *any*

$$\alpha < -\lambda + d + n + 1 = -\lambda + \lfloor \lambda + \tfrac{d}{p} \rfloor + 1 - \tfrac{d}{p} + \tfrac{d}{p} = 1 - \{\lambda + \tfrac{d}{p}\} + \tfrac{d}{p}$$

and, in particular, for *some* $\alpha > \frac{d}{p}$ (recall that $0 < \{\lambda + \frac{d}{p}\} < 1$ by assumption). This proves that the tail of $\mathrm{U}_{c,n}^{-\lambda}\phi(x)$ decays faster than $|x|^{-\frac{d}{p}}$ also for $n \ge 0$.

Thus, in all cases, $\mathrm{U}_{c,n}^{-\lambda}\phi$ decays faster than $|x|^{-\frac{d}{p}}$ at ∞ and is also locally p-integrable, as we already showed. We therefore have $\mathrm{U}_{c,n}^{-\lambda}\phi \in \mathrm{L}_p(\mathbb{R}^d)$ for all $\phi \in \mathscr{D}(\mathbb{R}^d)$.

(b) $\mathrm{U}_{c,n}^{-\lambda}$ *is continuous* $\mathscr{D}(\mathbb{R}^d) \to \mathrm{L}_p(\mathbb{R}^d)$. Take any sequence $\phi_l \to \phi$ in $\mathscr{D}$.

First, observe that for any $x \neq 0$,

$$u_x(y) := \rho_c^{\lambda-d}(x-y) - \sum_{|k|\le n} (-y)^k \frac{\partial_k \rho_c^{\lambda-d}(x)}{k!}$$

is a distribution in $\mathscr{D}'(\mathbb{R}^d)$ (in fact in $\mathscr{S}'(\mathbb{R}^d)$) and we may write

$$\mathrm{U}_{c,n}^{-\lambda}\phi_l(x) = \langle \phi_l, u_x \rangle.$$

This proves that $\mathrm{U}_{c,n}^{-\lambda}\phi_l(x) \to \mathrm{U}_{c,n}^{-\lambda}\phi_l(x)$ point-wise in $\mathbb{R}^d \backslash \{0\}$.

Secondly, since ϕ_l converges in $\mathscr{D}(\mathbb{R}^d)$, there exists $K > 0$ such that all $\phi_l(y)$s vanish outside $|y| \leq K$. From the discussion leading to (2.30) it then follows that from some l onwards, $\mathrm{U}_{c,n}^{-\lambda}\phi_l(x)$ is bounded by a term like $A_1|x|^{-\frac{d}{p}-\epsilon_1}$ for some $\epsilon_1 \in (0,1)$ in the region $|x| > 2K$ (A_1 is a constant that eventually dominates the A_{ϕ_l}s of (2.30)). One similarly shows that for $|x| \leq 2K$, $\mathrm{U}_{c,n}^{-\lambda}\phi_l(x)$ is eventually (in l) bounded by some $A_2|x|^{-\frac{d}{p}+\epsilon_2}$ with $\epsilon_2 \in (0,1)$. Thus, all in all, the sequence $\{\mathrm{U}_{c,n}^{-\lambda}\phi_l\}_l$ is eventually bounded by the function

$$A_1|x|^{-\frac{d}{p}-\epsilon_1}\mathbb{1}_{|x|>2K}(x) + A_2|x|^{-\frac{d}{p}+\epsilon_2}\mathbb{1}_{|x|\leq 2K}(x),$$

which is p-integrable. From there, by Lebesgue's dominated convergence theorem, we have $\lim_{l\to\infty}\|\mathrm{U}_{c,n}^{-\lambda}\phi_l\|_p = \|\lim_{l\to\infty}\mathrm{U}_{c,n}^{-\lambda}\phi_l\|_p = \|\mathrm{U}_{c,n}^{-\lambda}\phi\|_p$, proving L_p convergence. ∎

2.ar Remark. For $\operatorname{Re}\lambda < d - \frac{d}{p}$ where in the above theorem we have $n < 0$, the L_p-continuous operator introduced above corresponds to a convolution and is translation-invariant, but clearly not otherwise. In the former case, these translation-invariant L_p-bounded operators are proportional to Riesz potentials and derivatives discussed (for instance) in Stein [Ste70] and Mikhlin and Prössdorf [MP86].

2.as Remark. In the case of $p = 2$, one-dimensional versions of the above operators were introduced by Blu and Unser [BU07] without specific mention of their continuity. We generalized their definition to multiple dimensions in Tafti *& al.* [TVU09, TU10b]. A further extension for $p \neq 2$, comparable to the one described here, appears in Sun and Unser [SU]. In all of the cited works, the operator is defined in terms of a singular Fourier domain integral, whose correct interpretation for $p \neq 2$ is not always straightforward. In contrast, here we have given the primary definition of the operator and proved its continuity in the spatial domain (we have also generalized the theory to complex λ since the approach permitted it).

2.at Remark. Using (2.28) and Lemma **2.aj**, we find the adjoint of $\mathrm{U}_{c,n}^{-\lambda}$ on $\mathscr{D}(\mathbb{R}^d)$ to be given by the expression

$$\mathrm{U}_{c,n}^{-\lambda *}\phi(x) = \mathrm{U}_c^{-\lambda}\phi(x) - \sum_{|k|\leq n} x^k \frac{\partial_k \mathrm{U}_c^{\lambda}\phi(0)}{k!} = \mathrm{U}_c^{-\lambda}\phi(x) - T_{\mathrm{U}_c^{-\lambda}\phi,n}(x).$$

Here $T_{\mathrm{U}_c^{-\lambda}\phi,n}$ denotes the polynomial corresponding to the nth degree Taylor expansion of $\mathrm{U}_c^{-\lambda}\phi$ at $x = 0$. We therefore see that $\mathrm{U}_{c,n}^{-\lambda *}\phi$ and its derivatives up to order n vanish at the origin.

By the previous theorem, the above adjoint extends to a continuous operator $\mathrm{L}_{p'}(\mathbb{R}^d) \to \mathscr{D}'(\mathbb{R}^d)$ with n as above and $p' = p/(p-1)$.

§2.3 Rotation-invariant homogeneous operators: Vector case

We shall now develop the vector field parallel of the previous section. In this case our spaces of test functions are $\mathscr{D}^d(\mathbb{R}^d)$ and $\mathscr{S}^d(\mathbb{R}^d)$ with duals $(\mathscr{D}^d)'(\mathbb{R}^d) = (\mathscr{D}')^d(\mathbb{R}^d)$ and $(\mathscr{S}^d)'(\mathbb{R}^d) = (\mathscr{S}')^d(\mathbb{R}^d)$. $\mathscr{D}^d(\dot{\mathbb{R}}^d)$, $\widehat{\mathscr{D}}^d(\dot{\mathbb{R}}^d)$, and other varieties of $\dot{\mathbb{R}}^d$ spaces are defined in similar fashion to **2.o**, as the dth Cartesian power of the correponding spaces introduced there. As in the previous section, we shall use the symbol $\mathscr{E}$ to refer either to $\mathscr{D}$ or $\mathscr{S}$. For the purpose of matrix multiplication, a vector (generalized) function $f = (f_1, \dots, f_d)$ is understood to be a column vector (the same goes for variables such as $x = (x_1, \dots, x_d)$, test functions $\phi = (\phi_1, \dots, \phi_d)$, and all other vectors).

In what concerns us, the distinction between scalar and vector theories is primarily due to the different realization of rotations in the vector setting (see (G3) of **2.n**), plus the fact that the integration kernels of (2.5) and (2.6) become matrices in $(\mathscr{E}')^{d\times d}(\mathbb{R}^d \times \mathbb{R}^d)$ and $(\mathscr{E}')^{d\times d}(\mathbb{R}^d)$ respectively. We recall the law of rotation of a vector field by an orthogonal matrix $\omega \in O(d)$ from **2.n**:

$$\mathrm{R}_{\omega,v} : f \mapsto \omega f(\omega^{\mathrm{T}}\cdot).$$

One easily verifies that a translation-invariant vector field operator is rotation-invariant in the vector sense iff its convolution kernel K_{U} fulfils the condition

$$K_{\mathrm{U}}(\omega\cdot) = \omega K_{\mathrm{U}}(\cdot)\omega^{\mathrm{T}} \tag{2.31}$$

for any $\omega \in O(d)$ (or, equivalently, if the Fourier transform of K_{U} satisfies a similar condition). Hence, in the vector setting our task is to identify homogeneous distributions in $(\mathscr{E}')^{d\times d}(\mathbb{R}^d)$ that fulfil (2.31).

§2.3.1 Homogeneous and rotation-invariant vector distributions

2.au Notation. In this section, we shall introduce a number of different parametrizations of a family of matrix-valued distributions in $(\mathscr{E}')^{d\times d}(\mathbb{R}^d)$ that are homogeneous and rotation-invariant as per (2.31). Members of this family we shall represent as $P^{\lambda}_{\underline{\bullet}}$, with λ denoting the homogeneity order and $\underline{\bullet}$ serving as a place-holder for any one of the following three parametrizations:

$$\underline{s} = (s_1, s_2), \quad \underline{r} = (r_1, r_2), \quad \text{or} \quad \underline{k} = (k_1, k_2)$$

(the underline is there to emphasize that $\underline{s}, \underline{r}, \underline{k}$ denote a pair of scalar parameters and not a single parameter).

Thus, in writing $P^\lambda_{\underline{s}}$, $P^\lambda_{\underline{r}}$, or $P^\lambda_{\underline{k}}$ to denote an element of the family, the respective symbols $\underline{s}$, $\underline{r}$, and $\underline{k}$ are an integral part of the notation that distinguish the different parametrizations. Accordingly, when assigning numerical values to the parameters we shall write, e.g., $P^\lambda_{\underline{r}}$ or $P^\lambda_{(r_1,r_2)}$ and indicate that $(r_1, r_2) = (1, 1)$, rather than refer to an ambiguous $P^\lambda_{(1,1)}$. We shall, however, use variations of the notation involving the same letter to denote instances of the same parametrization but with a different value for the parameters; thus, for instance, we may write $P^\lambda_{\underline{k}}$, $P^\lambda_{\underline{k}'}$, or $P^\lambda_{\widehat{\underline{k}}}$, all to refer to the third (k) parametrization but with potentially different values for the parameters $\underline{k}$, $\underline{k}'$, and $\widehat{\underline{k}}$.

The unfortunate and superficial complexity that is associated with the use of several parametrizations instead of a single one is partly balanced by the fact that a variety of important properties of the family are simple to formulate in one parametrization but cumbersome in another.

2.av Definitions and conversion formulae. We define $P^\lambda_{\underline{s}}$, $P^\lambda_{\underline{r}}$, $P^\lambda_{\underline{k}}$ below as matrices in $(\mathscr{E}')^{d\times d}$ by indicating the form of their ijth entry, $1 \le i, j \le d$. We then provide the formulae that relate each parametrization to the others.

For $\underline{s} = (s_1, s_2) \in \mathbb{C}^2$ and $\lambda \in \mathbb{C}$,

$$\begin{aligned}
[P^\lambda_{\underline{s}}]_{ij} &:= \partial_{ij}|x|^2\rho^\lambda_{s_1} + x_i x_j\,\Delta\,\rho^\lambda_{s_2} \\
&= -\,\widehat{\lambda}\partial_{ij}\rho^{\lambda+2}_{s_1} + \lambda x_i x_j \rho^{\lambda-2}_{s_2} \\
&= -\,\widehat{\lambda}\partial_{ij}\rho^{\lambda+2}_{s_1} - \lambda\left(\partial_{ij}\rho^{\widehat{\lambda+2}}_{s_2}\right)^\wedge \\
&= -\,\widehat{\lambda}(\lambda+2)\partial_{ij}\,\Delta^{-1}\,\rho^\lambda_{s_1} - \lambda(\widehat{\lambda}+2)\frac{x_i x_j}{|x|^2}\rho^\lambda_{s_2}
\end{aligned} \tag{2.32}$$

(the equivalence of the different formulae is due to **2.af**). $P^\lambda_{\underline{s}}$ fulfils the requirements of λ-homogeneity and rotation-invariance as per (2.31) (that the s_2 term fulfils (2.31) is evident; for the s_1 term it becomes clear in the Fourier domain). We also define:

$$\begin{aligned}
[P^\lambda_{\underline{r}}]_{ij} &:= \partial_{ij}\,\Delta^{-1}\,\rho^\lambda_{r_1} + \left(\delta_{ij} - \partial_{ij}\,\Delta^{-1}\right)\rho^\lambda_{r_2} \\
&= \delta_{ij}\rho^\lambda_{r_2} + \frac{1}{\lambda+2}\partial_{ij}\rho^{\lambda+2}_{r_1-r_2};
\end{aligned} \tag{2.33}$$

$$\begin{aligned}
[P^\lambda_{\underline{k}}]_{ij} &:= \frac{x_i x_j}{|x|^2}\rho^\lambda_{k_1} + \left(\delta_{ij} - \frac{x_i x_j}{|x|^2}\right)\rho^\lambda_{k_2} \\
&= \delta_{ij}\rho^\lambda_{k_2} - \frac{1}{\widehat{\lambda}+2}x_i x_j \rho^{\lambda-2}_{k_1-k_2}.
\end{aligned} \tag{2.34}$$

Using the relations given in Paragraph **2.af**, it is seen that we have the equality $P^\lambda_{\underline{s}} = P^\lambda_{\underline{r}} = P^\lambda_{\underline{k}}$ when the parameters $\underline{s}$, $\underline{r}$, $\underline{k}$, and λ are related by the following

formulae:

$$\begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \begin{pmatrix} -\widehat{\lambda}(\lambda+2) & (\widehat{\lambda}+1)(\widehat{\lambda}+2) \\ 0 & \widehat{\lambda}+2 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = \frac{1}{\widehat{\lambda}} \begin{pmatrix} -\lambda-1 & 1-d \\ -1 & \widehat{\lambda}+1 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \end{pmatrix};$$
$$\begin{pmatrix} k_1 \\ k_2 \end{pmatrix} = \begin{pmatrix} (\lambda+1)(\lambda+2) & -\lambda(\widehat{\lambda}+2) \\ \lambda+2 & 0 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = \frac{1}{\lambda} \begin{pmatrix} -\widehat{\lambda}-1 & 1-d \\ -1 & \lambda+1 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}.$$

Therefore, $P^\lambda_{\underline{r}}$ and $P^\lambda_{\underline{k}}$ represent alternative parametrizations of the same family, and for a given $\underline{r}$ or $\underline{k}$, the respective distributions enjoy the same invariances as previously noted for $P^\lambda_{\underline{s}}$.

For completeness, we also state the following converse result concerning invariant distributions in the vector setting (this result is the vector counterpart of Theorem **2.ab**).

2.aw Theorem. Distributions in $(\mathscr{E}')^{d\times d}(\mathbb{R}^d)$ that fulfil (2.31) and are homogeneous of order λ are of the form $P^\lambda_{\underline{s}}$ for some $\underline{s} = (s_1, s_2) \in \mathbb{C}^2$.

Proof. Let us denote a candidate distribution by F.

First, observe that (2.31) implies

$$F(-\cdot) = (-\mathrm{I})\,F(\cdot)(-\mathrm{I}) = F(\cdot). \tag{2.35}$$

Next, fix $x \neq 0$ where F is regular and let $x_i^\perp$, $1 \le i \le d-1$, be $d-1$ pairwise orthogonal vectors in $\mathbb{R}^d$ all perpendicular to x and with $|x_i^\perp| = |x|$. We define the rotation matrices

$$\omega_i = \mathrm{I} - 2\frac{xx^{\mathrm{T}}}{|x|^2} - 2\frac{x_i^\perp (x_i^\perp)^{\mathrm{T}}}{|x|^2}, \qquad 1 \le i \le d-1.$$

Each ω_i is a simple rotation by π in the $x \wedge x_i^\perp$ plane. In particular, $\omega_i x = -x$. We also define, for $i \neq j$, the $\frac{\pi}{2}$ rotation matrix

$$\omega_{ij} = \mathrm{I} - \frac{x_i^\perp (x_i^\perp)^{\mathrm{T}}}{|x|^2} - \frac{x_j^\perp (x_j^\perp)^{\mathrm{T}}}{|x|^2} + \frac{x_j^\perp (x_i^\perp)^{\mathrm{T}}}{|x|^2} - \frac{x_i^\perp (x_j^\perp)^{\mathrm{T}}}{|x|^2}.$$

ω_{ij} maps

$$x_i^\perp \mapsto x_j^\perp \mapsto -x_i^\perp$$

and leaves x fixed (in this proof only, ω_{ij} denotes a whole matrix and *not* the ijth element of ω; similarly, only for the purpose of this proof, $x_i^\perp$ denotes a complete vector and *not* the ith element of some unknown vector $x^\perp$).

Note that the matrices ω_i, $1 \le i \le d-1$, commute pairwise; also, by (2.35) and (2.31),

$$F(x)\omega_i = F(-x)\omega_i = F(\omega_i x)\omega_i = \omega_i F(x),$$

which shows that the ω_is commute with $F(x)$ as well. In other words, $\omega_1, \ldots, \omega_{d-1}$ and $F(x)$ together form a system of commuting matrices. Since, for $d > 2$, the vectors $x, x_1^\perp, \ldots, x_{d-1}^\perp$ are precisely the common eigenvectors of $\omega_1, \ldots, \omega_{d-1}$, they must be eigenvectors of $F(x)$ as well, in particular x is an eigenvector of $F(x)$. Denote its corresponding eigenvalue by $\mu_1 = \mu_1(x)$. By taking the transpose of (2.31) and applying the same argument, we can show that x is also an eigenvector of $F(x)^{\mathrm{T}}$. Its corresponding eigenvalue, temporarily denoted by $\mu_1' = \mu_1'(x)$, must be the same as $\mu_1(x)$ since

$$\mu_1' x^{\mathrm{T}} x = x^{\mathrm{T}} F(x) x = \mu_1 x^{\mathrm{T}} x.$$

We similarly denote the eigenvalue corresponding to $x_i^\perp$ by $\mu_{2,i}(x)$.

Alternatively, to find the eigenvectors of $F(x)$ we might note that $F(x)$ commutes with all ω_{ij}s, $i \neq j$:

$$F(x)\omega_{ij} = F(\omega_{ij}x)\omega_{ij} = \omega_{ij}F(x),$$

and since x is an eigenvector of all ω_{ij}s with eigenvalue 1 (it is their only common eigenvector), $F(x)x$ must be a common eigenvector of all ω_{ij}s, thus $F(x)x = \mu_1(x)x$ for some scalar eigenvalue $\mu_1(x)$. Then, to show that the $x_i^\perp$s are also eigenvectors of $F(x)$, we observe that

$$\begin{aligned} F(x)x_i^\perp = F(-x)x_i^\perp &= F(\omega_i x)x_i^\perp \\ &= \omega_i F(x)\omega_i^{-1}x_i^\perp = \omega_i F(x)(-x_i^\perp) = -\omega_i F(x)x_i^\perp; \end{aligned}$$

whereby,

$$(\mathrm{I} + \omega_i)F(x)x_i^\perp = 0.$$

This shows that $F(x)x_i^\perp$ lies in the kernel of $\mathrm{I} + \omega_i$. But the kernel of $\mathrm{I} + \omega_i$ corresponds exactly to the span of $\{x, x_i^\perp\}$. We can therefore write

$$F(x)x_i^\perp = \alpha_i x + \mu_{2,i}x_i^\perp$$

for some $\alpha_i, \mu_{2,i}$. But then

$$x^{\mathrm{T}}F(x)x_i^\perp = \alpha_i |x|^2;$$

we also have

$$x^{\mathrm{T}}F(x)x_i^\perp = \mu_1 x^{\mathrm{T}}x_i^\perp = 0.$$

The last two equations show that $\alpha_i = 0$, that is, we have $F(x)x_i^\perp = \mu_{2,i}(x)x_i^\perp$. $F(x)$ therefore has $x, x_1^\perp, \ldots, x_{d-1}^\perp$ as d eigenvectors with respective eigenvalues $\mu_1(x), \mu_{2,1}(x), \ldots, \mu_{2,d-1}(x)$.

Next, we show that all $\mu_{2,i}$s are equal to some $\mu_2 = \mu_2(x)$: by (2.31),

$$\begin{aligned}\mu_{2,j}x_j^\perp = F(x)x_j^\perp = F(\omega_{ij}x)x_j^\perp &= \omega_{ij}F(x)\omega_{ij}^{-1}x_j^\perp \\ &= \omega_{ij}F(x)x_i^\perp = \mu_{2,i}\omega_{ij}x_i^\perp = \mu_{2,i}x_j^\perp,\end{aligned}$$

proving that all $\mu_{2,i}(x)$s are equal as claimed. Putting everything together, we find that $F(x)$ has the orthogonal eigenvectors $x, x_1^\perp, \ldots, x_{d-1}^\perp$ with eigenvalue $\mu_1(x)$ for x and eigenvalue $\mu_2(x)$ for the remaining vectors. Its eigen-decomposition is therefore of the form

$$F(x) = xx^{\mathrm{T}}\mu_1(x) + \left(|x|^2\mathrm{I} - xx^{\mathrm{T}}\right)\mu_2(x)$$

(for $d = 2$ we can make a similar demonstration of the above decomposition by working with the reflection matrix with axis x instead of the ω_{ij}s).

Finally, note that the scalar distribution $|x|^4\mu_1(x) = x^{\mathrm{T}}F(x)x$ is by assumption homogeneous of order $\lambda+2$ and by (2.31) rotation-invariant in the scalar sense. This implies that $\mu_1(x)$ is rotation-invariant and homogeneous of order $\lambda-2$, which in turn implies the same about $\mu_2(x)$ in view of $F(x)$'s being λ-homogeneous and satisfying (2.31). Using Theorem **2.ab** and (2.19) we thus find that $\mu_1(x) = |x|^{-2}\rho^\lambda_{k_1}(x)$ and $\mu_2(x) = |x|^{-2}\rho^\lambda_{k_2}(x)$ for some arbitrary, yielding the form of $P^\lambda_{\underline{k}}$ for the solution, which we can rewrite as $P^\lambda_{\underline{s}}$ using **2.af**, thus completing the proof. ∎

2.ax Remark. A version of the above theorem for dimensions $d = 2, 3$ and using the $P^\lambda_{\underline{k}}$ parameterization can be found in Arigovindan [Ari05].

2.ay The Fourier transform of $P^\lambda_{\underline{\bullet}}$. As in the scalar case, using the properties of the Fourier transform it is easy to verify that the Fourier transform of $P^\lambda_{\underline{\bullet}}$ is $\widehat{\lambda}$-homogeneous and satisfies the same rotation formula as $P^\lambda_{\underline{\bullet}}$, i.e. (2.31). We therefore expect the Fourier transform of $P^\lambda_{\underline{\bullet}}$ to be of the form $P^{\widehat{\lambda}}_{\underline{\widehat{\bullet}}}$, where $\underline{\widehat{\bullet}}$ depends on $\underline{\bullet} = \underline{s}, \underline{r}, \underline{k}$ and possibly on λ as well. This is also evident directly from the definition of $P^\lambda_{\underline{s}}$ given in (2.32). Specifically, using (2.32) and the formulae derived previously we obtain the

2.az Lemma. The parametric family $\{P^\lambda_{(s_1,s_2)}\}$ is closed under the Fourier transform with

$$(P^\lambda_{\underline{s}})^\wedge = P^{\widehat{\lambda}}_{\underline{\widehat{s}}},$$

where

$$\widehat{\lambda} := -d - \lambda \quad \text{and} \quad \underline{\widehat{s}} := (\widehat{s}_1, \widehat{s}_2) = (s_2, s_1).$$

This in turn shows that for the alternative parametrizations $P^\lambda_{(k_1,k_2)}$ and $P^\lambda_{(r_1,r_2)}$ we similarly have $P^{\widehat{\lambda}}_{\underline{\widehat{r}}} = (P^\lambda_{\underline{r}})^\wedge$ and $P^{\widehat{\lambda}}_{\underline{\widehat{k}}} = (P^\lambda_{\underline{k}})^\wedge$, with, respectively,

$$\underline{\widehat{r}} := \begin{pmatrix} \widehat{r}_1 \\ \widehat{r}_2 \end{pmatrix} = \frac{1}{\widehat{\lambda}} \begin{pmatrix} -\lambda - 1 & 1-d \\ -1 & \widehat{\lambda}+1 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} \tag{2.36}$$

and

$$\underline{\widehat{k}} := \begin{pmatrix} \widehat{k}_1 \\ \widehat{k}_2 \end{pmatrix} = \frac{1}{\lambda} \begin{pmatrix} -\widehat{\lambda} - 1 & 1-d \\ -1 & \lambda+1 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \end{pmatrix}. \tag{2.37}$$

2.ba Products and convolutions. Using previous results about the derivatives of ρ_c^λ in combination with **2.ag**, we find the identities

$$P^\lambda_{\underline{k}} P^{\lambda'}_{\underline{k}'} = P^{\lambda+\lambda'}_{\underline{k}''} \quad \text{with} \quad \underline{k}'' = (k_1'', k_2'') = \frac{\Gamma(-\frac{\widehat{\lambda+\lambda'}}{2})}{\Gamma(-\frac{\widehat{\lambda}}{2})\Gamma(-\frac{\widehat{\lambda'}}{2})} (k_1 k_1', k_2 k_2')$$

and

$$P^{\widehat{\lambda}}_{\underline{\widehat{r}}} * P^{\widehat{\lambda'}}_{\underline{\widehat{r}}'} = P^{\widehat{\lambda+\lambda'}}_{\underline{\widehat{r}}''} \quad \text{with} \quad \underline{\widehat{r}}'' = (\widehat{r''}_1, \widehat{r''}_2) = \frac{\Gamma(-\frac{\widehat{\lambda+\lambda'}}{2})}{\Gamma(-\frac{\widehat{\lambda}}{2})\Gamma(-\frac{\widehat{\lambda'}}{2})} (\widehat{r}_1 \widehat{r'}_1, \widehat{r}_2 \widehat{r'}_2),$$

valid on $\widehat{\mathscr{E}}^{d\times d}(\dot{\mathbb{R}}^d)$ (convolution and multiplication of distributions are conducted per the matrix product order, namely, $[A*B]_{ij} = \sum_k [A]_{ik} * [B]_{kj}$ and similarly for the product). The second result can be seen as a consequence of the first and the convolution-multiplication exchange formula.

As in **2.ag**, these results extend to $\mathscr{E}^{d\times d}(\mathbb{R}^d)$ at least for $\operatorname{Re}\lambda \geq 0$ and $\operatorname{Re}\lambda' \geq -\operatorname{Re}\lambda$ (with the suitable order of application); notably, we have

$$P^\lambda_{(k_1^{-1},k_2^{-1})} P^{-\lambda}_{(k_1,k_2)} = H_\lambda \mathrm{I}$$

and

$$P^{\widehat{\lambda}}_{(\widehat{r}_1^{-1},\widehat{r}_2^{-1})} * P^{\widehat{-\lambda}}_{(\widehat{r}_1,\widehat{r}_2)} = H_\lambda \mathrm{I}\, \delta \tag{2.38}$$

with I denoting the $d\times d$ identity matrix and $H_\lambda = 1/\Gamma(\frac{d+\lambda}{2})\Gamma(\frac{d-\lambda}{2})$ denoting the constant defined in (2.24).

2.bb Corollary of Theorem 2.aw. Continuous linear operators $\mathscr{E}^d(\mathbb{R}^d) \to (\mathscr{E}')^d(\mathbb{R}^d)$ that are simultaneously translation- and vector-rotation-invariant, and ho-

mogeneous of order $-\lambda \in \mathbb{C}$ in the sense of (2.4), are of the form [6]

$$\underline{\mathrm{U}}^{\lambda}_{\bullet} : \phi \;\mapsto\; \widehat{P^{\lambda}_{\bullet}} * \phi = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\langle \cdot, \xi \rangle}\; P^{\lambda}_{\bullet}(\xi)\; \widehat{\phi}(\xi)\; \mathrm{d}\xi,$$

where $\bullet$ stands for any of the three different parametrizations by $\underline{s}$, $\underline{r}$, or $\underline{k}$ as for the corresponding distributions.

2.bc For later reference, we give a different representation of the operator identified in the above corollary. Recall that in **2.ac** we defined the scalar operator U^{λ}_{c} with $c \in \mathbb{C}$. We shall denote its coordinate-wise application to a vector field $\phi \in \mathscr{E}^d$ as the operator $\underline{\mathrm{U}}^{\lambda}_{c}$, meaning that we have

$$\left[\underline{\mathrm{U}}^{\lambda}_{c} \phi\right]_i := \mathrm{U}^{\lambda}_{c} \phi_i.$$

We define a new operator

$$\mathrm{GD} : \phi \mapsto \mathrm{Grad}\,\mathrm{Div}\,\phi = \left[\sum_j \partial_{ij} \phi_j\right]$$

and note that it is associated with the Fourier multiplier matrix $[-\xi_i \xi_j]$. Using (2.33), we may then rewrite $\underline{\mathrm{U}}^{\lambda}_{(r_1, r_2)}$ as

$$\begin{aligned} \underline{\mathrm{U}}^{\lambda}_{\underline{r}} &= \mathrm{U}^{\lambda}_{r_2} + \frac{1}{\widehat{\lambda} + 2} \mathrm{U}^{\lambda-2}_{r_1 - r_2}\, \mathrm{GD} \\ &= \mathrm{U}^{\lambda}_{r_2} + \frac{1}{\widehat{\lambda} + 2}\, \mathrm{GD}\, \mathrm{U}^{\lambda-2}_{r_1 - r_2}. \end{aligned} \tag{2.39}$$

§2.3.2 L_p-continuous homogeneous and rotation-invariant operators

We proved in **§2.2.2** that the operator $\mathrm{U}^{-\lambda}_{c,n}$ with $n = \lfloor \mathrm{Re}\,\lambda + \frac{d}{p} \rfloor - d$ and $\mathrm{Re}\,\lambda + \frac{d}{p} \notin \mathbb{N}$ exhibits the same form of rotation and scale invariance as $\mathrm{U}^{-\lambda}_{c}$, and, also similar to $\mathrm{U}^{-\lambda}_{c}$, it is proportional to a left inverse of U^{λ}_{c} for $\mathrm{Re}\,\lambda > 0$, $\lambda \neq d, d+2, \ldots$. The distinguishing property of $\mathrm{U}^{-\lambda}_{c,n}$, captured in Theorem **2.aq**, is that it maps $\mathscr{D}(\mathbb{R}^d)$ continuously into $\mathrm{L}_p(\mathbb{R}^d)$. In this section we shall propose a similar modification of the vector operator $\underline{\mathrm{U}}^{-\lambda}_{\bullet}$ which shares its homogeneity and vectorial rotation-invariance properties and is at the same time continuous $\mathscr{D}^d(\mathbb{R}^d) \to \mathrm{L}^d_p(\mathbb{R}^d)$.

6. We use an underline visually to distinguish the new vector-to-vector operators from similar-looking scalar-to-scalar ones introduced previously (cf. $\underline{\mathrm{U}}^{\lambda}_{\underline{r}}$ vs U^{λ}_{c} and, later, $\underline{\mathrm{Reg}}^{\lambda}_{\underline{r},n}$ vs $\mathrm{Reg}^{\lambda}_{c,n}$).

2.bd We shall make use of Theorem **2.aq** as follows. By the said theorem, the vector to vector operator

$$\underline{\mathrm{U}}_{c,n}^{-\lambda} : [\phi_i] \mapsto \left[\mathrm{U}_{c,n}^{-\lambda}\phi_i\right],$$

$c \in \mathbb{C}$, which corresponds to the coordinate-wise application of the scalar operator $\mathrm{U}_{c,n}^{-\lambda}$, is continuous $\mathscr{D}^d(\mathbb{R}^d) \to \mathrm{L}_p^d(\mathbb{R}^d)$, $1 \le p < \infty$, for $n = \lfloor \mathrm{Re}\,\lambda + \frac{d}{p} \rfloor - d$. Furthermore, we have the

2.be **Lemma.** $\underline{\mathrm{U}}_{c,n}^{-\lambda}$ is $(-\lambda)$-homogeneous and and vector-rotation-invariant.

Proof. By Corollary **2.am**, $\mathrm{U}_{c,n}^{-\lambda}$ is $(-\lambda)$-homogeneous and scalar-rotation-invariant. The homogeneity of $\underline{\mathrm{U}}_{c,n}^{-\lambda}$ immediately follows. To prove vector-rotation-invariance, take an arbitrary orthogonal matrix $\omega = [\omega_{ij}]$; we have

$$\begin{aligned}
\left[\underline{\mathrm{U}}_{c,n}^{-\lambda}\mathrm{R}_{\omega,v}\phi\right]_i &= \mathrm{U}_{c,n}^{-\lambda}\Big(\sum_j \omega_{ji}\mathrm{R}_{\omega,s}\phi_j\Big) \\
&= \sum_j \omega_{ji}\mathrm{U}_{c,n}^{-\lambda}\mathrm{R}_{\omega,s}\phi_j \qquad \text{by linearity,} \\
&= \sum_j \omega_{ji}\mathrm{R}_{\omega,s}\mathrm{U}_{c,n}^{-\lambda}\phi_j \qquad \text{by Corollary 2.am,} \\
&= \left[\mathrm{R}_{\omega,v}\underline{\mathrm{U}}_{c,n}^{-\lambda}\phi\right]_i \qquad \text{by the definition of } \mathrm{R}_{\omega,v} \text{ (2.n).}
\end{aligned}$$

∎

Note that the operator $\mathrm{GD} = \mathrm{Grad}\,\mathrm{Div}$ with Fourier matrix $[-\xi_i\xi_j]$ (defined in **2.bc**) is continuous on $\mathscr{E}^d(\mathbb{R}^d)$ (derivatives are continuous on $\mathscr{D}(\mathbb{R}^d)$ and $\mathscr{S}(\mathbb{R}^d)$). In addition, GD is clearly both homogeneous of order 2 in the sense of (2.4) and rotation-invariant in the vector sense. Putting these facts together with the previous lemma, we obtain the

2.bf **Theorem.** The operator

$$\underline{\mathrm{U}}_{\underline{r},n}^{-\lambda} := \underline{\mathrm{U}}_{r_2,n}^{-\lambda} + \frac{1}{\widehat{-\lambda}+2}\underline{\mathrm{U}}_{r_1-r_2,n+2}^{-\lambda-2}\,\mathrm{GD}$$

with $n = \lfloor \mathrm{Re}\,\lambda + \frac{d}{p} \rfloor - d$, $\mathrm{Re}\,\lambda + \frac{d}{p} \notin \mathbb{N}$, is $(-\lambda)$-homogeneous, vector-rotation-invariant, and continuous $\mathscr{D}^d(\mathbb{R}^d) \to \mathrm{L}_p^d(\mathbb{R}^d)$.

2.bg The formula that we used above to define $\underline{\mathrm{U}}_{\underline{r},n}^{-\lambda}$ can be compared to (2.39). We next intend to show that $\underline{\mathrm{U}}_{\underline{r},n}^{-\lambda}$ can also be seen as a modified (corrected) version of the vector operator $\underline{\mathrm{U}}_{\underline{r}}^{-\lambda}$ introduced in the previous subsection, in the same way that the scalar operator $\mathrm{U}_{c,n}^{-\lambda}$ was defined in **2.am** as a correction of $\mathrm{U}_c^{-\lambda}$. To this end, we first define the vector operator $\underline{\mathrm{Reg}}_{c,n}^{-\lambda}$, $c \in \mathbb{C}$, as

the coordinate-wise application of the scalar operator $\mathrm{Reg}_{c,n}^{-\lambda}$ of **2.ai**. Next, we define

$$\underline{\mathrm{Reg}}_{\underline{r},n}^{-\lambda} = \underline{\mathrm{Reg}}_{r_2,n}^{-\lambda} + \frac{1}{-\lambda+2}\,\underline{\mathrm{Reg}}_{r_1-r_2,n+2}^{-\lambda-2}\,\mathrm{GD}\,.$$

This permits us to write

$$\underline{\mathrm{U}}_{\underline{r},n}^{-\lambda} = \underline{\mathrm{U}}_{\underline{r}}^{-\lambda} - \underline{\mathrm{Reg}}_{\underline{r},n}^{-\lambda}\,. \tag{2.40}$$

$\underline{\mathrm{Reg}}_{\underline{r},n}^{-\lambda}$ is null for functions in $\widehat{\mathscr{E}}^d(\dot{\mathbb{R}}^d)$ (whose moments of any order vanish), whence we obtain the

2.bh Lemma. The restriction of $\underline{\mathrm{U}}_{\underline{r},n}^{\lambda}$ to $\widehat{\mathscr{E}}^d(\dot{\mathbb{R}}^d)$ coincides with the translation-invariant operator $\underline{\mathrm{U}}_{\underline{r}}^{\lambda}$.

In addition, the above relations, in combination with Lemmata **2.ak**, **2.ao**, and **2.ap**, directly prove the following parallel of Lemma **2.ao**.

2.bi Lemma. Let $\underline{\mathrm{D}}_H$ be the vector-to-vector operator corresponding to the coordinate-wise application of the scalar symmetric finite difference D_H with $H = \{h_0, \ldots, h_n\}$, defined in **2.ao**. Then

$$\underline{\mathrm{Reg}}_{\underline{r},n}^{-\lambda}\ \underline{\mathrm{D}}_H = \underline{\mathrm{D}}_H\ \underline{\mathrm{Reg}}_{\underline{r},n}^{-\lambda} = 0.$$

Consequently, $\underline{\mathrm{U}}_{\underline{r},n}^{-\lambda}\underline{\mathrm{D}}_H = \underline{\mathrm{D}}_H\underline{\mathrm{U}}_{\underline{r},n}^{-\lambda} = \underline{\mathrm{D}}_H\underline{\mathrm{U}}_{\underline{r}}^{-\lambda}$, which is translation-invariant.

Using the previous lemmata and (2.38) and **2.ak**, we obtain the next result (compare with **2.ap**).

2.bj Lemma.

$$\underline{\mathrm{U}}_{(r_1,r_2),n}^{-\lambda}\underline{\mathrm{U}}_{(r_1^{-1},r_2^{-1})}^{\lambda}\,\phi = H_\lambda\phi$$

for all $\phi \in \widehat{\mathscr{E}}(\dot{\mathbb{R}}^d)$ (reminder: the constant $H_\lambda = 1/\Gamma(\frac{d+\lambda}{2})\Gamma(\frac{d-\lambda}{2})$ was defined in (2.24)).

More generally, the above relation holds for $\phi \in \mathscr{E}(\mathbb{R}^d)$ and $n < \mathrm{Re}\,\lambda$.

§2.4 The vector calculus of rotation-invariant homogeneous operators

2.bk Notational convention. In this section we shall find it convenient to make use of a loose version of *Einstein's index convention,* according to which repeated indices in a product are summed upon (typically over the range $1, 2, \ldots, d$); thus, for instance, the product of matrices $[a_{ij}]$ and $[b_{ij}]$ is the matrix $[c_{ij}]$ with

$$c_{ij} = a_{ik}b_{kj}.$$

§2.4.1 Review of vector calculus

We begin by recalling some familiar operators from vector calculus.

2.bl Definitions. In d dimensions, the *gradient* operator, which maps a scalar to a d-dimensional vector, is defined as

$$[\operatorname{Grad} f]_i = \partial_i f.$$

Its adjoint is the negative *divergence,* mapping d-dimensional vectors to scalars and defined as

$$\operatorname{Grad}^* f = -\operatorname{Div} f = -\partial_i f_i$$

(recall that per Einstein's convention, the last formula denotes a sum over i).

We next give a definition of the d-dimensional *curl* operator, which maps a d-dimensional vector to a $d \times d$ anti-symmetric tensor:

$$[\operatorname{Curl} f]_{ij} = \tfrac{1}{\sqrt{2}}(\partial_i f_j - \partial_j f_i)$$

The curl adjoint is the operator

$$[\operatorname{Curl}^* f]_i = \tfrac{1}{\sqrt{2}}\partial_j(f_{ij} - f_{ji})$$

which carries a $d \times d$ anti-symmetric tensor to a d-dimensional vector (in the latter equation j is a summation index, following Einstein's convention).

2.bm Since a $d \times d$ anti-symmetric tensor has $\frac{1}{2}d(d-1)$ degrees of freedom, we may also look at the curl operator as a map from d-dimensional vectors to $\frac{1}{2}d(d-1)$-dimensional *bi-vectors.* In the common 3-dimensional situation (i.e. when $d = 3$), we have $\frac{1}{2}d(d-1) = 3 = d$, and the curl operator can then be given the same formulaic form as its adjoint, which partly explains why the curl and its adjoint are sometimes confused in common usage. [7] Explicitly, with the above definition of $\operatorname{Curl} f$, in 3 dimensions we have

$$\operatorname{Curl}\begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} = \tfrac{1}{\sqrt{2}}\begin{pmatrix} 0 & \partial_1 f_2 - \partial_2 f_1 & \partial_1 f_3 - \partial_3 f_1 \\ \partial_2 f_1 - \partial_1 f_2 & 0 & \partial_2 f_3 - \partial_3 f_2 \\ \partial_3 f_1 - \partial_1 f_3 & \partial_3 f_2 - \partial_2 f_3 & 0 \end{pmatrix},$$

7. A 3-dimensional bi-vector behaves like an ordinary vector under proper rotations (those with determinant 1), but flips sign under improper rotations (with determinant -1); cf. the right-hand rule in physics.

and since the anti-symmetric matrix given above is determined by its upper-diagonal half, we can define a compressed curl operator,

$$\diamond \operatorname{Curl} \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} = \begin{pmatrix} \partial_3 f_2 - \partial_2 f_3 \\ \partial_1 f_3 - \partial_3 f_1 \\ \partial_2 f_1 - \partial_1 f_2 \end{pmatrix};$$

whence $\diamond$ can be understood formally as the map

$$\diamond : \tfrac{1}{\sqrt{2}} \begin{pmatrix} 0 & -g_3 & g_2 \\ g_3 & 0 & -g_1 \\ -g_2 & g_1 & 0 \end{pmatrix} \mapsto \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix},$$

which relates 3×3 anti-symmetric tensors to their bi-vector equivalent with $\frac{1}{2}d(d-1) = 3$ components.

2.bn Some familiar identities. As is readily verified, for a scalar field f

$$\operatorname{Grad}^* \operatorname{Grad} f = -\Delta f$$

where Δ is the scalar Laplacian. Also,

$$\operatorname{Curl} \operatorname{Grad} f = 0$$

and, dually, for a bi-vector field f,

$$\operatorname{Grad}^* \operatorname{Curl}^* f = -\operatorname{Div} \operatorname{Curl}^* f = 0.$$

Additionally, for a vector field f,

$$\operatorname{Curl}^* \operatorname{Curl} f = -\operatorname{Grad} \operatorname{Grad}^* f - \Delta f$$

with Δ now denoting the vector Laplacian. This yields the familiar decomposition of the vector Laplacian,

$$\Delta f = -\operatorname{Grad} \operatorname{Grad}^* f - \operatorname{Curl}^* \operatorname{Curl} f,$$

which is sometimes used to 'define' Δf.

2.bo The Helmholtz decomposition. For test functions $\phi \in \widehat{\mathscr{E}}^d(\dot{\mathbb{R}}^d)$, let $\underline{\mathrm{E}}$ denote the operator

$$\underline{\mathrm{E}} : \phi \mapsto \partial_{ij} \Delta^{-1} \phi_j$$

associated with the Fourier multiplier matrix with ijth entry $\xi_i \xi_j / |\xi|^2$, $1 \le i, j \le d$. (We shall also denote the adjoint of $\underline{\mathrm{E}}$ on $(\widehat{\mathscr{E}}^d(\dot{\mathbb{R}}^d))'$ by the same

symbol $\underline{\mathrm{E}}$ since it shares the same Fourier domain representation.) $\underline{\mathrm{E}}$ is idempotent, in other words a projection (i.e. $\underline{\mathrm{E}}^2 = \underline{\mathrm{E}}$), as is its *complement* $id - \underline{\mathrm{E}}$ (id denotes the identity operator). It follows that

$$\underline{\mathrm{E}}\,(id - \underline{\mathrm{E}}) = (id - \underline{\mathrm{E}})\underline{\mathrm{E}} = 0.$$

Moreover, we have,

$$\mathrm{Curl}\ \underline{\mathrm{E}}\,\phi = \underline{\mathrm{E}}\ \mathrm{Curl}\ \phi = 0 \quad \text{and} \quad \mathrm{Div}\,(id - \underline{\mathrm{E}})\,\phi = (id - \underline{\mathrm{E}})\,\mathrm{Div}\ \phi = 0.$$

Thus, the *Helmholtz decomposition* of $\phi \in \widehat{\mathscr{E}}^d(\dot{\mathbb{R}}^d)$ into the sum of a curl-free term and a divergence-free one is given as

$$\phi = \underline{\mathrm{E}}\phi + (id - \underline{\mathrm{E}})\phi.$$

§2.4.2 The Helmholtz decomposition of rotation-invariant homogeneous operators

2.bp For $\phi \in \widehat{\mathscr{E}}^d(\dot{\mathbb{R}}^d)$ it is straightforward to see that we have

$$\begin{aligned}\underline{\mathrm{U}}^\lambda_{\underline{r}}\phi &= \underline{\mathrm{U}}^\lambda_{\underline{r}}\underline{\mathrm{E}}\phi + \underline{\mathrm{U}}^\lambda_{\underline{r}}(id - \underline{\mathrm{E}})\phi \\ &= \underline{\mathrm{E}}\underline{\mathrm{U}}^\lambda_{\underline{r}}\phi + (id - \underline{\mathrm{E}})\underline{\mathrm{U}}^\lambda_{\underline{r}}\phi = \underline{\mathrm{U}}^\lambda_{(r_1,0)}\phi + \underline{\mathrm{U}}^\lambda_{(0,r_2)}\phi\end{aligned}$$

(term-by-term equality) with

$$\mathrm{Curl}\,\underline{\mathrm{U}}^\lambda_{(r_1,0)}\phi = 0 \quad \text{and} \quad \mathrm{Div}\,\underline{\mathrm{U}}^\lambda_{(0,r_2)}\phi = 0. \tag{2.41}$$

The last two identities hold more generally for $\phi \in \mathscr{E}^d(\mathbb{R}^d)$, thus proving that

$$\underline{\mathrm{U}}^\lambda_{\underline{r}}\phi = \underline{\mathrm{U}}^\lambda_{(r_1,0)}\phi + \underline{\mathrm{U}}^\lambda_{(0,r_2)}\phi \tag{2.42}$$

is the Helmholtz decomposition of ϕ into the sum of a curl-free and a divergence-free field (in that order), with parameters r_1, r_2 controlling the curl-free and divergence-free contributions respectively (setting $r_1 = 0$ therefore produces a divergence-free field while $r_2 = 0$ generates a curl-free one).

In duality to (2.41), we also have, for $\phi_1 \in \mathscr{E}^{d\times d}$ and $\phi_2 \in \mathscr{E}$,

$$\underline{\mathrm{U}}^\lambda_{(r_1,0)}\,\mathrm{Curl}^*\,\phi_1 = 0 \quad \text{and} \quad \underline{\mathrm{U}}^\lambda_{(0,r_2)}\,\mathrm{Grad}\,\phi_2 = 0. \tag{2.43}$$

§2.4.3 The Helmholtz decomposition of the L_p-continuous operators

Finally, we are able to make the following connection between the 'corrected' continuous operators $\underline{\mathrm{U}}^{-\lambda}_{\underline{r},n}$ of **§2.3.2** and curls and divergences (cf. (2.43)).

2.bq Lemma.

$$\underline{\mathrm{U}}^{-\lambda}_{\underline{r},n}\,\mathrm{Curl}^* = \underline{\mathrm{U}}^{-\lambda}_{(0,r_2),n}\,\mathrm{Curl}^* \quad \text{and} \quad \underline{\mathrm{U}}^{-\lambda}_{\underline{r},n}\,\mathrm{Div}^* = \underline{\mathrm{U}}^{-\lambda}_{(r_1,0),n}\,\mathrm{Div}^*.$$

Proof. It is sufficient to show that the above relations hold with $\underline{\mathrm{U}}^{-\lambda}_{\underline{r},n}$ replaced by $\underline{\mathrm{Reg}}^{-\lambda}_{\underline{r},n}$. The rest then follows from (2.43) and (2.40).

To prove the claims about $\underline{\mathrm{Reg}}^{-\lambda}_{\underline{r},n}$, first, note that for $\phi \in \mathscr{E}^{d\times d}$,

$$\big[\mathrm{GD}\,\mathrm{Curl}^*\,\phi\big]_i = \tfrac{1}{\sqrt{2}}\partial_{ij}(\partial_l\phi_{jl} - \partial_l\phi_{lj}) = \tfrac{1}{\sqrt{2}}(\partial_{ijl}\phi_{jl} - \partial_{ijl}\phi_{lj}) = 0$$

by a simple change of variables ($l \leftrightarrow j$ in the second expression). Also, this time for $\phi \in \mathscr{E}$,

$$\big[\mathrm{GD}\,\mathrm{Div}^*\,\phi\big]_i = -\partial_{ij}\partial_j\phi = -\,\Delta\,\partial_i\phi.$$

More concisely put,

$$\mathrm{GD}\,\mathrm{Curl}^* = 0 \quad \text{and} \quad \mathrm{GD}\,\mathrm{Div}^* = \Delta\,\mathrm{Div}^*\,.$$

Writing

$$\begin{aligned}\underline{\mathrm{Reg}}^{-\lambda}_{\underline{r},n} &= \underline{\mathrm{Reg}}^{-\lambda}_{r_2,n} + \frac{1}{\widehat{-\lambda}+2}\,\underline{\mathrm{Reg}}^{-\lambda-2}_{r_1-r_2,n+2}\,\mathrm{GD}\\ &= \underline{\mathrm{Reg}}^{-\lambda}_{r_2,n} - \frac{1}{\widehat{-\lambda}+2}\,\underline{\mathrm{Reg}}^{-\lambda-2}_{r_2,n+2}\,\mathrm{GD} + \frac{1}{\widehat{-\lambda}+2}\,\underline{\mathrm{Reg}}^{-\lambda-2}_{r_1,n+2}\,\mathrm{GD}\end{aligned}$$

we therefore see that

$$\begin{aligned}\underline{\mathrm{Reg}}^{-\lambda}_{\underline{r},n}\,\mathrm{Curl}^* &= \underline{\mathrm{Reg}}^{-\lambda}_{r_2,n}\,\mathrm{Curl}^*\\ &= \underline{\mathrm{Reg}}^{-\lambda}_{r_2,n}\,\mathrm{Curl}^* - \frac{1}{\widehat{-\lambda}+2}\,\underline{\mathrm{Reg}}^{-\lambda-2}_{r_2,n+2}\,\mathrm{GD}\,\mathrm{Curl}^*\\ &= \underline{\mathrm{Reg}}^{-\lambda}_{(0,r_2),n}\,\mathrm{Curl}^*\end{aligned}$$

which proves the first half of (**2.bq**). Also, for the second half of (**2.bq**) we have

$$\begin{aligned}\underline{\mathrm{Reg}}^{-\lambda}_{\underline{r},n}\,\mathrm{Div}^* &= \underline{\mathrm{Reg}}^{-\lambda}_{r_2,n}\,\mathrm{Div}^* - \frac{1}{\widehat{-\lambda}+2}\,\underline{\mathrm{Reg}}^{-\lambda-2}_{r_1-r_2,n+2}\,\Delta\,\mathrm{Div}^*\\ &= \underline{\mathrm{Reg}}^{-\lambda}_{r_2,n}\,\mathrm{Div}^* - \underline{\mathrm{Reg}}^{-\lambda}_{r_1-r_2,n}\,\mathrm{Div}^*\\ &= \underline{\mathrm{Reg}}^{-\lambda}_{r_1,n}\,\mathrm{Div}^*\end{aligned}$$

using **2.ak** and (2.20) together in the penultimate step. ∎

2.br The results of this section, most importantly Lemma **2.be** about vectorial rotation-invariance and homogeneity, and Theorem **2.bf** about L_p^d-continuity, mirror the results we obtained in **§2.2.2** for scalar L_p-continuous operators and scalar rotation-invariance. We also made a connection between the introduced L_p^d-continuous operators and curl and divergence operators in Lemma **2.bq**, which has no parallel in the scalar case. In the next chapter, we shall use these results and the operators we constructed above in order to define random vector fields with statistical rotation-invariance and self-similarity properties and the possibility to control their irrotational vs solenoidal tendencies by virtue of Lemma **2.bq**.

— 3 —

Fractional Brownian Vector Fields and Some Non-Gaussian Extensions

We begin this chapter with an overview of a few modelling contexts and assumptions that give rise to fractional Brownian models. We see a few different ways of characterizing fractional Brownian fields, among which we pay more attention to their representation in the innovation modelling framework of Chapter 1. We then extend the definition of fractional Brownian motions to the vector setting by importing the vector operators of the previous chapter into the innovation modelling framework, and derive some of the key properties of the new models. Some of these properties have direct equivalents in the scalar setting, while others are specific to the vector case with no scalar counterpart. Finally, we find some further extensions of the models that are consistent with the fundamental properties derived previously, by considering non-Gaussian innovations and introducing the notion of subspace-independence.

§3.1 A few characterizations of fractional Brownian motion

3.a Axiomatic characterization of (non-fractional) Brownian motion. To a large extent, the motivation behind the mathematical definition of the (non-fractional) Brownian motion process $B(x)$, $x \in \mathbb{R}$, comes from physics. One is interested in modelling the type of phenomenon, imagined as the limit of a symmetric random walk starting at some arbitrary point called the origin, with the properties that its increments $B(x) - B(y)$ are Gaussian, independent from one another, and identically distributed for the same step size. In addition, one normally requires the process to be continuous in some sense

(we shall assume the weakest form that suits our needs). We summarize these properties below.

(B1) $B(0) = 0$.

(B2) $B(x)$ is a Gaussian process with zero mean.

(B3) $\mathrm{D}_h B(x) := B(x + \frac{h}{2}) - B(x - \frac{h}{2})$ is stationary for fixed h. (stationary increments)

(B4) $B(x)$ is continuous in probability, i.e., for any convergent sequence $x_n \to x$, we have $B(x_n) \xrightarrow{w} B(x)$ (weak convergence). (continuity)

(B5) $\mathbb{E}\{|B(x) - B(z)|^2\} = \mathbb{E}\{|B(x) - B(y)|^2\} + \mathbb{E}\{|B(y) - B(z)|^2\}$ for any increasing triple $x, y, z \in \mathbb{R}$. (uncorrelated increments)

In mathematical modelling applications, the 'derivative' of B, which we may imagine as a collection of infinitesimal independent increments or *innovations,* provides an intuitive picture of *white Gaussian noise* (we shall develop this idea more fully later on).

Putting aside for the moment the question of the well-definition of B (i.e. the existence of a corresponding probability measure on some suitable function space), from the above properties we see that for any $n \in \mathbb{Z}$ and $x \in \mathbb{R}$,

$$\begin{aligned}
\mathbb{E}\{|B(nx)|^2\} &= \mathbb{E}\{|B(nx) - B(0)|^2\} && \text{by (B1)},\\
&= \mathbb{E}\{|B(|n|x) - B(0)|^2\} && \text{by (B3)},\\
&= \sum_{0 \le k < |n|} \mathbb{E}\{|B(kx + x) - B(kx)|^2\} && \text{by (B5)},\\
&= |n|\, \mathbb{E}\{|B(x) - B(0)|^2\} && \text{by (B3)},\\
&= |n|\, \mathbb{E}\{|B(x)|^2\} && \text{by (B1)},
\end{aligned}$$

form where it follows that for rational x,

$$\mathbb{E}\{|B(x)|^2\} = k|x| \tag{3.1}$$

where $k = \mathbb{E}\{|B(0)|^2\}$. Since, by (B4), the above variance function is continuous, (3.1) holds for all $x \in \mathbb{R}$. Using (B1) and (B5) we then obtain the identity

$$\mathbb{E}\{|B(x) - B(y)|^2\} = k|x - y| \tag{3.2}$$

for all $x, y \in \mathbb{R}$ and some $k \in \mathbb{R}_+$. The last equation identifies the *variogram* structure of B. It also shows that the process is mean-square continuous.

3.b Turning things around, we may *assume* (3.2) together with (B1) and (B2) and prove the other properties. Indeed, in view of (3.2), (B5) is trivial, while (B3) follows from (B2) and the fact that the Gaussian process $\mathrm{D}_h B$ has constant variance $k|h|$ and is completely determined by its second order statistics. Finally, (B4) is a consequence of the mean-square continuity of B.

3.c 'Fractional' Brownian motion and other generalizations. The above characterization of Brownian motion by means of its variogram lends itself to several immediate generalizations. One may, following Lévy [Lé54, Lé65], take x, y in (3.2) to belong to $\mathbb{R}^d$ instead of $\mathbb{R}$ in order to define Brownian *fields*. One may also replace $k|x-y|$ in (3.2) by a fractional power function $k|x-y|^{2H}$ (thus breaking (B5)) in order to obtain *fractional* Brownian motion (fBm). Another possibility is to relax Gaussianity and replace (B5) by incremental independence, whence, in the non-fractional case, one recovers random entitities known as *Lévy processes*. Finally, of particular interest to us is the generalization from a scalar-valued random field on $\mathbb{R}^d$ to a *vector-valued* one, by taking $|B(x)-B(y)|^2$ to denote the squared-magnitude of the vector $B(x)-B(y)$. Naturally, these generalizations may be combined as well.

Apart from the non-Gaussian extension, the other generalizations suggests above can be combined in the following equation, which defines the variogram of a speculative *fractional Brownian field* B_H:

$$\mathbb{E}\{|B_H(x)-B_H(y)|^2\} \equiv k|x-y|^{2H} \tag{3.3}$$

(k is a normalization parameter). When consideration is given to vector-field models, we may reasonably require that such vector generalization of fBm as we develop here also be consistent with (3.3) where $|B_H(x)-B_H(y)|$ is to be replaced by the magnitude of the d-dimensional vector $B_H(x)-B_H(y)$ (note, however, that this leaves the cross-correlation of the vector coordinates unspecified).

3.d Operator-based 'innovation' characterization of fBm using invariances. The previous description of fBm, while deductive and intuitive, does not directly emphasize the invariance and self-similarity properties of the field; although these properties (rotation invariance and homogeneity of order H in law) do follow quite easily from the above definition.

For the above reason and, importantly, also in order to facilitate the application of analytical and distribution-theoretic techniques (such as those used freely in Chapter 2), in this thesis we take a different approach to characterizing fBm and introducing its vectorial extensions. This approach, which was adopted by Blu and Unser [BU07] and to which we refer as 'innovation modelling' (Chapter 1), puts the invariances of fBm in evidence, and also ensures that the corresponding probability measure is always well-defined. These combined goals are achieved as follows: we define B_H as the image of some spatially-independent rotation-invariant and homogeneous generalized random field W (called the 'innovation') under a rotation-invariant and homogeneous operator U^*:

$$B_H = \mathrm{U}^* W. \tag{3.4}$$

To W is associated a 'cylindrical' probability measure, which is transformed into a proper probability measure for B_H by the operator U* (see Chapter 1 for an overview and Appendix A for some of the more technical details). By a suitable choice of W and the operator U* we may derive (3.3) as a property, thus proving consistency with the previous axiomatic description. Moreover, by postulating the innovation W to be spatially independent, we effectively separate the randomness of the model (corresponding to the distribution of W), from its structural aspect (captured in the 'mixing' operator U*). This idea easily generalizes; for instance, in the vector setting, we shall replace U* by an operator with the appropriate 'vector' invariances, and use a vector innovation field for W. Replacing the Gaussian innovation W by a non-Gaussian one yields a different generalization.

To remind the reader of how we shall implement the above idea, we repeat the three steps involved in characterizing a random field in the innovation framework from **1.m**:

(IM1) Characterize the innovation W as a generalized random field over some test function space $\mathcal{X}$ by specifying its (normalized, continuous, and non-negative-definite) characteristic functional $\widehat{\mathscr{P}}_W$.

(IM2) Identify a continuous linear operator $\mathrm{U} : \mathcal{E} \to \mathcal{X}$ with prescribed properties, where the space $\mathcal{E}$ is nuclear.

(IM3) Define the random model B_H as the generalized random field with characteristic functional

$$\widehat{\mathscr{P}}_{B_H} := \widehat{\mathscr{P}}_W \circ \mathrm{U}. \tag{3.5}$$

The above functional uniquely identifies the sotchastic law (probability measure) $\mathscr{P}_{B_H}$ associated with B_H, to which it is related by the identity

$$\widehat{\mathscr{P}}_{B_H}(\phi) = \mathbb{E}_{B_H}\{\mathrm{e}^{\mathrm{i}\langle \phi, F\rangle}\} = \int_{\mathcal{E}'} \mathrm{e}^{\mathrm{i}\langle \phi, f\rangle} \mathscr{P}_{B_H}(\mathrm{d}f).$$

B_H can be interpreted as the transformation of W by the adjoint $\mathrm{U}^* : \mathcal{X}' \to \mathcal{E}'$ of the operator (U) identified in the previous step, thus giving sense to (3.4).

3.e In the next two sections, we shall consider the application of the operators defined in Chapter 2 to scalar and vector Gaussian innovations introduced in **1.u**, and study the properties of the random fields thus obtained in accordance with innovation modelling formula (3.4). Later, in **§3.5**, we generalize these models in two ways by replacing Gaussian innovations with α-stable ones, and considering the super-position of independent homogeneous random fields belonging to different subspaces.

§3.2 Fractional Brownian scalar fields

3.f As an example of the procedure outlined in the previous paragraph and detailed in Chapter 1, we now give an innovation model corresponding to fractional Brownian scalar fields. We note in passing, and without proof, that the random field described subsequently is consistent with variogram formula (3.3), as well as being rotation-invariant and homogeneous (self-similar) in law. The reason we omit the proofs of these statements is that similar but more complex properties will be proved later on for the vector field counterpart of the scalar model considered here.

3.g Step (IM1): The innovation. The standard Gaussian generalized random field W_G on $\mathscr{X} = \mathrm{L}_2(\mathbb{R}^d)$ is the cylindrical probability with characteristic functional

$$\widehat{\mathscr{P}_{W_G}}(\phi) = \mathrm{e}^{-\frac{1}{2}\|\phi\|_2^2}, \quad \phi \in \mathrm{L}_2(\mathbb{R}^d).$$

From the above characterization it is clear that W_G is rotation-invariant and $(-\frac{d}{2})$-homogeneous (self-similar) in probability (**1.x**).

It is therefore possible to obtain a Gaussian random model that is rotation-invariant and homogeneous of some given order H by transforming W_G using an operator that is rotation-invariant and homogeneous of order $-H-\frac{d}{2}$ in the sense of (2.4). In order for the procedure described in **1.m** (reproduced above in steps (IM1)–(IM3) of **3.d**) to be applicable, this operator also needs to be continuous from some nuclear space $\mathscr{E}$ into $\mathrm{L}_2(\mathbb{R}^d)$.

3.h Step (IM2): The operator. In Corollary **2.am** and Theorem **2.aq**, we identified an operator fulfilling the above requirements, with $\mathscr{D}(\mathbb{R}^d)$, the space of compactly supported smooth test functions, serving as our nuclear space $\mathscr{E}$. We recall the definition of this operator:

$$\mathrm{U}_{c,n}^{-\lambda} : \phi \mapsto \rho_c^{\lambda-d} * \phi - \sum_{|k|\le n} \frac{\partial_k \rho_c^{\lambda-d}}{k!} \int_{\mathbb{R}^d} (-y)^k \phi(y)\, \mathrm{d}y,$$

where $-\lambda$ is the homogeneity order, c is an arbitrary normalization, and we must have $n = \lfloor \mathrm{Re}\,\lambda + \frac{d}{p} \rfloor - d$, $\mathrm{Re}\,\lambda + \frac{d}{p} \notin \mathbb{N}$, in order to have $\mathscr{D} \to \mathrm{L}_p$ continuity (recall also that ρ_c^γ is the homogeneous distribution defined in (2.15)).

Consequently, in our case with Gaussian $(-\frac{d}{2})$-homogeneous innovation in L_2 ($p=2$), in order to have an H-homogeneous model we need $-\lambda = -H - \frac{d}{2}$ and $n = \lfloor H \rfloor$, $H \notin \mathbb{N}$.

3.i Step (IM3): The random field. We use the operator $\mathrm{U}_{c,\lfloor H \rfloor}^{-H-\frac{d}{2}}$ to define the scalar Gaussian fractional Brownian motion $B_{H,c}$ as the generalized random field

in $\mathscr{D}'(\mathbb{R}^d)$ with characteristic functional

$$\begin{aligned}\widehat{\mathscr{P}}_{B_{H,c}}(\phi) &= \widehat{\mathscr{P}}_{W_G}(\mathrm{U}_{c,\lfloor H\rfloor}^{-H-\frac{d}{2}}\phi)\\ &= \exp\quad -\frac{1}{2}\left\|\rho_c^{H-\frac{d}{2}} * \phi(x) - \sum_{|k|\le\lfloor H\rfloor}\frac{\partial_k \rho_c^{H-\frac{d}{2}}(x)}{k!}\int_{\mathbb{R}^d}(-y)^k\phi(y)\,\mathrm{d}y\right\|_2^2\\ &= \exp -\frac{1}{2(2\pi)^{\frac{d}{2}}}\left\|\int_{\mathbb{R}^d}\mathrm{e}^{\mathrm{i}\langle x,\xi\rangle}\rho_c^{-H-\frac{d}{2}}(\xi)\left(\widehat{\phi}(\xi) - \sum_{|k|\le\lfloor H\rfloor}\frac{\xi^k}{k!}\partial_k\widehat{\phi}(0)\right)\mathrm{d}\xi\right\|_2^2.\end{aligned}$$

The last equality uses Parseval's identity.

The above construction is equivalent, in the sense of **1.n**, to defining $B_{H,c}$ by the innovation equation

$$B_{H,c} = \mathrm{U}_{c,\lfloor H\rfloor}^{-H-\frac{d}{2}*}W_G,$$

where $\mathrm{U}_{c,n}^{-\lambda *}$ denotes the adjoint of $\mathrm{U}_{c,n}^{-\lambda}$ formally identified in Remark **2.at**.

3.j Remarks. In adopting the above innovation framework, we are following more or less directly in the footsteps of Blu and Unser [BU07]. Also, for $0 < H < 1$ and Gaussian innovations, the above characterizazion of fBm is essentially the same as the so-called 'harmonizable' representation, a one-dimensional version of which can be found for example in Samorodnitsky and Taqqu [ST94]. In the general case (arbitrary $H \notin \mathbb{N}$, Gaussian or α-stable statistics), a similar characterization of the random fields using regularized integral transforms of white noise appears in Benassi and Istas [BI02], which builds upon the characterization given in Benassi, Jaffard, and Roux [BJR97] for the Gaussian case; the non-Gaussian variants were also discussed in Benassi, Cohen, and Istas [BCI02]. Earlier, Yaglom and Pinsker [YP53, Pin55] and Dobrushin [Dob79] had studied Gaussian self-similar random fields with stationary $n+1$st order increments including similar characterizations, with non-Gaussian extensions also appearing in the last reference. The name 'fractional Brownian motion' originates with Mandelbrot and Van Ness [MV68], who used it to refer to processes defined as fractional derivatives or integrals of the Wiener process; Kolmogorov [Kol40] had already considered the same processes many years earlier. Also among the earlier works, in a more applied context, the British hydrologist Harold Edwin Hurst had proposed essentially the same models to account for the phenomenon of long-range dependence in reservoir capacities [Hur51].

The non-fractional counterpart of these processes is associated with the names of Wiener [Wie24] (in 1D) and Paul Lévy [Lév54, Lév65], among others.

A notable extension of these models, which is of interest in applications, but which we shall not discuss here, is the 'multi-fractional' Brownian motion, in which the exponent H is no longer a constant but instead itself a function of the parameter of the process (the original reference is Peltier and Lévy Véhel [PLV95]; for an overview of the applications of these models in image processing see Pesquet-Popescu and Lévy Véhel [PPL02]).

Naturally, to make use of these models in practice, it is important to be able to estimate their parameters, most notably the self-similarity exponent H, also known as the *Hurst exponent.* Some of the earliest examples of related estimates appear in the work of the eponymous Hurst [Hur56]. Peltier and Lévy Véhel [PLV94] proposed a strongly consistent estimator of H using fractal dimension analysis (see also references therein for other earlier work). In the context of time-scale and wavelet analysis of fBm and estimators derived from such analyses, which leverage on the scaling properties of wavelets and self-similar processes, we should cite the early papers of Flandrin [Fla89] and Mallat [Mal89]. More refined and sophisticated wavelet estimators were proposed some time later by Veitch and Abry [VA99] (see also our later paper [TVU09] with a multi-dimensional wavelet-based estimator (reproduced in Appendix B), and references therein). Among other works on the connection between wavelets and fractal processes (for analysis as well as synthesis) we mention those of Flandrin [Fla92], Wornell and Oppenheim [WO92], Masry [Mas93], Meyer, Sellan, and Taqqu [MST99], and the volume edited by Abry, Gonçalves, and Lévy Véhel [AGLV09].

§3.3 Fractional Brownian vector fields

3.k Next, we replicate the previous construction in the vector field context, going into more detail and pointing out the differences (some of which are fundamental).

3.l Step (IM1): The innovation. We introduced the standard Gaussian vector field $\underline{W}_G$ (note the underscore) in **1.u**. To remind the reader, $\underline{W}_G$ is the generalized random field in $\mathrm{L}_2^d(\mathbb{R}^d)$ (the dth Cartesian power of L_2) whose stochastic law is determined by the characteristic functional

$$\widehat{\mathscr{P}_{\underline{W}_G}}(\phi) = \mathrm{e}^{-\frac{1}{2}\|\phi\|_2^2}, \quad \phi = (\phi_1, \ldots, \phi_d) \in \mathrm{L}_2^d(\mathbb{R}^d),$$

where $\|\phi\|_2$ now denotes the L_2^d norm of the vector field ϕ defined thus:

$$\|\phi\|_2^2 = \int_{\mathbb{R}^d} \phi^{\mathrm{H}}(x)\phi(x)\,\mathrm{d}x = \sum_{1 \le i \le d} \|\phi_i\|_2^2$$

(ϕ^{H} is the Hermitian conjugate of the column vector ϕ). From the fact that $\widehat{\mathscr{P}}_{\underline{W}_G}$ can be factorized into a product of characteristic functionals of scalar Gaussian innovations per

$$\widehat{\mathscr{P}}_{\underline{W}_G}(\phi) = \prod_{1 \le i \le d} \widehat{\mathscr{P}}_{W_G}(\phi_i),$$

it follows that each coordinate of $\underline{W}_G$ is a standard Gaussian scalar field and is independent from the other coordinates.

$\underline{W}_G$ is rotation-invariant and homogeneous of order $-\frac{d}{2}$ in law, as is clear from the identity

$$\widehat{\mathscr{P}}_{\underline{W}_G}(\mathrm{S}_\sigma \mathrm{R}_{\omega,v}\phi) = \widehat{\mathscr{P}}_{\underline{W}_G}(\phi)$$

for arbitrary scale $\sigma > 0$ and any $d \times d$ orthogonal matrix ω (Lemma **1.w**). Note that the vector rotation operator invoked here ($\mathrm{R}_{\omega,v}$) is different from the scalar rotation that we considered in the previous section. Specifically, $\mathrm{R}_{\omega,v}$ is defined so (**2.n**):

$$\mathrm{R}_{\omega,v} : \phi \mapsto \omega\phi(\omega^{\mathrm{T}}\cdot).$$

3.m Step (IM2): The operator. In **2.bf**, we introduced a parametric family of homogeneous and vector rotation-invariant operators that map $\mathscr{D}^d(\mathbb{R}^d)$ continuously into $\mathrm{L}^d_p(\mathbb{R}^d)$ for some given $p \ge 1$ (note that the space $\mathscr{D}^d(\mathbb{R}^d)$, being the product of d nuclear spaces, is itself nuclear). The continuous dual of $\mathscr{D}^d(\mathbb{R}^d)$, i.e. the space $(\mathscr{D}^d)'(\mathbb{R}^d) = (\mathscr{D}')^d(\mathbb{R}^d)$, defines the space of (generalized) vector fields in which our random vector field models will be characterized.

As seen in (2.40), these continuous operators are of the form

$$\underline{\mathrm{U}}^{-\lambda}_{\underline{r},n} = \underline{\mathrm{U}}^{-\lambda}_{\underline{r}} - \underline{\mathrm{U}}^{-\lambda}_{\underline{r}} - \underline{\mathrm{Reg}}^{-\lambda}_{\underline{r}},$$

where $\underline{\mathrm{U}}^{-\lambda}_{\underline{r}}$ is a convolution with the homogeneous matrix distribution $P^{\lambda-d}_{\widehat{\underline{r}}}$ introduced in §**2.3.1** with

$$\widehat{\underline{r}} = \begin{pmatrix} \widehat{r}_1 \\ \widehat{r}_2 \end{pmatrix} = \frac{1}{\lambda - d} \begin{pmatrix} \lambda - 1 & 1 - d \\ -1 & \lambda - d + 1 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$$

(using (2.36) with $-\lambda$ instead of λ), and is homogeneous of order $-\lambda$, while $\underline{\mathrm{Reg}}^{-\lambda}_{\underline{r}}$ is a 'correction' such that having $n = \lfloor \mathrm{Re}\,\lambda + \frac{d}{p} \rfloor - d$ guarantees the corrected operator to be $\mathscr{D}^d \to \mathrm{L}^d_p$ continuous for $\lambda + \frac{d}{p} \notin \mathbb{N}$. Thus, in addition to the homogeneity index, we now have two new independent parameters $(r_1, r_2) = \underline{r}$ in the definition, which control the interaction of $\underline{\mathrm{U}}^{-\lambda}_{\underline{r},n}$ with curl and divergence operators (see **2.bq**).

As in the scalar case, in order to apply the operator to the Gaussian innovation recalled in the previous paragraph and obtain an H-homogeneous random vector field, we need to choose $-\lambda = -H - \frac{d}{2}$ and $n = \lfloor H \rfloor$, $H \notin \mathbb{N}$. Hence, in this case, our dual mixing operator will be $\underline{\mathrm{U}}_{\underline{r},\lfloor H \rfloor}^{-H-\frac{d}{2}}$. (We shall see later that the homogeneity order H and λ are related differently for α-stable fields with $\alpha \neq 2$; that is why we shall give most expressions in terms of λ instead of H.)

3.n Step (IM3): The random field. Finally, we define the random vector field $\underline{B}_{H,\underline{r}}$ in $(\mathscr{D}')^d(\mathbb{R}^d)$ as the probability measure on $(\mathscr{D}')^d$ corresponding to the characteristic functional

$$\widehat{\mathscr{P}}_{\underline{B}_{H,\underline{r}}}(\phi) = \widehat{\mathscr{P}}_{\underline{W}_G}(\underline{\mathrm{U}}_{\underline{r},n}^{-\lambda}\phi), \tag{3.6}$$

with $\lambda = H + \frac{d}{2}$ and $n = \lfloor H \rfloor$. We may interpret this random vector field as the outcome of the innovation model

$$\underline{B}_{H,\underline{r}} = \underline{\mathrm{U}}_{\underline{r},n}^{-\lambda *}\underline{W}_G, \tag{3.7}$$

where $\underline{\mathrm{U}}_{\underline{r},n}^{-\lambda *}$ denotes the adjoint of $\underline{\mathrm{U}}_{\underline{r},n}^{-\lambda}$.

We shall refer to these random fields interchangeably as *fractional Brownian vector fields* or *vector fBms.*

3.o Remark. For a related vector extension of fBm based on second-order correlation considerations see Yaglom [Yag87]. Other $\mathbb{R}^m$-valued generalizations of fractional Brownian fields, called by the name of 'operator-self-similar' Gaussian fields, are discussed in Mason and Xiao [MX01].

§3.4 Some properties of fractional Brownian vector fields

In this section we shall establish some of the main properties of the random fields defined in the previous section.

3.p Self-similarity and rotation invariance. Fractional Brownian vector fields are statistically self-similar (fractal) and rotation-invariant, meaning their probability law does not change under rotation or scaling (with proper re-normalization). This can be proved quite easily by using the corresponding properties of the innovation $\underline{W}_G$ and the operator $\underline{\mathrm{U}}_{\underline{r},n}^{-\lambda}$ to show that the characteristic functional of $\underline{B}_{H,\underline{r}}$ is invariant under the noted transformations. To wit, for $\mathrm{M} = \mathrm{R}_{v,\omega}$ (a vector rotation) or $\mathrm{M} = \sigma^{H+\frac{d}{2}}\mathrm{S}_\sigma$, with $\mathrm{R}_{v,\omega}$ and S_σ as

defined in **2.n**,

$$\begin{aligned}
\widehat{\mathscr{P}}_{(\mathrm{M}^*\underline{B}_{H,\underline{r}})}(\phi) &= \widehat{\mathscr{P}}_{\underline{B}_{H,\underline{r}}}(\mathrm{M}\,\phi) && \text{by definition,}\\
&= \widehat{\mathscr{P}}_{\underline{W}_G}(\underline{\mathrm{U}}_{\underline{r},n}^{-\lambda}\,\mathrm{M}\,\phi) && \text{also by definition,}\\
&= \widehat{\mathscr{P}}_{\underline{W}_G}(\mathrm{M}\,\underline{\mathrm{U}}_{\underline{r},n}^{-\lambda}\,\phi) && \text{by 2.be,}\\
&= \widehat{\mathscr{P}}_{\underline{W}_G}(\underline{\mathrm{U}}_{\underline{r},n}^{-\lambda}\,\phi) && \text{by 1.w,}\\
&= \widehat{\mathscr{P}}_{\underline{B}_{H,\underline{r}}}(\phi) && \text{by definition,}
\end{aligned}$$

with (again, for Gaussian fields) $\lambda = H + \frac{d}{2}$ and $n = \lfloor H \rfloor$. This proves that

$$\mathrm{M}^*\ \underline{B}_{H,\underline{r}} = \underline{B}_{H,\underline{r}} \quad \text{in law}$$

or, put more explicitly,

$$\underline{B}_{H,\underline{r}}(\omega\cdot) = \omega\underline{B}_{H,\underline{r}}(\cdot) \quad \text{and} \quad \underline{B}_{H,\underline{r}}(\sigma\cdot) = \sigma^H\underline{B}_{H,\underline{r}}(\cdot) \quad \text{in law,}$$

for any orthogonal matrix $\omega \in \mathbb{R}^{d\times d}$ and scalar $\sigma > 0$.

3.q Non-stationarity and stationary $n+1$st-order increments. Vector fBms with $H > 0$ are non-stationary since the operator that defines them is not translation-invariant. On the other hand, as a direct consequence of the stationarity of the innovation $\underline{W}_G$ and Lemma **2.bi**, we can state the following: Let $Y = \{y_0, \dots, y_n\}$ be a set of $n+1$ vectors in $\mathbb{R}^d$, and let $\underline{\mathrm{D}}_Y$ denote, as in **2.bi**, the symmetric finite difference operator defined recursively by the relations

$$\begin{aligned}
\mathrm{D}_{y_0} f &= f(\cdot + \tfrac{1}{2}y_0) - f(\cdot - \tfrac{1}{2}y_0),\\
\mathrm{D}_Y f &= \mathrm{D}_{y_n}\mathrm{D}_{Y\setminus\{y_n\}} f.
\end{aligned}$$

Then the random field $\underline{\mathrm{D}}_Y B_{H,\underline{r}}$ is stationary.

3.r The variogram and correlation form of vector fBm. As we saw in the previous paragraph, for $0 < H < 1$ the random field $\underline{B}_{H,\underline{r}}$ has stationary first-order increments. In this case we may define its variogram (or second-order structure function) as the correlation matrix of the stationary increment $\underline{B}_{H,\underline{r}}(x) - \underline{B}_{H,\underline{r}}(y) = \underline{\mathrm{D}}_{x-y}\underline{B}_{H,\underline{r}}(\frac{x+y}{2})$ (cf. 3.2). More specifically, we shall consider

$$\begin{aligned}
2\gamma_{\underline{B}_{H,\underline{r}}}(x,y) &= \mathbb{E}\big\{[\underline{B}_{H,\underline{r}}(x) - \underline{B}_{H,\underline{r}}(y)][\underline{B}_{H,\underline{r}}(x) - \underline{B}_{H,\underline{r}}(y)]^{\mathrm{H}}\big\} \qquad (3.8)\\
&=: \mathbb{E}\big\{I(0)[I(0)]^{\mathrm{H}}\big\},
\end{aligned}$$

with $\underline{I} := \underline{\mathrm{D}}_{x-y}\underline{B}_{H,\xi}(\cdot - \frac{x+y}{2})$.

We first obtain the correlation form of $\underline{I}$ from its characteristic functional by identification, using the following general relationship between the characteristic functional $\widehat{\mathscr{P}_{\underline{I}}}$ of the zero-mean Gaussian random field $\underline{I}$ and its correlation form $\langle\!\langle \phi, \psi \rangle\!\rangle_{\underline{I}} := \mathbb{E}_{\underline{I}}\{\overline{\langle \phi, X\rangle}\langle \psi, X\rangle\}$, $\phi, \psi \in \mathscr{E}^d$:

$$\widehat{\mathscr{P}_{\underline{I}}}(\phi) = \mathrm{e}^{-\frac{1}{2}\langle\!\langle \phi, \phi \rangle\!\rangle_{\underline{I}}}.$$

With $\underline{I} = \underline{\mathrm{D}}_{x-y}\underline{B}_{H,\underline{r}}(\cdot - \frac{x+y}{2})$, and with the characteristic functional of $\underline{B}_{H,\underline{r}}$ given by (3.6), we have:

$$\begin{aligned}
\langle\!\langle \phi, \phi \rangle\!\rangle_{\underline{I}} &= -2\log \widehat{\mathscr{P}}_{\underline{B}_{H,\underline{r}}}\big(\underline{\mathrm{D}}_{y-x}\phi(\cdot + \tfrac{x+y}{2})\big) \\
&= -2\log \widehat{\mathscr{P}}_{\underline{W}_G}\big(\underline{\mathrm{D}}_{y-x}\underline{\mathrm{U}}_{\underline{r}}^{-\lambda}\phi(\cdot + \tfrac{x+y}{2})\big) && \text{by } \mathbf{2.bi}, \\
&= \| P_{\widehat{\underline{r}}}^{\lambda-d} * \phi(\cdot - y) - P_{\widehat{\underline{r}}}^{\lambda-d} * \phi(\cdot - x)\|_2^2 && \text{by definition}, \\
&= \iint_{\mathbb{R}^d\times\mathbb{R}^d} \phi(s-x)^{\mathrm{H}}\, P_{\widehat{\underline{r}'}}^{2\,\mathrm{Re}\,\lambda-d}(s-t)\ \phi(t-x)\,\mathrm{d}s\,\mathrm{d}t \\
&\quad + \iint_{\mathbb{R}^d\times\mathbb{R}^d} \phi(s-y)^{\mathrm{H}}\, P_{\widehat{\underline{r}'}}^{2\,\mathrm{Re}\,\lambda-d}(s-t)\ \phi(t-y)\,\mathrm{d}s\,\mathrm{d}t \\
&\quad - 2\,\mathrm{Re}\iint_{\mathbb{R}^d\times\mathbb{R}^d} \phi(s-x)^{\mathrm{H}}\, P_{\widehat{\underline{r}'}}^{2\,\mathrm{Re}\,\lambda-d}(s-t)\ \phi(t-y)\,\mathrm{d}s\,\mathrm{d}t && \text{by } \mathbf{2.ba},
\end{aligned}$$

with

$$\widehat{\underline{r}'} = (\widehat{r'}_1, \widehat{r'}_2) = \frac{\Gamma(\frac{d}{2} - \mathrm{Re}\,\lambda)}{\Gamma(\frac{d-\lambda}{2})\Gamma(\frac{d-\overline{\lambda}}{2})}(|\widehat{r}_1|^2, |\widehat{r}_2|^2) \tag{3.9}$$

and $\lambda = H + \frac{d}{2}$.

We therefore see, by the kernel theorem (cf. foonote 6 on p. 21), that the random field $\underline{I}$ has the distributional covariance matrix

$$C_{\underline{I}}(s,t) = 2P_{\widehat{\underline{r}'}}^{2H}(s-t) - 2P_{\widehat{\underline{r}'}}^{2H}(s-t-x+y).$$

Thus, in particular,

$$2\gamma_{\underline{B}_{H,\underline{r}}}(x,y) = \mathbb{E}\big\{\underline{I}(0)[\underline{I}(0)]^{\mathrm{H}}\big\} = -2P_{\widehat{\underline{r}'}}^{2H}(y-x) = -2P_{\widehat{\underline{r}'}}^{2H}(x-y)$$

(the negative sign is only apparent, and is cancelled by the negative sign of the numerator of (3.9) for $0 < H < 1$, the rest of the latter equation being positive). We record this result as a

3.s Proposition. The random field $\underline{B}_{H,\underline{r}}$ with $0 < H < 1$ has stationary increments and its variogram can be expressed as

$$2\gamma_{\underline{B}_{H,\underline{r}}}(x,y) = -2P_{\widehat{\underline{r}'}}^{2H}(x-y),$$

where the dependence of $\widehat{\underline{r}'} = (\widehat{r'}_1, \widehat{r'}_2)$ on $\widehat{\underline{r}} = (\widehat{r}_1, \widehat{r}_2)$ is dictated by (3.9) and $\widehat{\underline{r}}$, in turn, depends on $\underline{r} = (r_1, r_2)$ by (2.36).

By taking the trace of the above variogram, we obtain the following

3.t Corollary. For $0 < H < 1$ we have

$$\mathbb{E}\{|\underline{B}_{H,\underline{r}}(x) - \underline{B}_{H,\underline{r}}(y)|^2\} \propto |x-y|^{2H}.$$

This shows that the random fields described herein are consistent with (3.3) and it is therefore appropriate to call them fractional Brownian motions.

By developing (3.8) and applying the above theorem, and also using the property that $\underline{B}_{H,\underline{r}}$ vanishes at 0, we can also arrive at the correlation matrix of vector fBm for $0 < H < 1$:

$$\begin{aligned} 2\,\mathrm{Re}\,\mathbb{E}\{\underline{B}_{H,\underline{r}}(x)\underline{B}_{H,\underline{r}}(y)^{\mathrm{H}}\} &= 2\gamma_{\underline{B}_{H,\underline{r}}}(x,0) + 2\gamma_{\underline{B}_{H,\underline{r}}}(y,0) - 2\gamma_{\underline{B}_{H,\underline{r}}}(x,y) \\ &= 2P^{2H}_{\widehat{\underline{r}'}}(x) + 2P^{2H}_{\widehat{\underline{r}'}}(y) - 2P^{2H}_{\widehat{\underline{r}'}}(x-y). \end{aligned}$$

3.u Helmholtz decomposition. Recall that the operator $\underline{\mathrm{U}}^{-\lambda}_{\underline{r},n}$ can be decomposed as

$$\underline{\mathrm{U}}^{-\lambda}_{\underline{r},n} = \underline{\mathrm{U}}^{-\lambda}_{(r_1,0),n} + \underline{\mathrm{U}}^{-\lambda}_{(0,r_2),n}.$$

Referring to Lemma **2.bq**, and taking adjoints, we may thus rely on (3.7) to write:

$$\mathrm{Curl}\,\underline{B}_{H,\underline{r}} = \mathrm{Curl}\,\underline{B}_{H,(0,r_2)} \qquad \text{in law,}$$

and also,

$$\mathrm{Div}\,\underline{B}_{H,\underline{r}} = \mathrm{Div}\,\underline{B}_{H,(r_1,0)} \qquad \text{in law.}$$

In particular, the above equations imply that a vector fBm $\underline{B}_{H,(r_1,r_2)}$ with $r_2 = 0$ is curl-free, while one with $r_1 = 0$ has zero divergence, thus clarifying the significance of the parameters r_1 and r_2. For instance, a divergence-free random vector field may be of interest in the modelling of incompressible flow.

3.v To illustrate the previous point, in Figure 3.1 we present realizations of $\underline{B}_{H,\underline{r}}$ in two dimensions. These were generated in `MATLAB` (MathWorks, Natick, US-MA) by Fourier-domain filtering of pseudo-random discrete white Gaussian innovations, and visualized using `Mathematica` (Wolfram Research, Champaign, US-IL). The visualization technique used to generate these figures is known as Line Integral Convolution [CL93], and consists in local directional smoothing of an underlying image (typically white noise) in the direction of flow.

3.w We note, in passing, that analysis with kernels that incorporate Helmholtz projections can be used to estimate parameters r_1 and r_2 (see references Tafti and Unser [TU09] and Tafti *& al.*[TDGSU10] in Appendix B).

§3.5 Fractional stable vector fields and subspace-independent extensions

3.x Fractional stable vector fields. We next turn to some non-Gaussian generalizations of the previously defined random fields. We obtain these generalizations by replacing the Gaussian innovation $\underline{W}_G$ in (3.7) by a non-Gaussian one. Since our interest remains focused on self-similar random fields, we shall restrict ourselves to a subset of homogeneous innovations, more specifically to the symmetric α-stable ones introduced in **1.u** with $\alpha \geq 1$ (technically, the α-stable family includes the Gaussian innovation too, but in this section we shall use the term α-stable to distinguish from Gaussian innovations). To go through with the procedure outlined in **3.d** and detailed in Chapter 1, all that is needed is to replace the Gaussian innovation in (3.7) with an α-stable one from **1.u**, and the $\mathscr{D}^d(\mathbb{R}^d) \to \mathrm{L}_2^d(\mathbb{R}^d)$ continuous operator appearing therein (in adjoint form) by one in the $\underline{\mathrm{U}}_{\underline{r},n}^{-\lambda}$ family of Theorem **2.bf** that is continuous $\mathscr{D}^d(\mathbb{R}^d) \to \mathrm{L}_\alpha^d(\mathbb{R}^d)$. We can represent these *fractional stable vector fields* symbolically as

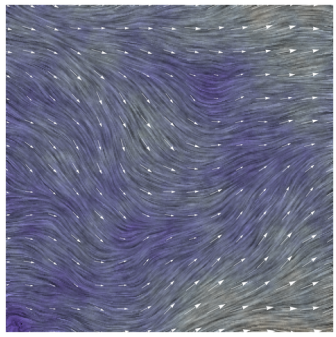
$$\underline{L}_{H,\underline{r}}^{\alpha} := \underline{\mathrm{U}}_{\underline{r},n}^{-\lambda *} \underline{W}_{S,\alpha} = \underline{\mathrm{U}}_{(r_1,0),n}^{-\lambda *} \underline{W}_{S,\alpha} + \underline{\mathrm{U}}_{(0,r_2),n}^{-\lambda *} \underline{W}_{S,\alpha} \tag{3.10}$$

with $n = \lfloor \lambda + \frac{d}{\alpha} \rfloor - d$. It follows from Lemma **1.w** and the $-\lambda$-homogeneity of $\mathrm{U}_{\underline{r},n}^{-\lambda}$ that in order for $\underline{L}_{H,\underline{r}}^{\alpha}$ to be homogeneous of order H, in this case we need $\lambda = H + d - \frac{d}{\alpha}$; from where follows, in turn, that we must once again have $n = \lfloor H \rfloor$.

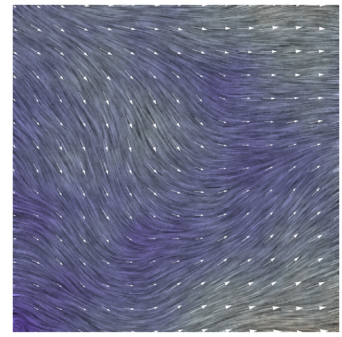
3.y Subspace-independent extensions. We shall go a step further and extend the above construction by considering compound fields obtained by superposing independent random fields belonging to the above family, with different values for the parameter vector $\underline{r}$ as well as α (we could also vary H, but this would break self-similarity). The most interesting elementary example of such a random field is perhaps the following one, which we call the *subspace-independent* fractional stable vector field.

$$\underline{L}_{H,\underline{r}}^{(\alpha_1,\alpha_2)} := \underline{\mathrm{U}}_{(r_1,0),\lfloor H \rfloor}^{-H-d+\frac{d}{\alpha_1} *} \underline{W}_{S_1,\alpha_1} + \underline{\mathrm{U}}_{(0,r_2),\lfloor H \rfloor}^{-H-d+\frac{d}{\alpha_2} *} \underline{W}_{S_2,\alpha_2}.$$

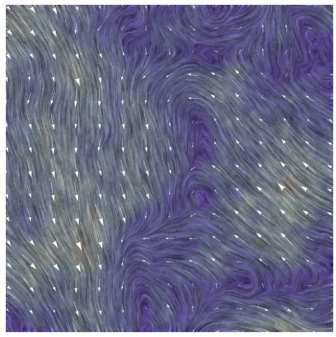
Note that in spite of the superficial symbolic similarity, even when $\alpha_1 = \alpha_2 = \alpha$, $\underline{L}_{H,\underline{r}}^{(\alpha_1,\alpha_2)}$ still defers from $\underline{L}_{H,\underline{r}}^{\alpha}$, since the latter uses the same innovation for both components while the former uses two independent ones. Thus, by



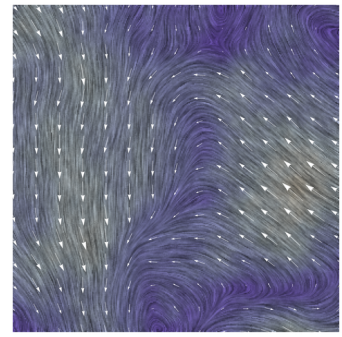
(a) $H = 0.6$, $r_1 = r_2$



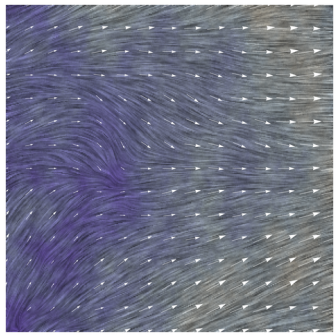
(b) $H = 0.9$, $r_1 = r_2$



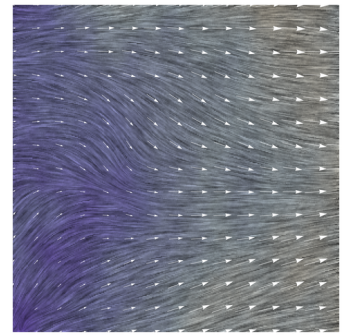
(c) $H = 0.6$, $r_1 = 0$



(d) $H = 0.9$, $r_1 = 0$



(e) $H = 0.6$, $r_2 = 0$



(f) $H = 0.9$, $r_2 = 0$

Figure 3.1: LIC visualization of synthesized fractional Brownian vector fields (see text); local amplitudes are coded inversely by saturation levels. Smoothness increases with increasing H, while r_1, r_2 determine directional behaviour ($r_1 = 0$ imposes zero divergence, $r_2 = 0$ characterizes irrotational models, and $r_1 = r_2$ defines fields with uncorrelated vector components).

virtue of Lemma **2.bq**, $\underline{L}_{H,\underline{r}}^{(\alpha_1,\alpha_2)}$ has independent curl and divergence (hence the name subspace-independent).

***3.z** To randomize even further, we might have, if we were so inclined, considered the sequences of parameters α and $(H, \underline{r})$ in a compound field, and even their number, to be random, and viewed them as additional sources of innovation in a hierarchical picture where the first innovation determines the number of fields to be superimposed, the second set of innovations determine the operator $(H, \underline{r})$ and type of spatial innovation (α) for each of the constituent independent random fields, and finally each of these component fields is itself associated with a spatial innovation field $\underline{W}_\alpha$ independent from all others. But we shall leave it at this.

3.aa Properties. Among the properties discussed in §3.4, statistical homogeneity and rotation-invariance (**3.p**) clearly hold for fractional stable vector fields as well, with essentially the same proofs. The stationary n + 1st-order increment property (**3.q**) also remains valid for the same reason as before. On the other hand, since the variance of α-stable fields with $\alpha < 2$ is not finite, we can no not define the variogram and correlation form of fractional stable fields, as we did for their Gaussian counterparts in **3.q**. With regard to the Helmholtz decomposition (**3.u**), as we saw in the previous paragraph, the subspace-independent fields introduced therein are by construction decomposable into a sum of independent irrotational and solenoidal random vector fields, and the directional behaviour of the field can be controlled by parameters r_1 and r_2.

— 4 —

Variational Reconstruction of Vector Fields

In this chapter, we study the practical problem of recovering a vector field from imperfect (and possibly indirect) measurements, and its refinement or enhancement. We shall be guided by the same principles of invariance and decoupling/innovation modelling as in the previous chapters, but this time around, they will serve as heuristics that will direct us to practically useful schemes. In the final part of this chapter we propose a practical vector field denoising and enhancement algorithm, which is shown experimentally to improve upon the state of the art.

§4.1 Methodology

§4.1.1 General considerations

4.a We shall be dealing with methods for reconstructing an object from a multitude of measurements or observations collectively denoted by Y.

Behind the previous, seemingly innocuous, statement hides a fundamental assumption or hypothesis, namely that underneath the possibly distorted and imperfect measurements referred to by Y, there exists a 'true' object (let's denote it by $f^\star$) that we strive to recover, or, failing that, approximate as closely as possible. In essence, our aim is to simplify away processes ('distortions') that are of less or no interest to us, while at the same time extracting whatever information we can about $f^\star$ from our measurements. We therefore distinguish between two types of loss of information: the 'systematic' variety, which is not helped by repeating the observations, and the 'random' kind, the effect of which can be reduced (at least hypothetically) by repeating the

measurements. [1] To include both effects, we model our observation vector Y as

$$Y = e(f^\star, \nu), \tag{4.1}$$

where $f^\star$, as already noted, is the 'true' object, ν is a 'random' object of sufficient complexity to account for modelling error due to simplifications or unknown factors, and e is a deterministic map that models both the systematic ambiguity about $f^\star$ and the way observations are influenced by ν.

4.b Another clarification is now in order. When it is said that we wish to recover or reconstruct $f^\star$, this requires us to have a model of what $f^\star$ *can be,* in the form of (at least) a set $\mathscr{F}$ of possible answers. [2] This set is parametrized in a fundamental sense by a minimal system of questions we can ask uniquely to distinguish $f^\star$ from every other element of $\mathscr{F}$, although the number of parameters may not be finite (or even countable). (Uniqueness here means that what we cannot distinguish based on these questions, we *consider* to be the same.) Naturally, the more restrictive the class of objects we consider, and the better our understanding of them, the more we can get out of fewer parameters. [3]

4.c Theoretically, it is imaginable that a kind of probabilistic model may be assigned to $f^\star$ as well, and when this is done (in the form of a probability on $\mathscr{F}$), this model is typically referred to as the 'prior'. More generally, instead of a probability measure in its standard mathematical sense, we may concoct some other criterion for quantifying our preference for one answer over another. We formalize this criterion as a non-negative functional $\mathfrak{R} : \mathscr{F} \to \mathbb{R}$, such that $\mathfrak{R}(f) < \mathfrak{R}(g)$ iff, all else being equal, f is considered preferable to g (as is standard, we shall call $\mathfrak{R}$ the *regularity criterion*, although in our view *simplicity criterion* might have been a better name, hinting at an application of Occam's razor or law of parsimony).

4.d With regard to the above, it is important to bear in mind that, random or not, in many practical problems we deal with observations made of a *single* (hypothetical) object $f^\star$ (or relatively few of them), rather than a representative ensemble of such. Thus, in ascribing a probability model to $f^\star$, we effectively *invent* our ensemble. (Even if it is not always admitted, at a fundamental level, all science is tautological).

1. This is not the only possible interpretation of randomness. From a different perspective, we may say that whatever we do not model systematically and deterministically (i.e. what we cannot or choose not to predict deterministically), we attribute to randomness.

2. This is not strictly true. In reality, when forming an opinion of the world, our set of ideas may not be fixed beforehand.

3. John von Neumann is said to have observed that with four parameters he could fit an elephant, and with five he could make him wiggle his trunk [Dys04].

This fabrication can occur in a variety of ways. Let us illustrate this with a simple example. Suppose that our true object $f^\star$ is the painting 'Mona Lisa' by Leonardo Da Vinci (1452–1519). Now, depending on what we know in advance about $f^\star$, we might take $\mathscr{F}$ to be one of the following:

– the set of paintings hanging on the walls of the Louvre in Paris on some specific day;
– a parametric set of representations of human portraits as combinations of simple geometric shapes;
– a matrix of numerical colour values at some specified resolution; *&c.*

We may then assign probabilities to each of the above sets in some imaginative fashion. For instance, in the first example, we may take these probabilities to be proportional to the number of tourists visiting each of the paintings hanging in the Louvre on that day. In the second case, we may assign probability distributions to the parameters of the portrait (position, elongation and orientation of the head, *&c.*) in one way or another. Finally, in the last example, we may create artificial diversity (randomness) by dividing the single complex entity (the Mona Lisa) into an ensemble of less complex elements (such as small neighbourhoods), and making further assumptions about stationarity and/or interdependencies, in order to estimate the statistical parameters of the model. (To an ant walking on the surface of the Mona Lisa, it does indeed *appear* as a random distribution of colours; said differently, the notion of randomness is not fundamental and inherent in the picture itself, but lies instead in our way of looking at it.)

Among the above three models, the last is the least specific, i.e. the most generic. For this reason, it is more likely to be useful if our image happens to be something other than the Mona Lisa (or a portrait, or a painting at all). At the same time, it gives us the least insight about the actual content and meaning of the object.

4.e The kinds of models we considered in the previous chapters, which involved independent spatial innovations, are closest in spirit to the third of the above examples. It is on this kind of idea that we shall base our solution methods in this chapter as well. While guided by this quasi-principle, we shall, however, refrain from making the statistical connection precise, that is, we shall not try to relate our methods of solution (i.e. our choice of the regularity criterion $\mathfrak{R}$ introduced in **4.c**) to specific probabilistic models. Although finding this connection is of interest and is something we hope eventually to do, we do have some reasons for the omission.

4.f For one thing, making such a connection would require adopting a paradigm of statistical optimality, the choice of which is somewhat arbitrary.

Moreover, supposing the former choice were somehow made, establishing a theoretical connection would then mean one of two things: either we would find a method first and then identify the probabilistic model for which it is optimal, or first come up with a probabilistic model in some fashion and then identify the method that is optimal for it. The latter approach might seem more reasonable at first glance, until one is reminded that in practice, the choice of method is dictated to a large extent by algorithmic and numerical considerations regarding what can be implemented with relative efficiency.

At any rate, the actual sequence of events in our case was that we took inspiration from the stochastic models of the previous chapters and formulated some algorithms that were seen to improve upon the state of the art in experiments. In regard of this, going back to the first of the two approaches sketched above, starting from a working method and reverse-engineering a probability model seemed in itself logically circular and of uncertain value (since we already knew, from experience, that the method would perform relatively well). We do not mean to discount the possibility that the exercise could potentially guide us somehow to modifications of the method that might be found to perform better in practice. But we did not plan or have the occasion to undertake such modifications and extensions in the present thesis.

Also, with respect to the second of the above two approaches (i.e. starting from a probabilistic model and deriving the scheme from it), in addition to the fact that practical considerations might and did limit our choice of method, there exists the fundamental problem of how to validate the stochastic model in the first place. Specifically, we were faced with the fact that we had relatively few datasets available to us (some of which were synthetic), and fitting the data with a probabilistic model which would later be used for processing the same data did not constitute good practice, to say the least, even though it would, automatically, produce self-confirmatory results.

4.g Having said all this, we shall, nevertheless, refer freely to the principles and models developed in the earlier chapters for motivation and guidance in choosing our practical course.

§4.1.2 Formulation of the problem

After the above philosophical prelude, let us now concentrate on the formal presentation of the problem and our adopted framework for its resolution, which consists in formulating the outcome as the solution of a mathematical optimization problem.

4.h Observation model. We assume to have at our disposal a finite number of observations of some phenomenon $f^\star$, i.e. a sequence $Y = \{Y[m]\}$ (with m in some finite index set which we denote by M). [4] It is supposed that beneath each observation $Y[m]$ hides a 'true' value $Y^\star[m]$, which is corrupted by the addition of some noise or distortion $N[m]$, that is,

$$Y[m] = Y^\star[m] + N[m], \quad m \in M.$$

We assume the values $N[m]$ to be random, identically distributed, and mutually independent for different m, and further suppose that they do not depend on Y either.

Additionally (and independently of the previous probabilistic assumptions), we suppose that $Y^\star$ depends linearly and continuously on the 'true' object ($f^\star$), that is,

$$Y^\star = \Phi f^\star$$

for some continuous linear map $\Phi : \mathscr{F} \to \mathbb{R}^M$ (the *sampling operator*). Equivalently, for each $m \in M$, the sample $Y[m]$ (the mth coordinate of $\Phi f^\star$) is the result of the application of a continuous linear functional ϕ_m to $f^\star$:

$$Y^\star[m] = \langle \phi_m, f^\star \rangle.$$

Thus we arrive at the following specialization of (4.1), with N taking the place of ν:

$$Y = \Phi f^* + N.$$

4.i The measurement fidelity/error criterion. It is typically supposed that Y^* lies at some 'distance' μ of Y as measured by some *fidelity* or *error* criterion (optimistic vs pessimistic naming) which we denote by $\operatorname{dist}(Y^\star; Y)$ (despite the notation, $\operatorname{dist}(\cdot;\cdot)$ need not be a metric). Thus we write

$$\operatorname{dist}(Y^\star; Y) = \mu. \tag{4.2}$$

If the criterion is taken to be the square error,

$$\operatorname{dist}(Y^*; Y) := \sum_m |Y[m] - Y^\star[m]|^2,$$

(4.2) then corresponds to fixing the sample noise variance. Note that in practice μ may not be known, at least precisely, in which case it can be

4. Enclosing the index in square brackets as in $Y[m]$ instead of subscripting or parenthesizing it like Y_m or $Y(m)$ is standard in signal processing for discrete signals (i.e. where m comes from a discrete set).

seen as a parameter that permits us systematically to search and explore the solution space. If instead of $Y^\star$ we seek the $f^\star$ from which it originates, we may rewrite the criterion as

$$\operatorname{dist}(\Phi f^\star; Y) = \mu. \tag{4.3}$$

4.j Choosing a solution: The 'regularity' criterion. Since in any non-trivial problem there exist many (even an infinity of) candidate solutions $f^\star \in \mathscr{F}$ that fulfil (4.3), we need some additional criterion to choose among them. This criterion may be thought of as a non-negative functional $\mathfrak{R} : \mathscr{F} \to \mathbb{R}$ with the property that in the absence of any other distinguishing factor, $\mathfrak{R}(f) < \mathfrak{R}(g)$ iff f is preferable to g as a solution. We may then formulate the task of choosing a solution (or a class of them) in the set defined by (4.3) as a minimization problem. The set of solutions is then defined as the argument $f^\star$ of (cf. **4.c**)

$$\min_{f^\star : \operatorname{dist}(\Phi f^\star; Y) = \mu} \mathfrak{R}(f^\star). \tag{4.4}$$

We emphasize that $f^\star$, hitherto used to denote the true but unknown object of interest, now (and henceforth) plays the role of the optimization variable that runs through the space of candidate solutions.

As we already noted in **4.c**, $\mathfrak{R}$ is traditionally referred to as a *regularization* functional or criterion.

4.k In the event, instead of (4.4), we may solve the following Lagrange-like relaxation of it, which, under some assumptions, is equivalent to the original problem for a suitable choice of α, but which may also be seen, more pragmatically, as a parametric scheme with α as its tuning parameter that permits us to search the solution space for a more 'desirable' solution.

$$\min_{f^\star} \operatorname{dist}(\Phi f^\star; Y) + \alpha \mathfrak{R}(f^\star). \tag{4.5}$$

4.l The place of invariances. The choice of $\mathfrak{R}$ in a specific problem is guided by practical considerations (unicity of the solution, ease of computation and optimization on a computer, *&c.*), as well as what prior assumptions we might have about the kind of solution we are seeking. In particular, if, in the absence of observations, we have *no reason to prefer* some solution $f^\star$ to its transformation $\mathrm{T}f^\star$ by a map T, this lack of bias can be built into the criterion by choosing an $\mathfrak{R}$ that is *invariant* under T:

$$\mathfrak{R}(\mathrm{T}f^\star) = \mathfrak{R}(f^\star).$$

We may relax the above strict form of invariance somewhat by choosing, instead of a single functional $\mathfrak{R}$, a parametric family $\mathfrak{R}_{\underline{\alpha}}$ ($\underline{\alpha}$ being a finite

vector), and requiring that there exist, for any $\underline{\alpha}$, an $\underline{\alpha}'$ such that

$$\mathfrak{R}_{\underline{\alpha}}(\mathrm{T}f^\star) = \mathfrak{R}_{\underline{\alpha}'}(f^\star) \tag{4.6}$$

independently of $f^\star$.

4.m Ultimately, whatever the considerations for choosing the form of $\mathfrak{R}$, any such choice needs to be vindicated by producing 'high-quality' (or at least 'acceptable') solutions with respect to some other, independently derived, performance criterion or figure of merit. The choice of such secondary criteria is itself somewhat (but not entirely) subjective and arbitrary. In academic and engineering contexts, it is common to use general-purpose metrics (such as the mean square error) that can be transposed easily from one problem to another, are easy to compute, and fulfil convenient mathematical properties. In contrast, in specific real-world applications, a solution may be deemed 'better' if it facilitates some practical task that would have been difficult (or even impossible) using some other solution method. [5]

4.n As a heuristic to guide our conception of formulae for $\mathfrak{R}$, we interpret $\mathfrak{R}$ as some sort of energy functional. This interpretation involves decomposing $\mathfrak{R}$ into two maps:

$$\mathfrak{R}(f^\star) \equiv \Xi\big(\mathrm{R}f^\star\big),$$

5. To give a concrete example, among two solutions for the problem of constructing a medical image from measurements, one may be considered better than another in a specific context if it simplifies the detection of certain features or defects that are hard to detect (or even missed) in the second reconstruction. On the other hand, in academic publications, different image construction methods are habitually compared in terms of their quadratic error (or some other simple metric) with respect to a known ground truth, since qualitative assessments such as the previous one are typically difficult to quantify and highly application-specific.

In the latter context, how the ground truth is 'known' is a further point of distinction. To continue with our example, one might actually acquire an image of a physical phantom which has been made to specifications and whose characteristics have been independently measured, and use these measurements for reference (quite good). Alternatively, the process of measurement and corruption by noise may be simulated on a computer, by using a *more precise* model of the measurement process than the one used in finding the solution (less good). Note that this extra precision can be in more accurate representation of systematic, as well as stochastic, characteristics of the measurement apparatus. Finally, it is common to see the same model being used for simulation and solution, in which case the numerical computer model is audaciously assumed to be perfect (some call this 'criminal' [Wir04]; although it is necessary as a 'sanity check': the method should work *at least* when all of its assumptions about the model are fulfilled!).

The problems considered in the experimental section of this chapter involve the simplest of all measurement models, consisting solely of corruption by independent additive noise. Our assessment of the methods presented therein relies on the first (qualitative, real-world) and last (perfect measurement model) kinds of experiments. To the extent that it concerns the systematic part of the model (represented here by the operator Φ in (4.3)), the consequences of assuming a perfect model are less severe when Φ is close to the identity (as in our case) or otherwise well-conditioned.

where:

– Ξ is a 'potential' functional with the property of being *additive with respect to independent arguments.* Mathematically, this is achieved by considering the argument of Ξ (let's call it g) as a function over some 'fundamental domain' Ω such that the values of g inside an arbitrary set $\Omega_1 \subset \Omega$ do not influence those inside any other disjoint set $\Omega_2 \subset \Omega$, $\Omega_1 \cap \Omega_2 = \varnothing$. Ξ is then conceived as an integral

$$\Xi(g) = \int_{\Omega} \xi\big(g(u)\big)\,\mu(\mathrm{d}u) \tag{4.7}$$

involving an elementary potential ξ and a measure μ that determines the relative importance of different subsets of the domain Ω.
In circumstances where a natural equivalence relation can be defined in Ω, the measure μ is required to be invariant with respect to this equivalence relation. For instance, if the fundamental domain Ω is taken to be (a subset of) $\mathbb{R}^d$ and translates of a subset of Ω are considered equivalent to it in terms of their contribution to Ξ, then μ is necessarily a multiple of the Lebesgue measure (when possible, it is convenient to normalize μ so as to have $\int_{\Omega} 1\mu(\mathrm{d}u) = 1$).
Note that the above integral becomes a sum if the domain Ω is finite or countable, i.e. if R separates $f^\star$ into a finite or countable number of independent constituents. In this case μ reduces to a collection of weights associated with each term of the sum. When all contributions are assumed to be equally important, μ disappears altogether or is reduced to an overall normalization factor (1 over the total number of contributions).

– R plays the role of a 'decoupling operator' that removes (or reduces) the redundancy in $f^\star$, so that the overall energy functional, $\mathfrak{R}(f^\star) = \Xi(\mathrm{R}f^\star)$, will also be (approximately) additive with respect to independent 'contributions'. It is helpful to keep in mind that R can be anything from a differential operator on $\mathbb{R}^d$ (mapping $f^\star$ to supposedly independent pointwise contributions), to a projection of $f^\star$ to its coefficients in some basis (such as a wavelet or Fourier basis), where they are assumed to be independent. [6]

6. The 'contributions' mentioned here can be compared intuitively to the innovations of Chapter 1. The answer to the question of whether or how this intuitive connection can be made a formal one by relating $\mathfrak{R}$ to a probability measure for $f^\star$ depends on the choice of probabilistic/statistical paradigm (typical examples: maximum a posteriori (MAP), minimum risk, *&c.*).

4.o In seeming more generality, one may consider compound potentials of the form

$$\mathfrak{R}(f^\star) = \sum_{1\le i\le N} \mathfrak{R}_i(f^\star) = \sum_{1\le i\le N} \int_{\Omega_i} \xi_i\big(\mathrm{R}_i f^\star(u)\big)\mu_i(\mathrm{d}u).$$

The generalization, however, is only apparent since we may re-express the above sum as a single integral of the form given in (4.7) on the larger space

$$\Omega = \{(i, \Omega_i)\}$$

by defining

$$\mu = \sum_i \delta_i \times \mu_i,$$
$$\mathrm{R}f^\star = (\mathrm{R}_1 f^\star, \dots, \mathrm{R}_N f^\star),$$
and
$$\xi(g_1, \dots, g_N) = \sum_i \xi_i(g_i)$$

(δ_i in the above denotes the Dirac measure on $\{1, \dots, N\}$ concentrated at i).

§4.1.3 Specialization to vector fields

4.p In our applications, $f^\star$ is a vector field. When the domain of definition of $f^\star$ is a continuum, to have the same kind of invariance to scaling and vector rotation as we studied in Chapter 2, we might consider curls and divergences, as well as the homogeneous and rotation-invariant operators $\underline{\mathrm{U}}^{\lambda *}_{\underline{r}}$ defined in Chapter 2, as our decoupling operators R_i.

4.q **Matrix L_p norms.** In addition, we may take the elementary potentials $\xi_i(g)$ (where, we recall, g may be scalar-, vector-, or tensor-valued) to be of the form

$$\xi_i(g) = \mathrm{Tr}\big((g^{\mathrm{H}} g)^{\frac{p}{2}}\big),$$

in which case the corresponding Ξ_i defines a rotation-invariant $\mathrm{L}_p(\Omega_i)$ norm for (scalar-, vector-, or tensor-valued) g per

$$\|g; \mathrm{L}_p(\Omega)\| = \|g\|_{p,\Omega} := \begin{cases} \left(\int_{\Omega_i} \mathrm{Tr}\big((g^{\mathrm{H}} g)^{\frac{p}{2}}\big)\, \mathrm{d}u\right)^{\frac{1}{p}}; & p < \infty, \\ \|\rho(g^{\mathrm{H}} g)^{\frac{1}{2}}\|_\infty. & p = \infty \end{cases}$$

with ρ denoting the spectral radius (we have taken the μ_is to be the Lebesgue measure on $\Omega \subset \mathbb{R}^d$; we shall, moreover, drop the Ω index from the norm

when the domain is all of $\mathbb{R}^d$).[7]

4.r The general form of energy functionals on $\mathbb{R}^d$. The previous combination of rotation- and scale-invariant operators and matrix L_p norms leads us to energy functionals of the form

$$\mathfrak{R}_{\underline{\alpha}}(f^\star) = \alpha_c \|\operatorname{Curl} f^\star\|_{p_c}^{p_c} + \alpha_d \|\operatorname{Div} f^\star\|_{p_d}^{p_d} + \sum_{1 \le i \le N} \alpha_i \|\underline{\mathrm{U}}_{\underline{r}_i}^{\lambda_i} f^\star\|_{p_i}^{p_i}$$

with multiplicative parameters $\alpha_c, \alpha_d, \alpha_1, \ldots, \alpha_N$, which are translation-, scale-, and rotation-invariant in the sense of (4.6).

4.s Discretization. In practice, either $f^\star$ is assumed to be discrete from the very beginning, in which case the previous operators are replaced by some form of discrete equivalents (most simply by replacing derivatives with finite differences), or, otherwise, we discretize R_i by writing it as

$$\mathrm{R}_i = \Pi' \mathrm{U} \Pi$$

where U is a continuously-defined operator and Π and Π' are finite-dimensional projections. In addition, the above integral L_p norms are replaced by finite sums in the obvious way.

4.t Discrete notation. In our implementation we adopted an elementary approach and simply discretized the operators by replacing partial derivatives by finite differences, as explained below. Regarding notation, we shall henceforth replace $f^\star$ by $F^\star = (F_1^\star, \ldots, F_d^\star)$ so as to be consistent in using capital letters for discrete quantities in the rest of this chapter, and we shall also drop the $\star$ as it has become an unnecessary ornament for some time now. All of our quantities (F, Y, and the like), will henceforth be defined on a centred hyper-rectangular Cartesian grid in $\mathbb{R}^d$. For simplicity, we shall take the grid to be uniform (with the same unit length in all principal directions), although the scheme can be generalized quite easily to non-uniform grids. In this way we can identify the grid with a subset M of $\mathbb{Z}^d$.

Thus, to take an example, $F_2[m]$ denotes the 2nd coordinate of the discrete vector field F at a point $m \in M$.

4.u Discretization of the basic operators of vector calculus. To begin with, we assume as given the shift operators s_i^k, and their corresponding adjoints s_i^{k*}, $1 \le i \le$

7. The above definition of L_p norms is motivated by a matrix Young inequality due to Ando [And95]:

$$\operatorname{Tr}\left((Y^{\mathrm{H}} X^{\mathrm{H}} X Y)^{\frac{1}{2}}\right) \le \tfrac{1}{p} \operatorname{Tr}\left((X^{\mathrm{H}} X)^{\frac{p}{2}}\right) + \tfrac{1}{p'} \operatorname{Tr}\left((Y^{\mathrm{H}} Y)^{\frac{p'}{2}}\right)$$

where $\frac{1}{p} + \frac{1}{p'} = 1$, which can be used to prove Hölder's inequality for these L_p spaces.

d, $k \in \mathbb{Z}^d$. s_i^k is to be compared to the shift operator

$$F \mapsto F[\cdot - k\widehat{e}_i],$$

where $\widehat{e}_i$ is the unit vector in direction i, but it also takes into account the relevant beoundary conditions for the problem (s_i^{k*} then implements the adjoint boundary conditions).[8] We drop the k superscript when it equals 1.

We can then define the backward finite difference operators

$$\delta_i : F \mapsto F - \mathrm{s}_i F,$$

and their corresponding adjoints

$$\delta_i^* : F \mapsto F - \mathrm{s}_i^* F, \quad 1 \leq i \leq d.$$

With the above definitions, a simple discretization of vector differential operators is obtained by replacing partial derivatives with finite differences:

$$\begin{aligned} [\mathrm{Grad}_\delta F]_i &= \delta_i F, \\ \mathrm{Div}_\delta F &= -\mathrm{Grad}_\delta^* F = \delta_i^* F_i, \\ [\mathrm{Curl}_\delta F]_{ij} &= \tfrac{1}{\sqrt{2}}(\delta_i F_j - \delta_j F_i), \\ \text{and} \qquad [\mathrm{Curl}_\delta^* f]_i &= \tfrac{1}{\sqrt{2}}\delta_j^*(F_{ij} - F_{ji}) \end{aligned}$$

(we are again using Einstein's convention as introduced in **2.bk**).

§4.2 Denoising and enhancement of vector fields with Div-Curl regularization

§4.2.1 Formulation

4.v To demonstrate the general method described above in an application, we now specialize to the problem of regularizing a discrete vector field $Y = (Y_1, \ldots, Y_d)$ of measurements, using the L_p norms of the discrete curl and divergence of the solution as regularization energies. We shall compare the solutions for $p = 1, 2$. Our two cost functions corresponding to $p = 1, 2$ will thus be of the form

$$\mathfrak{J}_{p,\alpha_c,\alpha_d}(F; Y) = \|F - Y\|_2^2 \; + \; \alpha_c \|\mathrm{Curl}_\delta F\|_p^p \; + \; \alpha_d \|\mathrm{Div}_\delta F\|_p^p. \qquad (4.8)$$

8. These so-called 'boundary conditions' may in fact consist of other kinds of linear constraints of appropriate dimensionality, and need not be associated with the boundaries. We call them boundary conditions for the lack of a better term. One of the most useful choices in image processing is the reflexive one.

Our computational task is to find the minimizer over $F \in \mathbb{R}^M$ of the above costs numerically. In the above problem, F and Y are defined on the same grid M and the final solution can be interpreted as a 'corrected' or 'enhanced' version of Y with more regularity.

4.w Bibliographical note. For $p = 2$, the above form of regularization with Div and Curl functionals has been considered by Amodei and Benbourhim [AB91], Dodu and Rabut [DR04], and Arigovindan *& al.*[Ari05, ASJ+07]. Suter and Chen [Sut94, SC00] also considered quadratic regularization with other types of differential operators for vector fields. Separately, regularization with scalar L_1 norms has come largely to replace L_2 regularization in many applications in scalar image processing, were it was initially proposed on account of its better edge-preserving properties compared to the latter [ROF92, GR92], but we are not aware of its application to vector fields.

4.x Let us begin by writing $\mathfrak{J}_p(F; Y)$ explicitly as

$$\mathfrak{J}_\delta^{(p)}(F; Y) = \sum_m |F[m] - Y[m]|^2 + \alpha_c \sum_m \big(\sqrt{|\operatorname{Curl}_\delta F[m]|^2}\big)^p + \alpha_d \sum_m \big(\sqrt{|\operatorname{Div}_\delta F[m]|^2}\big)^p,$$

where we have the following relations for the point-wise magnitudes of $\operatorname{Curl}_\delta F$ and $\operatorname{Div}_\delta F$:

$$|\operatorname{Curl}_\delta F[m]|^2 = \sum_{1 \le i < j \le d} (\delta_i F_j[m] - \delta_j F_i[m])^2;$$
$$|\operatorname{Div}_\delta F[m]|^2 = \sum_{1 \le i,j \le d} \delta_i F_i[m] \delta_j F_j[m].$$

§4.2.2 Numerical resolution

4.y For $p = 2$, the quadratic cost defined in (4.8) can be minimized quite efficiently by using an iterative linear solver to find the F for which its gradient vanishes (our implementation uses the conjugate gradient method). The minimization problem remains convex for $p = 1$, which suggests that it should still be solvable with relative efficiency, although not necessarily as easily as in the previous case. To minimize (4.8) for $p = 1$ we follow Figueiredo *& al.* [FDON06], and form a sequence of quadratic upper bounds (majorizers) of the cost function, which we then minimize (or only reduce) in sequence.

4.z Specifically, we can use the inequality $\sqrt{|a|} \le \sqrt{|a'|} + \frac{1}{2}(|a| - |a'|)/\sqrt{|a'|}$ to define the following tight upper bound (up to a constant term in F) on the

cost at a fixed F':

$$\begin{aligned}\mathfrak{Q}(F, F'; Y) := \quad & \sum_m \sum_{1\le i\le d} F_i[m]^2 - \sum_m \sum_{1\le i\le d} 2F_i[m]Y_i[m] \\ & + \alpha_c \sum_m c_m^{-1} |\operatorname{Curl}_\delta F[m]|^2 \\ & + \alpha_d \sum_m d_m^{-1} |\operatorname{Div}_\delta F[m]|^2,\end{aligned}$$

with

$$c_m := \sqrt{|\operatorname{Curl}_\delta F'[m]|^2} \quad \text{and} \quad d_m := \sqrt{|\operatorname{Div}_\delta F'[m]|^2}.$$

Next, consider the sequence

$$\tilde{F}_{(n)} := \arg\min_F \mathfrak{Q}(F, \tilde{F}_{(n-1)}; Y),$$

defined recursively with some initialization such as $\tilde{F}_{(0)} = 0$. For a given $\tilde{F}_{(n-1)}$ (corresponding to the intermediate solution obtained at the end of the $n-1$st outer iteration), the above minimization falls in the category of quadratic re-weighted least square problems, and is therefore solvable using a linear solver. Note that minimizing $\mathfrak{Q}(F, F; Y)$ is equivalent to minimizing $\mathfrak{J}_{p,\alpha_c,\alpha_d}(F; Y)$. Furthermore, we have

$$\mathfrak{Q}(\tilde{F}_{(n)}, \tilde{F}_{(n)}; Y) \le \mathfrak{Q}(\tilde{F}_{(n)}, \tilde{F}_{(n-1)}; Y) < \mathfrak{Q}(\tilde{F}_{(n-1)}, \tilde{F}_{(n-1)}; Y)$$

which shows that, with increasing n, the $\mathfrak{J}_\delta^{(1)}(\tilde{F}_{(n)}; Y)$'s form a decreasing sequence. (This is not a strict proof of convergence, but we have made so many approximations and heuristic arguments already that one more would not matter! In fact, in practice, we may not even *want* strictly to minimize the cost function, since stopping at an intermediate stage might produce a sufficiently good solution for a smaller computational budget, and there is no practical reason that this intermediate solution might not even be preferable to the strict minimizer of the cost according to our quality criterion.)

The final procedure for $p = 1$ is given in Algorithm 1. In actual implementation the local minimization in the last step of the loop is performed using an iterative linear solver such as conjugate gradient, which may however be stopped before full convergence while still reducing the global cost.

Algorithm 1: Algorithm for L_1 regularization

input: Y;
$\tilde{F}_{(0)} \leftarrow 0$;
repeat
  $n \leftarrow n + 1$;
  for *all data coordinates* m **do**
    $c_m \leftarrow \sqrt{|\operatorname{Curl}_\delta F'[m]|^2}$; $d_m \leftarrow \sqrt{|\operatorname{Div}_\delta F'[m]|^2}$;
  end
  $\tilde{F}_{(n)} \leftarrow \arg\min_F \Omega(F, \tilde{F}_{(n-1)}; Y)$;
until *stopping criteria are met*;
return $\tilde{F}_{(n)}$.

§4.3 Experiments

§4.3.1 Simulated denoising experiments

4.aa We implemented the scheme described previously for $p = 1$ and a linear solver for the quadratic ($p = 2$) case in MATLAB (MathWorks, Natick, US-MA) for denoising and enhancement in two and three dimensions. To test the method, we first used it to denoise simulated phantoms corrupted by different levels of additive white Gaussian noise, optimizing the algorithm parameters α_c and α_d for best mean squared error (MSE) performance.

In computer simulations, the true MSE can be calculated since the ground truth is known. In real-world applications, the ground truth is typically not accessible, although, depending on the circumstances, it might still be possible to estimate the MSE for the purpose of optimizing algorithm parameters. For example, if the distortion in the measurements can realistically be modelled as independent additive white Gaussian noise, then the Monte Carlo techniques proposed by Ramani *& al.* [RBU08] produce reliable estimates (see also Girard [Gir95]).

However, the noted assumptions about measurement noise can frequently fail in practice, and/or the computational cost of running the algorithm many times with different sets of parameters for optimization may become prohibiting. Other than that, the MSE might simply not be a suitable indicator of the usefulness of the reconstruction in the application of interest. In these cases choosing suitable values for the parameters becomes a matter of trial and error (usually not too many, due to the computational cost), and inspired guessing, guided by the physical significance of the curl and divergence regu-

larization terms (for instance, using a relatively large value for α_d pushes the solution towards being more-or-less divergence-free, which is a key property of homogeneous incompressible flow).

4.ab Phantoms. The first of our two 3D phantoms, presented in Figure 4.3, consists of the gradient field of the potential

$$\phi_{3D}(x_1, x_2, x_3) = x_1 x_2 e^{-|x|^2}$$

over a $41 \times 41 \times 41$ grid. The second 3D phantom, pictured in Figure 4.4, is a model of fully-developed laminar flow (with a parabolic profile) inside a tube, which is encircled by a constant flow in a torus. It is defined on a grid of size $26 \times 26 \times 26$.

Each of these phantoms was corrupted by different levels of additive Gaussian noise, so as to have a signal-to-noise ratio (SNR) of 0, 10, or 20 decibels (dBs). The SNR for a given signal F compared to the ground truth F_{true} is computed according to the formula below:

$$\text{SNR}(F; F_{\text{true}}) = 10 \log_{10}(\|F_{\text{true}}\|_2^2) - 10 \log_{10}(\|F - F_{\text{true}}\|_2^2).$$

4.ac Experiments. In order to compare the performance of L_1 and L_2 regularization per (4.8), we used the conjugate gradient method to solve the quadratic (L_2) problem until convergence. For the L_1 problem, we used Algorithm 1 with 8 external cycles and 600 internal conjugate gradient iterations per cycle to minimize the local quadratic bounds. Some results are shown graphically in Figures 4.1 and 4.2 (amplitude cross-sections) and also in Figures 4.3 and 4.4 (3D 'glyph' visualization generated using `ParaView` 3.8.0 [AGL05]). Quantitative comparisons in terms of SNR improvement (equiv. to MSE improvement) in dB are given in Table 4.1. In the same table, we also provide mean angular error figures. The latter performance measure, which we have adopted from Barron, Fleet, and Beauchemin [BFB94], is defined as the average point-wise angle between the ground truth and the output of the method.

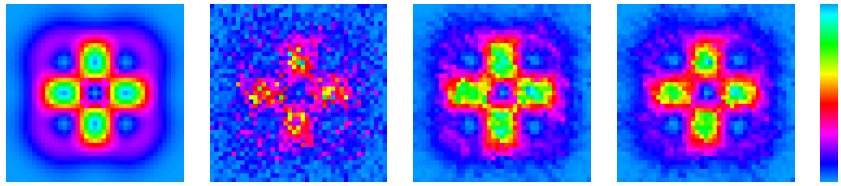
As suggested earlier, in each instance, we used a bracketing search method to optimize the algorithm parameters (α_c, α_d) for best MSE performance (but *not* taking into account the mean angular error). It should be noted that, in practice, the L_1 algorithm is terminated before full convergence in order to limit the amount of computation. The time or criterion for stopping the iterations can in fact be seen as an additional parameter of the algorithm, which we fixed in advance as indicated above. As a result of the different states of convergence in different trials with variable α_c, α_d, the numerical value of the performance criterion (the MSE) can fluctuate as a function of parameters α_c, α_d around its optimum, and a bracketing search may therefore

yield sub-optimal values for α_c, α_d. This does not invalidate the results (even the opposite), since it shows that the method performs well even with potentially sub-optimal parameter values.

4.ad Results and discussion. Both quantitative and qualitative results suggest that L_1 regularization performs better than its quadratic counterpart for the second phantom, which exhibits discontinuities in the flow, while the results of the two kinds of regularization are comparable for the first phantom (quadratic regularization performs slightly better in MSE terms in this case, which is not unexpected given that the first phantom is smooth and free of discontinuities; but L_1 regularization still produces favourable results in terms of mean angular error). In addition, visual inspection indicates that L_1 regularization is better at preserving small details and sharp discontinuities at flow boundaries, which are smoothed when using quadratic regularization.

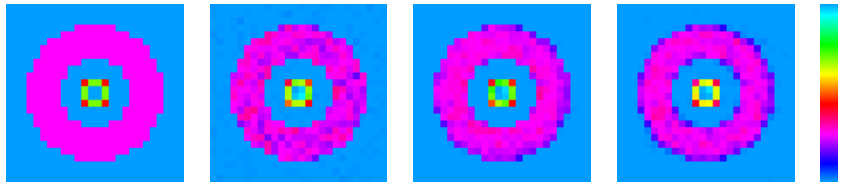
Table 4.1: Comparison of L_1 and L_2 regularization for denoising.

gradient field		
input SNR [dB] angular error [deg.]	SNR improvement [dB] angular error (mean ± stdev) [deg.]	
	L_1	L_2
0 (59.12° ± 39.93°)	**11.70** (**28.61°** ± 31.46°)	11.04 (31.84° ± 33.95°)
10 (37.81° ± 36.74°)	7.50 (**16.90°** ± 23.05°)	**7.78** (20.87° ± 28.31°)
20 (20.22° ± 28.11°)	4.49 (**10.03°** ± 15.80°)	**4.89** (12.40° ± 21.25°)
tube and torus		
input SNR [dB] angular error [deg.]	SNR improvement [dB] angular error (mean ± stdev) [deg.]	
	L_1	L_2
0 (12.11° ± 7.29°)	**8.03** (5.97° ± 4.04°)	6.37 (**5.95°** ± 3.93°)
10 (3.82° ± 2.21°)	**7.96** (**2.58°** ± 1.82°)	2.55 (3.16° ± 2.42°)
20 (1.21° ± 0.70°)	**6.67** (**0.99°** ± 0.71°)	0.51 (1.25° ± 0.85°)



(a) Original (b) Noisy (0 dB SNR) (c) L_1 denoised (11.70 dB improvement) (d) L_2 denoised (11.04 dB improvement)

Figure 4.1: Amplitude cross-sections, 'gradient' phantom, comparing L_1 and L_2 denoising.



(a) Original (b) Noisy (10 dB SNR) (c) L_1 denoised (7.96 dB improvement) (d) L_2 denoised (2.55 dB improvement)

Figure 4.2: Amplitude cross-sections, 'tube and torus' phantom, comparing L_1 and L_2 denoising.

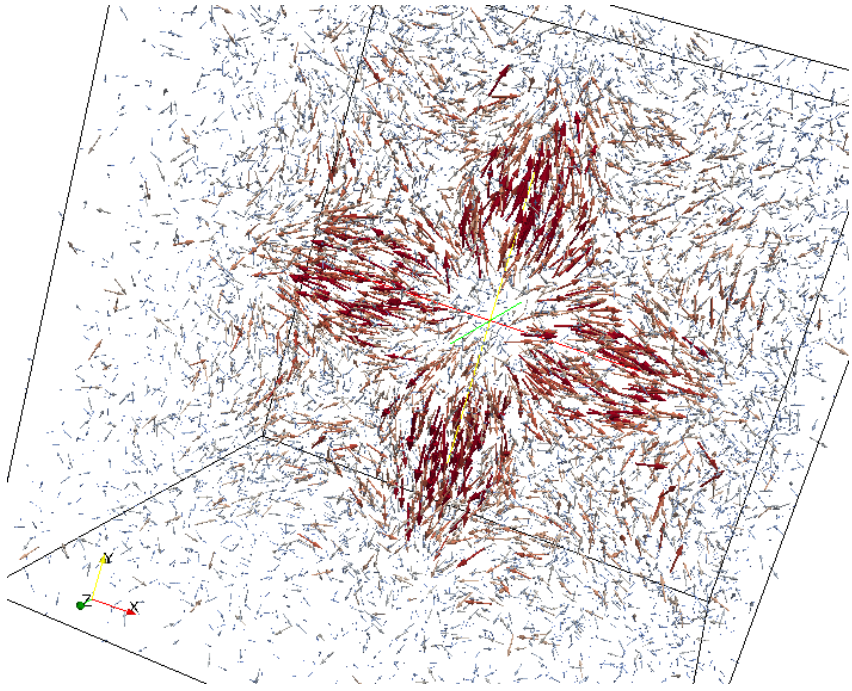
§4.3.2 Enhancement of flow-sensitive MRI

4.ae Experiment. In a more realistic application of the method, we used it to enhance directional features of measurements of blood flow in the thoracic aorta of a healthy human subject. These measurements were obtained using flow-sensitive Magnetic Resonance Imaging (MRI) in three dimensions plus time with ECG and respiratory gating, as described in Stalder [Sta09], Markl *& al.*[MHB+07], and Frydrychowicz *& al.*[FAH+08]. A brief description of the parameters of the MRI apparatus used in the experiment and standard systematic corrections applied to the measurements can be found in our conference paper [TDGSU10].

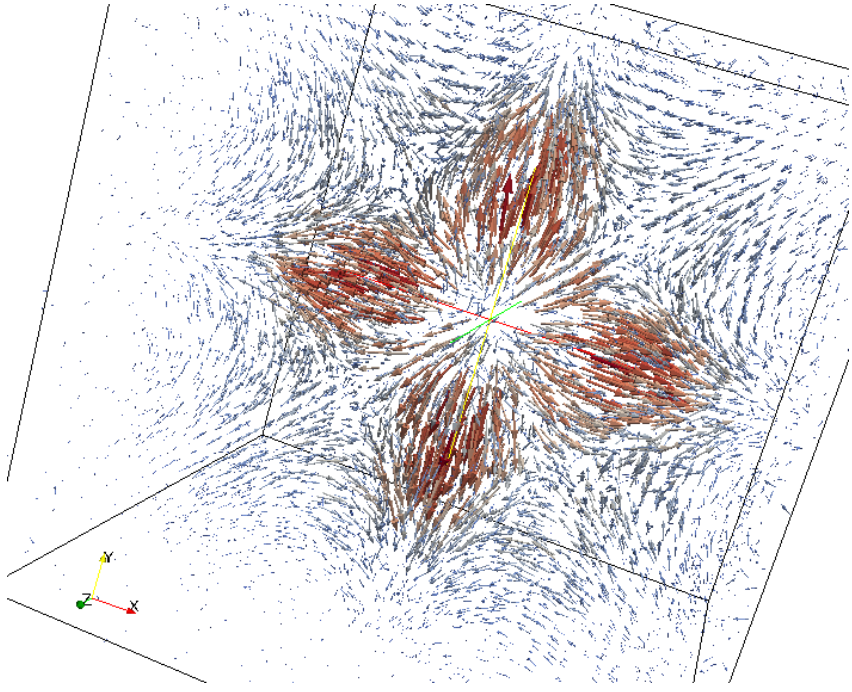
We applied the proposed L_1 algorithm to this dataset. Since the Gaussianity and independence of additive measurement error were not realistic assumptions, and also because our performance criterion was qualitative (namely, better expression of certain features of the flow), we adjusted the parameters

of the algorithm manually (in about two trials), with $\lambda_d \sim 150\lambda_c$ to enforce small divergence.

4.af Results and discussion. Evaluation of the effectiveness of the scheme was in this experiment judged on the basis of the quality and informativeness of flow visualizations that were produced using the enhanced dataset, compared to those generated from the original set. The difference is clearly visible in the pair of 'before/after' pathline visualizations given in Figure 4.5. The pathlines visible in this Figure are meant computationally to approximate the trajectory of massless particles inside the flow (data and visualizations are courtesy of Dr Aurélien F. Stalder; the visualizations were produced using the commercial software package `EnSight` (CEI, US-NC)). In particular, after applying the correction, many more pathlines remain within the lumen volume (the inside space of the artery) and complete their trajectory to the descending aorta.

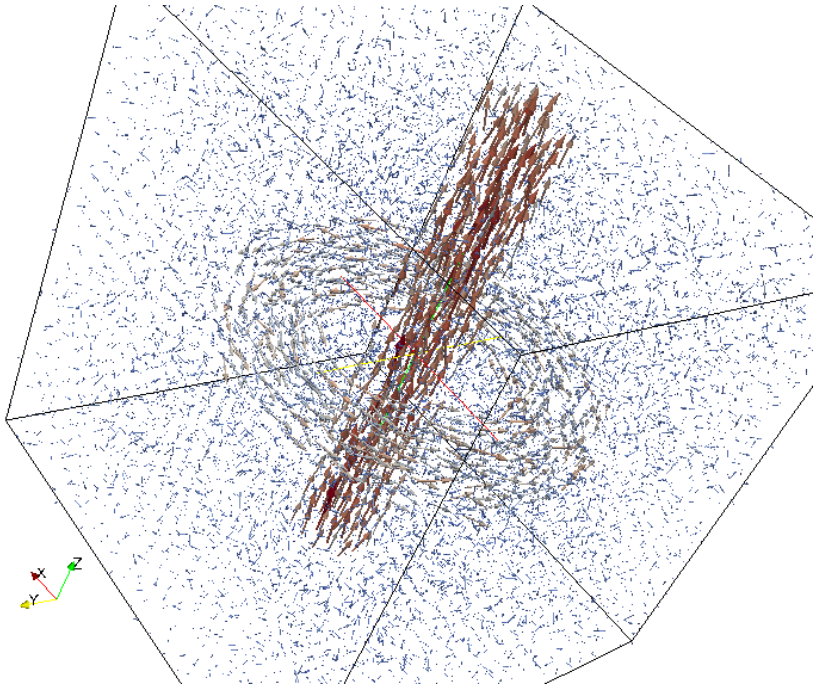


(a) Noisy field (0 dB SNR)

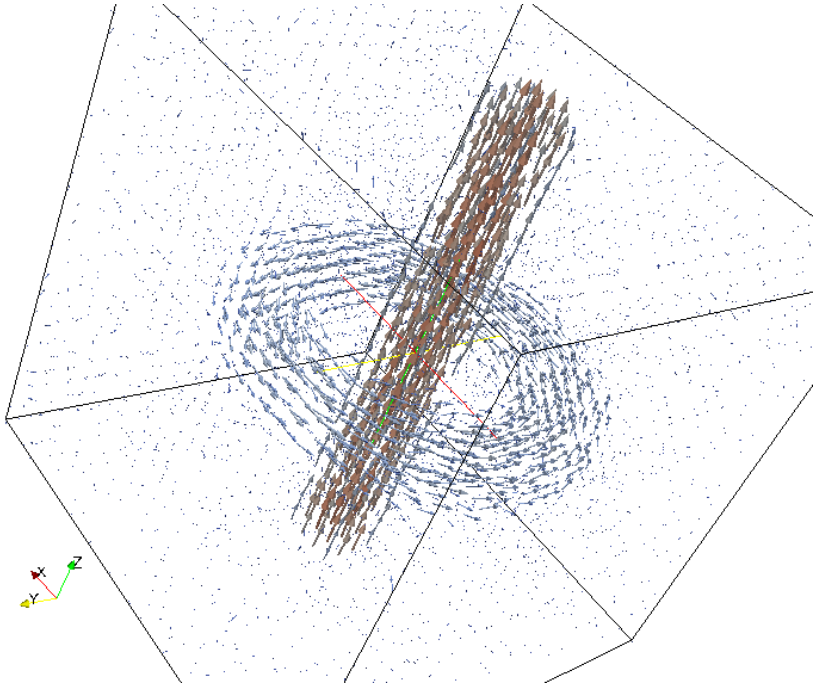


(b) Denoised field, using L_1 regularization (11.70 dB SNR)

Figure 4.3: 'Gradient' phantom; see text for a description of the experiment.

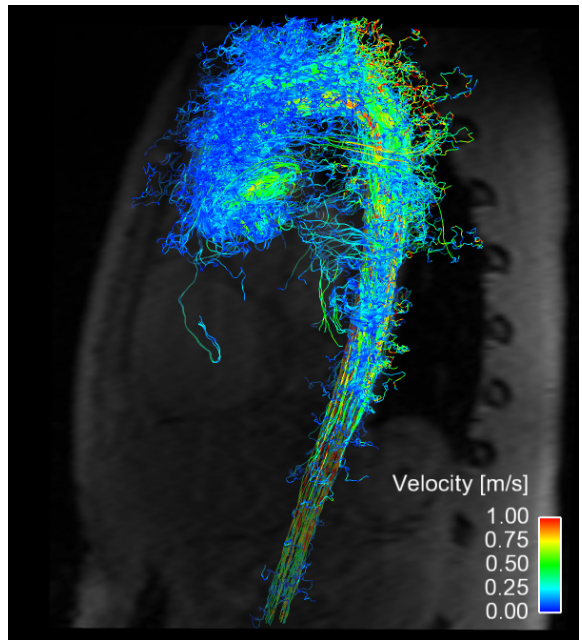


(a) Noisy field (0 dB SNR)

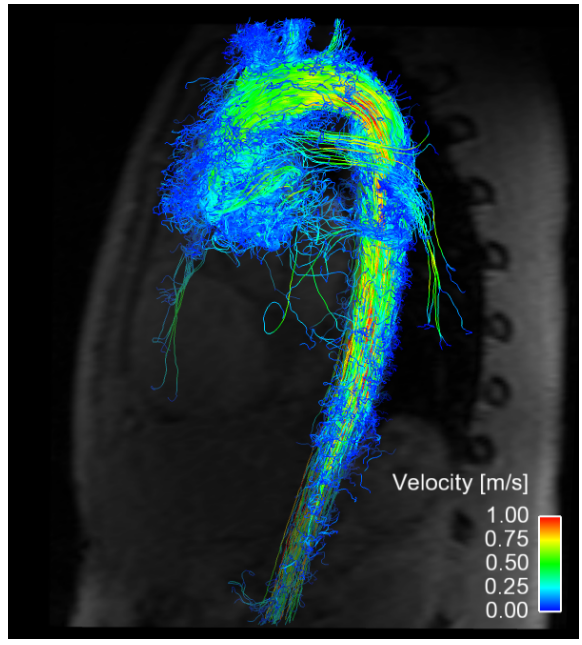


(b) Denoised field, using L_1 regularization (9.01 dB SNR)

Figure 4.4: 'Tube and torus' phantom; see text for a description of the experiment.



(a) Original pathlines



(b) Pathlines after enhancement with L_1 regularization

Figure 4.5: Flow-sensitive MRI recordings of blood flow in the aorta before and after denoising; data and visualizations courtesy of Dr Aurélien F. Stalder.

— 5 —

Conclusion and Outlook

In this thesis we presented new stochastic models for vector fields that are statistically self-similar and rotation-invariant in the sense that is appropriate for physical vectors. We defined these models using a powerful distribution-theoretic formalism that enabled us to characterize the application of certain singular integral operators to different categories of innovations (white noises). The noted stochastic models, introduced in Chapter 3, extend scalar models known as fractional Brownian motions (in the Gaussian case) and fractional stable motions (in the more general α-stable case), of which we also gave a new characterization amenable to such extension. At the same time, our proposed vector models permit us to take account of, and represent, specific directional properties of vector fields (such as irrotational and solenoidal tendencies), which have no counterpart in the scalar case.

For the purpose of formulating the mentioned characterizations, we put forward a general framework for stochastic modelling (outlined in Chapter 1) that builds upon, extends, and gives rigour to the idea of innovation modelling employed in signal processing. Using this framework, constructing a stochastic model becomes a relatively simple matter of choosing a probability model for independent innovations, and an operator that combines them and produces the specific dependency structure that is relevant for the modelling application of interest.

Next, in Chapter 2, we presented a novel family of matrix-valued (tensor) distributions that can be considered the equivalents of homogeneous isotropic distributions in $\mathscr{D}'$, and which share their properties of being closed under scaling, rotation, the Fourier transform, and multiplication and convolution (the latter two when permitted). We also emphasized certain directional properties of these distributions, important in applications and related to

their interaction with Helmholtz-type decompositions of vector fields into curl- and divergence-free components, which are unique to the vector setting.

In the same chapter, we investigated singular integral operators that arise from convolutions with the said distributions and their scalar parallels, employing techniques from complex analysis and distribution theory going back to Schwartz and Gel'fand *& al*. We then introduced the so-called L_p-continuous modifications of these convolution operators, which we later used in the characterization of the random vector fields noted in the first paragraph.

Complementing the theoretical development described above, in Chapter 4 we took inspiration from the same principles of invariance and innovation modelling in order to formulate a general framework and method of solution for the problem of reconstructing a vector field from imperfect observations. Once the general framework was put in place, we specialized to the problem of vector field denoising and enhancement, for which purpose we presented an algorithm that defines the present state of the art in its area of application. We concluded the chapter with experimental validation of the algorithm on synthetic numerical phantoms as well as real-world data from flow-sensitive magnetic resonance imaging (MRI) of blood flow in the thoracic aorta.

Finally, in two appendices, we included some thoughts on probability theory and, separately, reproduced some of our related publications that were not covered in the thesis, or whose presentation differed noticeably from the one given in the main text. In particular, the reader can find there an account of our work on the theory of splines and wavelets that go along with and approximate the operators of Chapter 2, and their applications in estimating the parameters of the models.

While the development in each of the above areas appears to us to have reached a satisfactory state, there exist nevertheless many possibilities for future work, both theoretical and applied. We enumerate a few of these.

It would be desirable, in principle, that the theory of probability on topological vector spaces that we employed in this thesis be developed and refined further in order to permit one to study probabilistic questions that are not countably-determined (such as continuity, boundedness, *&c.*). Although some work in this connection exists (see, for instance, König [Kön97, Kön06c] and also Schwartz [Sch73, Sch81]), this, nevertheless, is a major undertaking for which we did not feel fully qualified. (On a related note, the author has a vaguely-suppressed affinity for more 'constructive' theories of probability, logic, and topology based on algebraic and order-theoretic structures such as lattices but also more elementary partially-ordered sets, which he sees as alternatives to the set-theoretic foundations of topology and measure theory.)

Following up on the previous point, in this thesis we limited ourselves to random models defined by linear transformation of innovations affected by continuous linear operators. A more general theory of non-linear transformations would offer many more possibilities, but the convenient and powerful linear techniques of the theories of topological vector spaces and distributions would no longer apply.

Separately, it would be of interest to find the connection between the operators and distributions introduced in Chapter 2 (in what concerns their invariance properties, more specifically rotation-invariance) and differential forms, which would allow us to extend the theory from the scalar and vector cases considered here to tensors of arbitrary order. We did not have the necessary technical background for doing so.

Finally, it would be desirable to formalize the link(s) between the applied schemes introduced in Chapter 4 and stochastic models in the spirit of those introduced in Chapter 3. More practically, the framework presented in Chapter 4 is fairly general, and many engineering problems can be formulated in it, of which the ones we considered there (vector field denoising and enhancement), are only examples. Other applications, such as more complex forms of vector field reconstruction from indirect measurements, optical flow estimation, resolution refinement, and the like, could easily be conceived and implemented in the same framework. For the numerical implementation of such methods, other algorithms beside the one we presented in Chapter 4 (for instance those utilizing primal-dual methods and advanced techniques of convex analysis), could also be employed.

— 𝔄 —

A Probability Calculus

One cannot inquire into the foundations and nature of mathematics without delving into the question of the operations by which the mathematical activity of the mind is conducted. If one failed to take that into account, then one would be left studying only the language in which mathematics is represented rather than the essence of mathematics.

—L.E.J. Brouwer

This chapter serves to outline the basic probabilistic reasoning and results used in the thesis (mostly presented without proof). To be specific, the results referenced directly in the main text take up no more than a few pages and are all to be found in **§A.6**. The rest of this chapter consists mostly of digressions and occasional pseudo-philosophical musings, in which we have permitted ourselves to indulge on account of the auxiliary nature of the chapter and its relegation to the appendix. We intend, in particular, to draw attention to the justification (sometimes uncertain) of some of the basic tenets and constructions of probability theory in connection with quantifying our knowledge about the physical world.

We emphasize that in this chapter, we plan merely to *revisit* already familiar notions, occasionally from a slightly different perspective, and not to develop them from scratch. Fremlin [Fre04a, Fre03, Fre04b, Fre06, Fre08] and Bogachev [Bog07] achieve the latter purpose remarkably (and exhaustively). Other important references include Schwartz [Sch73], where the general the-

ory of Radon measures on arbitrary topological spaces is developed, and König [Kön08], which takes an unconventional approach aimed to be more naturally adapted to uncountable probability theory. The previous references notwithstanding, in an essential way, all of the probabilistic/measure theoretical results that we depend on in the rest of this thesis (specifically, those of **§A.6**) were already developed and exposed systematically by Gel'fand and Vilenkin in their 1961 volume (English trans. 1964) [GV64].

§A.1 Probability in the language of set theory

A.a We adopt a view of probability theory as the deductive calculus of observation-based (experimental) science, in the same fashion that logic can be viewed as the deductive calculus of mathematical proofs. Thus, in the same manner that logic is concerned with computing the *truth values* of compound statements based on an initial set of known facts, in probability theory we resolve to compute the *probabilities* of derived statements starting from a collection of probabilistic assumptions. We shall not dwell upon the *meaning* of probability; different schools of thought have emerged in this connection, which favour objective vs subjective interpretations and all, but the question is ultimately not mathematical. We shall therefore simply suppose that we are dealing with a model of *incomplete knowledge* as described below (and let philosophers debate the meaning and means of acquisition of *knowledge*).

There exists, in our model of reality, a fundamental set Ω of binary [1] (true or false), *mutually exclusive* statements (*atoms*), such that any other binary statement we might be inclined to consider can be construed as the logical "or" [2] of a number of fundamental statements, i.e. as a subset of Ω. (You can think of these fundamental statements as enumerating distinct *states* of the system under investigation.) $\mathfrak{P}(\Omega)$, the power set of Ω, is therefore the set of all (compound) statements one is able to evaluate based on the truth values of the fundamental ones. [3] We have *incomplete knowledge,* in the sense that we do not know the truth values of all statements with certainty, but, instead, for only some (compound) statements $S \in \mathfrak{S} \subset \mathfrak{P}(\Omega)$, we have at our disposal a number $\mathscr{P}(S)$ in the range $[0, 1]$, which we interpret as the *probability* that S is true (the question of where this number comes from and what it means is a source of debates and divisions among statisticians;

1. We could imagine nominally more general n-ary statements but the outcome would be equivalent.

2. An equivalent dual formulation with logical "and" is also possible.

3. The set of all statements has cardinality $2^{|\Omega|}$. This simple observation is fundamental in information theory.

we address the question by ignoring it). A theory of probability can then be defined as a system of procedures and rules of computation (a *calculus*) that allow us to assign, in an unambiguous way, a numerical value (probability) $\mathscr{P}(T)$ to (some) statements T not in the original set system $\mathfrak{S}$. These rules of computation are the *axioms* of our theory.

A probability is therefore a special kind of real-valued set function, defined on a set system and satisfying a number of axioms.

A.b Set systems. A *set system* in Ω is a non-empty subset of $\mathfrak{P}(\Omega)$. $\mathfrak{P}(\Omega)$, which is itself a set system, is an example of a Boolean algebra with distributive, binary, meet and join operations $\cap$ and $\cup$, a unary complement operation $\complement : T \mapsto T^{\complement}$, a bottom $\varnothing$, a top Ω, and a partial order $\subset$. Accordingly, most set systems that we shall encounter have some order-theoretic structure. Examples of set systems we shall see and use include lattices, algebras, and σ-algebras of sets (all defined below), as well as systems of open, closed, and compact sets in a topological space Ω, which we denote respectively as $\mathrm{Open}(\Omega)$, $\mathrm{Closed}(\Omega)$, and $\mathrm{Comp}(\Omega)$.

A.c Maps between set systems. Given two spaces Ω, Υ, a map $f : \Omega \to \Upsilon$ defines a set map $\mathfrak{P}(\Omega) \to \mathfrak{P}(\Upsilon)$ (also denoted by f) thus:

$$f : S \mapsto \{f(\omega) \in \Upsilon : \omega \in S\}.$$

The above map has an inverse $\mathfrak{P}(\Upsilon) \to \mathfrak{P}(\Omega)$ given by

$$f^{-1} : T \mapsto \{\omega \in \Omega : f(\omega) \in T\}.$$

$f(S)$, $S \subset \Omega$, and $f^{-1}(T)$, $T \subset \Upsilon$, are called, respectively, the *image* and *inverse image* (or *pre-image*) of S and T under f.

Given set systems $\mathfrak{S} \subset \mathfrak{P}(\Omega)$ and $\mathfrak{T} \subset \mathfrak{P}(\Upsilon)$, we say that f is *forward-compatible* $(\mathfrak{S}, \mathfrak{T})$ iff

$$S \in \mathfrak{S} \;\Rightarrow\; f(S) \in \mathfrak{T}$$

and *inverse-compatible* $(\mathfrak{S}, \mathfrak{T})$ iff

$$f(S) \in \mathfrak{T} \;\Rightarrow\; S \in \mathfrak{S}.$$

Clearly, f is forward-compatible $(\mathfrak{S}, \mathfrak{T})$ iff its inverse f^{-1} is inverse-compatible $(\mathfrak{T}, \mathfrak{S})$. f is said to *preserve* or be *fully-compatible* $(\mathfrak{S}, \mathfrak{T})$ iff it is both forward- and inverse-compatible $(\mathfrak{S}, \mathfrak{T})$ or, what is the same, iff

$$S \in \mathfrak{S} \;\Leftrightarrow\; f(S) \in \mathfrak{T}.$$

In particular, a map $f : \Omega \to \Omega$ is said to preserve $\mathfrak{S} \subset \mathfrak{P}(\Omega)$ iff it is fully-compatible $(\mathfrak{S}, \mathfrak{S})$ (compare with structure-preserving maps in ***2.j** and measurable maps below).

A.d Some examples of questionable relevance. In a topological space Ω, a map that is inverse-compatible with either Open(Ω) or Closed(Ω) is precisely a continuous map. A map that is forward-compatible Closed(Ω) is called a closed map, and one similarly defines an open map as being forward-compatible Open(Ω). A *homeomorphism* on Ω is a bijection that preserves Open(Ω), or, equivalently, a bijection that preserves Closed(Ω). Similarly, we define a *k-homeomorphism* as a bijection that preserves Comp(Ω). Every homeomorphism is also a k-homeomorphism (the continuous image of a compact set is compact), and if Ω is a Hausdorff k-space,[4] then every k-homeomorphism on Ω is a homeomorphism, as we show below.

We shall later define a *measurable map* $\Omega \to \Upsilon$ as one that is inverse-compatible with the pair of σ-algebras at hand.

***A.e Proposition.** A bijection h on a Hausdorff k-space is a homeomorphism if and only if for any set $K \subset \Omega$,

$$K \in \text{Comp}(\Omega) \;\Leftrightarrow\; h(K) \in \text{Comp}(\Omega).$$

Proof. The 'only if' part is clear (continuous maps preserve compact sets). To prove the 'if' part, take any compactum $K \subset \Omega$. Then $h(K)$ is compact by assumption. Thus any closed set $C \subset h(K)$ is also compact, being a closed subset of a compactum. Therefore $h^{-1}(C)$ is compact by assumption, and hence closed (Ω is Hausdorff). Thus the inverse image of any closed set C under the restriction $h|_K$ is closed, i.e. $h|_K$ is continuous. It is known that a map on a Hausdorff k-space is continuous iff its restriction to an arbitrary compactum is continuous. It therefore follows from the continuity of $h|_K$ that h itself is continuous. Similarly, one shows that h^{-1} is continuous, completing the proof. ∎

§A.2 Finitely additive probability measures

A.f Elementary axioms. We produce below a basic system of axioms based on logical intuition.

(P1) $\mathscr{P}(S \cup T) + \mathscr{P}(S \cap T) = \mathscr{P}(S) + \mathscr{P}(T)$; (modularity)

4. A Hausdorff space is a space in which any two distinct points are separated by disjoint neighbourhoods. Compacta of a Hausdorff space are closed. A k-space is a topological space verifying the condition

$$S \in \text{Closed}(\Omega) \;\Leftrightarrow\; \forall K \in \text{Comp}(\Omega)\ S \cap K \in \text{Closed}(\Omega).$$

A map $f : \Omega \to \Upsilon$, where Ω is a k-space, is continuous iff its restriction to every compact set of Ω is continuous [Bro64].

(P2) $S \subset T \;\Rightarrow\; \mathscr{P}(S) \leq \mathscr{P}(T)$; (isotonity)

(P3) $\mathscr{P}(\varnothing) = 0$; (neutral element)

(P4) $\mathscr{P}(\Omega) = 1$. (normalization)

The latter two axioms indicate our belief that *some* (usu. unknown) statement in Ω is true, and set the convention to assign the numerical values 0 and 1 to almost surely false and true statements respectively (it goes without saying that an axiom is *a priori* meaningful only if $\mathscr{P}$ is defined for the sets appearing in the axiom formula).

It is of course necessary for our initial set of probabilities on $\mathfrak{S}$ to be *consistent* with respect to the system of axioms, meaning that applying the axioms to the probabilities does not lead to contradictions or different probabilities for the same statement.

The probability assignment $\mathscr{P}$ is therefore a function with values in $[0,1]$ initially defined on some set system $\mathfrak{S} \subset \mathfrak{P}(\Omega)$ and satisfying the axioms; the object of the calculus of probabilities is consistently to *extend* the domain of $\mathscr{P}$ to some larger set system $\mathfrak{T} \supset \mathfrak{S}$, in such a way that the axioms remain true for the extension. In practice, one usually fixes the domain $\mathfrak{S}$ of the function $\mathscr{P} : \mathfrak{S} \to [0,1]$ and the set of axioms it is required to satisfy, and then seeks larger set systems $\mathfrak{T} \supset \mathfrak{S}$ to which $\mathscr{P}$ admits a *unique* extension consistent with the axioms.

A.g Lattices and algebras of sets. We shall take as our minimum set of axioms the above four. It is assumed in probability theory that the domain $\mathfrak{S}$ has, as a minimum, the structure of a *bounded lattice,* that is, it is closed under $\cap\cup$ (finite intersection and union) and has a top (Ω) and a bottom ($\varnothing$). This assumption is not always valid in practice: depending on the nature of our experimental data we may be able to assign empirical probabilities to two statements S, T separately, but not to their co-occurence $S \cap T$. In fact, even $\Omega \in \mathfrak{S}$ may be doubtful, since we cannot always be sure to have identified the entire set of fundamental statements (possibilities); hence we cannot always know the negation (Ω complement) of statements in $\mathfrak{S}$ (it is interesting to compare this last situation with the state of affairs in intuitionistic logic, where one disallows proofs based on the principle of excluded middle). Nevertheless, the bounded lattice assumption is virtually universally adopted by measure theorists and probability theorists alike, and we shall adopt it too, because it is convenient if unrealistic.

Once we suppose that $\mathfrak{S}$ is a lattice (or, at least, closed under intersection), and that $\mathscr{P}$ is consistent with the axioms on $\mathfrak{S}$, we can then extend $\mathscr{P}$ consistently and uniquely to the *algebra* generated by $\mathfrak{S}$, in symbols $\mathrm{Alg}(\mathfrak{S})$, defined as the $\cap \cup \complement$ closure of $\mathfrak{S}$ ($\complement$ stands for taking complements in Ω).

This extension is obtained by iterative application of the following formulae (derived from the axioms) to assign the left-hand side probabilities:

$$\begin{aligned}\mathscr{P}(S \cup T) &= \mathscr{P}(S) + \mathscr{P}(T) - \mathscr{P}(S \cap T),\\ \mathscr{P}(S^{\complement}) &= 1 - \mathscr{P}(S),\\ \mathscr{P}(S^{\complement} \cap T) &= \mathscr{P}(S \cup T) - \mathscr{P}(S) = \mathscr{P}(T) - \mathscr{P}(S \cap T).\end{aligned}$$

This proves consistency and uniqueness, if the original system is consistent, because the above formulae are propositions when restricted to $\mathfrak{S}$ and they allow us to compute the probability of any set in $\mathrm{Alg}(\mathfrak{S})$ in a finite number of steps (one should also verify that the values thus obtained lie in $[0, 1]$, but this is an easy consequence of Axioms (P2)–(P4); in particular, it is only here that Axiom (P2) is needed); for the full construction see Hausdorff [Hau14, pp. 451, 452] or Pettis [Pet51, Theorem 1.2].

We record this result as a small

A.h Finite extension theorem. A function $\mathscr{P} : \mathfrak{S} \to [0, 1]$ satisfying Axioms (P1)–(P4), where $\mathfrak{S}$ is a bounded lattice, has a unique extension $\mathrm{Alg}(\mathfrak{S}) \to [0, 1]$ that fulfils Axioms (P1)–(P4) on $\mathrm{Alg}(\mathfrak{S})$.

Since, by the above theorem, an extension of $\mathscr{P}$ from a bounded lattice to an algebra is always available and unique, we shall henceforth *always* assume that the initial set $\mathfrak{S}$ over which $\mathscr{P}$ is defined is an algebra. This motivates the

A.i Definition. A set function $\mathfrak{S} \to [0, 1]$ where $\mathfrak{S}$ is an algebra of sets is called an *elementary* (or *finitely-additive*) *probability measure* iff it satisfies Axioms (P1)–(P4). We shall also use the term *elementary stochastic model* to refer to the same. The set $\Omega = \bigcup_{S \in \mathfrak{S}} S$ equipped with the algebra $\mathfrak{S} \subset \mathfrak{P}(\Omega)$ is called an *elementary probability space.*

A.j More definitions. Let $\mathscr{P}, \mathscr{Q}$ be two elementary probabilities with respective domains $\mathfrak{S}, \mathfrak{T}$ and with $\Omega_{\mathscr{P}} = \bigcup_{S \in \mathfrak{S}} S = \bigcup_{T \in \mathfrak{T}} T = \Omega_{\mathscr{Q}}$. $\mathscr{P}, \mathscr{Q}$ are then *comparable* on $\mathfrak{S} \cap \mathfrak{T}$ (note that the intersection is again an algebra, and the restrictions of $\mathscr{P}, \mathscr{Q}$ to it are elementary probabilities). $\mathscr{P}$ is said to be *(strictly) richer* than $\mathscr{Q}$ iff $\mathfrak{T} \subsetneq \mathfrak{S}$ (it is a richer model because it can evaluate more statements). A probability $\mathscr{P} : \mathfrak{S} \to [0, 1]$ that is richer than some $\mathscr{Q} : \mathfrak{T} \to [0, 1]$ is said to be an *extension* of $\mathscr{Q}$ (and $\mathscr{Q}$ a *restriction* of $\mathscr{P}$) iff $\mathscr{P}(T) = \mathscr{Q}(T)$ for all $T \in \mathfrak{T} \subset \mathfrak{S}$. Note that, trivially, the restriction of an elementary probability $\mathscr{P}$ to *any* sub-algebra $\mathfrak{T} \subset \mathfrak{S}_{\mathscr{P}}$ is always a well-defined elementary probability in its own right.

Given $U \in \mathfrak{S}$ with $\mathscr{P}(U) > 0$, let $\mathfrak{T} = \{S \in \mathfrak{S} : S \subset U\}$. $\mathfrak{T}$ is then an algebra

in U. For $T \in \mathfrak{T}$ define

$$\mathscr{P}(T|U) = \frac{\mathscr{P}(T)}{\mathscr{P}(U)}$$

and extend $\mathscr{P}(\cdot|U)$ to all $S \in \mathfrak{S}$ by

$$\mathscr{P}(S|U) := \mathscr{P}(S \cap U|U).$$

One can then directly verify that $\mathscr{P}(\cdot|U)$ satisfies Axioms (P1)–(P4) on $\mathfrak{T}$ and by extension on $\mathfrak{S}$, meaning that $\mathscr{P}(\cdot|U)$ is a valid elementary stochastic model on $\mathfrak{S}$ (the *conditioning* of the probability $\mathscr{P}$ on U).

§A.3 Countably-additive probability measures

A.k In models with a finite number of fundamental statements (and, consequently, a finite number of compound statements), the above 'finitary' axioms and rules of computation seem quite adequate, and using them we can essentially reduce probability theory to combinatorics and counting;[5] including, perhaps, combinatorial asymptotics if we wish to study the large-scale behaviour of systems (where by large-scale we mean statements concerning a great many fundamental statements at the same time). It is unclear to the author if such a finite theory should not suffice in practice (or even if anything *but* a finite theory should be accepted on logical and philosophical, as well as practical, grounds), but, at any rate, due to the commonly perceived need for a formalism that can deal with infinite—even uncountable—cardinalities arising out of axiomatic set theory, we shall very soon (pretend to) put our philosophical doubts and misgivings aside, switch hats, and become well-mannered formalists happy to play the game of infinity for the better part of this thesis.

But before quite doing so, let us warn the reader that the current state of the theory of probability, in its application to uncountable infinite systems of mathematical interest, is not entirely satisfactory (and we shall not attempt to improve it). In what concerns us, the difficulty is essentially this. There exist two mathematical theories, namely measure theory and point-set topology, that deal with set systems and maps between them. On the surface, the two theories look very similar: both start from the notion of a set system that is closed under some specific set-theoretical operations (σ-algebras vs

5. The careful reader might ask if there is even a need for an infinitum, much less a continuum, of probability values in $[0, 1]$ in the finitist picture. Indeed there is none; a finite set of probability values would have sufficed. But we do not wish additionally to complicate the presentation by following up on this point.

topologies), and study maps that are *compatible* with this structure (measurable functions vs continuous functions). Quite naturally, therefore, one wonders if sets that appear in topology can be *measured* by measure theory. In attempting to answer this seemingly innocent question, all hell breaks loose. The fundamental problem is that the basic methods and procedures of abstract measure theory are *sequential* (countable), while topological sets are frequently defined by way of *non-sequential* (uncountable) procedures. [6] There exist some partial solutions, but they typically require us to restrict the class of topological spaces and set systems we work with.

A.I Limit axioms. The first and most commonly referenced axiomatization of probability theory was put forward by Kolmogorov [Kol33] using the language of measure and integration, and the previous observations therefore apply to its interaction with uncountable procedures and topology. In order to introduce infinite procedures (meaning limits) in his theory, Kolmogorov extended the set of axioms given previously [7] by introducing an axiom of *continuity*. Let us present here two such axioms for a probability $\mathscr{P} : \mathfrak{S} \to [0, 1]$:

(P_5) For any countable increasing sequence (S_k) of sets in $\mathfrak{S}$ with limit $\bigcup_k S_k = S \in \mathfrak{S}$, $\lim_{k\to\infty} \mathscr{P}(S_k) = \mathscr{P}(S)$. (upward σ-continuity)

(P_5') For any countable decreasing sequence (S_k) of sets in $\mathfrak{S}$ with limit $\bigcap_k S_k = S \in \mathfrak{S}$, $\lim_{k\to\infty} \mathscr{P}(S_k) = \mathscr{P}(S)$. (downward σ-continuity)

In the appellation of the above axioms, σ stands for *sequential* or *countable*. Similar notions can be defined in the *non-sequential* or *uncountable* case (typically referred to using the letter τ), by replacing sequences with arbitrary directed set systems. We shall not bother the reader with these generalizations.

With regard to the practical relevance and empirical justification of these additional axioms we quote Kolmogorov himself [Kol56]:

> Since the new axiom is essential for infinite fields of probability only, it is almost impossible to elucidate its empirical meaning [as has been done for previous axioms]. For, in describing any observable random process we can obtain only finite fields of probability. Infinite fields of probability occur only as idealized models of real random processes. *We limit ourselves, arbitrarily, to only*

6. König [Kön97, Kön06c] candidly identifies these imperfections better than we can hope to have done here, and goes on to propose a new (but compatible) measure theory that attempts to overcome these limitations. We shall, however, restrict ourselves to standard measure theory.

7. His set of axioms was in fact equivalent but not exactly the same.

> *those models which satisfy [the continuity axiom].* This limitation has been found expedient in researches of the most diverse sort.

It is straightforward to verify that, given the previous axioms, the new axioms (P_5) and (P_5') are in effect equivalent:

A.m Lemma. Let $\mathscr{P} : \mathfrak{S} \to [0,1]$ satisfy Axioms (P_1)–(P_4). Then

$$(P_5) \;\Leftrightarrow\; (P_5').$$

Proof. We show the forward direction. The other direction is proved similarly. Let (S_k) be any decreasing sequence with limit S. Then $(S_k^{\complement})$ is increasing with limit $S^{\complement}$, and from $\mathscr{P}(S_k) = 1 - \mathscr{P}(S_k^{\complement})$ and upward σ-continuity we have $\lim_{k\to\infty} \mathscr{P}(S_k) = 1 - \lim_{k\to\infty} \mathscr{P}(S_k^{\complement}) = 1 - \mathscr{P}(S^{\complement}) = \mathscr{P}(S)$; hence $\mathscr{P}$ is downward σ-continuous. ∎

The principal utility of the above axioms lies in that they allow us to extend the calculus of probabilities to countably-constructed set systems. Let us denote by $\sigma\,\mathrm{Alg}(\mathfrak{S})$ the σ-algebra generated by $\mathfrak{S}$ (where, we recall, $\mathfrak{S}$ is an algebra); by this we mean that $\sigma\,\mathrm{Alg}(\mathfrak{S})$ is the smallest set system in $\mathfrak{P}(\Omega)$ that includes $\mathfrak{S}$ and is closed under *countable* unions and intersections, as well as taking complements in Ω. In this case we have the following

A.n Countable extension theorem (Hahn–Kolmogorov). Let $\mathscr{P} : \mathfrak{S} \to [0,1]$ be a finitely-additive probability measure on an algebra $\mathfrak{S}$ and assume that $\mathscr{P}$ additionally fulfils Axiom (P_5). Then $\mathscr{P}$ can be extended to the σ-algebra generated by $\mathfrak{S}$ while retaining Axioms (P_1)–(P_5), and this extension is unique.

We shall not give a proof of the above well-known theorem as such a proof can be found in most standard references on measure theory. Instead, once again we quote Kolmogorov [Kol56] on the empirical significance and relevance of this extension:

> Even if the sets (events) [S of $\mathfrak{S}$] can be interpreted as actual and (perhaps only approximately) observable events, it does not, of course, follow from this that the sets of the extended field [$\sigma\,\mathrm{Alg}(\mathfrak{S})$] reasonably admit of such an interpretation.
>
> Thus there is the possibility that while a field of probability [$\mathfrak{S}$] may be regarded as the image (idealized, however) of actual random events, the extended field of probability [$\sigma\,\mathrm{Alg}(\mathfrak{S})$] will still remain merely a mathematical structure.
>
> Thus sets of [$\sigma\,\mathrm{Alg}(\mathfrak{S})$] are generally merely ideal events to which nothing corresponds in the outside world. However, if reason-

> ing which utilizes the probabilities of such ideal events leads us to a determination of the probability of an actual event of $[\mathfrak{S}]$, then, from an empirical point of view also, this determination will automatically fail to be contradictory.

By the previous theorem, we can always extend a real set function defined on an algebra and satisfying Axioms (P1)–(P5) to the σ-algebra generated by $\mathfrak{S}$ in a unique fashion, while maintaining consistency with the Axioms. This motivates the next

A.o Definition. A set function $\mathfrak{S} \to [0,1]$ satisfying Axioms (P1)–(P5), where $\mathfrak{S}$ is now a σ-algebra, is called a *non-elementary,* or *countably-additive, probability measure* or *stochastic model.* In the absence of indications to the contrary, the unqualified term *probability measure* is henceforth understood to refer to a non-elementary probability.

Note that this new definition is a specialization of the previous one given in **A.i**: any probability in the sense just described is automatically also an elementary probability (since every σ-algebra is also an algebra).

Non-elementary equivalents of the notions introduced in **A.j** are defined similarly.

Having answered the question of extending a probability measure from an algebra $\mathfrak{S}$ to the σ-algebra $\sigma\,\mathrm{Alg}(\mathfrak{S})$ generated by it with the aid of Axioms (P5), (P5') and Theorem **A.n**, let us now entertain extensions of (countably-additive) probability measures to larger σ-algebras. The next theorem, which we cite from Bogachev [Bog07, Theorem 1.12.14] without proof, provides important insight into the questions of existence and uniqueness of such extensions. We shall need two definitions first.

A.p Definition. Given a finitely-additive probability measure $\mathscr{P} : \mathfrak{S} \to [0,1]$ on a set system $\mathfrak{S}$ with $\bigcup_{S\in\mathfrak{S}} S = \Omega$, we define set functions $\mathscr{P}_*, \mathscr{P}^* : \mathfrak{P}(\Omega) \to [0,1]$ by the identities

$$\mathscr{P}_*(T) = \sup\{\mathscr{P}(S) : \mathfrak{S} \ni S \subset T\};$$
$$\mathscr{P}^*(T) = \inf\{\mathscr{P}(S) : \mathfrak{S} \ni S \supset T\}.$$

More generally, given a set system $\mathfrak{M} \subset \mathfrak{S}$ we define

$$\mathscr{P}_{*\mathfrak{M}}(T) = \sup\{\mathscr{P}(S) : \mathfrak{M} \ni S \subset T\};$$
$$\mathscr{P}^{*\mathfrak{M}}(T) = \inf\{\mathscr{P}(S) : \mathfrak{M} \ni S \supset T\}.$$

The former two set functions are related by the following well-known formula.

A.q Lemma (duality formula). With the same notation as in the above definition,

$$\mathscr{P}_*(T) + \mathscr{P}^*(T^{\complement}) = 1, \quad \text{for all } T \subset \Omega.$$

Proof. $\mathscr{P}_*(T) = \sup\{\mathscr{P}(S) : \mathfrak{S} \ni S \subset T\} = \sup\{1 - \mathscr{P}(S^{\complement}) : \mathfrak{S} \ni S \subset T\} = 1 - \inf\{\mathscr{P}(S^{\complement}) : \mathfrak{S} \ni S^{\complement} \supset T\} = 1 - \mathscr{P}^*(T)$. ∎

We also have the following rather obvious lemma.

A.r Lemma. Let $\mathscr{P}$ be a finitely-additive probability measure on some algebra $\mathfrak{S}$ and let $\mathfrak{M}, \mathfrak{N} \subset \mathfrak{S}$. Then,

$$\mathfrak{M} \subset \mathfrak{N} \quad \Rightarrow \quad \mathscr{P}_{*\mathfrak{M}} \leq \mathscr{P}_{*\mathfrak{N}} \text{ and } \mathscr{P}_{*\mathfrak{M}} \geq \mathscr{P}_{*\mathfrak{N}}.$$

A.s Theorem (existence of extensions). Given a probability measure $\mathscr{P}$ with domain $\mathfrak{S}$ and a set $T \notin \mathfrak{S}$, there exists probability measure $\mathscr{P}'$ on $\sigma\,\mathrm{Alg}(\mathfrak{S} \cup \{T\})$ extending $\mathscr{P}$ with

$$\mathscr{P}'(T) = \gamma,$$

for any γ such that

$$\mathscr{P}_*(T) \leq \gamma \leq \mathscr{P}^*(T).$$

Conversely, any extension $\mathscr{P}'$ of $\mathscr{P}$ to $\sigma\,\mathrm{Alg}(\mathfrak{S} \cup \{T\})$ must satisfy

$$\mathscr{P}_*(T) \leq \mathscr{P}'(T) \leq \mathscr{P}^*(T).$$

The converse part does not appear in Bogachev [Bog07, Theorem 1.12.14] but it is easy enough to show: we have $\mathscr{P}'_* \geq \mathscr{P}_*$, $\mathscr{P}'^* \leq \mathscr{P}^*$, and $\mathscr{P}'_* \leq \mathscr{P}' \leq \mathscr{P}'^*$ (the last relation being meaningful only on the domain of $\mathscr{P}'$ which is $= \sigma\,\mathrm{Alg}(\mathfrak{S} \cup \{T\})$), from where the desired result immediately follows. We now look at some of the consequences of the above theorem.

A.t Corollary. With $\mathscr{P}, \mathfrak{S}, T$ as in the above theorem, $\mathscr{P}$ has a unique extension to $\sigma\,\mathrm{Alg}(\mathfrak{S} \cup \{T\})$ iff

$$\mathscr{P}_*(T) = \mathscr{P}^*(T).$$

In this case, the extension satisfies $\mathscr{P}(T) = \mathscr{P}_*(T) = \mathscr{P}^*(T)$.

A.u Corollary. Given a probability measure $\mathscr{P}$ on $\mathfrak{S}$ and a set system $\mathfrak{T}$ such that

$$\mathscr{P}_*(T) = \mathscr{P}^*(T)$$

for all $T \in \mathfrak{T}$, there exists at most a unique extension of $\mathscr{P}$ to $\sigma\,\mathrm{Alg}(\mathfrak{S} \cup \mathfrak{T})$. This extension, if it exists, satisfies $\mathscr{P}(T) = \mathscr{P}_*(T) = \mathscr{P}^*(T)$ for all $T \in \mathfrak{T}$.

If the family $\mathfrak{T}$ is well-ordered (which it always is if we admit choice), it is possible to show that the extension hypothesized in the last corollary always exists provided that the sets in $\mathfrak{T}$ are disjoint.

The above results motivate the following

A.v Definition. After König [Kön97], given a finitely-additive probability measure $\mathscr{P}$ on $\mathfrak{S}$ (where $\mathfrak{S}$ is an algebra but not necessarily a σ-algebra), and set systems $\mathfrak{M}, \mathfrak{T} \subset \mathfrak{S}$, the probability $\mathscr{P}$ is said to be *inner regular* $\mathfrak{M}$ *at* $\mathfrak{T}$ iff

$$\mathscr{P}(T) = \mathscr{P}_{*\mathfrak{M}}(T) \quad \text{for all } S \in \mathfrak{T}.$$

Similarly, $\mathscr{P}$ is said to be *outer regular* $\mathfrak{M}$ *at* $\mathfrak{T}$ iff

$$\mathscr{P}(T) = \mathscr{P}^{*\mathfrak{M}}(T) \quad \text{for all } S \in \mathfrak{T}.$$

In the usual case of $\mathfrak{T} = \mathfrak{S}$ we call $\mathscr{P}$ simply *inner/outer regular* $\mathfrak{M}$. Note that a (finitely- or countably-additive) probability measure on $\mathfrak{S}$ is always inner and outer regular $\mathfrak{S}$.

The next lemma is obvious.

A.w Lemma. Let $\mathfrak{M} \subset \mathfrak{N} \subset \mathfrak{S}$. Then,

$$\mathscr{P} \text{ inner/outer regular } \mathfrak{M} \;\Rightarrow\; \mathscr{P} \text{ inner/outer regular } \mathfrak{N}.$$

In addition, the same reasoning applied in the proof of **A.q** also proves the

A.x Lemma. Let $\complement\mathfrak{M}$ denote the set system $\{M^\complement : M \in \mathfrak{M}\}$ (note that $\complement\complement\mathfrak{M} = \mathfrak{M}$). Then,

$$\mathscr{P} \text{ inner regular } \mathfrak{M} \;\Leftrightarrow\; \mathscr{P} \text{ outer regular } \complement\mathfrak{M}.$$

A.y Corollary. Assume $\mathfrak{M}, \mathfrak{N} \subset \mathfrak{S}$, $\mathfrak{N} \supset \complement\mathfrak{M}$. Then,

$$\mathscr{P} \text{ inner regular } \mathfrak{M} \;\Rightarrow\; \mathscr{P} \text{ outer regular } \mathfrak{N};$$
$$\mathscr{P} \text{ outer regular } \mathfrak{M} \;\Rightarrow\; \mathscr{P} \text{ inner regular } \mathfrak{N}.$$

The following examples are important.

A.z Borel and Radon measures. Let Ω be a Hausdorff topological space with topology $\mathfrak{O} = \mathrm{Open}(\Omega)$. We define the *Borel σ-algebra* on Ω as

$$\mathrm{Borel}(\Omega) = \sigma\,\mathrm{Alg}(\mathfrak{O}).$$

A probability measure defined on $\mathrm{Borel}(\Omega)$ is in turn called *Borel.* A Borel probability measure is *Radon* iff it is inner regular $\mathfrak{K} = \mathrm{Comp}(\Omega)$, where $\mathrm{Comp}(\Omega)$ denotes the system of compact sets in the topological space Ω. It follows from the previous corollary that a Radon measure is outer regular $\mathfrak{O}$, since in a Hausdorff space every compactum is closed. If, in addition, the space Ω is second-countable and locally compact, or, more generally, Polish (separable and completely metrizable), then every Borel probability measure on Ω is automatically Radon. This is a consequence of the fact that in such a space the Borel σ-algebra is identical to the σ-algebra generated by compact sets (in particular, Borel measures on $\mathbb{R}^n$ are Radon). Further generalizations of this result are possible but we shall omit them.

§A.4 Some additional topics

A.aa Transformation of measures by mappings. We now describe an important mechanism for using existing probability measures to define new ones. Given a probability measure $\mathscr{P}$ on $\mathfrak{S} \subset \mathscr{P}(\Omega)$ and a map $f : \Omega \to \Upsilon$, one may wish to evaluate statements of the form $\{f(\omega) \in T\}$ for subsets T of Υ on the basis of the probabilities assigned to statements in $\mathfrak{S}$. This is possible for any T whose inverse image under f,

$$f^{-1}(T) = \{\omega \in \Omega : f(\omega) \in T\}$$

belongs to $\mathfrak{S}$. The probability of such a statement T is given by $\mathscr{P} \circ f^{-1}(T)$. This motivates the following

A.ab Definition. Let

$$\mathfrak{T} = \{T \in \mathfrak{P}(\Upsilon) : f^{-1}(T) \in \mathfrak{S}\}$$

be the set system of all statements in Υ whose inverse images under f belong to $\mathfrak{S}$ (note that $\mathfrak{T}$ is a (σ-) algebra if $\mathfrak{S}$ is one). The *push-forward* of the measure $\mathscr{P}$ through f is the probability measure $f(\mathscr{P})$ on $\mathfrak{T}$ defined by the identity

$$f(\mathscr{P})(T) = \mathscr{P} \circ f^{-1}(T) = \mathscr{P}(f^{-1}(T)).$$

Somewhat differently, one may be given the algebras $\mathfrak{S} \subset \mathfrak{P}(\Omega)$ and $\mathfrak{T} \subset \mathscr{P}(\Upsilon)$ in advance, in which case a function $f : \Omega \to \Upsilon$ is said $(\mathfrak{S}, \mathfrak{T})$-*measurable* iff for any $T \in \mathfrak{T}$ we have $f^{-1}(T) \in \mathfrak{S}$. The push-forward $f(\mathscr{P})$ of a probability $\mathscr{P}$ on $\mathfrak{S}$ is then well-defined as a probability measure on $\mathfrak{T}$. The previous notions are defined similarly for elementary measures.

The above notions appear, for instance, in the definition of

A.ac Marginals and product measures. Assume given two σ-algebras $\mathfrak{S}_1, \mathfrak{S}_2$ with $\bigcup_{S \in \mathfrak{S}_1} S = \Omega_1$ and $\bigcup_{S \in \mathfrak{S}_2} S = \Omega_2$. Let $\mathfrak{S}$ be a σ-algebra on $\Omega_1 \times \Omega_2$ such that each canonical projection

$$\pi_i : \Omega_1 \times \Omega_2 \to \Omega_i : (\omega_1, \omega_2) \mapsto \omega_i,$$

$i = 1, 2$, is $(\mathfrak{S}, \mathfrak{S}_i)$-measurable. The smallest such σ-algebra on $\Omega_1 \times \Omega_2$ is called the *product* of $\mathfrak{S}_1$ and $\mathfrak{S}_2$ and denoted as $\mathfrak{S}_1 \otimes \mathfrak{S}_2$. Let $\mathscr{P}$ be a probability measure on $\mathfrak{S} \supset \mathfrak{S}_1 \otimes \mathfrak{S}_2$. The push-forward of $\mathscr{P}$ through π_i, $i = 1, 2$, is then the *marginal* of $\mathscr{P}$ on Ω_i.

Given the measures $\mathscr{P}_1, \mathscr{P}_2$ respectively on $\mathfrak{S}_1$ and $\mathfrak{S}_2$, one can always construct a measure $\mathscr{P}_1 \otimes \mathscr{P}_2$, called the *product measure,* with $\mathscr{P}_1, \mathscr{P}_2$ as its marginals, thus:

$$\mathscr{P}_1 \otimes \mathscr{P}_2(S_1 \times S_2) = \mathscr{P}_1(S_1)\mathscr{P}_2(S_2), \quad \text{for any pair } S_i \in \mathfrak{S}_i,\ i = 1, 2.$$

As a minimum, $\mathscr{P}_1 \otimes \mathscr{P}_2$ can be extended to $\mathfrak{S}_1 \otimes \mathfrak{S}_2$ (Hahn-Kolmogorov).

Note that, as a general rule, the product measure is simply one among a multitude of measures that can be defined on $\mathfrak{S} \supset \mathfrak{S}_1 \otimes \mathfrak{S}_2$ so as to have $\mathscr{P}_1, \mathscr{P}_2$ as their marginals.

Conversely, if for some measure $\mathscr{P}$ on $\mathfrak{S} \subset \mathfrak{P}(\Omega_1 \times \Omega_2)$ one can find measures $\mathscr{P}_1, \mathscr{P}_2$, respectively on Ω_1 and Ω_2, such that

$$\mathscr{P}(S_1 \times S_2) = \mathscr{P}_1(S_1)\mathscr{P}_2(S_2)$$

for any pair $S_i \in \mathfrak{S}_i$ $i = 1, 2$, then the marginals $\mathscr{P}_1, \mathscr{P}_2$ of $\mathscr{P}$ are said to be *independent.*

The above definitions generalize trivially to finite product spaces. Later, in **§A.5** and **§A.6**, we shall revisit some of them in the context of infinite-dimensional spaces.

A.ad Mathematical expectations. *Integrals* with respect to probability measures can be defined in the usual manner of measure theory (i.e. by means of limits of integrals of simple functions), and the basic theorems of integration theory also apply here. Details can be found in any text on measure theory.[8] We simply recall the basic definitions here. Given a σ-algebra $\mathfrak{S}$ with $\bigcup_{S\in\mathfrak{S}} S = \Omega$, for any finite set system $\mathfrak{T} \subset \mathfrak{S}$ we define the class of $\mathfrak{T}$-*simple functions* as

$$\mathrm{SF}_{\mathfrak{T}} = \Big\{ \sum_{T\in\mathfrak{T}} s_T \mathbb{1}_T : s_T \in \mathbb{R} \text{ for each } T \in \mathfrak{T} \Big\} \subset \mathbb{R}^{\Omega},$$

where $\mathbb{1}_T$ denotes the indicator function of the set T. The integral of a $\mathfrak{T}$-simple function $s = \sum_{T\in\mathfrak{T}} s_T \mathbb{1}_T \in \mathrm{SF}_{\mathfrak{T}}$ with respect to a measure $\mathscr{P}$ on $\mathfrak{S}$ is defined as

$$\int_\Omega s(\omega)\mathscr{P}(\mathrm{d}\omega) = \sum_{T\in\mathfrak{T}} s_T \mathscr{P}(T).$$

8. The reader may—rightfully—be wondering why we have discussed the elementary axioms of probability theory in some detail, but made at best only brief mention of important notions such as expectations and conditioning. The reasons are two-fold. First, and more innocently, our interest in this thesis is primarily in *constructing* and *characterizing* probability measures; our overview of probability theory is therefore limited to describing and motivating some of the schemes and procedures that are useful for the purpose of extending elementary definitions of probabilities to full-fledged probability measures, while those aspects of the classical theory that are omitted here are covered in any textbook on the subject. Secondly, we intend at least to raise, in the reader's mind, the question of how relevant these assumptions and basic axioms are for the purpose of modelling physical reality, especially when one moves from finite to infinite constructions. Mechanical re-derivation of standard results did not seem to us to contribute essentially towards this purpose.

Next, for any non-negative $\mathfrak{S}$-measurable function $f : \Omega \to \mathbb{R}_+$ we define its integral w.r.t. $\mathscr{P}$ as

$$\int_\Omega f(\omega)\, \mathscr{P}(\mathrm{d}\omega) = \sup\Big\{ \int_\Omega s(\omega)\mathscr{P}(\mathrm{d}\omega) \ : \\ \mathrm{SF}_{\mathfrak{T}} \ni s \leq f \text{ for some finite } \mathfrak{T} \subset \mathfrak{S} \Big\}.$$

For signed f, we define $\int_\Omega f = \int_\Omega f_+ - \int_\Omega f_-$ where $f_+ := f\mathbb{1}_{f>0}$ and $f_- := -f\mathbb{1}_{f<0}$. In probability theory, the integral of f w.r.t. a probability measure $\mathscr{P}$ is often referred to as its *expectation* and denoted as $\mathbb{E}_{\mathscr{P}}\{f\}$ (we shall normally drop the subscript where no ambiguity should arise).

A.ae Transformation of measures by integration. Integration provides another means of constructing new measures from existing ones. To wit, given a non-negative $\mathfrak{S}$-measurable function f with $\mathbb{E}_{\mathscr{P}}\{f\} = 1$, the identity

$$\mathscr{Q}(S) = \mathbb{E}_{\mathscr{P}}\{\mathbb{1}_S f\} = \int \Omega_S(\omega) f(\omega) \mathscr{P}(\mathrm{d}\omega)$$

defines a new probability measure $\mathscr{Q}$ on $\mathfrak{S}$ (continuity follows from Lebesgue's monotone convergence theorem).

A.af Definition. Let $\mathscr{P}, \mathscr{Q}$ be probability measures (more generally, two σ-finite positive measures) on a σ-algebra $\mathfrak{S}$. We say that $\mathscr{Q}$ is *absolutely continuous* with respect to $\mathscr{P}$, in symbols $\mathscr{Q} \leq \mathscr{P}$, iff for any $S \in \mathfrak{S}$

$$\mathscr{P}(S) = 0 \ \Rightarrow \ \mathscr{Q}(S) = 0.$$

In contrast, $\mathscr{Q}$ is said to be *singular* with respect to $\mathscr{P}$, in symbols $\mathscr{Q} \perp \mathscr{P}$, iff there exists $S \in \mathfrak{S}$ such that

$$\mathscr{P}(S) = \mathscr{Q}(S^{\complement}) = 0.$$

The following results, which we shall not prove, point to a number of possibilities for characterizing measures using integrals.

A.ag Theorem (Lebesgue decomposition). Given probability measures $\mathscr{P}, \mathscr{Q}$ on the same σ-algebra $\mathfrak{S}$, there exist unique measures $\mathscr{Q}_{\mathrm{ac}}, \mathscr{Q}_{\mathrm{s}}$ on $\mathfrak{S}$ such that $\mathscr{Q} = \mathscr{Q}_{\mathrm{ac}} + \mathscr{Q}_{\mathrm{s}}$ and $\mathscr{Q}_{\mathrm{ac}} \ll \mathscr{P}$, $\mathscr{Q}_{\mathrm{s}} \perp \mathscr{P}$. Furthermore, $\mathscr{Q}_{\mathrm{ac}} \perp \mathscr{Q}_{\mathrm{s}}$.

A.ah Theorem (Radon-Nikodým). With $\mathscr{P}, \mathscr{Q}, \mathfrak{S}$ as above, $\mathscr{Q} \ll \mathscr{P}$ precisely when there exists a non-negative $(\mathfrak{S}, \mathrm{Borel}(\mathbb{R}))$-measurable function f such that

$$\mathscr{Q}(S) = \mathbb{E}_{\mathscr{P}}\{\mathbb{1}_S f\} = \int_\Omega \mathbb{1}_S(\omega) f(\omega) \mathscr{P}(\mathrm{d}\omega).$$

The function f is unique modulo $\mathscr{P}$ (i.e. any two such functions defer only on a set of $\mathscr{P}$-measure 0). Members of the equivalence class of functions thus identified are referred to as *Radon-Nikodým derivatives* of $\mathscr{Q}$ w.r.t. $\mathscr{P}$. In practice, we often work with some representative of this class, which we (abusively but harmlessly) call *the* Radon-Nykodým derivative.

A.ai Corollary. For any pair of probability measures $\mathscr{P}, \mathscr{Q}$ defined on some σ-algebra $\mathfrak{S}$, there exists a non-negative $(\mathfrak{S}, \text{Borel}(\mathbb{R}))$-measurable function f_{ac} and a set $S_{\text{s}} \in \mathfrak{S}$ with $\mathscr{P}(S) = 0$ such that for any $S \in \mathfrak{S}$,

$$\mathscr{Q}(S) = \int_{\Omega} f_{\text{ac}}(\omega)\mathscr{P}(\mathrm{d}\omega) + \mathscr{Q}(S \cap S_{\text{s}}).$$

The function f_{ac} and the set S_s are unique modulo $\mathscr{P}$.

A.aj The above results generalize directly to the case of $\mathscr{P}$ being a σ-finite (rather than finite) measure, and admit various other generalizations as well, but the only place we shall make direct use of any such generalization is in the following example. The reader seeking the details or wishing to go further is referred to the treatises by Fremlin [Fre03] and Bogachev [Bog07] as well as the research monograph and papers by König [Kön97, Kön06a, Kön06b].

A.ak Probability measures on the Euclidean space. Let us end this section by discussing two ways of characterizing Borel probability measures on the Euclidean space. In this case we have $\Omega = \mathbb{R}^m$, and the domain of the probability measure $\mathscr{P}$ is taken to be the Borel σ-algebra of $\mathbb{R}^m$, denoted $\text{Borel}(\mathbb{R}^m)$, which (we recall) is defined as the smallest σ-algebra $\subset \mathfrak{P}(\mathbb{R}^m)$ that includes all open sets of the topology of $\mathbb{R}^m$. Probability measures on $\text{Borel}(\mathbb{R}^m)$ are characterized by Lebesgue's Decomposition Theorem: given any Borel probability measure $\mathscr{P}$ on $\mathbb{R}^m$, there exists a non-negative function $f_{\mathscr{P}} \in \text{L}_1(\mathbb{R}^m)$, called the *distribution* of (the continuous part of) $\mathscr{P}$, with $\|f_{\mathscr{P}}\|_1 \leq 1$, and a smallest (possibly empty) set $S \in \text{Borel}(\mathbb{R}^m)$ with $\mathscr{P}(S) > 0$, called the *singularity* set of $\mathscr{P}$, such that for any $B \in \text{Borel}(\mathbb{R}^m)$ we have

$$\mathscr{P}(B) = \int_B f_{\mathscr{P}}(x)\ \lambda(\mathrm{d}x) + \mathscr{P}(B \cap S) \quad \text{and} \quad \int_S \lambda(\mathrm{d}x) = 0,$$

where λ denotes the Lebesgue measure in $\mathbb{R}^m$ (see, e.g., Pinsker [Pin64, §1.3]).

A.al Characteristic functions. In the previous paragraph, we described how a Borel probability measure on $\mathbb{R}^m$ can be identified by its singular part and its distribution with respect to the Lebesgue measure on $\mathbb{R}^m$. We now look at another way of characterizing Borel (and hence Radon) probabilities on $\mathbb{R}^m$.

Given a probability measure $\mathscr{P}$ on $\mathbb{R}^m$, its *characteristic function* is defined as

$$\widehat{\mathscr{P}}(\xi) = \int_{\mathbb{R}^m} \mathrm{e}^{\mathrm{i}\langle x,\xi\rangle}\,\mathscr{P}(\mathrm{d}x) = \mathbb{E}_{\mathscr{P}}\{\mathrm{e}^{\mathrm{i}\langle\cdot,\xi\rangle}\}, \quad \xi \in \mathbb{R}^m. \tag{A.1}$$

It should be noted that $\widehat{\mathscr{P}}$ is essentially nothing but the Fourier transform of the measure $\mathscr{P}$. It is necessarily continuous, in particular continuous at $\xi = 0$, and *positive-definite* (more precisely *non-negative* definite) in the sense that

$$\sum_{i,j} \overline{\zeta}_i \zeta_j \widehat{\mathscr{P}}(\xi_i - \xi_j) \geq 0$$

for all finite vectors $(\zeta_1, \ldots, \zeta_n) \in \mathbb{C}^n$ and $(\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$ of any arbitrary length $n \in \mathbb{N}$. The latter property follows easily from the positivity of $\mathscr{P}$. We also have $\widehat{\mathscr{P}}(0) = 1$, corresponding to the fact that the probability measure of the entire domain, i.e. $\mathbb{R}^m$, is 1.

Conversely, we have

A.am Bochner's Theorem. Any positive-definite function $\widehat{\mathscr{P}} : \mathbb{R}^m \to \mathbb{C}$ that is continuously equal to 1 at 0 uniquely identifies a (Radon) probability measure $\mathscr{P}$ on $\mathbb{R}^m$ by identity (A.1).

Hence, Bochner's theorem suggests another way of characterizing Borel probability measures on $\mathbb{R}^m$, by identifying them with continuous positive-definite functions $\mathbb{R}^m \to \mathbb{C}$.

The next theorem relates convergence of sequences of characteristic functions to the convergence of the corresponding probability measures. To state it, we first state a

A.an Definition. A sequence $\{\mathscr{P}_\nu\}_\nu$ of Borel probability measures on a metrizable space Ω is said to *converge weakly* to some measure $\mathscr{P}$, in symbols

$$\mathscr{P}_\nu \xrightarrow{w} \mathscr{P}, \qquad \text{iff} \qquad \int_\Omega \phi(x)\mathscr{P}_\nu(\mathrm{d}x) \to \int_\Omega \phi(x)\mathscr{P}(\mathrm{d}x)$$

for any bounded continuous function $\phi : \Omega \to \mathbb{R}$.

A.ao Lévy's continuity theorem. Let $\mathscr{P}_\nu$ be a sequence of Borel measures on $\mathbb{R}^m$, and suppose that there exists a function $\widehat{\mathscr{P}} : \mathbb{R}^m \to \mathbb{C}$ such that

$$\widehat{\mathscr{P}_\nu} \to \widehat{\mathscr{P}} \quad \text{point-wise.}$$

Then $\widehat{\mathscr{P}}$ itself is the characteristic function of some Borel probability measure $\mathscr{P}$ on $\mathbb{R}^m$ iff $\widehat{\mathscr{P}}$ is continuous at 0.

§A.5 Probabilities on infinite-dimensional product spaces

The theory of *stochastic processes* is a branch of probability theory that concerns itself with probabilities depending on a parameter, which most often represents time or spatial position, or a combination thereof. What sets this theory apart from the one developed thus far is the assumption that the parameter can take on infinitely many different values. We are consequently faced with the task of defining probabilities on *infinite-dimensional* spaces. This notion is clarified below.

§A.5.1 Algebraic and topological preliminaries

A.ap By saying that a certain space Ω is infinite-dimensional we mean the following.

There exists a set $\mathscr{I}$ (the *index set*), and, for every $m \in \mathbb{N}$ and any *finite* subset I of $\mathscr{I}$ with $|I| = m$, there exists a *canonical projection* $\pi_I : \Omega \to \Omega_I$ where Ω_I is an m-dimensional space, which, for concreteness, we take to be $\mathbb{R}^I$ (we shall write ι instead of $\{\iota\}$ where no ambiguity should arise). Furthermore, for finite $I_1, I_2 \subset \mathscr{I}$, the axioms below are required.

(Π1) $I_2 \subset I_1 \;\Rightarrow\; \pi_{I_2} f$ is completely determined by $\pi_{I_1} f$, i.e. there exists a map $\pi_{I_1 I_2} : \Omega_{I_1} \to \Omega_{I_2}$ such that

$$\pi_{I_2} = \pi_{I_1 I_2} \circ \pi_{I_1}. \qquad \text{(projection)}$$

(Π2) $f \in \Omega$ is uniquely determined by its projections $\pi_I f$, $I \subset \mathscr{I}$ finite. Put slightly differently, the *evaluation map*

$$f \mapsto \{(\iota, \pi_\iota f) : \iota \in \mathscr{I}\}$$

is an injection. $\pi_\iota f$ is more familiarly denoted as $f(\iota)$. (separation)

It is both empirically sensible and mathematically expedient to limit oneself to finite projections, as we have done above. It is empirically sensible because, in reality, we can never make more than a finite number of observations of any phenomenon. It is mathematically expedient because we can make consistent deductions and computations on the basis of the above axioms.

The above axioms imply that every π_I is completely determined by the one-dimensional projections $\pi_\iota : \Omega \to \Omega_\iota$, $\iota \in I$. Moreover, by the second axiom, Ω can be identified with a subspace of the product space $\prod_{\iota \in \mathscr{I}} \Omega_\iota$ and we may then identify π_I with (the restriction to Ω of) the canonical projection from the infinite product space onto $\Omega_I = \prod_{\iota \in I} \Omega_\iota$. Similarly, $\pi_{I_1 I_2}$, $I_2 \subset I_1$, is

identified with the canonical projection from $\prod_{\iota\in I_1}\Omega_\iota$ onto $\prod_{\iota\in I_2}\Omega_\iota$. If the Ω_ιs are topological spaces, $\Omega\subset\prod_{\iota\in\mathscr{I}}\Omega_\iota$ can be equipped with the *projective topology* with respect to the maps $\pi_I:\Omega\to\Omega_I$, as defined below.

A.aq Definition. Given a space $\mathscr{E}$ and a family of functions $f_\alpha:\mathscr{E}\to\mathscr{E}_\alpha$, $\alpha\in A$, where the $\mathscr{E}_\alpha$ are topological spaces, the *projective topology* on $\mathscr{E}$ with respect to the family $\{f_\alpha\}_{\alpha\in A}$ is defined as the coarsest (smallest) topology on $\mathscr{E}$ with respect to which all f_αs are continuous.

As noted, in this section we have assumed $\Omega_\iota=\mathbb{R}$ and $\Omega_I=\mathbb{R}^I$ with its standard topology, from where $\Omega\subset\prod_{\iota\in I}\Omega_\iota=\mathbb{R}^{\mathscr{I}}$.

§A.5.2 Cylindrical probabilities and their extension to measures

A.ar We now consider the question of defining probability measures on the infinite-dimensional space Ω. As is always the case with infinities, the construction should involve taking limits in some way. In this instance, these will be the *projective limits* of families of finite-dimensional probability measures.

Let us for a moment imagine that a Borel probability measure $\mathscr{P}$ on the infinite-dimensional space Ω is already given. In this case, any finite-dimensional projection of $\mathscr{P}$ through some π_I, $I\subset\mathscr{I}$ finite, defines a Borel measure on $\mathbb{R}^I$ which is the push-forward $\pi_I(\mathscr{P})=\mathscr{P}\circ\pi_I^{-1}$. We thus define,

$$\mathscr{P}_I(B)=\pi_I(\mathscr{P})(B)=\mathscr{P}(\pi_I^{-1}B)$$

for any $B\in\text{Borel}(\Omega_I)$. In view of and in analogy with Axiom (Π1), the following *consistency* property holds for the family of all such *finite-dimensional marginals*.

(C1) Given finite $I_1,I_2\subset\mathscr{I}$,

$$I_2\subset I_1\quad\Rightarrow\quad\mathscr{P}_{I_2}=\pi_{I_1I_2}(\mathscr{P}_{I_1})=\mathscr{P}_{I_1}\circ\pi_{I_1I_2}^{-1}.$$

Going in the other direction, to find some sort of a probabilistic equivalent of Axiom (Π2), one may consider the situation where the probability measure $\mathscr{P}$ on Ω is not given *a priori*, but instead one is provided only with a family $\mathscr{P}_I$, $I\subset\mathscr{I}$ finite, of Borel probability measures satisfying (C1). Under these circumstances, one then wishes to know if the said family of finite-dimensional marginals uniquely identifies a probability measure $\mathscr{P}$ on Ω (i.e., informally speaking, if the 'limit' of the $\mathscr{P}_I$s as I goes to the entire set $\mathscr{I}$ exists).

A.as With regard to the above converse, one first notes that the collection of all sets of the form

$$C_{B,I}:=\pi_I^{-1}B,\quad B\in\text{Borel}(\Omega_I)$$

(called *Borel cylinder sets* because $C_{B,I}$ corresponds to the *cylinder* in Ω with *base* $B \subset \Omega_I = \mathbb{R}^I$), forms an algebra

$$\mathfrak{C}_\Omega := \{C_{B,I} \ : \ I \subset \mathscr{I} \text{ finite}, B \in \text{Borel}(\Omega_I)\} \tag{A.2}$$

in $\mathfrak{P}(\Omega)$, named the *algebra of Borel cylinder sets* (it is straightforward to verify that complements and finite unions/intersections of sets in $\mathfrak{C}_\Omega$ also belong in $\mathfrak{C}_\Omega$). On this algebra, one may then define an elementary (finitely-additive) probability measure $\mathscr{P}$ by means of the identity

$$\mathscr{P}(C_{B,I}) := \mathscr{P}_I(B). \tag{A.3}$$

Turning things around, we may now look at the $\mathscr{P}_I$s as the finite-dimensional marginals of the elementary probability $\mathscr{P}$. By assumption, $\mathscr{P}$ has σ-additive finite-dimensional marginals. On the other hand, it is not given that $\mathscr{P}$ itself should be (or extend to) a σ-additive probability measure.

A.at Definition. A finitely-additive probability measure $\mathscr{P}$ on $\mathfrak{C}_\Omega$ whose finite-dimensional marginals $\mathscr{P}_I := \pi_I(\mathscr{P}) = \mathscr{P} \circ \pi_I^{-1}$, $I \subset \mathscr{I}$ finite, are σ-additive, is called a *cylindrical probability (measure).* The said marginals then necessarily fulfil the consistency condition (C1). We may also refer to the family of probability measures $\{\mathscr{P}_I\}$ fulfilling (C1) as a cylindrical probability measure, knowing that $\mathscr{P}$ can then be recovered from the marginals by (A.3).

In view of the above, the fundamental question of the theory of stochastic processes boils down to this:

A.au Question. Given a cylindrical probability measure $\mathscr{P}$ on $\mathfrak{C}_\Omega$, is it possible (and under what conditions) uniquely to extend $\mathscr{P}$ to a σ-additive probability measure on some σ-algebra $\supset \mathfrak{C}_\Omega$?

A.av Definition. The σ-additive extension of $\mathscr{P}$ to $\sigma\,\text{Alg}(\mathfrak{C}_\Omega)$, if it exists, is called the *projective limit* of the cylindrical probability $\mathscr{P}$ (or of the family $\{\mathscr{P}_I\}$).

The usual course taken in standard textbooks on stochastic theory to answer the above question is to invoke Kolmogorov's extension theorem, which takes the space Ω to be *equal* to the product space $\prod_\iota \Omega_\iota$ and then gives a number of mild conditions for the σ-additive extension of $\mathscr{P}$ to exist and be unique. In the context of Radon measures, an alternative if not entirely unrelated approach involves using Prokhorov's theorem to define Radon probability measures on spaces Ω that can be smaller than the product space $\prod_\iota \Omega_\iota$. We state a version of each of these theorems below (the respective references for the following versions are Bogachev [Bog07] and Schwartz [Sch73]).

The statement of Kolmogorov's theorem we shall give below is somewhat more general than Kolmogorov's original version. It uses the following

*A.aw **Definition.** A set system $\mathfrak{K}_0$ is said to have the *finite intersection property* iff the intersection of any finite number of its elements is non-empty. We call a set system $\mathfrak{K}$ a *sequentially compact class* iff for any countable subset $\mathfrak{K}_0$ of $\mathfrak{K}$,

$$\mathfrak{K} \text{ has the finite intersection property } \Leftrightarrow \bigcap_{K\in\mathfrak{K}} K = \varnothing$$

(c.f. Bogachev [Bog07, Definition 1.4.1]).

Note that the $\Leftarrow$ direction of the above definition is always true.

The definition is motivated by the observation that, in a Hausdorff topological space, a subfamily of topologically compact sets is sequentially compact.

*A.ax **Kolmogorov's Extension Theorem.** Let $\mathscr{P}$ be a cylindrical probability measure on the algebra $\mathfrak{C}_\Omega$ defined in (A.2), with the additional property that for each finite set $I \subset \mathscr{I}$ there exists a sequentially compact class $\mathfrak{K}_I$ in Ω_I (see ***A.aw**), such that the finite-dimensional marginal $\mathscr{P}_I$ is inner regular $\mathfrak{K}_I$ in the sense of **A.v**. Then $\mathscr{P}$ extends uniquely to a probability measure on $\sigma\,\mathrm{Alg}(\mathfrak{C}_\Omega)$ (also denoted as $\mathscr{P}$).

A.ay **Corollary.** A cylindrical probability measure on $\mathfrak{C}_\Omega$ whose finite-dimensional marginals are all Borel-Radon extends uniquely to a countably-additive probability measure on $\sigma\,\mathrm{Alg}(\mathfrak{C}_\Omega)$.

A.az Since, by **A.z**, any Borel measure on $\mathbb{R}^I$, I finite, is Radon, by the above corollary we find that a consistent family of finite-dimensional Borel probability measures on the spaces $\mathbb{R}^I$, $I \subset \mathscr{I}$ finite, uniquely defines a countably-additive probability measure on (the σ-algebra of Borel cylinder sets in) $\mathbb{R}^{\mathscr{I}}$. This latter form corresponds to Kolmogorov's original statement of the theorem [Kol33].

A.ba As noted, it may be of interest in some applications to define a probability measure on a space Ω that is smaller than the product space $\prod_\iota \Omega_\iota$. The following theorem, originally due to Prokhorov, suggests a way to verify if this is possible for a particular choice of Ω (the version given here is adapted from Schwartz [Sch73, Ch. I, Theorem 22]; see also Fremlin [Fre04b, Theorem 418M]).

A.bb **Theorem (Prokhorov).** The projective limit of a consistent family $\{\mathscr{P}_I\}$ of finite-dimensional Radon probability measures as defined above exists and is Radon iff the following condition is fulfilled:

(K) For every $\epsilon > 0$ there exists a compactum $K_\epsilon \subset \Omega$ such that for all finite $I \subset \mathscr{I}$, $\mathscr{P}_I(\pi_I K_\epsilon) \geq 1 - \epsilon$.

We do not duel any longer on the above results since we shall not use any of them directly.

§A.6 Probabilities on infinite-dimensional topological vector spaces

Next, we turn our attention to the question of defining probability measures on an infinite-dimensional topological vector space. We shall adapt the definitions of the previous section to this new context.

§A.6.1 Algebraic and topological preliminaries

A.bc In this scenario, we have a dual system $\langle \mathscr{E}, \mathscr{F} \rangle$ consisting of two topological vector spaces $\mathscr{E}$ and $\mathscr{F}$ and a bilinear mapping

$$\mathscr{E} \times \mathscr{F} \to \mathbb{R} : (e, f) \mapsto \langle e, f \rangle =: \pi_e f,$$

called the *scalar product,* which fulfils the following *separation axioms:*

(S) $\langle e, f_1 \rangle = \langle e, f_2 \rangle$ for all $e \in \mathscr{E} \quad \Leftrightarrow \quad f_1 = f_2$;

(S') $\langle e_1, f \rangle = \langle e_2, f \rangle$ for all $f \in \mathscr{F} \quad \Leftrightarrow \quad e_1 = e_2$.

We shall take $\mathscr{F}$ as the space on which to construct probabilities. By the above axioms, we can identify each element f of $\mathscr{F}$ with a linear map on $\mathscr{E}$, as any such f is completely determined by its scalar products $\langle e, f \rangle$ with elements $e \in \mathscr{E}$.

A.bd **Compatible topologies.** The spaces $\mathscr{E}, \mathscr{F}$ are given topologies. We say (and henceforth require) the topology of $\mathscr{F}$ to be *compatible* with the bilinear form $\langle \cdot, \cdot \rangle$ (or with the dual system $\langle \mathscr{E}, \mathscr{F} \rangle$). Compatibility here means two things. First, that the maps

$$\mathscr{F} \ni f \mapsto \langle e, f \rangle \in \mathbb{R}, \qquad e \in \mathscr{E}, \tag{A.4}$$

are continuous in the topology of $\mathscr{F}$. Secondly, that *any* continuous linear map $\mathscr{F} \to \mathbb{R}$ is identified with a projection of the form given in (A.4). In other words, the topologies of $\mathscr{E}, \mathscr{F}$ are such that the continuous dual of $\mathscr{F}$ is equal to $\mathscr{E}$ (as a set), no more and no less.

Similar conditions may be formulated for the topology of $\mathscr{E}$ to be compatible with the dual system, in which case $\mathscr{F}$, as a set, may be identified the continuous dual of $\mathscr{E}$.

***A.be** **Weak and Mackey topologies.** We may, in general, define a range of topologies on $\mathscr{F}$ that are compatible with the bilinear form (similarly for $\mathscr{E}$). The range is determined by the smallest and largest topologies on $\mathscr{F}$ that are compatible with $\langle \mathscr{E}, \mathscr{F} \rangle$. By definition, the lower bound (smallest topology) of the range is precisely the projective topology on $\mathscr{F}$ with respect to the family $\{\pi_e\}_{e \in \mathscr{E}}$

(cf. **A.aq**). This topology, which is in this case called the *weak* topology and denoted as $\sigma(\mathscr{F}, \mathscr{E})$, is immediately seen to be locally convex, as it is induced by the family of semi-norms $|\pi_e \cdot |$, $e \in \mathscr{E}$.[9] Moreover, as a result of the first separation axiom above, the weak topology is automatically Hausdorff.

The upper end of the range is determined by the finest (richest) locally convex topology one can assign to $\mathscr{F}$ without enlarging the family of all continuous linear functionals $\mathscr{F} \to \mathbb{R}$ beyond the set $\{\pi_e\}_{e\in\mathscr{E}}$; in other words, the largest topology with respect to which every linear functional $\mathscr{F} \to \mathbb{R}$ that is discontinuous w.r.t. $\sigma(\mathscr{F}, \mathscr{E})$ remains discontinuous. This topology is known as the *Mackey* topology on $\mathscr{F}$ and is denoted as $\tau(\mathscr{F}, \mathscr{E})$.

Clearly, any locally convex topology on $\mathscr{F}$ that is compatible with $\langle \mathscr{E}, \mathscr{F} \rangle$ is no coarser than $\sigma(\mathscr{F}, \mathscr{E})$ and no finer than $\tau(\mathscr{F}, \mathscr{E})$.

A.bf We see therefore that, as a result of the separation axioms and the compatibility condition given above, to every *fi*nite *l*inearly-*i*ndependent (*fili*) set $E \subset \mathscr{E}$ there corresponds a linear map

$$\pi_E : \mathscr{F} \to \mathbb{R}^E : f \mapsto \{(e, \langle e, f \rangle) : e \in E\}$$

that is continuous and surjective. Conversely, *all* real finite-dimensional linear projections of $\mathscr{F}$ that are continuous and surjective can be constructed in the above fashion. We define a partial order $\leqq$ on fili subsets of $\mathscr{E}$ thus:

(Π1') $E_1 \leqq E_2$ iff there exists a (unique) continuous linear operator $\pi_{E_1 E_2}$ such that

$$\pi_{E_2} = \pi_{E_1 E_2} \circ \pi_{E_1}.$$

Moreover, by the separation axioms, the following analogue of (Π2) holds:

(Π2') $f \in \mathscr{F}$ is uniquely determined by its projections $\pi_E f$, $E \subset \mathscr{E}$ fili.

§A.6.2 Cylindrical probability measures on topological vector spaces

Having formulated the equivalents (Π1') and (Π2') of Axioms (Π1) and (Π2) for linear spaces(cf. **A.ap**), we shall now entertain adapted definitions of cylinder sets and cylindtrical measures in this new setting.

9. A topology $\mathfrak{O}$ on the space $\mathscr{F}$ is *locally convex* iff it is the projective topology on $\mathscr{F}$ with respect to some family (finite or infinite) of semi-norms on $\mathscr{F}$. In turn, a *semi-norm* on an $\mathbb{F}$-vector space $\mathscr{F}$ is a non-negative function $p : \mathscr{F} \to \mathbb{R}$ with the following two properties: $p(cf) = |c|p(f)$ for all $c \in \mathbb{F}$ (positive homogeneity), and $p(f+g) \leq p(f)+p(g)$ (sub-additivity a.k.a. the triangle inequality).

A.bg Definition. *Borel cylinder sets* on $\mathscr{F}$ are sets of the form

$$C_{B,E} := \pi_E^{-1} B,$$

where E is a fili subset of $\mathscr{E}$ and $B \in \text{Borel}(\mathbb{R}^E)$.

Once again, Borel cylinder sets define an algebra,

$$\mathfrak{C}_{\mathscr{F}} := \{C_{B,E} \ : \ E \subset \mathscr{E} \text{ fili}, \ B \in \text{Borel}(\mathbb{R}^E)\},$$

known as the *algebra of Borel cylinder sets* in $\mathscr{F}$.

A.bh Definition. A family of Borel (hence Radon) probability measures $\mathscr{P}_E$ on $\mathbb{R}^E$, $E \subset \mathscr{E}$ fili, is collectively called a *cylindrical probability* on $\mathscr{F}$, provided the following *consistency condition* is fulfilled for all fili sets $E_1, E_2 \subset \mathscr{E}$ (compare with (C1)):

(C1') Given fili sets $E_1, E_2 \subset \mathscr{E}$,

$$E_2 \leqq E_1 \quad \Rightarrow \quad \mathscr{P}_{E_2} = \mathscr{P}_{E_1} \circ \pi^{-1}_{E_1 E_2}. \qquad \text{(consistency)}$$

As before, we can identify the family $\{\mathscr{P}_E\}$, $E \subset \mathscr{E}$ fili, with a finitely-additive probability $\mathscr{P}$ on $\mathfrak{C}_{\mathscr{F}}$. This identification is achieved by the identities

$$\mathscr{P}(C_{B,E}) = \mathscr{P}_E(B) \qquad \text{and} \qquad \mathscr{P}_E(B) = \mathscr{P} \circ \pi_E^{-1}(B).$$

$\mathscr{P}$, while being only finitely-additive (at least initially), has σ-additive finite-dimensional marginals, which are precisely the measures $\mathscr{P}_E$, $E \subset \mathscr{E}$ fili. In view of the above identification of $\mathscr{P}$ with the family $\{\mathscr{P}_E\}$, we may refer to $\mathscr{P}$ itself as the cylindrical probability.

A.bi Continuous cylindrical measures. We shall later consider situations in which the topology of the space $\mathscr{E}$ (the first item of the dual pair) is already given. At least in the case of the spaces we shall be working with, it will be convenient and useful to limit our consideration to cylindrical measures $\mathscr{P}$ that are *continuous* in the sense of the following axiom:

(C2') Given a convergent sequence of fili sets $E_\nu \in \mathscr{E}^m$ with limit $E \in \mathscr{E}^m$, $m \in \mathbb{N}$,

$$\mathscr{P}_{E_\nu} \xrightarrow{w} \mathscr{P}_E \qquad \text{(continuity)}$$

where $\xrightarrow{w}$ indicates weak convergence as defined in **A.an** [10]

In view of Lévy's continuity theorem, the above property is equivalent to the point-wise convergence of the sequence $\{\widehat{\mathscr{P}_{E_\nu}}\}_\nu$ of characteristic functions to $\widehat{\mathscr{P}_E}$.

10. The topologically-inclined reader will note that the above definition is based on *sequential* continuity irrespective of whether $\mathscr{E}^m$ is first-countable or not.

§A.6.3 Characteristic functionals and the extension of cylindrical probabilities to measures

A.bj The question of whether a cylindrical probability $\mathscr{P}$ on $\mathscr{F}$ extends to a σ-additive probability measure is not entirely topological. Nevertheless, it is possible in some important cases to formulate topological conditions that are sufficient for such an extension to exist and be unique. In the setting that will be of interest to us, the said conditions turn out to be necessary as well.

We shall formulate the said topological conditions in terms of the continuity of a certain functional associated with the cylindrical probability measure. This functional, related to the one originally introduced by Kolmogorov [Kol35] as the Laplace transform of the probability, is essentially the infinite-dimensional equivalent of the characteristic function of a Borel measure on the Euclidean space $\mathbb{R}^m$ (see (A.1)). For this reason it is dubbed the *characteristic functional* of the cylindrical measure.

The utility of the characteristic functional in defining measures is due to an infinite-dimensional analogue of Bochner's theorem (**A.am**), which outlines sufficient (and sometimes necessary) conditions for the characteristic functional uniquely to define a σ-additive probability measure on $\mathscr{F}$ extending $\mathscr{P}$.

A.bk **Definition.** Given the dual system $\langle \mathscr{E}, \mathscr{F} \rangle$ as discussed above, we define the *characteristic functional* of a cylindrical measure $\mathscr{P}$ on $\mathscr{F}$ as the functional

$$\widehat{\mathscr{P}} : \mathscr{E} \to \mathbb{C} : e \mapsto \int_{\mathscr{F}} \mathrm{e}^{\mathrm{i}\pi_e f} \mathscr{P}(\mathrm{d}f) = \int_{\mathscr{F}} \mathrm{e}^{\mathrm{i}\langle e, f\rangle} \mathscr{P}(\mathrm{d}f) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}t} \mathscr{P}_e(\mathrm{d}t).$$

A.bl The significance of the above definition lies in the fact that we can obtain the chracteristic function $\widehat{\mathscr{P}_E}$ of any finite marginal $\mathscr{P}_E$, $E = \{e_1, \ldots, e_m\} \subset \mathscr{E}$ fili, from $\widehat{\mathscr{P}}$ as follows:

$$\begin{aligned}
\widehat{\mathscr{P}_E}(\xi_1, \ldots, \xi_m) &= \int_{\mathbb{R}^m} \mathrm{e}^{\mathrm{i}\sum_k \xi_k t_k} \quad \mathscr{P}_E(\mathrm{d}t) \\
&= \int_{\mathscr{F}} \mathrm{e}^{\mathrm{i}\sum_k \xi_k \langle e_k, f\rangle} \, \mathscr{P}(\mathrm{d}f) \\
&= \int_{\mathscr{F}} \mathrm{e}^{\mathrm{i}\langle \sum_k \xi_k e_k, f\rangle} \, \mathscr{P}(\mathrm{d}f) = \widehat{\mathscr{P}}\big(\sum_k \xi_k e_k\big).
\end{aligned}$$

Furthermore, if the characteristic functional is continuous in the topology of $\mathscr{E}$, so will be all $\widehat{\mathscr{P}_E}$s, hence by Lévy's continuity theorem the cylindrical measure $\mathscr{P}$ will satisfy continuity axiom (C2') of **A.bi**. The converse is also true, viz. if a cylindrical measure $\mathscr{P}$ is continuous in the sense of (C2'), then its characteristic functional $\widehat{\mathscr{P}}$ will be continuous in the topology of $\mathscr{E}$.

A.bm Definition. We call a functional $\widehat{\mathscr{P}} : \mathscr{E} \to \mathbb{C}$ *non-negative definite* (or, more loosely, *positive* definite) iff

$$\sum_{i,j} \overline{\zeta}_i \zeta_j \widehat{\mathscr{P}}(e_i - e_j) \geq 0$$

for all finite vectors $(\zeta_1, \ldots, \zeta_n) \in \mathbb{C}^n$ and $(e_1, \ldots, e_n) \in \mathscr{E}^n$ of arbitrary length $n \in \mathbb{N}$.

The characteristic functional of a cylindrical probability $\mathscr{P}$ is non-negative definite in the above sense, since we have

$$\sum_{i,j} \overline{\zeta}_i \zeta_j \widehat{\mathscr{P}}(e_i - e_j) = \int_{\mathscr{F}} \Big(\sum_{i,j} \overline{\zeta_i \mathrm{e}^{-\mathrm{i}\langle e_i, f\rangle}} \zeta_j \mathrm{e}^{-\mathrm{i}\langle e_j, f\rangle} \Big) \mathscr{P}(\mathrm{d}f),$$

which is non-negative because the integrand is of the form $u^{\mathrm{H}} \mathrm{A} u$ for some vector u and the non-negative definite matrix $\mathrm{A} = [1]_{ij}$, hence non-negative, and it is being integrated w.r.t. a positive cylindrical measure.

Going in the other direction, given an otherwise arbitrary functional $\widehat{\mathscr{P}} : \mathscr{E} \to \mathbb{C}$ that is non-negative definite in the above sense, the functionals $\widehat{\mathscr{P}_E}$, $E = \{e_1, \ldots, e_m\} \subset \mathscr{E}$ fili, derived from it by defining

$$\widehat{\mathscr{P}_E}(\xi_1, \ldots, \xi_m) = \widehat{\mathscr{P}}\big(\sum_k \xi_k e_k\big)$$

(cf. **A.bl**) are non-negative definite in the sense of **A.al**.

Putting the above observations together with those of **A.bl** and the two theorems of **A.al**, we obtain the

A.bn Proposition. Given a compatible dual pair $\langle \mathscr{E}, \mathscr{F} \rangle$, a functional $\widehat{\mathscr{P}} : \mathscr{E} \to \mathbb{C}$ is the characteristic functional of some continuous cylindrical probability measure $\mathscr{P}$ on $\mathscr{F}$ iff it is non-negative definite and continuous at 0 in the topology of $\mathscr{E}$ with $\widehat{\mathscr{P}}(0) = 1$.

A.bo Extension of cylindrical probabilities to measures. The previous proposition allows us to identify cylindrical probabilities with their characteristic functionals. It is therefore foreseeable that, at least in some cases, the question of extending a cylindrical measure $\mathscr{P}$ on $\mathscr{F}$ to a true (i.e. σ-dditive) probability measure can be answered by way of studying its characteristic functional. We conclude this chapter by citing a number of results that clarify this connection in some important special cases, including probabilities on the spaces we deal with in the present thesis.

Before we state the main result, we shall need to introduce a few topological notions. These may however be skipped as, to apply the main result, it

suffices simply to know that the spaces $\mathscr{D}(\mathbb{R}^d)$ and $\mathscr{S}(\mathbb{R}^d)$ of, respectively, compactly-supported and Schwartz test functions are *nuclear* (whatever that term means), as are their finite Cartesian powers. The reader is therefore encouraged to go directly to the final result (Corollary **A.bv**).

***A.bp Definition.** A *Hilbert-Schmidt operator* $\mathrm{T} : \mathscr{H}_1 \to \mathscr{H}_2$, where $\mathscr{H}_1, \mathscr{H}_2$ are Hilbert spaces, is a continuous linear operator $\mathscr{H}_1 \to \mathscr{H}_2$ such that for some orthonormal basis $\{e_k\}_{k\in K}$ of $\mathscr{H}_1$,

$$\|\mathrm{T}\|_{\mathrm{HS}} := \sum_{k\in K} \|\mathrm{T}e_k\|_{\mathscr{H}_2} < \infty.$$

It can then be shown that the above value, called the *Hilbert-Schmidt norm* of T, is indepdendent of the choice of the orthonormal basis $\{e_k\}_{k\in K}$.

***A.bq Definition.** An operator $\mathrm{N} : \mathscr{H}_1 \to \mathscr{H}_2$ is called *nuclear* or *trace class* iff it can be factorized as

$$\mathscr{H}_1 \xrightarrow{\mathrm{T}_1} \mathscr{H}_3 \xrightarrow{\mathrm{T}_2} \mathscr{H}_2$$

where $\mathscr{H}_3$ is some Hilbert space and $\mathrm{T}_1, \mathrm{T}_2$ are Hilbert-Schmidt operators. The definition generalizes to arbitrary locally convex spaces, but we shall not need this generalization.

The following definition of the Sazonov topology is taken from Smolyanov and Fomin [SF76, §4] (see also Kolmogorov [Kol59], Prokhorov [Pro61, §5], Schwartz [Sch73, Part II, §IV.3], and Bogachev [Bog07, §7.13]).

***A.br Definition.** Let $\mathscr{E}$ be a locally convex space whose topology is determined by a countable family of Hilbert norms

$$\|e\|_n = \sqrt{\langle e, e\rangle_n}, \quad n \in \mathbb{N}$$

(such a space is called *countably Hilbert*) and let $\overline{\mathscr{E}}_n$ denote the completion of $\mathscr{E}$ with respect to $\|\cdot\|_n$. Without loss of generality (for instance, by redefining the nth inner product as the sum of the first n inner products), we may assume the norms $\|\cdot\|_n$, $n \in \mathbb{N}$, to be increasing in the sense that

$$\|e\|_{n+1} \geq \|e\|_n, \quad \text{for all } e \in \mathscr{E},\ n \in \mathbb{N}.$$

Let $\mathscr{E}$ be a locally convex space with topology $\mathfrak{O}$. The *Sazonov topology* associated with the topology of the space $\mathscr{E}$, denoted as $\tau_N(\mathscr{E})$ or $\tau_{\mathrm{N}}(\mathfrak{O})$, is the locally convex topology associated with the family of all semi-norms p satisfying the following properties:

($\tau_N 1$) p is continuous with respect to the original topology of $\mathscr{E}$.

($\tau_N 2$) p is derived from a Hilbert semi-inner-product, i.e. there exists a semi-inner-product $\langle\cdot,\cdot\rangle_p$ on $\mathscr{E}\times\mathscr{E}$ with $p(e) \equiv \langle e, e\rangle_p$.

($\tau_N 3$) There exists a semi-norm q on $\mathscr{E}$ also fulfilling ($\tau_N 1$) and ($\tau_N 2$) and with $q(e) \geq p(e)$ for all $e \in \mathscr{E}$, such that the injection $\overline{\mathscr{E}}_q \to \overline{\mathscr{E}}_p$ is Hilbert-Schmidt. Here $\overline{\mathscr{E}}_p, \overline{\mathscr{E}}_q$ denote the completions of $\mathscr{E}$ with respect to p and q.

***A.bs** **Definition.** We call a locally convex space $\mathscr{E}$ *nuclear* iff its Sazonov topology $\tau_{\mathrm{N}}(\mathscr{E})$ coincides with its original topology.

A.bt **Remark.** The following spaces are nuclear [Pie72, Sch73]: the space $\mathscr{D}(\mathbb{R}^d)$ of compactly supported smooth test functions with its standard topology; the space $\mathscr{S}(\mathbb{R}^d)$ of rapidly-decaying Schwartz test functions; spaces of the form $\mathscr{E}^k$, $k \in \mathbb{N}$, where $\mathscr{E}$ is nuclear; closed subspaces of a nuclear space; the quotient of a nuclear space with respect to a closed subspace.

After these preliminaries, an infinite-dimensional generalization of Bochner's theorem, credited to Minlos and Sazonov, can be stated as follows.

***A.bu** **Theorem.** Let $\langle \mathscr{E}, \mathscr{F} \rangle$ be a dual system satisfying the separation axioms (S) and (S'). Denote by $\tau(\mathscr{E}, \mathscr{F})$ the Mackey topology on $\mathscr{E}$, i.e. the strongest topology on $\mathscr{E}$ which is compatible with the duality $\langle \mathscr{E}, \mathscr{F} \rangle$ in the sense of **A.bd**. Let $\widehat{\mathscr{P}} : \mathscr{E} \to \mathbb{C}$ be a non-negative definite functional satisfying $\widehat{\mathscr{P}}(0) = 1$, and assume, in addition, that $\widehat{\mathscr{P}}$ is continuous at 0 in the Sazonov topology $\tau_{\mathrm{N}}(\tau(\mathscr{E}, \mathscr{F}))$ associated with the Mackey topology on $\mathscr{E}$. $\widehat{\mathscr{P}}$ is then the characteristic functional of a cylindrical measure on $\mathscr{F}$ that extends uniquely to a σ-additive probability measure on the σ-algebra of Borel cylinder sets in $\mathscr{F}$.

Note that if $\widehat{\mathscr{P}}$ is continuous w.r.t. any Sazonov topology $\tau_{\mathrm{N}}(\mathfrak{O})$ where $\mathfrak{O}$ is some topology on $\mathscr{E}$ that is compatible with $\langle \mathscr{E}, \mathscr{F} \rangle$, then it is also continuous w.r.t. the (stronger) topology $\tau_{\mathrm{N}}(\tau(\mathscr{E}, \mathscr{F}))$, therefore the continuity requirement of the theorem is fulfilled in this case.

The above theorem was proved by Minlos [Min63] in the case of $\mathscr{E}$ being nuclear (confirming a conjecture of Gel'fand) and separately by Sazonov [Saz58] in the case where $\mathscr{E}$ is a Hilbert space. Kolmogorov clarified the connection between the two [Kol59]. Prokhorov [Pro56] had previously underlined the important role of nuclear operators in this context. The version of the result we have given above is taken from Smolyanov and Fomin [SF76, §4, Theorem 2].

By ***A.bs**, we obtain the following important

A.bv **Corollary (Minlos's Theorem).** Let $\mathscr{E}$ be a nuclear space and $\mathscr{F} = \mathscr{E}'$. A continuous non-negative definite functional $\widehat{\mathscr{P}} : \mathscr{E} \to \mathbb{C}$ with $\widehat{\mathscr{P}}(0) = 1$ defines a unique σ-additive probability measure $\mathscr{P}$ on the σ-algebra of Borel cylinder

sets in $\mathscr{F}$ such that

$$\widehat{\mathscr{P}}(e) = \int_{\mathscr{E}'} \mathrm{e}^{\mathrm{i}\langle e,f\rangle}\,\mathscr{P}(\mathrm{d}f).$$

The above corollary is about the only advanced probabilistic result that is used directly in the main text of the thesis.

— 𝔅 —

Some Related Publications

For the purpose of reference and comparison, in this appendix we have reproduced some of our related publications that were either not incorporated in the thesis or whose presentation differs significantly from the one given therein due to extensions and refinements subsequent to their publication. We have excluded other publications that were more directly assimilated in the main text.

P. D. Tafti, D. Van De Ville, and M. Unser, *'Invariances, Laplacian-like wavelet bases, and the whitening of fractal processes',* IEEE Trans. Image Process., vol. 18, no. 4, pp. 689-702, Apr. 2009.

333.0pt530.0pt

Invariances, Laplacian-Like Wavelet Bases, and the Whitening of Fractal Processes

Pouya Dehghani Tafti, *Member, IEEE*, Dimitri Van De Ville, *Member, IEEE*, and Michael Unser, *Fellow, IEEE*

***Abstract*—In this contribution, we study the notion of affine invariance (specifically, invariance to the shifting, scaling, and rotation of the coordinate system) as a starting point for the development of mathematical tools and approaches useful in the characterization and analysis of multivariate fractional Brownian motion (fBm) fields. In particular, using a rigorous and powerful distribution theoretic formulation, we extend previous results of Blu and Unser (2006) to the multivariate case, showing that polyharmonic splines and fBm processes can be seen as the (deterministic vs stochastic) solutions to an identical fractional partial differential equation that involves a fractional Laplacian operator. We then show that wavelets derived from polyharmonic splines have a behavior similar to the fractional Laplacian, which also turns out to be the whitening operator for fBm fields. This fact allows us to study the probabilistic properties of the wavelet transform coefficients of fBm-like processes, leading for instance to ways of estimating the Hurst exponent of a multiparameter process from its wavelet transform coefficients. We provide theoretical and experimental verification of these results. To complement the toolbox available for multiresolution processing of stochastic fractals, we also introduce an extended family of multidimensional multiresolution spaces for a large class of (separable and nonseparable) lattices of arbitrary dimensionality.**

***Index Terms*—Affine invariance, fractional Brownian motion (fBm), fractional partial differential equations, Hurst exponent, lattices, multidimensional wavelets, operator wavelets, polyharmonic splines, whitening.**

I. Introduction

THE notion of invariance plays a significant role in mathematical modeling. The development of fractals, for instance, is entirely based on the idea of self-similarity (i.e., scale-invariance up to a scalar factor) [1], [2]. This self-similarity can be deterministic—in which case we are led to deterministic fractals such as the famous Koch snowflake, or the elaborate Mandelbrot set—but it can also be understood in a statistical sense—leading to stochastic fractals, the prime examples of which are fractional Brownian motion (fBm) processes [3] (see also Chainais *et al.* [4] for a generalization based on the notion of scaling).

Fractional Brownian motion models generalize Lévy's Brownian motion [5] of Gaussian type. These processes have long been associated with the phenomenon of long-range dependence and $1/f^{\alpha}$-like power spectra that frequently appear in areas as diverse as hydrology, financial mathematics, network traffic analysis, terrain modeling, and image processing [1], [6]–[8]. In the case of the latter, the relevance of fBm processes in modeling images has been claimed on the basis of observations of scale-invariance and the associated power-law spectra in natural images [9]–[11].

A multivariate fBm field $\mathfrak{B}_H$ is a nonstationary Gaussian process[1] identified by a single parameter $0 < H < 1$—the *Hurst* parameter, after Harold Edwin Hurst (1880–1978), for his seminal contribution to the study of such processes in the context of hydrology [2], [12]—that characterizes its covariance up to a scalar normalization factor

$$\mathbf{E}\{\mathfrak{B}_H(\boldsymbol{x})\mathfrak{B}_H(\boldsymbol{x}')\} \propto \|\boldsymbol{x}\|^{2H} + \|\boldsymbol{x}'\|^{2H} - \|\boldsymbol{x}-\boldsymbol{x}'\|^{2H}.$$

Estimation of the Hurst parameter is important in practical applications, and is, e.g., used in image processing to classify different types of texture based on their second order statistics [13], [14].

Multiresolution analysis [15], [16] was identified early on in its development as a decidedly effective tool for the study of self-similarity [17]–[26]. Its utility in the estimation of parameters of self-similar processes (especially in the 1-D setting and for estimating the Hurst parameter) is, therefore, well documented [20], [27]–[29]. The essential observation in this regard is that the logarithm of the wavelet energy of an fBm process varies linearly with scale, with a slope that depends on the Hurst parameter H.

Intuitively, the above observation appears deceptively simple. After all, this would seem to be a straightforward consequence of the $1/f^{\alpha}$-like power spectrum of fBm and the logarithmic spectral partitioning afforded by the wavelet transform. A rigorous derivation of this result is, however, subtler, as fBm—being nonstationary—does not, in fact, have a power spectrum in the classical sense.

On account of this, one of our main motivations in writing this paper has been to propose a rigorous interpretation of the spectral characterization of multivariate isotropic fBm, in the sense of a whitening/innovation model (Section V). This distributional framework, which is deduced from basic invariance principles (Section III), provides a powerful formalism for defining and analysing fBm and similar processes. Our results here generalize those of Blu and Unser [30] who studied the single-variable case. Operator models for self-similar fields were also studied

Manuscript received October 09, 2007; revised October 20, 2008. Current version published March 13, 2009. This work was supported in part by the Swiss Science Foundation. This work was supported by the Swiss Science Foundation. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Minh N. Do.

The authors are with the Biomedical Imaging Group,École Polytechnique Fédérale de Lausanne (EPFL), Station 17, CH-1015 Lausanne, Switzerland (e-mail: p.d.tafti@ieee.org; dimitri.vandeville@epfl.ch; michael.unser@epfl.ch).

Digital Object Identifier 10.1109/TIP.2008.2011451

[1]In this paper, we do not distinguish between random *processes* and random *fields*, using both terms interchangeably to refer to multivariate random functions.

by Benassi *et al.* [31], [32], who focused on the link between operators and multivariate random fields and their relation to wavelets. The 1-D analysis of Wyss [33] is also relevant.

The said formulation also links the study of fBm processes to spline theory via providing a convenient and unifying interpretation of fBm processes and polyharmonic splines as stochastic vs deterministic solutions to the same (fractional) partial differential equation [34]. This, in the light of the fundamental relation between splines and wavelets (Section IV), allows us to derive interesting and general results concerning the wavelet analysis of fractional Brownian motion (Section V). We for instance show the quasi-whitening effect of a polyharmonic wavelet transform on fBm processes.

To complement the mathematical toolset for the analysis of multivariate fBm, we have included a comprehensive account of a general construction scheme for multidimensional polyharmonic spline multiresolution spaces, proving all essential properties for forming a multiresolution analysis. The generality of our construction (which extends the works of Rabut and Bacchelli *et al.* [35], [36] and Van De Ville *et al.* [37]) makes it suitable for multiresolution approximation in any number of dimensions and on virtually all sampling lattices of interest that display some form of isotropy.

The organization of the remainder of the paper, in brief, is as follows. In Section II, we review some mathematical preliminaries. We formalize the idea of isotropic affine invariance in Section III and use it to identify a family of fractional partial differential operators that appear in the characterization of both polyharmonic splines and fBm processes. The theory of multidimensional polyharmonic spline multiresolution is developed in Section IV. Next, in Section V, we provide a characterization of fBm based on an innovation model. We then exploit the link between splines and fBm processes in Section VI, to derive some characteristic results concerning polyharmonic wavelet analysis of fBm. Based on these results, estimation of the Hurst parameter is also discussed and a few experimental results are provided in Section VI-B. Some final remarks conclude the paper.

II. Mathematical Preliminaries and Notation

The theory of generalized random processes utilized in this paper is exposited in the works of Gel'fand *et al.* [38], [39]. For reference, some of the main definitions are summarized in this section. This section shall also serve to fix our basic notation and to recall some facts and definitions from the theory of lattices.

A. Some Notational Conventions

We use the MATLAB notation for row and column vectors and also follow the multi-index convention, according to which, given a vector $\boldsymbol{x} = [x_1; \ldots; x_d] \in \mathbb{R}^d$ and a multi-index $\boldsymbol{k} = [k_1; \ldots; k_d] \in \mathbb{Z}^d_{\geq 0}$ (d always denotes the dimensionality of the domain)

$$\boldsymbol{x}^{\boldsymbol{k}} \stackrel{\text{def}}{=} \prod_{1 \leq i \leq d} x_i^{k_i}, \quad \boldsymbol{k}! \stackrel{\text{def}}{=} \prod_{1 \leq i \leq d} k_i!, \quad \text{and} \quad |\boldsymbol{k}| \stackrel{\text{def}}{=} \sum_{1 \leq i \leq d} k_i.$$

Other notation is defined where first used.

B. Generalized Functions

A regular function u of a variable $\boldsymbol{x} \in \mathbb{R}^d$ is characterized by the value it assigns to its argument $\boldsymbol{x}$ (i.e., $u(\boldsymbol{x})$ for $\boldsymbol{x} \in \mathbb{R}^d$). In contrast, a *generalized function* or *distribution* f is specified in terms of inner-products[2] $\langle f, u \rangle$ with *test functions* u belonging to some inner-product space $\mathcal{K}$. Intuitively, these inner-products can be interpreted as linear observations or measurements of f. The advantage is that in this framework we can conceive of entities that need no longer be defined point-wise. The space of all generalized functions defined by their (bounded) inner-products with elements of $\mathcal{K}$ is identified with $\mathcal{K}'$, the continuous dual of $\mathcal{K}$.

Given an operator A with adjoint A^*, both defined on our space of test functions, we may extend the domain of A to the corresponding space of generalized functions ($\mathcal{K}'$) using the following defining identity

$$\langle \mathrm{A}f, u \rangle \stackrel{\text{def}}{=} \langle f, \mathrm{A}^* u \rangle.$$

Thus, e.g., for the shift operator we shall have

$$\langle f(\cdot - \boldsymbol{h}), u(\cdot) \rangle \stackrel{\text{def}}{=} \langle f(\cdot), u(\cdot + \boldsymbol{h}) \rangle, \text{ for all } u \in \mathcal{K}.$$

The Fourier transform defines a one-to-one mapping between a suitably chosen space $\mathcal{K}$ of test functions and the space $\hat{\mathcal{K}}$ of their Fourier transforms. With Parseval's identity in mind, the Fourier transform of a generalized function $f(\boldsymbol{x}) \in \mathcal{K}'$ can be defined as the generalized function $\hat{f}(\boldsymbol{\omega}) \in \hat{\mathcal{K}}'$ that satisfies the identity

$$\langle \hat{f}(\boldsymbol{\omega}), \hat{u}(\boldsymbol{\omega}) \rangle = (2\pi)^d \langle f(\boldsymbol{x}), u(\boldsymbol{x}) \rangle, \text{ for all } u \in \mathcal{K}.$$

If we choose $\mathcal{K}$ to be the Schwartz space of d-variate rapidly decaying smooth functions (denoted here by $\mathcal{S}(\mathbb{R}^d)$ or simply by $\mathcal{S}$), $\mathcal{K}$ and $\hat{\mathcal{K}}$ (and, therefore, $\mathcal{K}'$ and $\hat{\mathcal{K}}'$) coincide. A familiar example of a generalized function defined over $\mathcal{S}$ is Dirac's delta

$$\langle \delta, u \rangle \stackrel{\text{def}}{=} u(\boldsymbol{0}).$$

The Fourier transform of $\delta(\boldsymbol{x})$ is the constant 1, since $\langle 1, \hat{u} \rangle = \int \mathrm{d}\boldsymbol{\omega}\, 1 \hat{u}(\boldsymbol{\omega}) = (2\pi)^d u(\boldsymbol{0}) = (2\pi)^d \langle \delta, u \rangle$.

C. Generalized Random Processes and Random Fields

To generalize the notion of a random process a similar approach may be used, where one replaces point values by inner products. Accordingly, in the stochastic analysis of Gel'fand and Vilenkin [39], a generalized random process $\mathfrak{X}$ is defined as a random generalized function, which is to say that it corresponds to a family of random variables

$$\mathfrak{X}_u \stackrel{\text{def}}{=} \langle \mathfrak{X}, u \rangle, \quad u \in \mathcal{K}$$

characterized by the consistent specification of a joint probability measure for all finite sets of test functions u. This should

[2]What we shall here refer to as an inner-product is in more accurate (but perhaps less familiar) terms a duality pairing.

be compared with the definition of classical random processes, where point-wise random variables $\mathfrak{X}(\boldsymbol{x})$ replace the $\mathfrak{X}_u$s.

Characteristic Functional: A (real-valued) generalized random process $\mathfrak{X}$ can also be described by its characteristic functional

$$\mathbf{Z}_{\mathfrak{X}}(u) \stackrel{\text{def}}{=} \mathbf{E}\left\{\mathrm{e}^{\mathrm{j}\langle \mathfrak{X}, u\rangle}\right\}$$

(where $\mathbf{E}$ denotes the expectation functional). The characteristic functional is continuous and positive-definite, and is equal to 1 for $u \equiv 0$. It provides a complete description of the random process $\mathfrak{X}$. This is due to the fact that

$$\mathbf{Z}_{\mathfrak{X}}\left(\sum_{1\le k\le N} \omega_k u_k\right)$$

is a continuous and positive-definite function of ω_ks and, hence, by Bochner's theorem, corresponds to the Fourier transform of a probability measure—specifically, the joint probability measure of $\mathfrak{X}_{u_1}, \ldots, \mathfrak{X}_{u_N}$ [39, ch. III, §2.6].

(In comparison, in the classical theory $\mathbf{E}\left\{\mathrm{e}^{\mathrm{j}\sum_k \omega_k \mathfrak{X}(\boldsymbol{x}_k)}\right\}$ provides the Fourier transform of the joint probability measure of $\mathfrak{X}(\boldsymbol{x}_1), \ldots, \mathfrak{X}(\boldsymbol{x}_N)$. Informally, this would correspond to choosing $\sum_{1\le k\le N} \omega_k \delta(\cdot - \boldsymbol{x}_k)$ as the "test" function).

Correlation Form: The correlation form $\langle\!\langle u, v\rangle\!\rangle_{\mathfrak{X}}$ of the (real) random process $\mathfrak{X}$ is defined as

$$\langle\!\langle u, v\rangle\!\rangle_{\mathfrak{X}} \stackrel{\text{def}}{=} \mathbf{E}\{\mathfrak{X}_u \mathfrak{X}_v\}.$$

The following relationship exists between the generalized correlation form $\langle\!\langle u, v\rangle\!\rangle_{\mathfrak{X}}$ and the (generalized) correlation function $c_{\mathfrak{X}}(\boldsymbol{x}, \boldsymbol{x}')$ of a generalized random process

$$\langle\!\langle u, v\rangle\!\rangle_{\mathfrak{X}} = \int \mathrm{d}\boldsymbol{x}\, \mathrm{d}\boldsymbol{x}'\; c_{\mathfrak{X}}(\boldsymbol{x}, \boldsymbol{x}') u(\boldsymbol{x}) v(\boldsymbol{x}'). \tag{1}$$

In addition, for a Gaussian random process, the characteristic functional and the correlation form are related by the equation

$$\mathbf{Z}_{\mathfrak{X}}(u) = \mathrm{e}^{-(1/2)\langle\!\langle u, u\rangle\!\rangle_{\mathfrak{X}}}.$$

This shows that, as expected, a Gaussian process is fully characterized by its correlation form.

D. Lattices

A *lattice* $\pounds_0$ in $\mathbb{R}^d$ is the set of all integer linear combinations of d linearly independent vectors $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_d$; that is

$$\pounds_0 = \mathsf{Q}\mathbb{Z}^d$$

with $\mathsf{Q} = [\boldsymbol{q}_1, \ldots, \boldsymbol{q}_d]$[40], [41]. In general, there exist several *generator matrices* Q that lead to the same lattice. Yet, they all have the same absolute determinant $|\mathsf{Q}|$ (known as the *sampling density*). For simplicity, we shall assume the normalization $|\mathsf{Q}| = 1$.

A multidimensional lattice may be partitioned into so-called *cosets* that are translates of one another. This is a generalization of the concept of dividing the set of integers $\mathbb{Z}$ into even and odd numbers, or, more generally, into k equivalence classes modulo k. In the case of lattices in $\mathbb{R}^d$, such a partitioning is achieved by means of a *subsampling matrix* D, which plays the role of the integer k in the 1-D case. D is an integer $d \times d$ matrix with all eigenvalues strictly greater than 1 in the absolute. It is used to define a subsampling relation for lattices

$$\pounds_{n+1} = \mathsf{QDQ}^{-1} \pounds_n \stackrel{\text{def}}{=} \mathsf{D}_{\mathsf{Q}} \pounds_n. \tag{2}$$

From there

$$\pounds_n = \mathsf{QD}^n \mathbb{Z}^d = \mathsf{D}_{\mathsf{Q}}^n \pounds_0.$$

Similar to the partitioning of the integers modulo k, we find a two-scale relationship for the decomposition of $\pounds_n$ into $|\mathsf{D}|$ cosets, which are translated versions of the lower resolution lattice $\pounds_{n+1}$

$$\pounds_n = \left(\bigcup_{1\le i<|\mathsf{D}|} \pounds_{n+1} + \mathsf{QD}^n \boldsymbol{\zeta}_i\right) \cup \pounds_{n+1}. \tag{3}$$

Here, the multiinteger vectors $\boldsymbol{\zeta}_i$—taken to be of minimum length and dubbed *principal coset representatives*—are specified uniquely modulo $\mathsf{D}\mathbb{Z}^d$.

For a given lattice hierarchy $\pounds_n$, $n \in \mathbb{Z}$, the *dual* (or *reciprocal*) lattice hierarchy $\pounds_{-n}^*$ is defined by the relation

$$\boldsymbol{p}^{\mathsf{T}} \boldsymbol{q} \in \mathbb{Z}, \quad \text{for all } \boldsymbol{q} \in \pounds_n, \ \boldsymbol{p} \in \pounds_{-n}^*.$$

It follows that this hierarchy can be constructed using the matrix pair of $\mathsf{Q}^{-\mathsf{T}} \stackrel{\text{def}}{=} (\mathsf{Q}^{\mathsf{T}})^{-1}$ and D^{T}. Accordingly, we also define $\mathsf{D}_{\mathsf{Q}}^* \stackrel{\text{def}}{=} \mathsf{Q}^{-\mathsf{T}} \mathsf{D}^{\mathsf{T}} \mathsf{Q}^{\mathsf{T}}$.

We define the *lattice convolution operator* or *lattice filter* corresponding to a sequence $v[\boldsymbol{k}]$, $\boldsymbol{k} \in \mathbb{Z}^d$, as the operator

$$\mathrm{V}_{\mathsf{Q}} : f(\cdot) \mapsto \sum_{\boldsymbol{k}\in\mathbb{Z}^d} v[\boldsymbol{k}] f(\cdot - \mathsf{Q}\boldsymbol{k}).$$

Its Fourier expression is

$$\hat{\mathrm{V}}_{\mathsf{Q}}(\boldsymbol{\omega}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} v[\boldsymbol{k}] \mathrm{e}^{-\mathrm{j}\boldsymbol{k}^{\mathsf{T}}\mathsf{Q}^{\mathsf{T}}\boldsymbol{\omega}}.$$

Conversely, those and only those operators with Fourier expressions that can be written in the above form represent lattice convolutions. These Fourier expressions are in effect those that are $2\pi\pounds_0^*$-periodic (i.e., $2\pi\boldsymbol{p}$-periodic for any $\boldsymbol{p} \in \pounds_0^*$).

We also have a lattice version of the Poisson formula

$$\mathcal{F}\left\{\sum_{\boldsymbol{k}\in\mathbb{Z}^d} \delta(\boldsymbol{x} - \mathsf{Q}\boldsymbol{k})\right\} = \frac{(2\pi)^d}{|\mathsf{Q}|} \sum_{\boldsymbol{k}\in\mathbb{Z}^d} \delta(\boldsymbol{\omega} - 2\pi\mathsf{Q}^{-\mathsf{T}}\boldsymbol{k}). \tag{4}$$

Remark 1: The families of multiscale lattices that we shall consider in this work are restricted in two ways.

LAT–1. First, for our multiresolution construction we are interested in *self-similar multiscale lattices.* This means that the lattice coarsening matrix D_{Q}—and, consequently, its dual $\mathsf{D}_{\mathsf{Q}}^*$—should correspond to similarity transforms.

LAT–2. Second, we require the existence of a $d \times N$ integer matrix $\mathsf{Y} \stackrel{\text{def}}{=} [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N]$ ($N \ge d$), such that the lattice

vectors $\mathsf{Q}\boldsymbol{y}_1, \ldots, \mathsf{Q}\boldsymbol{y}_N$ generate $\mathcal{L}_0$, and constitute a tight frame for $\mathbb{R}^d$. The latter is equivalent to requiring that

$$\mathsf{Q}\mathsf{Y}\mathsf{Y}^{\mathsf{T}}\mathsf{Q}^{\mathsf{T}} = \sum_{1 \le i \le N} \mathsf{Q}\boldsymbol{y}_i\boldsymbol{y}_i^{\mathsf{T}}\mathsf{Q}^{\mathsf{T}} = \mu^2 \mathbf{I} \tag{5}$$

for some scalar μ. We furthermore assume $\{\mathsf{Q}\boldsymbol{y}_i\}$ to be simple, i.e., not to contain any pair of linearly dependent vectors.

We note that for *any* lattice, there exist infinitely many subsampling schemes that satisfy the first requirement. In addition, the second requirement is satisfied by virtually all lattices that are typically used in multidimensional multiresolution signal processing (such as the Cartesian, quincunx, and hexagonal lattices in $\mathbb{R}^2$, and the Cartesian, FCC, and BCC lattices in $\mathbb{R}^3$). For instance, for the Cartesian and quincunx lattices in $\mathbb{R}^2$ (both with $\mathsf{Q} = [1, 0; 0, 1]$), the matrices

$$\mathsf{Y} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathsf{Y}' = \begin{bmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

provide two examples of such systems. A similar system for the hexagonal lattice (with $\mathsf{Q} \propto [1, 0.5; 0, \sqrt{3}/2]$) uses the matrix

$$\mathsf{Y}'' = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}.$$

III. Operators and Invariances

The fundamental observation that underlies this work is that we can characterize specific classes of splines and stochastic processes as solutions to a fractional partial differential equation of the form

$$\mathrm{U}\{\text{solution}\} = \text{driving term}$$

where U is a fractional partial differential operator with certain properties, and the *driving term* is either a sum of Dirac deltas (in the deterministic formulation, leading to U-splines) or a white Gaussian noise process (in the stochastic formulation, leading to random processes *whitened* by U).

In this section we shall use invariance principles to define a particular family of such fractional partial differential operators that produce polyharmonic splines (Section IV) as deterministic solutions and also characterize isotropic multidimensional fractional Brownian motion (Section V) in the stochastic setting.

The link between the deterministic and stochastic formulations is later explored in Section VI, where we investigate the properties of polyharmonic wavelet analysis of fractional Brownian motion.

A. Scale- and Rotation-Invariant Operators

The invariances we shall consider are those under the scaling, shifts, and rotations of the coordinate system [1], with the first leading to self-similar fractal structures, and the latter two relieving us from the—uncomfortable and often arbitrary—choice of an origin and a set of preferred directions.

Specifically, we shall study a family of convolution operators with continuous Fourier expressions, which, in addition to shift-invariance (intrinsic to convolution), have the following invariance properties.

INV–1. Scale-invariance: The operators of interest commute with scaling operators (up to a constant that may vary continuously with scale) in order to allow multiscale constructions. In mathematical notation, we want

$$\mathrm{U} \circ \mathrm{S}_a = \alpha(a)\, \mathrm{S}_a \circ \mathrm{U}$$

where $\mathrm{S}_a : f(\boldsymbol{x}) \mapsto f(a^{-1}\boldsymbol{x}), a > 0$, represents the scaling operator and $\alpha(a)$ is a strictly positive continuous function.

INV–2. Rotation-invariance: The operators are in addition invariant under rotations of the coordinate system and, therefore, lead to isotropic models. In other words, the operators commute with rotations about the origin

$$\mathrm{U} \circ \mathrm{R}_{\boldsymbol{\theta}} = \mathrm{R}_{\boldsymbol{\theta}} \circ \mathrm{U}.$$

The following is a known result in the context of rotation- and scale-invariant quadratic functionals (in this case, $Q(f) \stackrel{\text{def}}{=} \|\mathrm{U}f\|_2^2$) [42]–[44].

Theorem 1: The (per assumption continuous) Fourier expression of a real operator U fulfilling requirements INV–1 and INV–2 has the following form for some $\gamma \ge 0$

$$\hat{\mathrm{U}}(\boldsymbol{\omega}) = c\|\boldsymbol{\omega}\|^{2\gamma}. \tag{6}$$

The normalized version of such an operator (with $c = 1$), which we denote by Δ^γ, can be considered the γth real (fractional) iterate of the Laplacian (albeit discarding a factor of $(-1)^\gamma$). The following are easy to check

$$\Delta^0 = \text{identity}; \qquad \Delta^\gamma \Delta^{\gamma'} = \Delta^{\gamma+\gamma'}. \tag{7}$$

The fractional Laplacian has a nontrivial null-space and, as a result, infinitely many inverses differing in terms from the null-space.

Remark 2: The null-space includes, for instance, certain functions with (generalized) Fourier transforms concentrated at the origin (i.e., at $\boldsymbol{\omega} = \mathbf{0}$). Since any such generalized Fourier symbol can be written as a finite sum of derivatives of $\delta(\boldsymbol{\omega})$[45, ch. II, §4.5, p. 119, Theorem], the corresponding members of the null-space are polynomial functions up to a certain degree. This, however, is not a complete characterization of the null-space in general.

B. Inverse Operators

Looking back at (7), one may be tempted to define the inverse of Δ^γ as the operator $\Delta^{-\gamma}$ with the Fourier expression

$$\|\boldsymbol{\omega}\|^{-2\gamma}.$$

It is immediately noticed, however, that this Fourier form has a nonsummable singularity at the origin for $2\gamma \ge d$; therefore, in general, the integral

$$\Delta^{-\gamma} f(\boldsymbol{x}) = (2\pi)^{-d} \int_{\mathbb{R}^d} \mathrm{d}\boldsymbol{\omega}\, \mathrm{e}^{\mathrm{j}\boldsymbol{x}^{\mathsf{T}}\boldsymbol{\omega}} \|\boldsymbol{\omega}\|^{-2\gamma} \hat{f}(\boldsymbol{\omega}) \tag{8}$$

needs to be properly interpreted, i.e., regularized.[3] Since regularization can be done in more than one way, $\Delta^{-\gamma}$ in fact represents a family of inverses rather than a single one.

Different regularizations essentially correspond to different (boundary or other) linear constraints on the solution of a fractional differential equation of the form

$$\Delta^{\gamma}\rho(\boldsymbol{x}) = f(\boldsymbol{x}).$$

These constraints may be satisfied by adding an appropriate term from the null-space of Δ^{γ} to a particular solution.

One of the possible inverse operators is the *left* inverse (introduced by Blu and Unser in the single-variable setting [30]; denoted by $\grave{\Delta}^{-\gamma}$ here), which is obtained by removing a sufficient number of lower order terms from the Taylor series expansion of $\hat{f}(\boldsymbol{\omega})$ at the origin

$$\grave{\Delta}^{-\gamma} f(\boldsymbol{x}) \stackrel{\text{def}}{=} (2\pi)^{-d} \times \int_{\mathbb{R}^d} \mathrm{d}\boldsymbol{\omega} \mathrm{e}^{\mathrm{i}\boldsymbol{x}^{\mathsf{T}}\boldsymbol{\omega}} \frac{\hat{f}(\boldsymbol{\omega}) - \sum_{|\boldsymbol{k}|\le\lfloor 2\gamma - d/2\rfloor} \hat{f}(\boldsymbol{k})(\boldsymbol{0}) \frac{\boldsymbol{\omega}^{\boldsymbol{k}}}{\boldsymbol{k}!}}{\|\boldsymbol{\omega}\|^{2\gamma}}. \quad (9)$$

It can be checked that

$$\grave{\Delta}^{-\gamma}\Delta^{\gamma} f = f$$

for any $f \in \mathcal{S}$, hence the name *left* inverse.

The adjoint of $\grave{\Delta}^{-\gamma}$ over $\mathcal{S}$ is the operator $\acute{\Delta}^{-\gamma}$ defined by

$$\acute{\Delta}^{-\gamma} f(\boldsymbol{x}) \stackrel{\text{def}}{=} (2\pi)^{-d} \times \int_{\mathbb{R}^d} \mathrm{d}\boldsymbol{\omega} \frac{\mathrm{e}^{\mathrm{i}\boldsymbol{x}^{\mathsf{T}}\boldsymbol{\omega}} - \sum_{|\boldsymbol{k}|\le\lfloor 2\gamma - d/2\rfloor} \frac{\mathrm{j}^{|\boldsymbol{k}|}\boldsymbol{x}^{\boldsymbol{k}}\boldsymbol{\omega}^{\boldsymbol{k}}}{\boldsymbol{k}!}}{\|\boldsymbol{\omega}\|^{2\gamma}} \hat{f}(\boldsymbol{\omega}). \quad (10)$$

It satisfies

$$\Delta^{\gamma}\acute{\Delta}^{-\gamma} f = f$$

for all $f \in \mathcal{S}$ and is called the *right* inverse. We can extend $\acute{\Delta}^{-\gamma}$ to a subset of $\mathcal{S}'$ by duality

$$\langle \acute{\Delta}^{-\gamma} f, u\rangle \stackrel{\text{def}}{=} \langle f, \grave{\Delta}^{-\gamma} u\rangle$$

wherever the r.h.s. is meaningful for all $u \in \mathcal{S}$.

While the above definitions may look arbitrary at first glance, they have intuitive interpretations. For example, supposing $f(\boldsymbol{x})$ to be a well-behaved test function whose moments vanish up to degree $\lfloor 2\gamma - d/2\rfloor$, (9) simply corresponds to a shift-invariant inverse (all the terms in the sum will be zero in this case), while (10) defines an inverse with all derivatives up to order $\lfloor 2\gamma - d/2\rfloor$ forced to be zero at the origin. This latter property is significant in the characterization of fractional Brownian motion as there, by definition, the process should equal zero at $\boldsymbol{x} = \boldsymbol{0}$.

It also bears mentioning that, unlike the fractional Laplacian, these inverse operators are in general *not* shift-invariant when applied to members of $\mathcal{S}$ (they are, however, scale- and rotation-invariant in the previously defined sense).

[3]"Regularization" here stands for a general way of assigning a value to an integral with a singular kernel, in a manner that would be consistent with what one would expect when evaluating the integral for a smooth function that vanishes in a neighborhood of the singularity (and for which the integral can be evaluated).

IV. Polyharmonic Splines and Wavelets

A. Splines and Operators

By differentiating a polynomial spline a sufficient number of times, one procures a sum of Dirac deltas located at the knots. This observation underlies a conceptual framework in which splines are defined as functions that are mapped to a sum of Dirac deltas by some suitably chosen operator U. This approach leads to interesting generalizations: one may for example use fractional derivatives to obtain splines of fractional order [34], [46].

Formally, in this framework, given a shift invariant operator U, we define a *lattice* U-*spline* as a function $s(\boldsymbol{x})$ for which

$$\mathrm{U}s(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} c[\boldsymbol{k}]\delta(\boldsymbol{x} - \mathsf{Q}\boldsymbol{k}) \quad (11)$$

with $c \in \ell_\infty(\mathbb{Z}^d)$ and where the points $\mathsf{Q}\boldsymbol{k}$ belong to a lattice. One may try to solve the equation

$$\mathrm{U}\varrho(\boldsymbol{x}) = \delta(\boldsymbol{x}) \quad (12)$$

for $\varrho(\boldsymbol{x})$ (Green's function) by finding an inverse operator. $s(\boldsymbol{x})$ can then be expressed in terms of $\varrho(\boldsymbol{x})$ and its lattice shifts, plus a term from the null-space of U; that is

$$s(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} c[\boldsymbol{k}]\varrho(\boldsymbol{x} - \mathsf{Q}\boldsymbol{k}) + s_0(\boldsymbol{x})$$

with $\mathrm{U}s_0 = 0$.

In practice, it is often of interest to limit oneself to splines $s(\boldsymbol{x}) \in \mathrm{L}_2(\mathbb{R}^d)$, in which case we consider a modified version of the above problem, where we introduce a *localization operator (filter)* V_{Q} and study the equation

$$\mathrm{U}\phi(\boldsymbol{x}) = \mathrm{V}_{\mathsf{Q}}\delta(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} v[\boldsymbol{k}]\delta(\boldsymbol{x} - \mathsf{Q}\boldsymbol{k}) \quad (13)$$

in place of (12). B-splines, which form spatially localized bases for square-integrable spline spaces, are in fact solutions to such equations [34], [37].

In the remainder of this section, we first introduce such localized (B-spline) bases for spaces of square-integrable *polyharmonic* splines, for which the operator U is a fractional Laplacian, and V_{Q} is its discretization over any one of the lattices introduced in Remark 1. Next, in Section IV-C, we show how these B-splines can act as scaling functions for a multiresolution analysis (Theorem 2). We follow this by the investigation of one of the main properties of wavelets derived from these B-splines, namely that polyharmonic wavelet kernels behave like low-frequency approximations of the fractional Laplacian (Theorem 3).

B. Polyharmonic B-Splines

If we take the operator U of the previous subsection to be the fractional Laplacian Δ^{γ}, solutions to (11) (which in this case is a polyharmonic equation) are called *polyharmonic splines* [56].

As noted in Section III-B, when the function $\hat{f}(\boldsymbol{\omega})$ has sufficiently many zeros at the origin, the fractional Laplacian can be inverted via (8) without difficulty. Indeed, one of the ways to deal with singular integrals is to multiply the integration kernel by a function that vanishes at the singularity.

It is, therefore, reasonable in our problem to first choose an appropriate localization filter $\mathrm{V_Q}$ whose Fourier symbol $\hat{\mathrm{V}}_\mathrm{Q}(\boldsymbol{\omega})$ approximates that of Δ^γ at its zero at the origin, thus cancelling the singularity of $\Delta^{-\gamma}$ and permitting us to solve the spline equation

$$\Delta^\gamma \phi_{2\gamma}(\boldsymbol{x}) = \mathrm{V_Q}\delta(\boldsymbol{x}) \tag{14}$$

in the Fourier domain, for the B-spline $\phi_{2\gamma}(\boldsymbol{x})$. Different choices of such an operator $\mathrm{V_Q}$ lead to different families of polyharmonic B-splines (quasi-isotropic, orthogonal, *etc.*) [37].

In the simplest case, the elementary localization filter $\mathrm{V_Q}$ corresponds (up to a factor of $(-1)^\gamma$) to the γth fractional iteration of an elementary discretization of the Laplacian. Specifically, for $\gamma = 1$ we define the elementary localization operator, Δ_Q, in the spatial domain as follows.

$$\Delta_\mathrm{Q} f \stackrel{\mathrm{def}}{=} \frac{1}{\mu^2} \sum_{1 \le i \le N} 2f(\cdot) - f(\cdot - \mathrm{Q}\boldsymbol{y}_i) - f(\cdot + \mathrm{Q}\boldsymbol{y}_i)$$

(see LAT–2 for the definition of $\boldsymbol{y}_i$s). Note also its Fourier symbol

$$\hat{\Delta}_\mathrm{Q}(\boldsymbol{\omega}) = \frac{4}{\mu^2} \sum_{1 \le i \le N} \sin^2\left(\frac{\boldsymbol{y}_i^\mathsf{T}\mathrm{Q}^\mathsf{T}\boldsymbol{\omega}}{2}\right). \tag{15}$$

For other values of $\gamma > 0$, we simply define

$$\hat{\Delta}_\mathrm{Q}^\gamma(\boldsymbol{\omega}) \stackrel{\mathrm{def}}{=} \left[\hat{\Delta}_\mathrm{Q}(\boldsymbol{\omega})\right]^\gamma.$$

This choice of the localization operator leads to a fractional generalization of Rabut's elementary γ-harmonic B-splines, here denoted $\phi_{2\gamma,\mathrm{el}}$ [35], [37].

More generally, the localization operator $\hat{\mathrm{V}}_\mathrm{Q}$ used in (14) can be any one with a Fourier symbol factorizable as

$$\hat{\mathrm{V}}_\mathrm{Q}(\boldsymbol{\omega}) = \hat{\Delta}_\mathrm{Q}^\gamma(\boldsymbol{\omega})\,\hat{\mathrm{T}}_\mathrm{Q}^\gamma(\boldsymbol{\omega})$$

where $\hat{\mathrm{T}}_\mathrm{Q}^\gamma(\boldsymbol{\omega}) \stackrel{\mathrm{def}}{=} \left[\hat{\mathrm{T}}_\mathrm{Q}(\boldsymbol{\omega})\right]^\gamma$ is the continuous Fourier expression of some lattice operator (filter), and is bounded from above and below with a strictly positive lower bound. We shall assume $\hat{\mathrm{T}}_\mathrm{Q}(\boldsymbol{\omega})$ to be normalized with $\hat{\mathrm{T}}_\mathrm{Q}(\mathbf{0}) = 1$.

Remark 3: The choice of $\hat{\mathrm{T}}_\mathrm{Q}^\gamma(\boldsymbol{\omega})$, apart from these constraints, is essentially arbitrary in so far as it corresponds to a discrete (lattice) filter, as all such choices lead to the same multiresolution subspaces. However, as will be seen shortly, different choices of $\hat{\mathrm{T}}_\mathrm{Q}$ do lead to different B-spline functions spanning the same spaces, and $\hat{\mathrm{T}}_\mathrm{Q}$ may be specifically selected so as to give these functions a desired correlation structure.

The solution to (14) can now be written explicitly in the Fourier domain as

$$\hat{\phi}_{2\gamma}(\boldsymbol{\omega}) = \frac{\hat{\mathrm{V}}_\mathrm{Q}(\boldsymbol{\omega})}{\|\boldsymbol{\omega}\|^{2\gamma}} = \hat{\mathrm{T}}_\mathrm{Q}^\gamma(\boldsymbol{\omega})\frac{\hat{\Delta}_\mathrm{Q}^\gamma(\boldsymbol{\omega})}{\|\boldsymbol{\omega}\|^{2\gamma}} = \hat{\mathrm{T}}_\mathrm{Q}^\gamma(\boldsymbol{\omega})\hat{\phi}_{2\gamma,\mathrm{el}}(\boldsymbol{\omega}) \tag{16}$$

where $\hat{\phi}_{2\gamma,\mathrm{el}}(\boldsymbol{\omega}) \stackrel{\mathrm{def}}{=} \hat{\Delta}_\mathrm{Q}^\gamma(\boldsymbol{\omega})/\|\boldsymbol{\omega}\|^{2\gamma}$ is the Fourier transform of the elementary γ-harmonic B-spline $\phi_{2\gamma,\mathrm{el}}$ that was mentioned before.

In order for the *polyharmonic B-spline* function $\phi_{2\gamma}(\boldsymbol{x})$ thus defined to be square-integrable we need to have

$$\gamma > \frac{d}{4}. \tag{17}$$

The following proposition summarizes the smoothness and integrability properties of $\phi_{2\gamma}(\boldsymbol{x})$.

Proposition 1: $\phi_{2\gamma}$, with $\gamma > d/4$, belongs to the Sobolev space $\mathcal{H}^s$ for any $s < 2\gamma - d/2$.

Proof: Using the Taylor expansion of $\hat{\Delta}_\mathrm{Q}^\gamma(\boldsymbol{\omega})$, we can immediately see that $\hat{\phi}_{2\gamma,\mathrm{el}}(\boldsymbol{\omega})$ tends to 1 as $\|\boldsymbol{\omega}\| \to 0$

$$\lim_{\|\boldsymbol{\omega}\|\to 0} \hat{\phi}_{2\gamma,\mathrm{el}}(\boldsymbol{\omega}) = \lim_{\|\boldsymbol{\omega}\|\to 0} \left|\frac{\sum_i \frac{4}{\mu^2}\sin^2\left(\frac{\boldsymbol{y}_i^\mathsf{T}\mathrm{Q}^\mathsf{T}\boldsymbol{\omega}}{2}\right)}{\|\boldsymbol{\omega}\|^2}\right|^\gamma$$
$$= \lim_{\|\boldsymbol{\omega}\|\to 0} \left|\frac{\boldsymbol{\omega}^\mathsf{T}\left(\sum_i \frac{\mathrm{Q}\boldsymbol{y}_i\boldsymbol{y}_i^\mathsf{T}\mathrm{Q}^\mathsf{T}}{\mu^2}\right)\boldsymbol{\omega}}{\boldsymbol{\omega}^\mathsf{T}\boldsymbol{\omega}}\right|^\gamma = 1$$

[cf. (5)]. In addition, both $\hat{\Delta}_\mathrm{Q}^\gamma(\boldsymbol{\omega})$ and $\hat{\mathrm{T}}_\mathrm{Q}^\gamma(\boldsymbol{\omega})$ are by definition continuous and bounded. What all this means is that $\hat{\phi}_{2\gamma}(\boldsymbol{\omega})$ is continuous and bounded everywhere and decays like $|\boldsymbol{\omega}|^{2\gamma}$ [cf. (16)]. It then follows from the Fourier-domain definition of the Sobolev space $\mathcal{H}^s$ that $\phi_{2\gamma} \in \mathcal{H}^s$ for all $s < 2\gamma - d/2$. ∎

As was already mentioned, the trivial choice of $\hat{\mathrm{T}}_\mathrm{Q}(\boldsymbol{\omega}) \equiv 1$ in (16) leads to elementary fractional polyharmonic B-splines. Among other possibilities, one can, e.g., opt for the orthogonal polyharmonic B-spline $\phi_{2\gamma}^\perp(\boldsymbol{x})$. In effect, starting from any localization operator $\mathrm{V_Q}$ and its corresponding B-spline $\phi_{2\gamma}(\boldsymbol{x})$, one can define the orthogonal localization operator $\mathrm{V_Q^\perp}$ as

$$\hat{\mathrm{V}}_\mathrm{Q}^\perp(\boldsymbol{\omega}) = \frac{\hat{\mathrm{V}}_\mathrm{Q}(\boldsymbol{\omega})}{\sqrt{\left|\hat{\mathrm{A}}_\mathrm{Q}(\boldsymbol{\omega})\right|}}$$

where we have introduced the autocorrelation filter

$$\hat{\mathrm{A}}_\mathrm{Q}(\boldsymbol{\omega}) \stackrel{\mathrm{def}}{=} \hat{\mathrm{A}}_\mathrm{Q}\{\phi_{2\gamma}\}(\boldsymbol{\omega}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} \left|\hat{\phi}_{2\gamma}(\boldsymbol{\omega} + 2\pi\mathrm{Q}^{-\mathsf{T}}\boldsymbol{k})\right|^2 \tag{18}$$

defined as the lattice Fourier transform of $a[\boldsymbol{k}] \stackrel{\mathrm{def}}{=} \langle\phi_{2\gamma}(\cdot - \mathrm{Q}\boldsymbol{k}), \phi_{2\gamma}(\cdot)\rangle$. Division by the square root of $\hat{\mathrm{A}}_\mathrm{Q}\{\phi_{2\gamma}\}(\boldsymbol{\omega})$ guarantees that $\langle\phi_{2\gamma}^\perp(\cdot - \mathrm{Q}\boldsymbol{k}), \phi_{2\gamma}^\perp(\cdot)\rangle = \delta_{\boldsymbol{k}}$. The above orthogonalization depends on the positivity and boundedness of $\hat{\mathrm{A}}_\mathrm{Q}$. The demonstration of these properties is included in the proof of Theorem 2.

C. Polyharmonic Multiresolution Analysis

The following theorem allows us to form a multiresolution analysis based on polyharmonic B-splines (in their different flavors).

Theorem 2: The polyharmonic B-splines defined in (16) have the following properties.

MRA–1. They form a partition of unity.

MRA–2. They fulfil a two-scale relation of the following form:

$$\phi_{2\gamma}(\mathsf{D}_{\mathsf{Q}}^{-1}\boldsymbol{x}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} h[\boldsymbol{k}]\phi_{2\gamma}(\boldsymbol{x} - \mathsf{Q}\boldsymbol{k}) \tag{19}$$

with $h \in \ell_1(\mathbb{Z}^d)$.

MRA–3. They generate a Riesz basis for their ℓ_2 span.

Proofs are given in Appendix I.

Properties MRA–1–3 are those necessary to form a Mallat-type multiresolution analysis [15], [16], [47]. The basic spline approximation subspace is defined as

$$\mathcal{V}_{2\gamma,0} \stackrel{\text{def}}{=} \left\{ \sum_{\boldsymbol{k}\in\mathbb{Z}^d} c[\boldsymbol{k}]\phi_{2\gamma}(\cdot - \mathsf{Q}\boldsymbol{k}) \;\middle|\; c \in \ell_2(\mathbb{Z}^d) \right\}.$$

More generally, the nth level multiresolution spline space is

$$\mathcal{V}_{2\gamma,n} \stackrel{\text{def}}{=} \left\{ f(\mathsf{D}_{\mathsf{Q}}^{-n}\cdot) \;\middle|\; f(\cdot) \in \mathcal{V}_{2\gamma,0} \right\}.$$

Note that because $\hat{\mathrm{T}}_{\mathsf{Q}}^{\gamma}(\boldsymbol{\omega})$ is bounded away from zero, the definition of the above spaces is independent of its particular choice. As a consequence of Theorem 2, these spaces are nested

$$\{0\} \subset \cdots \subset \mathcal{V}_{2\gamma,1} \subset \mathcal{V}_{2\gamma,0} \subset \mathcal{V}_{2\gamma,-1} \subset \cdots \subset \mathrm{L}_2$$

and the closure of their union is L_2.

The next result concerns the fractional derivatives and integrals of polyharmonic splines, which are polyharmonic splines in their own right, but of a different order (see Appendix I for the proof).

Proposition 2:

1. The γ_0th fractional Laplacian of a polyharmonic spline of order 2γ belonging to $\mathcal{V}_{2\gamma,0}$, with $\gamma > \gamma_0$, is a lower order spline in $\mathcal{V}_{2(\gamma-\gamma_0),0}$.
2. If $\Delta^{\gamma_0}s(\boldsymbol{x})$ is a polyharmonic spline of order 2γ, then $s(\boldsymbol{x})$ is a polyharmonic spline of order $2\gamma + 2\gamma_0$.

Polyharmonic Wavelets: Polyharmonic wavelets can be defined as basis functions that span the orthogonal complements in the series of nested approximation spaces. For a given multiresolution hierarchy, there will in general be $|\mathsf{D}| - 1$ distinct mother-wavelets $\psi_{2\gamma}^{i}$, $1 \le i < |\mathsf{D}|$ (we shall subsequently drop the index i as all arguments apply equally to all wavelets).

The semi-orthogonality condition imposed on the wavelet spaces forces the wavelets to have a behavior similar to the operator Δ^{γ} at low frequencies. This quality is encapsulated in the next theorem (a proof is given in Appendix I).

Theorem 3: A semi-orthogonal polyharmonic wavelet of order 2γ can be written as

$$\psi_{2\gamma}(\boldsymbol{x}) = \Delta^{\gamma}\eta(\boldsymbol{x})$$

where $\eta(\boldsymbol{x})$ (the *smoothing kernel*) is a polyharmonic spline of order 4γ that belongs to the Sobolev space $\mathcal{H}^s$ for any $s < 4\gamma - d/2$.

A special case of the general multiresolution construction studied in this section can be found in a previous paper [37], where an explicit construction scheme for the 2-D quincunx lattice (requiring the design of only one mother wavelet) was provided.

V. Characterization of Fractional Brownian Random Fields

A random field is said to be *self-similar* when applying a similarity transform to its domain does not change its stochastic behavior (apart from a possible renormalization factor). For a review of self-similar random fractals we refer the reader to Benassi and Istas [32]. Gaussian self-similar processes were also studied by Dobrushin in his 1979 paper [48].

Fractional Brownian motions form a subset of (continuous) self-similar fields distinguished by their Gaussian statistics and stationary increments [3]. Stochastic self-similarity and stationary increments in particular force the fields to have homogeneous (self-similar) variance functions. Given that fBms are Gaussian and, hence, are fully defined by their second-order statistics, one traditional way of characterizing them is by specifying their variogram, which, for a normalized fBm of Hurst exponent H, has the following form [49, ch. 18]:

$$\mathbf{E}\{|\mathfrak{B}_H(\boldsymbol{x}) - \mathfrak{B}_H(\boldsymbol{x}')|^2\} = 2\|\boldsymbol{x} - \boldsymbol{x}'\|^{2H}.$$

$\mathfrak{B}_H$ is additionally postulated to have zero mean and to be zero at $\boldsymbol{x} = \boldsymbol{0}$ almost surely. One remarks that the derived variance function is indeed homogeneous

$$\mathbf{E}\{|\mathfrak{B}_H(\boldsymbol{x})|^2\} = 2\|\boldsymbol{x}\|^{2H}.$$

Some of the other definitions of fBm fields are in terms of integrals of white noise [50] and by their spectral harmonizable representation [31], [51]. (The latter formulation is closely related to what we present in the sequel. See Remark 4).

An important approach to characterization often used in the analysis and synthesis of *stationary* random processes relies on the notion of *whitening*. In this formulation, an operator is sought after which *whitens* the process in question, i.e., maps it to white noise. Next, a suitable inverse operator needs to be identified, which can then be applied to white noise in order to recreate instances of the desired random process. While standard in the study of stationary processes, this scheme can be extended to certain nonstationary cases, and in particular to the definition of fBm, by adopting a distribution theoretic formalism. This will be demonstrated in this section.

In effect, in the sequel we will show that fractional Laplacians introduced previously whiten multivariate fBm fields of corresponding H exponent (as also discussed by Benassi *et al.* [31]); that is

$$\Delta^{H/2+d/4}\mathfrak{B}_H = \epsilon_H\mathfrak{W}$$

where $\mathfrak{W}$ is normalized white Gaussian noise and ϵ_H is a constant. We also show that an fBm field may be obtained by applying the *right inverse* (cf. Section III-B) to white Gaussian noise, which is to say that

$$\mathfrak{B}_H = \epsilon_H\Delta^{-H/2-d/4}\mathfrak{W}.$$

In addition to being conceptually interesting, the above characterization of multivariate fractional Brownian motion leads to

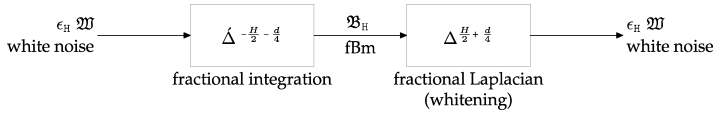


Fig. 1. Innovation model for multivariate fractional Brownian motion.

a natural generalization of the definition to values of H outside the (0, 1) range.

Furthermore, fractal properties of the process find their correspondent in the operator: the scale-invariance property imposed on the operator induces the statistical self-similarity of the process, while rotation-invariance entails its statistical isotropy.

These results all follow from a multivariate generalization of Theorem 1 of Blu and Unser [30], which provides a spectral characterization of fBm through its characteristic functional (cf. Section II-C).

Theorem 4: Let $0 < H < 1$. An fBm field with Hurst parameter H and variogram $2\|\boldsymbol{x}-\boldsymbol{x}'\|^{2H}$ has the following characteristic functional:

$$\mathbf{Z}_{\mathfrak{B}_H}(u) = \exp\left(-\frac{\epsilon_H^2}{2(2\pi)^d}\int \mathrm{d}\boldsymbol{\omega}\ \frac{|\hat{u}(\boldsymbol{\omega})-\hat{u}(\mathbf{0})|^2}{\|\boldsymbol{\omega}\|^{2H+d}}\right) \quad (20)$$

where

$$\epsilon_H^2 = -2^{2H+d}\pi^{d/2}\frac{\Gamma(H+\frac{d}{2})}{\Gamma(-H)}. \quad (21)$$

Proof: A complete proof can be found in Appendix II. The main step of the demonstration consists in showing that (20) defines a Gaussian process whose correlation function $c_{\mathfrak{B}_H}(\boldsymbol{x},\boldsymbol{x}')$ is that of an isotropic fractional Brownian motion with Hurst parameter H, that is, the function

$$c_{\mathfrak{B}_H}(\boldsymbol{x},\boldsymbol{x}') = \|\boldsymbol{x}\|^{2H} + \|\boldsymbol{x}'\|^{2H} - \|\boldsymbol{x}-\boldsymbol{x}'\|^{2H}.$$

∎

We recall the characteristic functional of the unit random field $\mathfrak{W}$ (a.k.a. white Gaussian noise)

$$\mathbf{Z}_{\mathfrak{W}}(u) = \exp\left(-\frac{1}{2}\int \mathrm{d}\boldsymbol{x}\ |u(\boldsymbol{x})|^2\right)$$
$$= \exp\left(-\frac{1}{2}(2\pi)^{-d}\int \mathrm{d}\boldsymbol{\omega}\ |\hat{u}(\boldsymbol{\omega})|^2\right).$$

From comparing this with (20) and by applying a duality argument, we can deduce that

$$\mathbf{Z}_{\mathfrak{B}_H}(u) = \mathbf{Z}_{\mathfrak{W}}(\epsilon_H \acute{\Delta}^{-\gamma_0} u) = \mathbf{Z}_{\epsilon_H \acute{\Delta}^{-\gamma_0}\mathfrak{W}}(u)$$

with $\gamma_0 = H/2 + d/4$. This means that the random field obtained by applying the right inverse $\acute{\Delta}^{-\gamma_0}$ to the unit (generalized) random field $\mathfrak{W}$ is a multivariate fBm with Hurst parameter H, i.e.,

$$\mathfrak{B}_H = \epsilon_H \acute{\Delta}^{-H/2-d/4}\mathfrak{W}. \quad (22)$$

Equation (22) is an alternative characterization of fractional Brownian motion, and can be used to extend the definition to noninteger $H > 1$. The covariance function of these extensions can be obtained with the aid of Lemma 1 of Appendix II.

It also follows that fractional Brownian motion is *whitened* by the fractional Laplacian operator

$$\Delta^{H/2+d/4}\mathfrak{B}_H = \epsilon_H\mathfrak{W}$$

a fact that leads to the innovation model depicted in Fig. 1.

Remark 4: For $0 < H < 1$, a related characterization of real fractional Brownian fields is by their *harmonizable* representation as the stochastic integral

$$\int_{\mathbb{R}^d} \frac{\mathrm{e}^{\mathrm{j}\boldsymbol{x}^T\boldsymbol{\omega}}-1}{\|\boldsymbol{\omega}\|^{H+d/2}}\widehat{W}(\mathrm{d}\omega)$$

where $\widehat{W}$ is a (Hermitian symmetric) complex random measure corresponding to the Fourier transform of real-valued white Gaussian noise (see Samorodnitsky and Taqqu [51] for an in-depth discussion of the single-parameter case). The integrand $(\mathrm{e}^{\mathrm{j}\boldsymbol{x}^T\boldsymbol{\omega}}-1)/\|\boldsymbol{\omega}\|^{H+d/2}$ is comparable to the spectral representation of the right inverse in (10), which reduces to the said integrand for $0 < H < 1$. The treatment of Benassi *et al.* [31] is also of direct pertinency, and includes similar extensions.

VI. Polyharmonic Wavelet Analysis of Multivariate fBm

Considering the inherent link between polyharmonic splines and fBms that has been emphasized throughout this article, it should not come as a surprise that a wavelet analysis of multivariate fBm would have interesting properties. We study some of these in the first part of this section. Next, we complement and verify our derivations through some experimental results.

A. Probability Distribution of Wavelet Coefficients

Proposition 3: The polyharmonic spline wavelet transform of order $2\gamma \geq 2\gamma_0$, with $\gamma_0 \stackrel{\text{def}}{=} H/2+d/4$, maps the nonstationary process $\mathfrak{B}_H$ into a series of stationary (discrete) Gaussian processes.

Proof: We can rely on Theorem 3 and the innovation model to see that, e.g., the wavelet coefficients at level $n = 0$ are stationary Gaussian processes obtained by filtering white noise

$$w_0[\boldsymbol{k}] = \langle \mathfrak{B}_H, \psi_{2\gamma}(\cdot - \mathrm{Q}\boldsymbol{k})\rangle = \langle \Delta^{\gamma_0}\mathfrak{B}_H, \Delta^{\gamma-\gamma_0}\eta(\cdot-\mathrm{Q}\boldsymbol{k})\rangle$$
$$= \langle \epsilon_H\mathfrak{W}, \Delta^{\gamma-\gamma_0}\eta(\cdot-\mathrm{Q}\boldsymbol{k})\rangle.$$

(Note that even though the polyharmonic spline $\Delta^{\gamma-\gamma_0}\eta(\cdot-\mathrm{Q}\boldsymbol{k})$ is not a Schwartz test function, its inner-product with the white

noise process is nonetheless well-defined as it is continuous and belongs to $\mathcal{H}^s$ for some $s > 0$; cf. Theorem 3). The demonstration for an arbitrary level n is similar, except that a scale-dependent normalization factor also appears. ∎

What this property means is that the $w_0[\boldsymbol{k}]$s correspond to the lattice samples of a *stationary* process with power spectrum $\epsilon_H^2 \|\boldsymbol{\omega}\|^{2\gamma-2\gamma_0} |\hat{\eta}(\boldsymbol{\omega})|^2$ (which is well-defined in the L_2 sense since $\eta \in \mathcal{H}^s$ for all $s < 4\gamma - d/2$). This relation is essentially scale-invariant up to a proportionality factor.

Proposition 4: The variance of the polyharmonic wavelet coefficients depends exponentially on the Hurst exponent and the scale n as per

$$\mathbf{E}\{w_n^2[\boldsymbol{k}]\} = |\mathsf{D}|^{(2H+d)n/d} \mathbf{E}\{w_0^2[\boldsymbol{k}]\}.$$

Proof: This property can be shown using the correlation form $\langle\!\langle \cdot, \cdot \rangle\!\rangle_{\mathfrak{B}_H}$. One has [cf. (28)]

$$\begin{aligned}
&\mathbf{E}\{w_n^2[\boldsymbol{k}]\} \\
&= \langle\!\langle |\mathsf{D}|^{-n/2} \psi_{2^\gamma}(\mathsf{D}_\mathsf{Q}^{-n}\boldsymbol{x} - \mathsf{Q}\boldsymbol{k}), |\mathsf{D}|^{-n/2} \psi_{2^\gamma}(\mathsf{D}_\mathsf{Q}^{-n}\boldsymbol{x} - \mathsf{Q}\boldsymbol{k}) \rangle\!\rangle_{\mathfrak{B}_H} \\
&= \frac{\epsilon_H^2}{(2\pi)^d} \int \mathrm{d}\boldsymbol{\omega} \, \frac{|\hat{\psi}_{2^\gamma}(\boldsymbol{\omega})|^2}{\|\mathsf{Q}^{-\mathsf{T}} \mathsf{D}^{-n\mathsf{T}} \mathsf{Q}^\mathsf{T} \boldsymbol{\omega}\|^{2H+d}} \\
&= |\mathsf{D}|^{(n/d)(2H+d)} \frac{\epsilon_H^2}{(2\pi)^d} \int \mathrm{d}\boldsymbol{\omega} \, \frac{|\hat{\psi}_{2^\gamma}(\boldsymbol{\omega})|^2}{\|\boldsymbol{\omega}\|^{2H+d}} \\
&= |\mathsf{D}|^{(n/d)(2H+d)} \langle\!\langle \psi_{2^\gamma}(\boldsymbol{x}), \psi_{2^\gamma}(\boldsymbol{x}) \rangle\!\rangle_{\mathfrak{B}_H} \\
&= |\mathsf{D}|^{(n/d)(2H+d)} \mathbf{E}\{w_0^2[\boldsymbol{k}]\}.
\end{aligned}$$ ∎

More generally, we have the following result.

Proposition 5: The covariance of intrascale wavelet coefficients is given by the relation

$$\begin{aligned}
&\frac{\mathbf{E}\{w_n[\boldsymbol{k}] w_n[\boldsymbol{k}']\}}{|\mathsf{D}|^{(n/d)(2H+d)}} \\
&= \frac{\epsilon_H^2}{2(2\pi)^d} \int \mathrm{d}\boldsymbol{\omega} \, \frac{|\hat{\psi}_{2^\gamma}(\boldsymbol{\omega})|^2}{\|\boldsymbol{\omega}\|^{2H+d}} \left(\left| 1 + \mathrm{e}^{\mathrm{j}(\boldsymbol{k}-\boldsymbol{k}')^\mathsf{T} \mathsf{D}^{n\mathsf{T}} \mathsf{Q}^\mathsf{T} \boldsymbol{\omega}} \right|^2 - 2 \right) \\
&= \mathbf{E}\{w_0[\mathsf{D}^n \boldsymbol{k}] w_0[\mathsf{D}^n \boldsymbol{k}']\}.
\end{aligned}$$

Proof: At scale 0 we have

$$\begin{aligned}
\mathbf{E}\{w_0[\boldsymbol{k}] w_0[\boldsymbol{k}']\} = \frac{1}{2} \Big[&\langle\!\langle \psi_{2^\gamma,\boldsymbol{k}} + \psi_{2^\gamma,\boldsymbol{k}'}, \psi_{2^\gamma,\boldsymbol{k}} + \psi_{2^\gamma,\boldsymbol{k}'} \rangle\!\rangle_{\mathfrak{B}_H} \\
&- \langle\!\langle \psi_{2^\gamma,\boldsymbol{k}}, \psi_{2^\gamma,\boldsymbol{k}} \rangle\!\rangle_{\mathfrak{B}_H} - \langle\!\langle \psi_{2^\gamma,\boldsymbol{k}'}, \psi_{2^\gamma,\boldsymbol{k}'} \rangle\!\rangle_{\mathfrak{B}_H} \Big].
\end{aligned}$$

The proposition is then proved using (28) and with a change of variables as in the previous proof. ∎

Remark 5: The above result can be compared with those obtained by Meyer *et al.* [23] in the 1-D setting. The wavelets proposed by Meyer *et al.* depend on the Hurst parameter H that is matched to the Hurst exponent of the 1-D fBm process in consideration (which should be known *a priori*). Independence of the wavelet coefficients (true whitening) is a consequence of this perfect match. This in fact corresponds to the wavelets being orthogonal in terms of the positive-definite form $\langle\!\langle \cdot, \cdot \rangle\!\rangle_{\mathfrak{B}_H}$. Since this design depends on the Hurst exponent being known, in the problem of estimating H a parameter higher than the true unknown value must be used, in which case the wavelet coefficients will again be correlated. Also note that the results provided in the present paper are general and concern *any* family of semi-orthogonal polyharmonic wavelets. In the actual implementation of wavelets for a given lattice, there is some room for incorporating certain desired behaviors in the design of the wavelet filter, which will in turn affect the smoothing function of Theorem 3.

As a demonstration of potential, the above results (Propositions 3 and 4 in particular) allow us to extend 1-D wavelet estimators of the Hurst exponent reported in the literature [18], [20], [27]–[29] to the multidimensional setting. In its simplest form, estimation can be based on the identity

$$\log_{\sqrt[d]{|\mathsf{D}|}} \left(\mathbf{E}\left\{ w_n^2[\boldsymbol{k}] \right\} \right) = (2H + d)n + C \tag{23}$$

where $C = \log_{\sqrt[d]{|\mathsf{D}|}} \left(\mathbf{E}\left\{ w_0^2[\boldsymbol{k}] \right\} \right)$ is a computable constant that depends on the choice of the wavelet (Proposition 4). This means that a linear regression of the estimates of the variance in each sub-band in the log scale provides an estimate of H.

An improved estimate may be obtained using a maximum-likelihood (ML) formulation. This is essentially a multidimensional adaptation of the ML-estimator of Wornell [27], [29].[4] The estimate is defined as the minimizer of a negative log-likelihood approximate (leaving out the constant term)

$$\ell(\boldsymbol{w}|\boldsymbol{\theta}) = \frac{1}{2} \sum_{n \in \mathcal{N}} N_n \log \sigma_n^2(\boldsymbol{\theta}) + \frac{E_n}{\sigma_n^2(\boldsymbol{\theta})}. \tag{24}$$

In the above formula $\boldsymbol{\theta} \stackrel{\text{def}}{=} (H, C')$—with C' a normalization factor—is the set of parameters to estimate; $\mathcal{N}$ is the set of levels used for estimation; N_n denotes the number of coefficients at level n

$$\sigma_n^2(\boldsymbol{\theta}) \stackrel{\text{def}}{=} \mathbf{E}\{w_n^2[\boldsymbol{k}]\} = C' |\mathsf{D}|^{(2H+d)n/d}$$

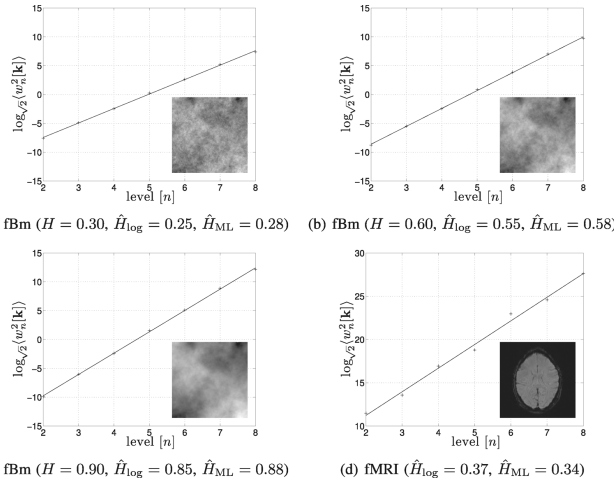
is the theoretical variance of level n wavelet coefficients (cf. Proposition 4); and, finally, E_n is the *observed* wavelet energy (i.e., the sum of coefficients squared) at level n. In the implementation we have used the previous regression estimate as an initial guess and applied Newton's method to the derivative of ℓ. This provides a fast (essentially real-time) way of producing an improved estimate of H.

MATLAB implementations of the above estimators are available via our website (http://bigwww.epfl.ch/demo/fractaldimension/).

B. Experimental Results

The estimation procedure outlined previously was applied to instances of (periodic, due to discretization) 2-D fBm, generated via Fourier domain filtering as per Section III-B [cf. (22) and Remark 4]. The wavelets used for analysis were isotropic polyharmonic wavelets of Van De Ville *et al.* [37], which have a fast FFT-based implementation. The order of the wavelets was chosen to exceed $H + d/2$ in order to satisfy the requirements of Proposition 3. We used a quincunx subsampling scheme, which

[4]Note that, as is the case for the cited estimators, the ML formulation is approximate where the wavelet is not specifically designed to exactly match the process, as the correlation between wavelet coefficients is not taken into consideration. We have provided formulae for the covariances, which could in principle be used to improve the estimate. This, however, would substantially complicate the estimator.



(a) fBm ($H = 0.30$, $\hat{H}_{\log} = 0.25$, $\hat{H}_{\mathrm{ML}} = 0.28$) (b) fBm ($H = 0.60$, $\hat{H}_{\log} = 0.55$, $\hat{H}_{\mathrm{ML}} = 0.58$)

(c) fBm ($H = 0.90$, $\hat{H}_{\log} = 0.85$, $\hat{H}_{\mathrm{ML}} = 0.88$) (d) fMRI ($\hat{H}_{\log} = 0.37$, $\hat{H}_{\mathrm{ML}} = 0.34$)

Fig. 2. (a)–(c) Regression plots for the estimation of Hurst exponent of discretized bivariate fBm for various values of the Hurst parameter, all generated from the same instance of pseudo-random noise; (d) regression plot for an fMRI image (original images are given as insets).

TABLE I
WAVELET-BASED ESTIMATION OF H (100 REALIZATIONS)

true value	log-regression estimate		ML estimate	
	mean	stdev	mean	stdev
0.3	0.290	0.007	0.293	0.004
0.6	0.590	0.008	0.593	0.004
0.9	0.890	0.008	0.893	0.005

offers a more gradual scale progression, thus furnishing more regression points for the estimation. Another advantage is that the quincunx design involves only a single mother-wavelet.

Hurst parameter estimation was performed on 100 instances of 512×512 fBm images for three different values of H (0.3, 0.6, and 0.9). Decomposition levels 2 to 8 were used for estimation. Examples of fBm images and corresponding regression curves can be seen in Fig. 2. The average and standard deviation of the estimated values, obtained by regression and ML estimation are given in Table I. In experiments we noticed very good fits and small standard deviations, which underline the robustness of the process.

Results of the same analysis applied to a single axial slice of a functional magnetic resonance image (fMRI) of the brain are also shown in Fig. 2. Boundary and background wavelet coefficients were discarded for the analysis in order to avoid boundary effects. The corresponding fractal dimension according to the improved estimate is $d + 1 - H = 2.66$.

It has been suggested that anatomical growth processes lead to fractal-like structures. In the case of the brain, Bullmore *et al.* [52] have argued that the boundary between the white matter and the cerebral cortex has a fractal-like shape. Additionally, based on recently made possible 3-D high-resolution imaging of the vasculature [53], the branching of the tree structure of the arteries appears to constitute a fractal organization in space. As fMR imaging of brain tissue indirectly measures the flow of oxygenated blood, these arguments can in a way account for the fractal behavior evidenced in Fig. 2(d).

VII. CONCLUSION

Our approach in this paper was based on the observation that certain families of splines and random processes can be characterized as deterministic vs stochastic solutions of the same fractional partial differential equation.

Motivated by the works of Duchon [43], Arigovindan [44], and Kybic *et al.* [42] on invariances, in this paper we focused on a particular class of such equations that is singled out by imposing certain fundamental invariance properties on the operator involved. This pointed us to a family of fractional differential operators that are invariant to the translation, rotation, and scaling of the coordinate system. We substantiated the following points.

- These operators (which turn out to be fractional iterations of the Laplacian) lead naturally to the definition of polyharmonic B-splines and multiresolution spline spaces over a large family of multidimensional lattices.
- The same operators whiten multivariate fractional Brownian motion, and can thus be used to rigorously characterize this important family of random fields.
- The relation between deterministic and stochastic formulations provides a natural framework for the analysis of fBm. In particular, a polyharmonic multiresolution analysis of fractional Brownian motion has interesting properties that

can be deduced from the parallelism between the two formulations. As an example, we showed an application of this observtion in the estimation of the Hurst parameter associated with fBm processes.

Our results relate, generalize, and formalize previous results of multiple authors, including those of Rabut *et al.* [35], [36] and Van De Ville *et al.* [37] (on polyharmonic splines and wavelets), Blu and Unser [30], [34] (on the distributional characterization of 1-D fBm), and Flandrin, Wornell, and Veitch and Abry [20], [27], [28] (on the wavelet analysis of 1-D fBm). In addition, given the generality of the approach, it opens an interesting avenue of research for the future investigation of any of these subjects.

Appendix I
Proofs of Theorems 2 and 3 and of Proposition 2

Proof of Theorem 2:

Proof of MRA–1: By (16), the zeros of $\hat{\phi}_{2\gamma}(\boldsymbol{\omega})$ are the same as those of $\hat{\Delta}_{\mathsf{Q}}^{\gamma}(\boldsymbol{\omega})$, with the exception of the zero at $\boldsymbol{\omega} = \mathbf{0}$ which disappears (see the proof of Proposition 1). From (15) we can see that $\hat{\Delta}_{\mathsf{Q}}^{\gamma}(\boldsymbol{\omega})$ is zero iff

$$\boldsymbol{y}_i^{\mathsf{T}} \mathsf{Q}^{\mathsf{T}} \frac{\boldsymbol{\omega}}{2\pi} \in \mathbb{Z} \quad \text{for all } i.$$

Since the vectors $\mathsf{Q}\boldsymbol{y}_i$ generate $\pounds_0$, by the definition of the dual lattice, the above condition is equivalent to

$$\frac{\boldsymbol{\omega}}{2\pi} \in \pounds_0^*.$$

Removing the zero at $\boldsymbol{w} = \mathbf{0}$ produces $2\pi\pounds_0^* \backslash \{\mathbf{0}\}$ as the set of zeros of $\hat{\phi}_{2\gamma}(\boldsymbol{\omega})$.

Property MRA–1 is then a consequence of the Poisson summation formula [cf. (4); also of direct relevance is Kolountzakis [54, Eqn (5)]]. ∎

Proof of MRA–2: Property MRA–2 can be verified by writing the Fourier expression of the refinement filter h as

$$\hat{\mathrm{H}}_{\mathsf{Q}}(\boldsymbol{\omega}) = |\mathsf{D}| \frac{\hat{\phi}_{2\gamma}(\mathsf{D}_{\mathsf{Q}}^{*}\boldsymbol{\omega})}{\hat{\phi}_{2\gamma}(\boldsymbol{\omega})} = |\mathsf{D}| \frac{\frac{\hat{\mathrm{V}}_{\mathsf{Q}}(\mathsf{D}_{\mathsf{Q}}^{*}\boldsymbol{\omega})}{||\mathsf{D}_{\mathsf{Q}}^{*}\boldsymbol{\omega}||^{2\gamma}}}{\frac{\hat{\mathrm{V}}_{\mathsf{Q}}(\boldsymbol{\omega})}{||\boldsymbol{\omega}||^{2\gamma}}} = |\mathsf{D}|^{1-2\gamma/d} \frac{\hat{\mathrm{V}}_{\mathsf{Q}}(\mathsf{D}_{\mathsf{Q}}^{*}\boldsymbol{\omega})}{\hat{\mathrm{V}}_{\mathsf{Q}}(\boldsymbol{\omega})}.$$

(The last step results from $\mathsf{D}_{\mathsf{Q}}^{*}$ being, per definition, a similarity transform matrix; cf. LAT–1). We observe that (i) the numerator and denominator of the last expression are, respectively, $2\pi\pounds_{-1}^*$- and $2\pi\pounds_0^*$-periodic; that (ii) the zeros of the numerator and the denominator happen respectively over the sets $2\pi\pounds_{-1}^*$ and $2\pi\pounds_0^*$ and are all of order 2γ; and finally, that (iii) both the numerator and the denominator are bounded.

We know from (3) that $2\pi\pounds_0^* \subset 2\pi\pounds_{-1}^*$. Therefore, first, from (i) it follows that $\hat{\mathrm{H}}_{\mathsf{Q}}(\boldsymbol{\omega})$ is $2\pi\pounds_0^*$-periodic. Secondly, from (ii) and (iii) one concludes that $\hat{\mathrm{H}}_{\mathsf{Q}}(\boldsymbol{\omega})$ is bounded, with its set of zeros being

$$\left\{\boldsymbol{\omega} \;\middle|\; \hat{\mathrm{H}}_{\mathsf{Q}}(\boldsymbol{\omega}) = 0\right\} = 2\pi\left(\pounds_{-1}^* \backslash \pounds_0^*\right) = 2\pi \bigcup_{1 \le i < |\mathsf{D}|} \left(\pounds_0^* + \mathsf{Q}^{-\mathsf{T}}\mathsf{D}^{-\mathsf{T}}\boldsymbol{\zeta}_i^*\right). \quad (25)$$

These observations establish that $\hat{\mathrm{H}}_{\mathsf{Q}}(\boldsymbol{\omega})$ is the lattice Fourier transform of a sequence $h \in \ell_1$. The two-scale relation therefore holds. ∎

Proof of MRA–3: Proving the existence of lower and upper Riesz bounds is equivalent to showing that the Fourier transform of the autocorrelation filter (18) is bounded away from zero.

Since $\hat{\mathrm{A}}_{\mathsf{Q}}(\boldsymbol{\omega})$ is $2\pi\pounds_0^*$-periodic, we can restrict our attention to the unit cell corresponding to the Voronoi region of $\mathbf{0}$ with respect to $2\pi\pounds_0^*$. Within this region, we rewrite (18), replacing $\hat{\phi}_{2\gamma}$ from (16) and noting the periodicity and boundedness of $\hat{\mathrm{V}}_{\mathsf{Q}}(\boldsymbol{\omega})$ to obtain

$$\begin{aligned}\hat{\mathrm{A}}_{\mathsf{Q}}(\boldsymbol{\omega}) &= \sum_{\boldsymbol{k}\in\mathbb{Z}^d} \frac{\hat{\mathrm{V}}_{\mathsf{Q},4\gamma}(\boldsymbol{\omega})}{|\boldsymbol{\omega} + 2\pi\mathsf{Q}^{-\mathsf{T}}\boldsymbol{k}|^{4\gamma}} \\ &= \frac{\hat{\mathrm{V}}_{\mathsf{Q},4\gamma}(\boldsymbol{\omega})}{|\boldsymbol{\omega}|^{4\gamma}} + \hat{\mathrm{V}}_{\mathsf{Q},4\gamma}(\boldsymbol{\omega}) \sum_{\boldsymbol{k}\in\mathbb{Z}^d\backslash\{\mathbf{0}\}} |\boldsymbol{\omega} + 2\pi\mathsf{Q}^{-\mathsf{T}}\boldsymbol{k}|^{-4\gamma}.\end{aligned}$$

The existence of a positive lower bound is then evident as $\hat{\mathrm{A}}_{\mathsf{Q}}(\boldsymbol{\omega})$ is bounded from below by $\hat{\phi}_{4\gamma}(\boldsymbol{\omega}) = \hat{\mathrm{V}}_{\mathsf{Q},4\gamma}(\boldsymbol{\omega})/|\boldsymbol{\omega}|^{4\gamma}$, which is strictly positive in the noted region.

Also, since we assumed $\gamma > d/4$, the second sum converges for all $\boldsymbol{\omega}$ in the unit cell, and is bounded from above (with both factors being bounded). This, in addition to the boundedness of $\hat{\phi}_{4\gamma}(\boldsymbol{\omega})$, confirms the existence of an upper bound and completes the proof of the Riesz property. ∎

Proof of Proposition 2:

Proof of 1: Any element $f(\boldsymbol{x})$ of $\mathcal{V}_{2\gamma,0}$ can be expressed in the Fourier domain as

$$\hat{\mathrm{C}}_{\mathsf{Q}}(\boldsymbol{\omega}) \frac{\hat{\Delta}_{\mathsf{Q}}^{\gamma}(\boldsymbol{\omega})}{||\boldsymbol{\omega}||^{2\gamma}}$$

where the $2\pi\pounds_0^*$-periodic and locally square integrable function $\hat{\mathrm{C}}_{\mathsf{Q}}(\boldsymbol{\omega})$ is the lattice Fourier transform of a sequence $c \in \ell_2$. By applying Δ^{γ_0} to f we shall have

$$\begin{aligned}\mathcal{F}\{\Delta^{\gamma_0} f\} &= \hat{\mathrm{C}}_{\mathsf{Q}}(\boldsymbol{\omega}) \hat{\Delta}_{\mathsf{Q}}^{\gamma_0}(\boldsymbol{\omega}) \frac{\hat{\Delta}_{\mathsf{Q}}^{(\gamma-\gamma_0)}(\boldsymbol{\omega})}{||\boldsymbol{\omega}||^{2(\gamma-\gamma_0)}} \\ &= \hat{\mathrm{C}}_{\mathsf{Q}}(\boldsymbol{\omega}) \hat{\Delta}_{\mathsf{Q}}^{\gamma_0}(\boldsymbol{\omega}) \hat{\phi}_{2(\gamma-\gamma_0),\mathrm{el}}(\boldsymbol{\omega}).\end{aligned}$$

Since $\hat{\mathrm{C}}_{\mathsf{Q}}(\boldsymbol{\omega}) \hat{\Delta}_{\mathsf{Q}}^{\gamma_0}(\boldsymbol{\omega})$ is also a $2\pi\pounds_0^*$-periodic and locally square integrable function (due to the periodicity and boundedness of the second factor), it corresponds to the Fourier transform of some ℓ_2 sequence c'. $\Delta^{\gamma_0} f$ can, therefore, be written in the form

$$\sum_{\boldsymbol{k}\in\mathbb{Z}^d} c'[\boldsymbol{k}] \phi_{2(\gamma-\gamma_0)}(\cdot - \mathsf{Q}\boldsymbol{k})$$

whereby $\Delta^{\gamma_0} f \in \mathcal{V}_{(\gamma-\gamma_0),0}$. ∎

Proof of 2: From the assumption, by the definition of polyharmonic splines [see (11)], we have

$$\Delta^{\gamma}\Delta^{\gamma_0} s(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} c[\boldsymbol{k}] \delta(\boldsymbol{x} - \mathsf{Q}\boldsymbol{k}).$$

Using (7), we can write

$$\Delta^{\gamma+\gamma_0} s(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} c[\boldsymbol{k}] \delta(\boldsymbol{x} - \mathsf{Q}\boldsymbol{k})$$

which, per definition, establishes $s(\boldsymbol{x})$ as a polyharmonic spline of order $2\gamma + 2\gamma_0$. ∎

Proof of Theorem 3:

Proof: The semi-orthogonality condition is equivalent to stating that

$$\langle \phi_{2\gamma}(\mathsf{D}_{\mathsf{Q}}^{-1}\boldsymbol{x}), \psi_{2\gamma}(\mathsf{D}_{\mathsf{Q}}^{-1}\boldsymbol{x} - \mathsf{Q}\boldsymbol{k})\rangle \equiv 0. \tag{26}$$

We replace the B-spline $\phi_{2\gamma}$ and the wavelet $\psi_{2\gamma}$ in the above equality by their higher resolution B-spline expansions, given in (19) for $\phi_{2\gamma}$ and below for $\psi_{2\gamma}$

$$\psi_{2\gamma}(\mathsf{D}_{\mathsf{Q}}^{-1}\boldsymbol{x}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} g[\boldsymbol{k}]\phi_{2\gamma}(\boldsymbol{x} - \mathsf{Q}\boldsymbol{k})$$

where $g \in \ell_1$ is the stable wavelet filter. The autocorrelation filter $a[\boldsymbol{k}] \stackrel{\text{def}}{=} \langle \phi_{2\gamma}(\cdot - \mathsf{Q}\boldsymbol{k}), \phi_{2\gamma}(\cdot)\rangle$ appears in the resulting equation. Using its symmetry, we can restate (26) as follows:

$$(\check{h} * a * g)[\mathsf{D}\boldsymbol{k}] \equiv 0$$

with $\check{h}[\boldsymbol{k}] \stackrel{\text{def}}{=} h[-\boldsymbol{k}]$.

Let us define $b \stackrel{\text{def}}{=} \check{h} * a * g$. The above relation then finds the following Fourier domain expression (cf. Viscito and Allebach [55])

$$\hat{\mathrm{B}}_{\mathsf{Q}}(\boldsymbol{\omega}) + \sum_{1\le i<|\mathsf{D}|} \hat{\mathrm{B}}_{\mathsf{Q}}(\boldsymbol{\omega} - 2\pi\mathsf{Q}^{-\mathsf{T}}\mathsf{D}^{-\mathsf{T}}\boldsymbol{\zeta}_i) = 0.$$

From the definition of b, we have

$$\hat{\mathrm{B}}_{\mathsf{Q}}(\boldsymbol{\omega}) = \overline{\hat{\mathrm{H}}_{\mathsf{Q}}(\boldsymbol{\omega})}\hat{\mathrm{A}}_{\mathsf{Q}}(\boldsymbol{\omega})\hat{\mathrm{G}}_{\mathsf{Q}}(\boldsymbol{\omega}). \tag{27}$$

Therefore

$$\hat{\mathrm{G}}_{\mathsf{Q}}(\boldsymbol{\omega}) = -\frac{\sum_{1\le i<|\mathsf{D}|} \hat{\mathrm{B}}_{\mathsf{Q}}(\boldsymbol{\omega} - 2\pi\mathsf{Q}^{-\mathsf{T}}\mathsf{D}^{-\mathsf{T}}\boldsymbol{\zeta}_i)}{\overline{\hat{\mathrm{H}}_{\mathsf{Q}}(\boldsymbol{\omega})}\hat{\mathrm{A}}_{\mathsf{Q}}(\boldsymbol{\omega})}.$$

We see from (25) and (27) that the numerator has an uncancelled (and isotropic) zero of degree 2γ at the origin. Since $\hat{\mathrm{G}}(\boldsymbol{\omega})$ is by definition bounded (as $g \in \ell_1$), this means that we can extract the symbol $||\boldsymbol{\omega}||^{2\gamma}$ (corresponding to Δ^{γ}) from the Fourier transform of the wavelet filter, and consequently from the Fourier transform of the wavelet itself. In other words, the function

$$\hat{\eta}(\boldsymbol{\omega}) \stackrel{\text{def}}{=} ||\boldsymbol{\omega}||^{-2\gamma}\hat{\psi}_{2\gamma}(\boldsymbol{\omega})$$

will be continuous at $\mathbf{0}$.

We also note that the wavelet, by construction, has the same Sobolev regularity as the B-splines; i.e., its Fourier transform decays like $||\boldsymbol{\omega}||^{-2\gamma}$, leading to a $||\boldsymbol{\omega}||^{-4\gamma}$-like decay for $\hat{\eta}(\boldsymbol{\omega})$. From this we deduce that $\eta(\boldsymbol{x})$ is of the claimed Sobolev regularity. That it is also a polyharmonic spline of order 4γ follows from the second part of Proposition 2. ∎

Appendix II
Proof of Theorem 4

As was mentioned in the introduction, the characteristic functional of a Gaussian field $\mathfrak{X}$ satisfies (see Gel'fand *et al.* [39, ch. III, §2.6])

$$\mathbf{Z}_{\mathfrak{X}}(u) = \exp\left(-\frac{1}{2}\langle\!\langle u, u\rangle\!\rangle_{\mathfrak{X}}\right).$$

Therefore, in our case, we need to show that for $0 < H < 1$

$$\langle\!\langle u, u\rangle\!\rangle_{\mathfrak{B}_H} = \frac{\epsilon_H^2}{(2\pi)^d}\int \mathrm{d}\boldsymbol{\omega}\, \frac{|\hat{u}(\boldsymbol{\omega}) - \hat{u}(\mathbf{0})|^2}{||\boldsymbol{\omega}||^{2H+d}}. \tag{28}$$

This correlation form is related to the (generalized) correlation function $c_{\mathfrak{B}_H}(\boldsymbol{x}, \boldsymbol{x}')$ thus [*ibid.*, ch. III, §2.1]

$$\langle\!\langle u, v\rangle\!\rangle_{\mathfrak{B}_H} = \int \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{x}'\, c_{\mathfrak{B}_H}(\boldsymbol{x}, \boldsymbol{x}')u(\boldsymbol{x})v(\boldsymbol{x}'). \tag{29}$$

The correlation function of a normalized fractional Brownian field with parameter $H, 0 < H < 1$, derived from its variogram, is

$$c_{\mathfrak{B}_H}(\boldsymbol{x}, \boldsymbol{x}') = \left(||\boldsymbol{x}||^{2H} + ||\boldsymbol{x}'||^{2H} - ||\boldsymbol{x} - \boldsymbol{x}'||^{2H}\right). \tag{30}$$

To show (28), we plug (30) into (29), and break the integral at the additions to get (after replacing $\boldsymbol{x}$ in the first, $\boldsymbol{x}'$ in the second, and $\boldsymbol{x}' - \boldsymbol{x}$ in the last integral, all by $\boldsymbol{x}$)

$$\begin{aligned}\langle\!\langle u, u\rangle\!\rangle_{\mathfrak{B}_H} &= \left\langle ||\boldsymbol{x}||^{2H}, \mathcal{F}^{-1}\left\{\overline{\hat{u}(\mathbf{0})}\hat{u}(\boldsymbol{\omega})\right\}\right\rangle \\ &\quad + \left\langle ||\boldsymbol{x}||^{2H}, \mathcal{F}^{-1}\left\{\overline{\hat{u}(\boldsymbol{\omega})}\hat{u}(\mathbf{0})\right\}\right\rangle \\ &\quad - \left\langle ||\boldsymbol{x}||^{2H}, \mathcal{F}^{-1}\left\{\overline{\hat{u}(\boldsymbol{\omega})}\hat{u}(\boldsymbol{\omega})\right\}\right\rangle \\ &= -\left\langle ||\boldsymbol{x}||^{2H}, \mathcal{F}^{-1}\{\hat{v}(\boldsymbol{\omega})\}\right\rangle \end{aligned} \tag{31}$$

where

$$\begin{aligned}\hat{v}(\boldsymbol{\omega}) &\stackrel{\text{def}}{=} \overline{\hat{u}(\mathbf{0})}\hat{u}(\boldsymbol{\omega}) + \overline{\hat{u}(\boldsymbol{\omega})}\hat{u}(\mathbf{0}) - \overline{\hat{u}(\boldsymbol{\omega})}\hat{u}(\boldsymbol{\omega}) \\ &= |\hat{u}(\boldsymbol{\omega}) - \hat{u}(\mathbf{0})|^2 - |\hat{u}(\mathbf{0})|^2\end{aligned}$$

is a linear combination of test functions and is, therefore, a valid test function itself.

The inner product in (31) can be evaluated in the Fourier domain by applying the Parseval equivalence

$$\langle ||\boldsymbol{x}||^{2H}, \mathcal{F}^{-1}\{\hat{v}(\boldsymbol{\omega})\}\rangle = -(2\pi)^{-d}\langle \epsilon_H^2 \mathrm{R}||\boldsymbol{\omega}||^{-2H-d}, \hat{v}(\boldsymbol{\omega})\rangle \tag{32}$$

valid for $2H \neq -d, -d-2, \ldots$ [38, p. 363]. Here $\mathrm{R}||\boldsymbol{\omega}||^{-2H-d}$ is a generalized function (distribution) that corresponds to a particular (canonical) regularization of the function $||\boldsymbol{\omega}||^{-2H-d}$. The canonical regularization is to be conducted according to the recipe given in Gel'fand and Shilov [38, §3.3], as detailed below.

We restate (32) in (hyper)spherical coordinates as

$$\langle\!\langle u, u\rangle\!\rangle_{\mathfrak{B}_H} = (2\pi)^{-d}\epsilon_H^2\Omega_d\langle \mathrm{R}\rho^{-2H-1}, S_{\hat{v}}(\rho)\rangle$$

where $\rho \stackrel{\text{def}}{=} ||\boldsymbol{\omega}||$, $\Omega_d \stackrel{\text{def}}{=}$ the area of the hypersphere in $\mathbb{R}^d$, and $S_{\hat{v}}(\rho)$ denotes the average of $\hat{v}(\boldsymbol{\omega})$ over the hypersphere of radius ρ centered at the origin. Also, $\mathrm{R}\rho^{-2H-1}$ denotes the particular regularization of ρ^{-2H-1} invoked in (33).

$S_{\hat{v}}(\rho)$ is a smooth and even function of ρ with rapid decay, with a Taylor series expansion of the form

$$S_{\hat{v}}(\rho) = \hat{v}(\mathbf{0}) + a_2\rho^2 + a_4\rho^4 + \cdots + a_{2k}\rho^{2k} + \mathrm{o}(\rho^{2k}).$$

For $0 < H < 1$, we have $-3 < -2H - 1 < 0$ and from there, by the definition of the generalized function $\mathrm{R}\rho^{-2H-1}$ (see [38, p. 363])

$$\langle \mathrm{R}\rho^{-2H-1}, S_b(\rho)\rangle = \int_0^\infty \mathrm{d}\rho\, \rho^{-2H-1}\left[S_b(\rho) - \hat{v}(\mathbf{0})\right] \quad (33)$$

(where the right-hand integral should be interpreted as a limit). By expanding $S_b(\rho)$ and returning to Cartesian coordinates, we can now write

$$\begin{aligned}\langle ||\boldsymbol{\omega}||^{-2H-d}, \hat{v}(\boldsymbol{\omega})\rangle &= \int \mathrm{d}\boldsymbol{\omega}\, ||\boldsymbol{\omega}||^{-2H-d}\left[\hat{v}(\boldsymbol{\omega}) - \hat{v}(\mathbf{0})\right] \\ &= \int \mathrm{d}\boldsymbol{\omega}\, ||\boldsymbol{\omega}||^{-2H-d}|\hat{u}(\boldsymbol{\omega}) - \hat{u}(\mathbf{0})|^2\end{aligned}$$

(using the definition of $\hat{v}(\boldsymbol{\omega})$). From combining this with (32), we arrive at the desired result, i.e., (28). ∎

Remark 6: The following lemma allows us to generalize the results given here for $0 < H < 1$ to the case of noninteger $H > 1$. The proof is technical and is not reproduced here.

Lemma 1: Let $v(\boldsymbol{x})$ be a test function and $H > 0$ be noninteger. Then, in the sense of generalized functions of Gel'fand and Shilov

$$\begin{aligned}\langle ||\boldsymbol{x}||^{2H}, v(\boldsymbol{x})\rangle = -\epsilon_H^2 \int \mathrm{d}\boldsymbol{\omega}\, ||\boldsymbol{\omega}||^{-2H-d} \\ \times \left(\hat{v}(\boldsymbol{\omega}) - \Gamma\Big(\frac{d}{2}\Big) \sum_{0 \le 2k \le \lfloor 2H \rfloor} \frac{\Delta^k \hat{v}(\mathbf{0})||\boldsymbol{\omega}||^{2k}}{2^{2k} k! \Gamma\left(k + \frac{d}{2}\right)}\right).\end{aligned}$$

ACKNOWLEDGMENT

The authors would like to thank K. Balać, who conducted the wavelet analysis of fMRI images. The authors would also like to thank Dr T. Blu and R. Delgado for interesting discussions, as well as the anonymous referees for their detailed and helpful reviews.

REFERENCES

[1] B. B. Mandelbrot, *Gaussian Self-Affinity and Fractals: Globality, The Earth $1/f$ Noise, and R/S*. New York: Springer, 2002, vol. H, Selecta (Old or New).

[2] B. B. Mandelbrot, *The Fractal Geometry of Nature*. New York: W. H. Freeman, 1982.

[3] B. B. Mandelbrot and J. W. Van Ness, "Fractional Brownian motions, fractional noises and applications," *SIAM Rev.*, vol. 10, no. 4, 1968.

[4] P. Chainais, R. Riedi, and P. Abry, "Warped infinitely divisible cascades: Beyond power laws," *Traitement du Signal*, vol. 22, no. 1, pp. 27–39, 2005.

[5] P. Lévy, *Processus stochastiques et mouvement brownien*, 2nd ed. France: Gauthier-Villars, 1965.

[6] M. S. Keshner, "$1/f$ noise," *Proc. IEEE*, vol. 70, no. 3, pp. 212–218, Mar. 1982.

[7] W. E. Leland, W. Willinger, M. S. Taqqu, and D. V. Wilson, "On the self-similar nature of ethernet traffic," *SIGCOMM Comput. Commun. Rev.*, vol. 25, no. 1, pp. 202–213, 1995.

[8] B. Pesquet-Popescu and J. Levy Véhel, "Stochastic fractal models for image processing," *IEEE Signal Process. Mag.*, vol. 19, no. 5, pp. 48–62, May 2002.

[9] A. P. Pentland, "Fractal-based description of natural scenes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 661–674, Nov. 1984.

[10] D. J. Field, "Scale-invariance and self-similar 'wavelet' transforms: An analysis of natural scenes and mamalian visual systems," in *Wavelets, Fractals, and Fourier Transforms*, M. Farge, J. C. R. Hunt, and J. C. Vassilicos, Eds. Oxford, U.K.: Clarendon, 1993.

[11] N. Sarshar and X. Wu, "On rate-distortion models for natural images and wavelet coding performance," *IEEE Trans. Image Process.*, vol. 16, no. 5, pp. 1383–1394, May 2007.

[12] H. E. Hurst, "Long-term storage capacity of reservoirs," *Trans. Amer. Soc. Civil Eng.*, vol. 116, pp. 770–808, 1951.

[13] T. Lundahl, W. J. Ohley, S. M. Kay, and R. Siffert, "Fractional Brownian motion: A maximum likelihood estimator and its application to image texture," *IEEE Trans. Med. Imag.*, vol. MI-5, no. 3, pp. 152–161, Sep. 1986.

[14] B. J. Super and A. C. Bovik, "Localized measurement of image fractal dimension using Gabor filters," *J. Vis. Commun. Image Represent.*, vol. 2, no. 2, pp. 114–128, 1991.

[15] S. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed. San Diego, CA: Academic, 1999.

[16] M. Unser and T. Blu, "Wavelet theory demystified," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 470–483, Feb. 2003.

[17] A. Arneodo, G. Grasseau, and M. Holschneider, "Wavelet transform of multifractals," *Phys. Rev. Lett.*, vol. 61, no. 20, Nov. 1988.

[18] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 673–693, Jul. 1989.

[19] P. Flandrin, "On the spectrum of fractional Brownian motions," *IEEE Trans. Inf. Theory*, vol. 35, no. 1, pp. 197–199, Jan. 1989.

[20] P. Flandrin, "Wavelet analysis and synthesis of fractional Brownian motion," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 910–916, Mar. 1992.

[21] M. Farge, "Wavelet transforms and their applications to turbulence," *Annu. Rev. Fluid Mech.*, vol. 24, pp. 395–457, 1992.

[22] G. W. Wornell and A. V. Oppenheim, "Estimation of fractal signals from noisy measurements using wavelets," *IEEE Trans. Signal Process.*, vol. 40, no. 3, pp. 611–623, Mar. 1992.

[23] Y. Meyer, F. Sellan, and M. Taqqu, "Wavelets, generalized white noise and fractional integration: The synthesis of fractional Brownian motion," *J. Fourier Anal. Appl.*, vol. 5, no. 5, pp. 465–494, 1999.

[24] Y. Meyer, *Wavelets, Vibrations and Scalings*. Providence, RI: Amer. Math. Soc., 1997, vol. 9, CRM Monograph Series.

[25] A. H. Tewfik and M. Kim, "Correlation structure of the discrete wavelet coefficients of fractional Brownian motion," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 904–909, Mar. 1992.

[26] E. Masry, "The wavelet transform of stochastic processes with stationary increments and its application to fractional Brownian motion," *IEEE Trans. Inf. Theory*, vol. 39, no. 1, pp. 260–264, Jan. 1993.

[27] G. W. Wornell, "Wavelet-based representations for the $1/f$ family of fractal processes," *Proc. IEEE*, vol. 81, no. 10, pp. 1428–1450, Oct. 1993.

[28] D. Veitch and P. Abry, "A wavelet-based joint estimator of the parameters of long-range dependence," *IEEE Trans. Inf. Theory*, vol. 45, no. 3, pp. 878–897, Apr. 1999.

[29] B. Ninness, "Estimation of $1/f$ noise," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 32–46, Jan. 1998.

[30] T. Blu and M. Unser, "Self-similarity: Part II—Optimal estimation of fractal processes," *IEEE Trans. Signal Process.*, vol. 55, no. 4, pp. 1364–1378, Apr. 2007.

[31] A. Benassi, S. Jaffard, and D. Roux, "Elliptic gaussian random processes," *Rev. Mat. Iberoamer.*, vol. 13, no. 1, pp. 19–90, 1997.

[32] A. Benassi and J. Istas, "Processus autosimilaires," in *Lois d'échelle, fractales et ondelettes* , P. Abry, P. Gonçalvès, and J. Levy Véhel, Eds. Paris, France: Hermes Lavoisier, 2002, vol. 2, Traité IC2.

[33] W. Wyss, "Fractional noise," *Found. Phys. Lett.*, vol. 4, pp. 235–246, 1991.

[34] M. Unser and T. Blu, "Self-similarity: Part I—Splines and operators," *IEEE Trans. Signal Process.*, vol. 55, no. 4, pp. 1352–1363, Apr. 2007.

[35] C. Rabut, "Elementary m-harmonic cardinal B-splines," *Numer. Alg.*, vol. 2, pp. 39–61, 1992.

[36] B. Bacchelli, M. Bozzini, C. Rabut, and M.-L. Varas, "Decomposition and reconstruction of multidimensional signals using polyharmonic pre-wavelets," *Appl. Comput. Harmon. Anal.*, vol. 18, no. 3, pp. 282–299, May 2005.

[37] D. Van De Ville, T. Blu, and M. Unser, "Isotropic polyharmonic B-splines: Scaling functions and wavelets," *IEEE Trans. Image Process.*, vol. 14, no. 11, pp. 1798–1813, Nov. 2005.

[38] I. M. Gel'fand and G. E. Shilov, *Properties and Operations* Transl. :Translated by Eugene Saletan. New York: Academic, 1964, vol. I, Generalized Functions.

[39] I. M. Gel'fand and N. Vilenkin, *Applications of Harmonic Analysis* Transl.:Translated by Amiel Feinstein. New York: Academic, 1964, vol. IV, Generalized Functions.

[40] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*. New York: Springer-Verlag, 1988.
[41] J. D. Gibson and K. Sayood, "Lattice quantization," *Adv.Electron. Electron Phys.*, vol. 72, pp. 259–330, 1988.
[42] J. Kybic, T. Blu, and M. Unser, "Generalized sampling: A variational approach—Part I: Theory," *IEEE Trans. Signal Process.*, vol. 50, no. 8, pp. 1965–1976, Aug. 2002.
[43] J. Duchon, "Splines minimizing rotation-invariant semi-norms in Sobolev spaces," in *Proc. Constructive Theory of Functions of Several Variables*, A. Dold and B. Eckmann, Eds. New York: Springer-Verlag, 1977, Lecture Notes in Mathematics, pp. 85–100.
[44] M. Arigovindan, "Variational Reconstruction of Vector and Scalar Images from Non-Uniform Samples," Ph.D. dissertation, EPFL, Biomedical Imaging Group, Lausanne, Switzerland, 2005.
[45] I. M. Gel'fand and G. E. Shilov, *Spaces of Fundamental and Generalized Functions* Transl.:Translated by Morris D. Friedman. New York: Academic, 1968, vol. II, Generalized Functions.
[46] M. Unser and T. Blu, "Fractional splines and wavelets," *SIAM Rev.*, vol. 42, no. 1, pp. 43–67, Mar. 2000.
[47] C. Micchelli, C. Rabut, and F. I. Utreras, "Using the refinement equation for the construction of pre-wavelets III: Elliptic splines," *Numer. Alg.*, vol. 1, no. 3, pp. 331–352, 1991.
[48] R. L. Dobrushin, "Gaussian and their subordinated self-similar random generalized fields," *Ann. Probab.*, vol. 7, no. 1, pp. 1–28, 1979.
[49] J.-P. Kahane, *Some Random Series of Functions*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1985.
[50] T. Lindstrøm, "Fractional Brownian fields as integrals of white noise," *Bull. London Math. Soc.*, vol. 25, no. 1, pp. 83–88, 1993.
[51] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. New York: Chapman & Hall, 1994.
[52] E. Bullmore, M. Brammer, I. Harvey, R. Persaud, R. Murray, and M. Ron, "Fractal analysis of the boundary between white matter and cerebral cortex in magnetic resonance images: A controlled study of schizophrenic and manic-depressive patients," *Psychol. Med.*, vol. 24, no. 3, pp. 771–781, 1994.
[53] S. Heinzer, T. Krucker, M. Stampanoni, R. Abela, E. P. Meyer, A. Schuler, P. Schneider, and R. Muller, "Hierarchical microimaging for multiscale analysis of large vascular networks," *NeuroImage*, vol. 32, no. 2, pp. 626–636, 2006.
[54] M. N. Kolountzakis, "The study of translational tiling with Fourier analysis," in *Fourier Analysis and Convexity*, L. Brandolini, L. Colzani, A. Iosevich, and G. Travaglini, Eds. Boston, MA: Birkhäuser, 2004, Applied and Numerical Harmonic Analysis, pp. 131–187.
[55] E. Viscito and J. P. Allebach, "The analysis and design of multidimensional FIR perfect reconstruction filter banks for arbitrary sampling lattices," *IEEE Trans. Circuits Syst.*, vol. 38, no. 1, pp. 29–41, Jan. 1991.
[56] W. R. Madych and S. A. Nelson, "Polyharmonic cardinal splines," *J. Approx. Theory*, vol. 50, no. 2, pp. 141–156, Feb. 1990.

Pouya Dehghani Tafti (M'00) was born in 1981. He received the B.Sc. degree from Sharif University of Technology, Tehran, Iran, and the M.A.Sc. degree from McMaster University, Hamilton, ON, Canada, both in electrical engineering. He is currently pursuing the Ph.D. degree in the Biomedical Imaging Group, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland.

His research interests include stochastic modeling and mathematical signal processing.

Dimitri Van De Ville (M'02) received the M.S. degree in engineering and computer sciences from Ghent University, Ghent, Belgium, in 1998, and the Ph.D. degree in 2002.

He received a grant as Research Assistant with the Fund for Scientific Research Flanders Belgium (FWO). In 2002, he joined Prof. M. Unser's Biomedical Imaging Group at the École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. Since December 2005, he has been heading the Signal Processing Unit at the University Hospital of Geneva, Geneva, Switzerland, as part of the Centre d'Imagerie Biomédicale (CIBM). His research interests include wavelets, statistical analysis, multidimensional splines, and applications in biomedical imaging, such as functional magnetic resonance imaging, spectroscopy, electro-encephalography, and microscopy.

Dr. Van De Ville serves as an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING (since February 2006) and was previously an Associate Editor for IEEE SIGNAL PROCESSING LETTERS (2004–2006). Since 2003, he has also been an Editor and Webmaster of *The Wavelet Digest*. He is co-chair of the Wavelets XII (2007) Wavelets XIII (2009) international conferences, together with V. Goyal and M. Papadakis.

Michael Unser (M'89–SM'94–F'99) received the M.S. (*summa cum laude*) and Ph.D. degrees in electrical engineering in 1981 and 1984, respectively, from the École Polytechnique Fédérale de Lausanne (EPFL), Switzerland.

From 1985 to 1997, he was a Scientist with the National Institutes of Health, Bethesda, MD. He is now full Professor and Director of the Biomedical Imaging Group, EPFL. His main research area is biomedical image processing. He has a strong interest in sampling theories, multiresolution algorithms, wavelets, and the use of splines for image processing. He has published over 150 journal papers on those topics and is one of ISI's Highly Cited authors in Engineering (http://isihighlycited.com).

Dr. Unser has held the position of associate Editor-in-Chief (2003–2005) for the IEEE TRANSACTIONS ON MEDICAL IMAGING and has served as Associate Editor for the same journal (1999–2002; 2006–2007), the IEEE TRANSACTIONS ON IMAGE PROCESSING (1992–1995), and the IEEE SIGNAL PROCESSING LETTERS (1994–1998). He is currently member of the editorial boards of *Foundations and Trends in Signal Processing*, the *SIAM Journal of Imaging Sciences*, and *Sampling Theory in Signal and Image Processing*. He co-organized the first IEEE International Symposium on Biomedical Imaging (ISBI2002). He was the founding chair of the technical committee of the IEEE-SP Society on Bio Imaging and Signal Processing (BISP). He received the 1995 and 2003 Best Paper Awards and the 2000 Magazine Award from the IEEE Signal Processing Society.

The following article originally appeared as

P. D. Tafti and M. Unser, *'Self-similar random vector fields and their wavelet analysis',* Proc. SPIE Wavelets XIII, vol. 7446, pp. 7446Y-1–8, San Diego, US-CA, Aug. 2009.

Self-similar Random Vector Fields and Their Wavelet Analysis

Pouya Dehghani Tafti and Michael Unser

ABSTRACT

This paper is concerned with the mathematical characterization and wavelet analysis of self-similar random *vector* fields. The study consists of two main parts: the construction of random vector models on the basis of their invariance under coordinate transformations, and a study of the consequences of conducting a wavelet analysis of such random models. In the latter part, after briefly examining the effects of standard wavelets on the proposed random fields, we go on to introduce a new family of Laplacian-like *vector* wavelets that in a way duplicate the covariant-structure and whitening relations governing our random models.

1. INTRODUCTION

Self-similar or *fractal* random processes and fields have been used for modelling a wide range of man-made and natural phenomena.[1] This is to a large extent due to their special form of scale-invariant behaviour (self-similarity) that is either strictly or approximately obeyed by the phenomena in question. This form of scale-invariance is also central to the definition of wavelets. Wavelets therefore come across as natural tools for the analysis of data obtained from the observation of fractal phenomena and for quantifying their correspondence to the said models.[2–7]

The most prominent example of self-similar random fields is fractional Brownian motion (fBm).[8] While classical fractional Brownian fields typically correspond to physical scalars, a vectorial extension of such models can be envisioned by bringing in certain laws of coordinate transformation that govern vector quantities in physics. In the present work we first introduce such a vector extension of fractional Brownian fields, and then review the analysis of these random models using existing, as well as novel, wavelets.

To obtain a physically relevant stochastic vector model, we shall focus on two important categories of coordinate transformations—namely rotations and scalings—and study $\mathbb{R}^d$-valued self-similar random fields that rotate according to the laws of coordinate transformation that hold true for vector fields. The desired properties are found in a pseudo-differential operator (an extension of the vector Laplacian), which is used to define the vector fBm field $\mathfrak{B}_{H,\boldsymbol{\xi}}$ via a *whitening* equation:

$$(-\boldsymbol{\Delta})^{\frac{H}{2}+\frac{d}{4}}_{\boldsymbol{\xi}}\mathfrak{B}_{H,\boldsymbol{\xi}} = \epsilon_H \mathfrak{W} \tag{1}$$

($\mathfrak{W}$ is a vector of independent and normalized white noise fields and ϵ_H is a constant).

The above equation is interpreted in Gel'fand and Vilenkin's framework for distribution-theoretical stochastic analysis.[9] In this formalism, random fields are defined by specifying the joint probability measures of their 'inner-products' with test functions. This process can also be understood as taking the random field to be a generalized function (distribution) chosen at random from a generalized function space.

In the following sections we first briefly review the definition of the extended fractional Laplacian operators mentioned above, and then use them to define vector fBm fields. Wavelet analysis of vector fBm takes up the remainder of the paper, and is divided in two parts: an analysis using standard wavelets comes first; this is then followed by a novel *vector* wavelet construction that is also used to analyze the random fields in question.

2. FRACTIONAL LAPLACIANS AND THEIR INVERSES

The Laplacian operator $(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}$ that appears in (1) is linked to the matrix-valued function

$$\boldsymbol{\Phi}^{\gamma}_{(\xi_1,\xi_2)}(\boldsymbol{x}) := \|\boldsymbol{x}\|^{2\gamma}\left[\mathrm{e}^{\xi_1}\frac{\boldsymbol{x}\boldsymbol{x}^{\mathsf{T}}}{\|\boldsymbol{x}\|^2} + \mathrm{e}^{\xi_2}\left(\mathbf{I} - \frac{\boldsymbol{x}\boldsymbol{x}^{\mathsf{T}}}{\|\boldsymbol{x}\|^2}\right)\right], \tag{2}$$

Wavelets XIII, edited by Vivek K. Goyal, Manos Papadakis, Dimitri Van De Ville, Proc. of SPIE
Vol. 7446, 74460Y · © 2009 SPIE · CCC code: 0277-786X/09/$18 · doi: 10.1117/12.824873

via the equation

$$(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}} \boldsymbol{f} := (2\pi)^{-d} \int_{\mathbb{R}^d} \mathrm{d}\boldsymbol{\omega}\ \mathrm{e}^{\mathrm{j}\boldsymbol{x}^{\mathsf{T}}\boldsymbol{\omega}} \boldsymbol{\Phi}^{\gamma}_{\boldsymbol{\xi}}(\boldsymbol{\omega}) \hat{\boldsymbol{f}}(\boldsymbol{\omega}),$$

for $\gamma \in \mathbb{R}_+$ and $\mathrm{e}^{\xi_1}, \mathrm{e}^{\xi_2} \in \mathbb{R}$.

To gain a better understanding of the action of the extended fractional Laplacian let us introduce the auxiliary operator

$$\mathbf{E} : \boldsymbol{f} \mapsto (2\pi)^{-d} \int_{\mathbb{R}^d} \mathrm{d}\boldsymbol{\omega}\ \mathrm{e}^{\mathrm{j}\boldsymbol{x}^{\mathsf{T}}\boldsymbol{\omega}} \frac{\boldsymbol{\omega}\boldsymbol{\omega}^{\mathsf{T}}}{\|\boldsymbol{\omega}\|^2} \hat{\boldsymbol{f}}(\boldsymbol{\omega});$$

whereby

$$(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}} = \mathrm{e}^{\xi_1} (-\boldsymbol{\Delta})^{\gamma}_{\mathbf{0}} \mathbf{E} + \mathrm{e}^{\xi_2} (-\boldsymbol{\Delta})^{\gamma}_{\mathbf{0}} (\mathrm{Id} - \mathbf{E}).$$

$\mathbf{E}$ and its complement $(\mathrm{Id} - \mathbf{E})$ project their argument onto the curl-free, and divergence-free, subspaces of $(\mathrm{L}^d)^d$ respectively. This observation clarifies the action of the extended Laplacian $(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}$, which can thus be said to consists of an ordinary fractional vector Laplacian—operating independently on each component of the field—combined with a Helmholtz-like decomposition and weighted recombination, with the weights depending directly on $\boldsymbol{\xi} = (\xi_1, \xi_2)$.

It can therefore be seen that the parameters ξ_1 and ξ_2 govern the vectorial behaviour of the solution to (1). As important special cases, in addition to the classical fractional Brownian motion obtained by choosing $\xi_1 = \xi_2$, divergence-free (solenoidal) and curl-free (irrotational) solutions to (1) can be produced by letting $\xi_1 \to \infty$ and $\xi_2 \to \infty$ respectively.

Our interest in extended fractional Laplacians is rooted in the fact that they interact with coordinate transformations in a specific way: given the transformation operators

$$\mathbf{S}_{\sigma} : \boldsymbol{f}(\bullet) \mapsto \boldsymbol{f}(\sigma^{-1}\bullet) \qquad \text{(scaling)}$$

and

$$\mathbf{R}_{\boldsymbol{\Omega}} : \boldsymbol{f}(\bullet) \mapsto \boldsymbol{\Omega}\boldsymbol{f}(\boldsymbol{\Omega}^{\mathsf{T}}\bullet), \qquad \text{(vector-field rotation)}$$

one can verify that the following 'quasi-invariances' hold.

$$(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}} \mathbf{S}_{\sigma} \boldsymbol{f} = \sigma^{2\gamma} \mathbf{S}_{\sigma} (-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}} \boldsymbol{f} \qquad \text{(self-similarity)}$$

and

$$(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}} \mathbf{R}_{\boldsymbol{\Omega}} \boldsymbol{f} = \mathbf{R}_{\boldsymbol{\Omega}} (-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}} \boldsymbol{f}. \qquad \text{(rotation-invariance)}$$

($\boldsymbol{\Omega}$ denotes an arbitrary rotation or reflection matrix.)

2.1 Inverse Operators

To solve (1) for a self-similar solution one needs to invert $(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}$ subject to zero boundary conditions at $\boldsymbol{x} = \mathbf{0}$. This process gives rise to the inverse operator

$$(-\acute{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}} : \boldsymbol{f} \mapsto \frac{1}{(2\pi)^d} \Big(\mathrm{e}^{\mathrm{j}\langle \boldsymbol{x}, \boldsymbol{\omega} \rangle} - \sum_{|\boldsymbol{k}| \leq \lfloor 2\gamma - \frac{d}{2} \rfloor} \frac{\mathrm{j}^{|\boldsymbol{k}|} \boldsymbol{x}^{\boldsymbol{k}} \boldsymbol{\omega}^{\boldsymbol{k}}}{\boldsymbol{k}!} \Big) \boldsymbol{\Phi}^{-\gamma}_{-\boldsymbol{\xi}}(\boldsymbol{\omega}) \hat{\boldsymbol{f}}(\boldsymbol{\omega})\ \mathrm{d}\boldsymbol{\omega}.$$

As already mentioned in the introduction, in our distribution-theoretical formalism the self-similar solution to (1), that is, the random field

$$\mathfrak{B}_{H,\boldsymbol{\xi}} := \epsilon_H (-\acute{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}} \mathfrak{W},$$

is defined in terms of 'inner products' $\langle \epsilon_H(-\acute{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}}\mathfrak{W}, g\rangle$ with test functions g. We shall demonstrate in the next section that an exact meaning can be attached to these inner products by transposing the operator to the right side. For this purpose, it is necessary to identify the dual of the above inverse operator; i.e., the operator $(-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\bar{\boldsymbol{\xi}}}$ satisfying

$$\langle (-\acute{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}} f, g\rangle = \langle f, (-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\bar{\boldsymbol{\xi}}} g\rangle$$

for all test functions f and g. This dual operator is given by the integral

$$(-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}} : f \mapsto (2\pi)^{-d} \int_{\mathbb{R}^d} \mathrm{d}\boldsymbol{\omega}\ \mathrm{e}^{\mathrm{j}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\omega}} \boldsymbol{\Phi}^{-\gamma}_{-\boldsymbol{\xi}}(\boldsymbol{\omega})[\mathbf{R}^{\gamma}\hat{f}](\boldsymbol{\omega})$$

with the *regularizer* $\mathbf{R}^{\gamma}$ defined by

$$\mathbf{R}^{\gamma} : \hat{f}(\bullet) \mapsto \hat{f}(\bullet) - \sum_{|\boldsymbol{k}| \le \lfloor 2\gamma - \frac{d}{2} \rfloor} \mathbf{T}_{\boldsymbol{k}}[\hat{f}](\bullet)^{\boldsymbol{k}},$$

where $\mathbf{T}_{\boldsymbol{k}}[\hat{f}]$ denotes the (vector) coefficient of $(\bullet)^{\boldsymbol{k}}$ in the Taylor series expansion of $\hat{f}(\bullet)$ around $\mathbf{0}$.

3. VECTOR FRACTIONAL BROWNIAN MOTION

The properties of statistical self-similarity and invariance to domain rotations are at the heart of the various definitions of vector fractional Brownian motion we shall now consider.

3.1 Vector Fractional Brownian Motion

Fractional Brownian motions (fBm's) are defined as specific types of non-stationary zero-mean Gaussian fields. As a result of their Gaussian probability law, knowledge of their second-order statistics suffices for their full characterization. A vector fBm $\mathfrak{B}_H$ with Hurst exponent H, $0 < H < 1$ is classically defined via its variogram, i.e., the expectation

$$\mathbf{E}\left\{\|\mathfrak{B}_H(\boldsymbol{x}) - \mathfrak{B}_H(\boldsymbol{y})\|^2\right\} \propto \|\boldsymbol{x} - \boldsymbol{y}\|^{2H}.$$

Self-similarity and rotation-invariance of $\mathfrak{B}_{H,\boldsymbol{\xi}}$ can be directly verified in the above formula.

The above definition leaves the cross-correlation of vector components unspecified. These components are normally assumed to be independent, in which case vector fBm becomes nothing but a vector of scalar fBm's.[10] When considering specific forms of vector behaviour (such as divergence-free flow), this independence assumption proves to be too restrictive and, consequently, a broader definition is to be sought.

Such a broad definition of vector fBm was given in the introduction by means of a whitening equation (Eqn (1)), that is, a stochastic fractional PDE of the form

$$(-\boldsymbol{\Delta})^{\frac{H}{2}+\frac{d}{4}}_{\boldsymbol{\xi}} \mathfrak{B}_{H,\boldsymbol{\xi}} = \epsilon_H \mathfrak{W};$$

where ϵ_H is a constant, $\mathfrak{W}$ is a vector of normalized white noise fields, and $\mathfrak{B}_{H,\boldsymbol{\xi}}$ denotes the desired solution. The extended Laplacian operator $(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}$ defined in the previous section accounts for the self-similarity and rotation-invariance of $\mathfrak{B}_{H,\boldsymbol{\xi}}$. To give an explicit definition of $\mathfrak{B}_{H,\boldsymbol{\xi}}$ the operator needs to be inverted with attention to the special boundary conditions imposed on the solution (zero boundary conditions at $\boldsymbol{x} = \mathbf{0}$). This is achieved using the first of the particular inverse operators introduced in §2.1:

$$\mathfrak{B}_{H,\boldsymbol{\xi}} := \epsilon_H (-\acute{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\boldsymbol{\xi}} \mathfrak{W}. \tag{3}$$

One has command over the order of self-similarity of the resulting random models via the parameter H (the Hurst exponent). Directional behaviour (i.e., solenoidal vs irrotational tendencies) can also be controlled by adjusting $\boldsymbol{\xi}$, as noted in §2. These models can therefore be adapted to the directional structure of the phenomenon under investigation.

3.2 Mathematical Interpretation

We shall at this point elaborate briefly on the mathematical interpretation of Eqn (3). In Gel'fand and Vilenkin's theory of stochastic analysis, one takes a random function such as $\mathfrak{W}$ to be an entity whose 'inner products' $\langle \mathfrak{W}, g\rangle$ with test functions g in some function space are ordinary random variables. $\mathfrak{W}$ is then fully defined once the joint probability measure of any finite number of such inner products is given in a consistent manner. If an operator such as $(-\acute{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\boldsymbol{\xi}}$ acts on $\mathfrak{W}$, this operator is carried over to the side of the test function by duality, yielding the equality

$$\langle (-\acute{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\boldsymbol{\xi}} \mathfrak{W}, g\rangle = \langle \mathfrak{W}, (-\grave{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\boldsymbol{\xi}} g\rangle$$

that holds by definition.

As noted previously, when dealing with zero-mean Gaussian random functions (as in the present case), knowledge of second-order statistics suffices. For the vector fBm $\mathfrak{B}_{H,\boldsymbol{\xi}}$ this information is captured by the *correlation form* defined by the equation

$$\langle\!\langle f, g\rangle\!\rangle_{\mathfrak{B}_{H,\boldsymbol{\xi}}} := \mathbf{E}\left\{\overline{\langle \mathfrak{B}_{H,\boldsymbol{\xi}}, f\rangle}\langle \mathfrak{B}_{H,\boldsymbol{\xi}}, g\rangle\right\}.$$

This quantity can be computed as follows:

$$\begin{aligned}
\mathbf{E}\left\{\overline{\langle \mathfrak{B}_{H,\boldsymbol{\xi}}, f\rangle}\langle \mathfrak{B}_{H,\boldsymbol{\xi}}, g\rangle\right\} &= \mathbf{E}\left\{\overline{\langle \epsilon_H(-\acute{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\bar{\boldsymbol{\xi}}} \mathfrak{W}, f\rangle}\langle \epsilon_H(-\acute{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\bar{\boldsymbol{\xi}}} \mathfrak{W}, g\rangle\right\} \\
&= |\epsilon_H|^2 \mathbf{E}\left\{\overline{\langle \mathfrak{W}, (-\grave{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\boldsymbol{\xi}} f\rangle}\langle \mathfrak{W}, (-\grave{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\boldsymbol{\xi}} g\rangle\right\} \\
&= |\epsilon_H|^2 \langle\!\langle (-\grave{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\boldsymbol{\xi}} f, (-\grave{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\boldsymbol{\xi}} g\rangle\!\rangle_{\mathfrak{W}} \\
&= |\epsilon_H|^2 \langle (-\grave{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\boldsymbol{\xi}} f, (-\grave{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\boldsymbol{\xi}} g\rangle \\
&= \frac{|\epsilon_H|^2}{(2\pi)^d} \int_{\mathbb{R}^d} \mathrm{d}\boldsymbol{\omega}\, [\mathbf{R}^{\frac{2H+d}{4}} \hat{f}]^{\mathsf{H}}(\boldsymbol{\omega}) \boldsymbol{\Phi}^{-H-\frac{d}{2}}_{-2\mathrm{Re}\,\boldsymbol{\xi}}(\boldsymbol{\omega}) [\mathbf{R}^{\frac{2H+d}{4}} \hat{g}](\boldsymbol{\omega}); \qquad \text{(Parseval)}
\end{aligned}$$

where we have used the following fact/definition regarding the correlation form of white Gaussian noise:

$$\langle\!\langle f, g\rangle\!\rangle_{\mathfrak{W}} = \langle f, g\rangle.$$

4. WAVELET ANALYSIS OF VECTOR FBM

As a consequence of the whitening relation dictated by (1), vector fBm fields can (in loose terms) be regarded as combinations of (non-stationary) fractional integrals of white noise, with fractional integration appearing in the inversion of $(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}$.

For the purpose of analyzing these fields, it is desirable to undo the integration process, in order to produce a *stationarized* random field. This can be achieved by the application of a wavelet transform to the field, as wavelets—owing to their vanishing moments—behave like smoothed differentiators. Another important property of such an analysis follows from the self-similarity of both the wavelets and the random fields under consideration, which results in a simple relation between the probability distributions of wavelet coefficients at different scales. Consequently, wavelet analysis offers an advantageous way of verifying the introduced models and estimating their parameters in practice.[2,3,6,7,11–14]

Stationarizing the models in the manner we have just described can be seen as an attempt to construct a multi-resolution approximation of the operator $(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}$ by using wavelets that have the operator embedded in them. Since in this process we are dealing with vector quantities and operators, the said wavelets are in general represented as matrices (or equivalently as groups of vector-valued functions).

4.1 Scalar Wavelet Analysis of Vector fBm

In a biorthogonal wavelet system the basis elements can be represented as $\psi_{n,\boldsymbol{k}} \in \mathrm{L}^2$ with dual elements $\tilde{\psi}_{n,\boldsymbol{k}} \in \mathrm{L}^2$ (where n denotes the resolution and $\boldsymbol{k}$ indexes a lattice in $\mathbb{R}^d$), so that the i-th component of $\mathfrak{B}_{H,\boldsymbol{\xi}}$ finds the representation

$$\big[\mathfrak{B}_{H,\boldsymbol{\xi}}\big]_i = \sum_{n,\boldsymbol{k}} \langle \big[\mathfrak{B}_{H,\boldsymbol{\xi}}\big]_i, \tilde{\psi}_{n,\boldsymbol{k}} \rangle \psi_{n,\boldsymbol{k}} = \sum_{n,\boldsymbol{k}} \langle \mathfrak{B}_{H,\boldsymbol{\xi}}, \hat{\boldsymbol{e}}_i \tilde{\psi}_{n,\boldsymbol{k}} \rangle \psi_{n,\boldsymbol{k}}.$$

In standard constructions, wavelets in a given sub-band at a given resolution n are lattice shifts of one another:

$$\psi_{n,\boldsymbol{k}}(\boldsymbol{x}) = \psi_{n,\boldsymbol{0}}(\boldsymbol{x} - \boldsymbol{k}).$$

Consider the discrete random process $w_{n,i}$ defined by

$$w_{n,i}[\boldsymbol{k}] := \langle \mathfrak{B}_{H,\boldsymbol{\xi}}, \hat{\boldsymbol{e}}_i \tilde{\psi}_{\boldsymbol{k}} \rangle.$$

Then

$$w_{n,i}[\boldsymbol{k}] = \langle \epsilon_H (-\dot{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\tilde{\boldsymbol{\xi}}} \mathfrak{W}, \hat{\boldsymbol{e}}_i \tilde{\psi}_{n,\boldsymbol{k}} \rangle = \langle \epsilon_H \mathfrak{W}, (-\dot{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\boldsymbol{\xi}} \hat{\boldsymbol{e}}_i \tilde{\psi}_{n,\boldsymbol{k}} \rangle. \tag{4}$$

Assuming that $\tilde{\psi}_{n,\boldsymbol{k}}$ has sufficiently many vanishing moments (Fourier-domain zeros at $\boldsymbol{\omega} = 0$) so that $\mathbf{R}^{\frac{2H+d}{4}} \hat{\boldsymbol{e}}_i \tilde{\psi}_{n,\boldsymbol{k}} = \hat{\boldsymbol{e}}_i \tilde{\psi}_{n,\boldsymbol{k}}$, we have

$$(-\dot{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\boldsymbol{\xi}} \hat{\boldsymbol{e}}_i \tilde{\psi}_{n,\boldsymbol{k}} = (2\pi)^{-d} \int_{\mathbb{R}^d} \mathrm{d}\boldsymbol{\omega}\, \mathrm{e}^{\mathrm{j}\boldsymbol{x}^{\mathsf{T}}\boldsymbol{\omega}} \big[\boldsymbol{\Phi}^{-\frac{2H+d}{4}}_{-\boldsymbol{\xi}}(\boldsymbol{\omega})\big]^{\mathsf{T}}_i \hat{\tilde{\psi}}_{n,\boldsymbol{k}}(\boldsymbol{\omega}) \in \mathrm{L}^2.$$

It therefore becomes clear that the above inverse operator is shift-invariant over this particular subspace of functions (in contrast to the general case of functions with non-vanishing moments for which it is not, due to the space-dependent operation of $\mathbf{R}^{\frac{2H+d}{4}}$). But then, since $\mathfrak{W}$ is stationary, the last argument and (4) together prove that in this case the discrete random process $w_{n,i}$ is stationary.

In other words, a wavelet analysis of vector fractional Brownian motion with wavelets whose moments of degrees up to $\lfloor H \rfloor$ vanish yields stationary coefficients at each resolution.

4.2 Extension to Vector Wavelets

Instead of using ordinary scalar wavelets to analyze vector fBm, it is possible to define *vector* wavelets that provide a multi-resolution approximation of the fractional Laplacian operator $(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}$ (in comparison, scalar wavelets approximate differentiators of some order). This purpose is achieved by constructing a multi-resolution hierarchy of vector *splines* and defining wavelets via the application of a fractional Laplacian to interpolating spline basis functions.

4.2.1 Vector Splines

In a manner similar to that in which polyharmonic splines and wavelets are defined,[14] we construct a multi-resolution analysis of $(\mathrm{L}^d)^d$ by first finding a localized version of the Green's function of $(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}$. Towards this end, we first recall that the said Green's function satisfies

$$(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}} \boldsymbol{\Gamma} = \delta \mathbf{I},$$

which yields the equation

$$\boldsymbol{\Phi}^{\gamma}_{\boldsymbol{\xi}}(\boldsymbol{\omega}) \hat{\boldsymbol{\Gamma}}(\boldsymbol{\omega}) = \mathbf{I}, \quad \text{almost everywhere.}$$

Given that the matrix inverse of $\boldsymbol{\Phi}^{\gamma}_{\boldsymbol{\xi}}(\boldsymbol{\omega})$ for $\boldsymbol{\omega} \neq \boldsymbol{0}$ is simply $\boldsymbol{\Phi}^{-\gamma}_{-\boldsymbol{\xi}}(\boldsymbol{\omega})$, we can conclude that the desired Green's function has the Fourier transform

$$\boldsymbol{\Phi}^{-\gamma}_{-\boldsymbol{\xi}}(\boldsymbol{\omega}) \quad \text{almost everywhere.}$$

Much in the same way as in scalar wavelet theory, a localized and interpolating version of the Green's function can be obtained via dividing its Fourier expression by its periodization. A Fourier-domain definition of the Lagrange basis function $\mathbf{L}_{\boldsymbol{\xi}}^{\gamma}$ is therefore

$$\hat{\mathbf{L}}_{\boldsymbol{\xi}}^{\gamma}(\boldsymbol{\omega}) := \boldsymbol{\Phi}_{-\boldsymbol{\xi}}^{-\gamma}(\boldsymbol{\omega}) \big[\sum_{\boldsymbol{k}\in\mathbb{Z}^d} \boldsymbol{\Phi}_{-\boldsymbol{\xi}}^{-\gamma}(\boldsymbol{\omega} - 2\pi\boldsymbol{k}) \big]^{-1},$$

which ensures that

$$\mathbf{L}_{\boldsymbol{\xi}}^{\gamma}(\boldsymbol{k}) = \delta_{\boldsymbol{k}}\mathbf{I} \quad \text{for } \boldsymbol{k} \in \mathbb{Z}^d$$

(with $\delta_{\boldsymbol{k}}$ denoting Kronecker's delta function).

With basis elements (multi-integer shifts of $\mathbf{L}_{\boldsymbol{\xi}}^{\gamma}$) given as matrix-valued functions in the vector setting, linear expansions take vectors as coefficients. The approximation space $\mathcal{V}_0$ is therefore defined thus:

$$\mathcal{V}_0 = \big\{ \sum_{\boldsymbol{k}\in\mathbb{Z}^d} \mathbf{L}_{\boldsymbol{\xi}}^{\gamma}(\bullet - \boldsymbol{k})\boldsymbol{c}[\boldsymbol{k}] \big| \boldsymbol{c} \in (\ell^2)^d \big\}.$$

One can verify that $\mathbf{L}_{\boldsymbol{\xi}}^{\gamma}$ is indeed refinable, with the Fourier transform of the refinement filter given by*

$$\big[\hat{\mathbf{L}}_{\boldsymbol{\xi}}^{\gamma}(\boldsymbol{\omega})\big]^{-1} \hat{\mathbf{L}}_{\boldsymbol{\xi}}^{\gamma}(2\boldsymbol{\omega}) = 2^{-2\gamma} \big[\sum_{\boldsymbol{k}\in\mathbb{Z}^d} \boldsymbol{\Phi}_{-\boldsymbol{\xi}}^{-\gamma}(\bullet - 2\pi\boldsymbol{k}) \big] \big[\sum_{\boldsymbol{k}\in\mathbb{Z}^d} \boldsymbol{\Phi}_{-\boldsymbol{\xi}}^{-\gamma}(2\bullet - 2\pi\boldsymbol{k}) \big]^{-1},$$

which is a $2\pi\mathbb{Z}^d$-periodic function over $\mathbb{R}^d$. Nested multi-resolution spaces are then defined by the relation

$$\mathcal{V}_n := \big\{ \boldsymbol{f}(2^{-n}\bullet) \big| \boldsymbol{f}(\bullet) \in \mathcal{V}_0 \big\}.$$

4.2.2 Vector Wavelets

Our intention in this subsection is to introduce wavelets that can be used to analyze a vector field $\boldsymbol{f}(\bullet)$ in such a way that analysis coefficients together can be seen as a multi-resolution representation of the fractional Laplacian of the field in question. In other words, a given wavelet coefficient $\boldsymbol{w}_n[\boldsymbol{k}]$ (with n encoding the resolution and $\boldsymbol{k}$ the position) is to represent the quantity

$$\int_{\mathbb{R}^d} \mathrm{d}\boldsymbol{x}\, \mathbf{K}(2^n\boldsymbol{x} - \boldsymbol{k})(-\boldsymbol{\Delta})_{\boldsymbol{\xi}}^{\gamma} \boldsymbol{f}(\boldsymbol{x})$$

where $\mathbf{K}$ is a localized matrix-valued function. The above integral can then be interpreted as a local measure of $(-\boldsymbol{\Delta})_{\boldsymbol{\xi}}^{\gamma}\boldsymbol{f}(\boldsymbol{x})$, the locality or resolution of which depends on n. One typically imposes the additional requirements on the wavelets that they be orthogonal to lower resolution approximation spaces and span their orthogonal complements in spaces of higher resolutions.

---

* The two-scale relation in this case takes the form

$$\mathbf{L}_{\boldsymbol{\xi}}^{\gamma}(2^{-1}\bullet) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} \mathbf{L}_{\boldsymbol{\xi}}^{\gamma}(\bullet - \boldsymbol{k})\boldsymbol{h}[\boldsymbol{k}],$$

which is to say that the refinement filter $\boldsymbol{h}$ corresponds to a sequence of vectors. The Fourier transform $\mathring{\boldsymbol{h}}$ of a vector filter $\boldsymbol{h}$ is given by the relation

$$\mathring{\boldsymbol{h}}(\boldsymbol{\omega}) := \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{j\boldsymbol{k}^{\mathsf{T}}\boldsymbol{\omega}} \boldsymbol{h}[\boldsymbol{k}].$$

A $2\pi\mathbb{Z}^d$-periodic vector-valued function over $\mathbb{R}^d$ can be taken to be the Fourier transform of a vector filter. As such, to show that $\mathbf{L}_{\boldsymbol{\xi}}^{\gamma}$ is refinable, we need to show that the refinement filter obtained by dividing the Fourier transform of $\mathbf{L}_{\boldsymbol{\xi}}^{\gamma}(2^{-1}\bullet)$ by that of $\mathbf{L}_{\boldsymbol{\xi}}^{\gamma}(\bullet)$ is $2\pi\mathbb{Z}^d$-periodic.

It turns out that the requirements sketched above can be satisfied by defining $2^d - 1$ mother wavelets (corresponding to the usual $2^d - 1$ Cartesian cosets) as follows.

$$\mathbf{W}^{\gamma}_{\boldsymbol{\xi},m}(\bullet) := (-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}} \mathbf{L}^{2\gamma}_{\boldsymbol{\xi}+\bar{\boldsymbol{\xi}}}(\bullet - \boldsymbol{\varsigma}_m),$$

with $\boldsymbol{\varsigma}_m$, $1 \leq m \leq 2^d - 1$, denoting the m-th coset identifier (i.e. the vector corresponding to the binary representation of m). A Fourier-domain characterization of these wavelets is given by the relation

$$\begin{aligned}\hat{\mathbf{W}}^{\gamma}_{\boldsymbol{\xi},m}(\boldsymbol{\omega}) &:= \boldsymbol{\Phi}^{\gamma}_{\boldsymbol{\xi}}(\boldsymbol{\omega}) \hat{\mathbf{L}}^{2\gamma}_{\boldsymbol{\xi}+\bar{\boldsymbol{\xi}}}(\boldsymbol{\omega}) \mathrm{e}^{-\mathrm{j}\boldsymbol{\omega}^{\mathsf{T}}\boldsymbol{\varsigma}_m} \\ &= \boldsymbol{\Phi}^{-\gamma}_{-\boldsymbol{\xi}}(\boldsymbol{\omega}) \big[\sum_{\boldsymbol{k}\in\mathbb{Z}^d} \boldsymbol{\Phi}^{-2\gamma}_{-(\boldsymbol{\xi}+\bar{\boldsymbol{\xi}})}(\boldsymbol{\omega} - 2\pi\boldsymbol{k}) \big]^{-1} \mathrm{e}^{-\mathrm{j}\boldsymbol{\omega}^{\mathsf{T}}\boldsymbol{\varsigma}_m}.\end{aligned}$$

4.2.3 Vector Wavelet Analysis of Vector fBm

Let us now assume a $\gamma \geq \frac{H}{2} + \frac{d}{4}$ and consider the vector wavelet coefficients of a vector fBm field $\mathfrak{B}_{H,\boldsymbol{\xi}}$:

$$\begin{aligned}\boldsymbol{w}_{0,m}[\boldsymbol{k}] = &\int_{\mathbb{R}^d} \mathrm{d}\boldsymbol{x} \, \big[\mathbf{W}^{\gamma}_{\bar{\boldsymbol{\xi}}}(\boldsymbol{x})\big]^{\mathsf{H}} \mathfrak{B}_{H,\boldsymbol{\xi}}(\boldsymbol{x}) \\ &= \epsilon_H \int_{\mathbb{R}^d} \mathrm{d}\boldsymbol{x} \, \big[(-\boldsymbol{\Delta})^{\gamma-\frac{H}{2}-\frac{d}{4}}_{\mathbf{0}} \mathbf{L}^{2\gamma}_{\boldsymbol{\xi}+\boldsymbol{\xi}'}(\boldsymbol{x})\big] \mathfrak{W}(\boldsymbol{x})\end{aligned}$$

which, similar to the previous analysis with scalar wavelets, shows that these wavelet analysis coefficients constitute a stationary vector-valued random field (defined over $\mathbb{Z}^d$) corresponding to filtered white noise.

5. CONCLUSION

We defined a broad family of fractional Brownian random vector fields by introducing a pseudo-differential operator that extends the fractional vector Laplacian, and positing it as the whitening operator of the said random fields. A wavelet analysis of the introduced random fields using standard wavelets was then studied, with emphasis put on the stationarizing effect of a wavelet transform on vector fBm.

Next, we turned our attention to the construction of multi-resolution approximation spaces and wavelets that are directly linked to the extended fractional Laplacian operator, in the sense that an analysis based on the said wavelets provides in effect a multi-resolution representation of the fractional Laplacian of the vector field under investigation. This property was used to demonstrate the stationarizing effect of these novel vector wavelets on fractional Brownian vector fields.

These newly introduced wavelets, and the random fields to which they are inherently related, have many interesting properties which we have not examined in this brief overview. These properties, and their interplay will be the subject of our future investigations. Moreover, in the present study, we have not addressed the practical implementation of the proposed wavelet analysis on a digital computer. Nevertheless, we are aware of an interesting connection between these wavelets and an important special function (namely the Epstein zeta function) for which fast evaluation algorithms have been proposed in the literature.

REFERENCES

1. B. B. Mandelbrot, *Gaussian Self-Affinity and Fractals: Globality, The Earth, $1/f$ Noise, and R/S*, vol. H of *Selecta (Old or New)*, Springer, New York, 2002.
2. S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.* **11**, pp. 674–693, July 1989.
3. P. Flandrin, "Wavelet analysis and synthesis of fractional Brownian motion," *IEEE Trans. Inf. Theory* **38**, pp. 910–916, March 1992.
4. G. W. Wornell, "Wavelet-based representations for the $1/f$ family of fractal processes," *Proc. IEEE* **81**, pp. 1428–1450, October 1993.
5. B. Ninness, "Estimation of $1/f$ noise," *IEEE Trans. Inf. Theory* **44**(1), pp. 32–46, 1998.

6. D. Veitch and P. Abry, "A wavelet-based joint estimator of the parameters of long-range dependence," *IEEE Trans. Inf. Theory* **45**, pp. 878–897, April 1999.
7. Y. Meyer, F. Sellan, and M. Taqqu, "Wavelets, generalized white noise and fractional integration: The synthesis of fractional Brownian motion," *J. Fourier Anal. Appl.* **5**(5), pp. 465–494, 1999.
8. B. B. Mandelbrot and J. W. Van Ness, "Fractional Brownian motions, fractional noises and applications," *SIAM Review* **10**(4), pp. 422–437, 1968.
9. I. M. Gel'fand and N. Vilenkin, *Applications of Harmonic Analysis*, vol. IV of *Generalized Functions*, Academic Press, 1964. translated by Amiel Feinstein.
10. J.-P. Kahane, *Some Random Series of Functions*, Cambridge University Press, 2nd ed., 1985.
11. M. Farge, "Wavelet transforms and their applications to turbulence," *Annu. Rev. Fluid Mech.* **24**, pp. 395–457, 1992.
12. E. Masry, "The wavelet transform of stochastic processes with stationary increments and its application to fractional Brownian motion," *IEEE Trans. Inf. Theory* **39**, pp. 260–264, January 1993.
13. T. Blu and M. Unser, "Self-similarity: Part II—Optimal estimation of fractal processes," *IEEE Trans. Signal Process.* **55**, pp. 1364–1378, April 2007.
14. P. D. Tafti, D. Van De Ville, and M. Unser, "Invariances, Laplacian-like wavelet bases, and the whitening of fractal processes," *IEEE Trans. Image Process.* **18**, pp. 698–702, April 2009.

P. D. Tafti, R. Delgado-Gonzalo, A. F. Stalder, and M. Unser, *'Fractal modelling and analysis of flow-field images',* Proc. 7th IEEE Int. Symp. on Biomed. Imaging (ISBI 2010), pp. 49–52, Rotterdam, The Netherlands, Apr. 2010.

FRACTAL MODELLING AND ANALYSIS OF FLOW-FIELD IMAGES

Pouya Dehghani Tafti[*], *Ricard Delgado-Gonzalo*[*], *Aurélien F. Stalder*[§‡], *and Michael Unser*[*]

[*]Biomedical Imaging Group, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
[§]Dept of Diagnostic Radiology – Medical Physics, University Hospital Freiburg, Germany
[‡]Dept of Radiology, Xuanwu Hospital, Capital Medical University, Beijing, China

ABSTRACT

We introduce stochastic models for flow fields with parameters that dictate the scale-dependent (self-similar) character of the field and control the balance between its rotational vs compressive behaviour. The development of our models is motivated by the availability of imaging modalities that measure flow vector fields (flow-sensitive MRI and Doppler ultrasound). To study such data, we formulate estimators of the model parameters, and use them to quantify the Hurst exponent and directional properties of synthetic and real-world flow fields (measured by means of phase-contrast MRI) in 3D.

***Index Terms*—** vector fractional Brownian motion, flow-sensitive MRI, wavelets, vector fields, Hurst exponent, curl, divergence.

1. INTRODUCTION

Stochastic fractal models are commonly used in a range of applications where some form of self-similarity or scale-invariance is observed (examples include image processing, seismology, and the study of growth processes [1, 2]). The quintessential stochastic fractal is the fractional Brownian motion (fBm)—so named by Mandelbrot and Van Ness [3] but already considered by Kolmogorov [4] and others before them—which can be defined by means of the structure of its variogram (variance of increments) [5]:

$$\mathbb{E}\{|B_H(\boldsymbol{x}) - B_H(\boldsymbol{y})|^2\} \propto \|\boldsymbol{x} - \boldsymbol{y}\|^{2H}$$

In the above equation $\mathbb{E}$ denotes the mathematical expectation operator; H is the Hurst exponent, named after H.E. Hurst who first used estimates of its value in the context of hydrology. The definition is not complete unless we also mention that B_H is a Gaussian process with zero mean which almost surely takes the value 0 at $\boldsymbol{x} = \boldsymbol{0}$.

Fractal behaviour is also observed in the study of flow and turbulence [6]. With the availability of new biomedical imaging techniques that allow of measurement of flow fields (e.g. Doppler ultrasound or flow-sensitive MRI [7, 8]), the question of the applicability of fractal models to these measured phenomena naturally arises. In order to address this question, it is necessary first to generalize the classical scalar fractal models to the vector setting, and then use statistical methods to compare these models against simulated and real-world data. In this paper we take steps in the mentioned directions.

To define the vector counterpart of fBm, we rely on the observation (also classical) that fBm may be regarded as the solution of a fractional differential equation involving fractional Laplacians, subject to zero boundary conditions at $\boldsymbol{x} = \boldsymbol{0}$ [9, 10]:

$$(-\Delta)^{\frac{H}{2}+\frac{d}{4}} B_H = \epsilon_H W \qquad (1)$$

This work was funded by the Swiss Science Foundation under grant 200020-121763.

(d is the dimension of the domain; W is a normalized white Gaussian noise process; and ϵ_H is a proportionality constant). A vector generalization is then obtained by replacing the scalar Laplacian in (1) with a(n extended) vector Laplacian (Section 2). In addition to the Hurst exponent, the new random model is indexed by two (dependent) additional parameters that control the balance between rotational and divergent or convergent tendencies in the field.

We shall employ wavelets to address the second aspect (statistical study of the models). Wavelet-based techniques have been used with efficacy in the past in the statistical analysis of fBm and, in particular, in the estimation of the Hurst exponent [2, 10–14]. The effectiveness of wavelet analysis for this purpose relies on two facts: First, wavelets essentially behave as low-frequency differentiators, and this, by virtue of Eqn (1), means that a wavelets analysis of fBm (which is non-stationary) yields coefficients that correspond to (stationary) filtered white noise. Second, the multi-scale nature of a wavelet analysis captures the self-similar structure of fBm.

Consequently, in order to move towards understanding the connection between vector fBm models and biomedical data, in Section 3 we develop a wavelet estimator of the parameters of the model, which we then apply to synthetic fields and measured phase-contrast MRI data (Section 4). A few remarks and observations conclude the paper (Section 5).

2. VECTOR FRACTIONAL BROWNIAN MOTIONS

The vector extension of Fractional Brownian Motion (fBm) we shall consider spans a new[1] family of random vector field models that are singled out by their special invariance and self-similarity properties with respect to changes of scale and rotations of the coordinate system. These random fields can be defined as solutions of the parametric fractional differential equation (a.k.a. whitening equation)

$$(-\boldsymbol{\Delta})^{\frac{H}{2}+\frac{d}{4}}_{\boldsymbol{\xi}} \boldsymbol{B}_{H,\boldsymbol{\xi}} = \epsilon_H \boldsymbol{W} \qquad (2)$$

where: $\boldsymbol{W}$ is a vector of independent white Gaussian noises; $(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}$ is an extended fractional Laplacian we shall define below; H denotes the Hurst exponent that is a measure of the dependence of the values of the random field at different locations; d is the number of spatial dimensions; $\boldsymbol{\xi} = (\xi_1, \xi_2)$ is a vector of (dependent) parameters that, as we shall see, capture the directional behaviour of the vector field; and ϵ_H is a special constant. The equation is to be solved by imposing zero boundary conditions at the origin.

The fractional vector Laplacian that appears in Eqn (2) is a combination of the fractional Laplacian (Riesz derivative) $(-\boldsymbol{\Delta})^{\gamma}$, defined in the Fourier domain by the symbol $\|\boldsymbol{\omega}\|^{2\gamma}$, and a rebalancing of the divergence-free and curl-free components of the

[1] To the best of the authors' knowledge.

operand, achieved by means of an operator $\mathbf{E}$ that projects its operand onto its curl-free component[2] and has Fourier symbol $\boldsymbol{\omega}\boldsymbol{\omega}^{\mathsf{T}}/\|\boldsymbol{\omega}\|^2$. In symbols:

$$(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}} = \big[\mathrm{e}^{\xi_1}\mathbf{E} + \mathrm{e}^{\xi_2}(\mathrm{Id} - \mathbf{E})\big](-\boldsymbol{\Delta})^{\gamma}$$

with

$$\begin{aligned}
(-\boldsymbol{\Delta})^{\gamma} &\quad\stackrel{\mathcal{F}}{\longleftrightarrow}\quad \|\boldsymbol{\omega}\|^{2\gamma} \\
\mathbf{E} &\quad\stackrel{\mathcal{F}}{\longleftrightarrow}\quad \frac{\boldsymbol{\omega}\boldsymbol{\omega}^{\mathsf{T}}}{\|\boldsymbol{\omega}\|^2} \\
(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}} &\quad\stackrel{\mathcal{F}}{\longleftrightarrow}\quad \|\boldsymbol{\omega}\|^{2\gamma}\Big[\mathrm{e}^{\xi_1}\frac{\boldsymbol{\omega}\boldsymbol{\omega}^{\mathsf{T}}}{\|\boldsymbol{\omega}\|^2} + \mathrm{e}^{\xi_2}\big(\mathbf{I} - \frac{\boldsymbol{\omega}\boldsymbol{\omega}^{\mathsf{T}}}{\|\boldsymbol{\omega}\|^2}\big)\Big] =: \hat{\Phi}^{\gamma}_{\boldsymbol{\xi}}(\boldsymbol{\omega})
\end{aligned}$$

(note the definition of the matrix valued function $\hat{\Phi}^{\gamma}_{\boldsymbol{\xi}}$; also note that, at least for $\gamma \geq 0$, the operators $(-\boldsymbol{\Delta})^{\gamma}$ and $\mathbf{E}$ commute).

The inverse of $(-\boldsymbol{\Delta})^{H/2+d/4}_{\boldsymbol{\xi}}$, taking into account the zero boundary conditions at $\boldsymbol{x} = \mathbf{0}$, is given by the integral operator

$$(-\boldsymbol{\Delta})^{-\gamma}_{-\boldsymbol{\xi}} : \boldsymbol{f} \mapsto \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\Big(\mathrm{e}^{\mathrm{j}\langle \boldsymbol{x},\boldsymbol{\omega}\rangle} - \sum_{|\boldsymbol{k}|\leq\lfloor H\rfloor}\frac{\mathrm{j}^{|\boldsymbol{k}|}\boldsymbol{x}^{\boldsymbol{k}}\boldsymbol{\omega}^{\boldsymbol{k}}}{\boldsymbol{k}!}\Big)\cdot \hat{\Phi}^{-\gamma}_{-\boldsymbol{\xi}}(\boldsymbol{\omega})\hat{\boldsymbol{f}}(\boldsymbol{\omega})\ \mathrm{d}\boldsymbol{\omega}$$

($\hat{\boldsymbol{f}}$ denotes the distributional Fourier transform of $\boldsymbol{f}$). The following identity is easy to establish.

$$(-\boldsymbol{\Delta})^{-\gamma}_{-\boldsymbol{\xi}} = (-\boldsymbol{\Delta})^{-\gamma}_{\mathbf{0}}\big[\mathrm{e}^{-\xi_1}\mathbf{E} + \mathrm{e}^{-\xi_2}(\mathrm{Id} - \mathbf{E})\big]. \tag{3}$$

The inverse operator that we just introduced has the following self-similarity properties with respect to changes of scale and rotations:

$$(-\boldsymbol{\Delta})^{-\gamma}_{-\boldsymbol{\xi}}\{\boldsymbol{f}(\sigma^{-1}\bullet)\} = \sigma^{-2\gamma}\big((-\boldsymbol{\Delta})^{-\gamma}_{-\boldsymbol{\xi}}\boldsymbol{f}\big)(\sigma^{-1}\bullet) \quad \text{(scale-invariance)}$$

$$(-\boldsymbol{\Delta})^{-\gamma}_{-\boldsymbol{\xi}}\{\boldsymbol{\Omega}\boldsymbol{f}(\boldsymbol{\Omega}^{\mathsf{T}}\bullet)\} = \boldsymbol{\Omega}\big((-\boldsymbol{\Delta})^{-\gamma}_{-\boldsymbol{\xi}}\boldsymbol{f}\big)(\boldsymbol{\Omega}^{\mathsf{T}}\bullet) \quad \text{(rotation-invariance)}$$

($\boldsymbol{\Omega}$ denotes an arbitrary rotation matrix in $\mathbb{R}^d$).

Using the above inverse operator, we shall now give a direct definition of the extended vector fBm with parameters H and $\boldsymbol{\xi}$:[3]

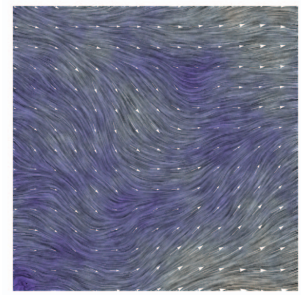
$$\boldsymbol{B}_{H,\boldsymbol{\xi}} := \epsilon_H(-\boldsymbol{\Delta})^{-\frac{H}{2}-\frac{d}{4}}_{-\boldsymbol{\xi}}\boldsymbol{W}. \tag{4}$$

The scale- and rotation-invariance of the operator $(-\boldsymbol{\Delta})^{-H/2-d/4}_{-\boldsymbol{\xi}}$ and the statistical invariance of the white noise field $\boldsymbol{W}$ with respect to changes of scale and rotations together mean that the vector fBm $\boldsymbol{B}_{H,\boldsymbol{\xi}}$ is also statistically invariant with respect to scalings and rotations of the system of coordinates. One may also note that, as a consequence of the factorization relation (3), $\boldsymbol{B}_{H,\boldsymbol{\xi}}$ becomes divergence-free (respectively curl-free) as $\xi_1 - \xi_2$ (resp. $\xi_2 - \xi_1$) approaches $+\infty$.

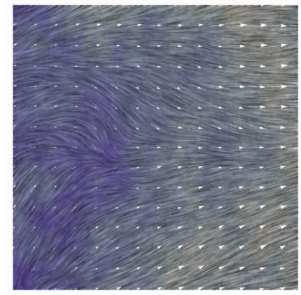
A few examples of computer-generated two-dimensional vector fBm are given in Figure 1.

[2] The complement, $\mathrm{Id} - \mathbf{E}$, is a projection onto the divergence-free component

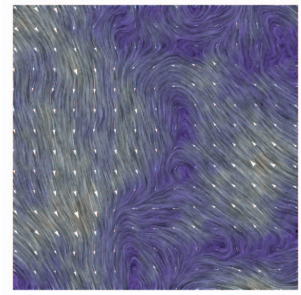
[3] The action of an operator on a random field finds a rigorous interpretation in the framework of the theory of generalized random processes of Gelfand and Vilenkin [15].



(a) $H = 0.60, \xi_1 = 0, \xi_2 = 0$



(b) $H = 0.60, \xi_1 = 0, \xi_2 = 100$



(c) $H = 0.60, \xi_1 = 100, \xi_2 = 0$

Fig. 1: Simulated vector fBm with $H = 0.6$ and variable ξ_1 and ξ_2, visualized using Mathematica's implementation of line integral convolution (LIC).

3. ESTIMATION OF VECTOR FBM PARAMETERS USING WAVELETS

It is a well-known fact that wavelet transform coefficients can be used to estimate the Hurst exponent of scalar fBm processes and fields [2, 10–14, 16]. This property of wavelets holds true also in the vector setting, in the following manner: A wavelet transform applied independently to each of the vectorial components of a fractional Brownian vector field can be used, in almost exactly the same fashion as in the scalar setting, to estimate H. We refer the reader to Tafti *& al.* [10] for the details of two such estimators.

As in the scalar case, in order for the wavelet to stationarize the random field, it has to incorporate a fractional Laplacian of sufficient order so as to cancel out the (non-shift-invariant) inverse operator in (4); which is to say that it is necessary that the mother wavelet can be written as $(-\boldsymbol{\Delta})^{\gamma}\Theta$, with $\gamma > \frac{H}{2} + \frac{d}{4}$, where Θ is a (matrix-valued) smoothing kernel. The component-wise analysis of a vector field described in the previous paragraph corresponds to the special case of Θ being a scalar matrix (i.e. a multiple of identity).

We shall now observe that such a diagonal wavelet transform is not sufficient for our purpose of estimating all parameters of a vector fBm. Indeed, while an independent component-wise analysis can provide estimates of H, such an analysis cannot differentiate between different choices of the directional parameters ξ_1 and ξ_2 in a statistically meaningful way, for the reason that it does not measure the interdependence of the vector components.

In order to estimate the directional characteristics of the field it is therefore appropriate to consider full matrix wavelets that capture the inter-component structure of the field. Such wavelets can be constructed by decomposing a scalar matrix wavelet $\Psi = (-\boldsymbol{\Delta})^{\gamma}\Theta$—where $\Theta = \theta\mathbf{I}$ is a scalar matrix function with diagonal θ—in the following fashion:

$$\begin{aligned}\Psi = (-\boldsymbol{\Delta})^{\gamma}\Theta &= (-\boldsymbol{\Delta})^{\gamma}\big[\mathbf{E} + (\mathrm{Id} - \mathbf{E})\big]\Theta \\ &= \underbrace{(-\boldsymbol{\Delta})^{\gamma}\mathbf{E}\Theta}_{\Psi_1} + \underbrace{(-\boldsymbol{\Delta})^{\gamma}(\mathrm{Id} - \mathbf{E})\Theta}_{\Psi_2}.\end{aligned}$$

Convolving Ψ_1 and Ψ_2 with $\boldsymbol{B}_{H,\boldsymbol{\xi}}$ then yields

$$\Psi_1 * \boldsymbol{B}_{H,\boldsymbol{\xi}} = \mathrm{e}^{-\xi_1}\big[(-\boldsymbol{\Delta})^{\gamma-\frac{H}{2}-\frac{d}{4}}\mathbf{E}\Theta\big] * \boldsymbol{W} =: \mathrm{e}^{-\xi_1}\boldsymbol{W}_1;$$

$$\Psi_2 * \boldsymbol{B}_{H,\boldsymbol{\xi}} = \mathrm{e}^{-\xi_2}\big[(-\boldsymbol{\Delta})^{\gamma-\frac{H}{2}-\frac{d}{4}}(\mathrm{Id} - \mathbf{E})\Theta\big] * \boldsymbol{W} =: \mathrm{e}^{-\xi_2}\boldsymbol{W}_2.$$

The random fields on the right-hand side of the above equations are stationary filtered-white-noise-type processes.

It is possible to construct a very simple estimator of the quantity $(\xi_1 - \xi_2)$ if one notes that the mathematical expectations $\mathbb{E}\{\|\boldsymbol{W}_1\|^2\}$ and $\mathbb{E}\{\|\boldsymbol{W}_2\|^2\}$ are constant functions of the spatial coordinates that depend solely on the choice of the smoothing kernel Θ.[4] The ratio

$$R := \frac{\mathbb{E}\{\|\boldsymbol{W}_1\|^2\}}{\mathbb{E}\{\|\boldsymbol{W}_2\|^2\}}$$

can therefore be pre-computed or, alternatively, estimated beforehand in a calibration step. Subsequently, for a realization of $\boldsymbol{B}_{H,\boldsymbol{\xi}}$ with unknown $\boldsymbol{\xi}$ we may estimate the ratio

$$\frac{\mathbb{E}\{\|\Psi_1 * \boldsymbol{B}_{H,\boldsymbol{\xi}}\|^2\}}{\mathbb{E}\{\|\Psi_2 * \boldsymbol{B}_{H,\boldsymbol{\xi}}\|^2\}} = \mathrm{e}^{2(\xi_2-\xi_1)}R$$

between the mean values of the energies of the two wavelet transforms; from where an estimate of $\xi_1 - \xi_2$ can be trivially obtained.

[4] Note that it is the difference of the parameters ξ_1 and ξ_2—and not their individual values as such—that determines the directional behaviour of the field.

Table 1: Estimation of the H parameter of synthesized vector fBm.

True value	average of local estimates	their variance
0.3	0.30	0.0074
0.6	0.58	0.0107
0.9	0.87	0.0140

4. EXPERIMENTS

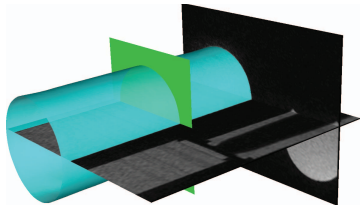
To verify the correctness of the estimation mechanism sketched in the previous section, $64 \times 64 \times 64$ volumes of discretized vector fBm were generated in MATLAB in accordance with the synthetic model of Section 2. Estimation of the Hurst exponent was performed over local neighbourhoods, using the first of the two estimation methods described in Tafti *& al.* [10]. To obtain a finer scale progression we replaced the discrete wavelet transform of Tafti *& al.* by a Laplacian of Gaussian continuous wavelet transform with the σ parameter of the Gaussian spanning the range 0.5 to 2 with steps of size 0.25. The results are summarized in Table 1.

Next, an estimation of $\xi_1 - \xi_2$ was performed using, as input, pseudo-random realizations of vector fBm with varying $\xi_1 - \xi_2$. The estimates were fairly accurate, with the correlation coefficient being virtually equal to 1 over the range -30 to 30 with step size 2.

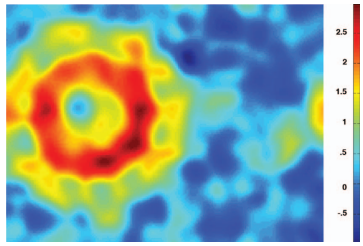
A similar analysis was applied to measured MRI data obtained from a phantom. We shall now briefly describe the set-up. The flow model—based on a rigid PVC tube with an inner diameter of 3.4cm—was connected to a clinical blood-pump system to produce a constant (non-pulsatile) fully developed flow. The fluid used was a solution of a Gadolinium chelate contrast agent in distilled water at 37°C. The flow model was imaged on a 3T MRI system using a 3D phase-contrast sequence [7]. The sequence relies on the difference of phase of spins moving along the direction of a magnetic field gradient to determine their velocities. The sequence, typically used for blood flow measurements in vivo [7, 8], allows the acquisition of three-directional velocities with a three-dimensional coverage. The MRI acquisition parameters were: voxel size [mm^3]: $0.4 \times 0.4 \times 0.6$, velocity encoding factor (venc) [cm/s]: 50, T E / T R [ms]: 4.62 / 8, bandwidth [Hz/pixel]: 440, flip angle (α) [degrees]: 13.

Phase-contrast MRI can assess flow velocities without restriction in anatomic coverage or direction but is limited by its relatively long imaging times, limited spatial and temporal resolutions, or limited signal to noise ratio (SNR). Errors limiting the SNR can be induced by intrinsic measurement errors, eddy currents, gradient field inhomogeneities, concomitant gradients, or acceleration errors [17].

Figure 2 schematically shows the imaging set-up, along with a colour-coded cross-section of local directional parameters $(\xi_1 - \xi_2)$; the cross-section was taken perpendicular to the direction of flow. (The significant part of the image is the circular disc on the left; the background—apart from some static structures—does not correspond to any flow and should be discarded.) A positive $\xi_1 - \xi_2$, as observed inside the tube (Figure 2b), indicates a divergence-free tendency, which is consistent with the incompressible nature of the fluid used in the experiment. Due to the non-turbulent (and essentially predictable) nature of this example, estimation of the Hurst exponent is not particularly meaningful in this case. Such an analysis would however be of great interest in studying flow fields of a more random and turbulent character. Further experiments and studies in this direction are still needed and will be the subject of our future research.



(a) Flow model (figure taken from Stalder [18])



(b) Local estimates of $\xi_1 - \xi_2$

Fig. 2: Differential analysis of a 3D flow (see text for details).

5. CONCLUSION

In this paper we proposed a model for stochastic fractal vector fields in the spirit of fractional Brownian motion models that was motivated by the wish to study biomedical flow-field measurements (in particular flow-sensitive MRI data). In addition to the usual Hurst exponent that quantify the fractality of the field, the vector models we introduced also have parameters to control the balance between the extremities of irrotational and solenoidal behaviour. Next, in order to study the relevance of these models in the analysis of measured data, we developed estimators of the different parameters of these models. We verified these estimators by applying them to synthesized vector fBm, and then used them to analyze 3D flow measurements obtained using phase-contrast MRI. The outcome of the analsyis was consistent with the known properties of the flow (i.e. incompressibility). Additional experiments will be directed at a better understanding of the significance of the estimated parameters and the study of flow fields with different structures.

6. REFERENCES

[1] B. B. Mandelbrot, *The Fractal Geometry of Nature*. New York: W. H. Freeman and Company, 1982.

[2] P. D. Tafti, D. Van De Ville, and M. Unser, "Innovation modelling and wavelet analysis of fractal processes in bio-imaging," in *Proc. 5th IEEE Int. Symp. Biomed. Imaging (ISBI'08)*, (Paris, France), pp. 1501–1504, 2008.

[3] B. B. Mandelbrot and J. W. Van Ness, "Fractional Brownian motions, fractional noises and applications," *SIAM Rev.*, vol. 10, no. 4, pp. 422–437, 1968.

[4] A. N. Kolmogoroff, "Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum," *C. R. (Doklady) Acad. Sci. URSS*, vol. 26, no. 2, pp. 115–118, 1940.

[5] J.-P. Kahane, *Some Random Series of Functions*. Cambridge University Press, 2nd ed., 1985.

[6] M. Farge, N. Kevlahan, V. Perrier, and E. Goirand, "Wavelets and turbulence," *Proc. IEEE*, vol. 84, no. 4, pp. 639–669, 1996.

[7] M. Markl, A. Harloff, T. A. Bley, M. Zaitsev, B. Jung, E. Weigang, M. Langer, J. Hennig, and A. Frydrychowicz, "Time-resolved 3D MR velocity mapping at 3T: Improved navigator-gated assessment of vascular anatomy and blood flow," *J. Magn. Reson. Imaging*, vol. 25, no. 4, pp. 824–831, 2007.

[8] A. F. Stalder, M. F. Russe, A. Frydrychowicz, J. Bock, J. Hennig, and M.Markl, "Quantitative 2d and 3d phase contrast MRI: Optimized analysis of blood flow and vessel wall parameters," *Magn. Reson. Med.*, vol. 60, no. 5, pp. 1218–1231, 2008.

[9] A. Benassi, S. Jaffard, and D. Roux, "Elliptic gaussian random processes," *Rev. Mat. Iberoamericana*, vol. 13, no. 1, pp. 19–90, 1997.

[10] P. D. Tafti, D. Van De Ville, and M. Unser, "Invariances, Laplacian-like wavelet bases, and the whitening of fractal processes," *IEEE T. Image. Process.*, vol. 18, no. 4, pp. 698–702, 2009.

[11] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE T. Pattern. Anal.*, vol. 11, no. 7, pp. 674–693, 1989.

[12] G. W. Wornell, "Wavelet-based representations for the $1/f$ family of fractal processes," *Proc. IEEE*, vol. 81, no. no. 10, pp. 1428–1450, 1993.

[13] D. Veitch and P. Abry, "A wavelet-based joint estimator of the parameters of long-range dependence," *IEEE T. Inform. Theory.*, vol. 45, no. 3, pp. 878–897, 1999.

[14] P. D. Tafti and M. Unser, "Self-similar random vector fields and their wavelet analysis," in *Proc. SPIE Conf. Math. Imaging: Wavelets XIII* (M. P. Vivek K. Goyal and D. V. D. Ville, eds.), vol. 7446, 74460Y, (San Diego CA, USA), pp. 74460Y–1–74460Y–8, 2009.

[15] I. M. Gel'fand and N. Vilenkin, *Applications of Harmonic Analysis*, vol. IV of *Generalized Functions*. Academic Press, 1964.

[16] R. Delgado-Gonzalo, P. D. Tafti, and M. Unser, "Fractional Laplacian pyramids," in *Proc. IEEE Int. Conf. Image Process. (ICIP 2009)*, 2009.

[17] E. Haacke, R. Brown, M. Thompson, and R. Venkatesan, *Magnetic Resonance Imaging Physics: Physical Principles and Sequence Design*. Wiley, 1999.

[18] A. F. Stalder, *Quantitative analysis of blood flow and vessel wall parameters using 4D-flow sensitive MRI*. PhD thesis, Universität Freiburg, Technische Fakultät (bisher: Fak. f. Angew. Wiss.), 2009.

The following paper originally appeared as

P. D. Tafti and M. Unser, *'Fractional Brownian vector fields',* SIAM Multiscale Model. Simul., vol. 8, no. 5, pp. 1645–70, 2010.

Multiscale Model. Simul.
Vol. 8, No. 5, pp. 1645–1670

FRACTIONAL BROWNIAN VECTOR FIELDS*

POUYA DEHGHANI TAFTI† AND MICHAEL UNSER†

Abstract. This work puts forward an extended definition of vector fractional Brownian motion (fBm) using a distribution theoretic formulation in the spirit of Gel'fand and Vilenkin's stochastic analysis. We introduce random vector fields that share the statistical invariances of standard vector fBm (self-similarity and rotation invariance) but which, in contrast, have dependent vector components in the general case. These random vector fields result from the transformation of white noise by a special operator whose invariance properties the random field inherits. The said operator combines an inverse fractional Laplacian with a Helmholtz-like decomposition and weighted recombination. Classical fBm's can be obtained by balancing the weights of the Helmholtz components. The introduced random fields exhibit several important properties that are discussed in this paper. In addition, the proposed scheme yields a natural extension of the definition to Hurst exponents greater than one.

Key words. fractional Brownian motion, random vector fields, self-similarity, invariance, Helmholtz decomposition, generalized random processes, Gel'fand–Vilenkin stochastic analysis

AMS subject classifications. 60G18, 60G20, 60G60, 60H40, 60H20, 35S30, 42B20

DOI. 10.1137/090752638

1. Introduction.

A one-dimensional fractional Brownian motion (fBm) $B_H(x)$, $x \in \mathbb{R}$, is a nonstationary zero-mean Gaussian random process satisfying $B_H(0) = 0$, with the characteristic property that for any fixed step size $|x - y|$ the increment $B_H(x) - B_H(y)$ is a stationary Gaussian process with a variance proportional to the step size:[1]

$$\mathbf{E}\{|B_H(x) - B_H(y)|^2\} = 2\alpha|x - y|^{2H}.$$

The parameter $H \in (0, 1)$ is known as the Hurst exponent, after Harold Edwin Hurst, a pioneer in the study of long-range statistical dependence [21, 53] (α is an arbitrary positive constant).

The above definition can be extended to the multivariate setting in the style of Lévy's characterization of multiparameter Brownian motion [29, 30], by making the parameter $\boldsymbol{x}$ a vector in $\mathbb{R}^d$ and defining multidimensional fBm as a Gaussian random field with a variogram of the form

$$\mathbf{E}\{|B_H(\boldsymbol{x}) - B_H(\boldsymbol{y})|^2\} = 2\alpha\|\boldsymbol{x} - \boldsymbol{y}\|^{2H}$$

(we shall use bold symbols to denote vector quantities).

FBm's are important examples of stochastic fractals: They are statistically self-similar in the sense that an fBm $B_H(\cdot)$ and its scaled version $\sigma^H B_H(\sigma\cdot)$ have the same statistics. FBm processes have been used to model natural and man-made phenomena in different areas of application including optics, fluid mechanics, seismology,

*Received by the editors March 13, 2009; accepted for publication (in revised form) June 10, 2010; published electronically September 16, 2010. This work was supported by Swiss National Science Foundation grant 200020-121763.
http://www.siam.org/journals/mms/8-5/75263.html
†Ecole Polytechnique Fédérale de Lausanne (EPFL), Biomedical Imaging Group, EPFL-STI-IMT-LIB, Bâtiment BM, Station 17, Lausanne VD CH-1015, Switzerland (pouya.tafti@a3.epfl.ch, michael.unser@epfl.ch).

[1]The second moment of the increment of a process, defined in the above fashion, is also known as its *variogram* or structure function.

financial mathematics, network traffic analysis, and image processing, among others [13, 14, 15, 24, 27, 33, 39, 40, 48, 54]. Since the notion of invariance that inspires the definition of fractals is also fundamental in physics, it is natural to expect that the scope of applications of such models will further expand with time. For the same reason, investigation of physical invariances and stochastic models characterized by them seems worthwhile.

Our goal in the present paper is to extend the definition of fBm to the vector (multicomponent) field setting. In this undertaking we draw our inspiration from two sources. One is the usual consideration of self-similarity. The second influence comes from physics, as we shall impose on the model a special form of rotation invariance that is compatible with the effect of rotations on coordinates of physical vector fields.

We remark that a trivial vectorial extension of scalar fBm satisfying the conditions of homogeneity and vector rotation invariance can be readily constructed by taking the components of the d-dimensional vector to be independent scalar fBm's. This extension is consistent with the variogram relation [22]

$$\mathbf{E}\{\|\boldsymbol{B}_H(\boldsymbol{x}) - \boldsymbol{B}_H(\boldsymbol{y})\|^2\} = 2\alpha'\|\boldsymbol{x} - \boldsymbol{y}\|^{2H} \tag{1.1}$$

(note that the absolute value has been replaced by a Euclidean norm in the argument of the expectation operator). But it should be emphasized that (1.1) in itself does not specify the cross-correlation structure of the components of $\boldsymbol{B}_H$, and the classical assumption of independent components is not exhaustive. Hence, in this paper we shall consider more general families of fractal vector fields satisfying (1.1) whose vector components can be correlated in ways that lead to a full range of vectorial comportment from fully solenoidal to completely irrotational.

This paper continues the line of reasoning adopted in Tafti, Van De Ville, and Unser [51] (where we considered scalar fBm fields) and more originally in Blu and Unser [4] (where one-dimensional fBm processes were studied). In keeping with these previous works, we shall characterize fBm vector fields as particular solutions of a stochastic fractional differential equation

$$\mathrm{U}\boldsymbol{B}_H = \boldsymbol{W} \tag{1.2}$$

subject to zero boundary conditions at $\boldsymbol{x} = \boldsymbol{0}$, where $\boldsymbol{W}$ denotes a vector of normalized and independent white noise fields (defined in subsection 3.1). The "whitening" operator U is chosen based on its specific invariance properties that carry over to the random vector field $\boldsymbol{B}_H$. U will turn out to be a generalization of the fractional vector Laplacian $(-\boldsymbol{\Delta})^{\gamma}$, with additional parameters that control the solenoidal versus irrotational tendencies of the solution. Rigorous interpretation and inversion of (1.2) are conducted in the framework of Gel'fand and Vilenkin's theory of generalized random processes and distributional stochastic analysis [18]. Some aspects of this theory that are relevant to our work are summarized in subsection 3.1.

Our characterization by means of a whitening equation gives mathematical meaning to inverse power-law spectra that are traditionally associated with self-similar processes, by providing the mechanics for resolving the singularity of the said spectra at $\boldsymbol{\omega} = \boldsymbol{0}$ (the noted processes, being nonstationary, do not have power spectra in the classical sense). We should, however, note that in a different approach to the mathematical modeling and simulation of self-similar physical phenomena, the introduction of a cut-off length can provide an alternative way of dealing with the frequency-domain singularity.

Previous related work in the direction of the present paper has appeared, for instance, in Yaglom [62], where a second-order analysis of random vector fields with similar invariance properties was given. Yaglom and others also considered, albeit separately, scalar random processes with stationary nth-order increments [11, 42, 63]. In addition to differences in formulation and approach—for example, in our consideration of singular operators and our focus on characteristic functionals—in the present paper we bring together these separate generalizations (cf. subsection 4.4). Furthermore, our approach is not limited, in essence, to the study of second-order statistics (even though this would have sufficed for the Gaussian fields considered here). This means that by using a similar approach it is possible, without too much difficulty, to construct and completely characterize other—non-Gaussian—models satisfying similar invariance properties, by driving (1.2) with different types of non-Gaussian white noise.

On the applied side, consideration of models in line with (1.1) and their relatives has a long history in fluid dynamics and specifically in the study of turbulence, although the emphasis and methodology are frequently different from ours (see, e.g., Monin and Yaglom [37, Chapter 8], Avellaneda and Majda [2], Carmona [7], Orszag [38], or Klyatskin, Woyczynski, and Gurarie [25]).

In the remainder of this paper we first turn our attention to the search for an operator U satisfying the required invariances (section 2). There, the question of inverting U—which is necessary for solving (1.2)—requires us to consider a particular regularization of singular Fourier integrals. Next, in section 3, we solve (1.2) and give a complete stochastic characterization of generalized vector fBm fields as particular solutions of this equation. A list of the main properties of these random fields is given in section 4. This is followed by computer simulations (section 5) and conclusions (section 6). Proofs of some intermediate results have been deferred until the appendices.

2. Vector operators invariant under rotation and scaling.

2.1. Generalized fractional Laplacians. Let $\boldsymbol{f}(\boldsymbol{u})$, $\boldsymbol{u} \in \mathbb{R}^d$, represent a vector field in terms of the standard coordinates $\boldsymbol{u} = (u_1, \dots, u_d)$. Consider a second coordinate system $\boldsymbol{x}$ related to $\boldsymbol{u}$ by means of a smooth invertible map $\boldsymbol{\phi} : \mathbb{R}^d \to \mathbb{R}^d$ as per

$$\boldsymbol{x} = \boldsymbol{\phi}(\boldsymbol{u}).$$

The coordinates of $\boldsymbol{f}$ in the second system are then given by the formula

$$\boldsymbol{f}_{\boldsymbol{\phi}}(\boldsymbol{x}) = \frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{u}}(\boldsymbol{u}) \boldsymbol{f}(\boldsymbol{u})$$

(this can be seen as a consequence of identifying vector fields with differential one-forms and applying the chain rule of differentiation; cf. Rudin [44, paragraphs 10.21, 10.42]).

In particular, for a linear coordinate transformation $\boldsymbol{x} = \mathbf{M}\boldsymbol{u}$, where $\mathbf{M}$ is an invertible $d \times d$ matrix, one has

$$\boldsymbol{f}_{\mathbf{M}}(\boldsymbol{x}) = \mathbf{M} \boldsymbol{f}(\mathbf{M}^{-1}\boldsymbol{x}).$$

It follows that if $\mathbf{M} = \boldsymbol{\Omega}$ is orthogonal (in particular, a rotation matrix), then

$$\boldsymbol{f}_{\boldsymbol{\Omega}}(\boldsymbol{x}) = \boldsymbol{\Omega} \boldsymbol{f}(\boldsymbol{\Omega}^{\mathsf{T}}\boldsymbol{x});$$

and if $\mathbf{M} = \sigma\mathbf{I}$ with $\sigma > 0$ (a scaling), then

$$\boldsymbol{f}_\sigma(\boldsymbol{x}) = \sigma \boldsymbol{f}(\sigma^{-1}\boldsymbol{x}).$$

We shall consider certain *convolution operators* acting on vector fields, as well as their inverses. By the former we mean those operators which can be written in terms of an inverse Fourier integral as per

$$\mathrm{U} : \boldsymbol{f} \mapsto (2\pi)^{-d} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{j}\langle \boldsymbol{x}, \boldsymbol{\omega}\rangle} \hat{\mathrm{U}}(\boldsymbol{\omega}) \hat{f}(\boldsymbol{\omega}) \, \mathrm{d}\boldsymbol{\omega}, \tag{2.1}$$

where $\hat{\mathrm{U}}$ is the (matrix-valued) Fourier expression for the operator U and $\hat{\boldsymbol{f}}$ is the Fourier transform of $\boldsymbol{f}$.

Operators of the above type appear in equations of the form

$$\mathrm{U}\boldsymbol{B}_H = \boldsymbol{W}, \tag{2.2}$$

which we shall use to model *statistically self-similar* (homogeneous) and *rotation-invariant* (isotropic) vector fields. These properties are imposed on the solution $\boldsymbol{B}_H$ of the above equation by requiring that the right inverse of U interact in a particular way with rotations and scalings of the coordinate system.[2]

The "invariance" properties the operator U is required to satisfy are the following:

$$\mathrm{U}\boldsymbol{f}_{\boldsymbol{\Omega}} = (\mathrm{U}\boldsymbol{f})_{\boldsymbol{\Omega}} \qquad \text{(rotation invariance)}; \tag{2.3}$$

$$\mathrm{U}\boldsymbol{f}_\sigma = \sigma^{2\gamma}(\mathrm{U}\boldsymbol{f})_\sigma \qquad \text{(degree } 2\gamma \text{ homogeneity)} \tag{2.4}$$

(γ relates to one of the main parameters of the family of the random solutions, namely the Hurst exponent, by the relation $H = 2\gamma - d/2$). Note that we shall assume invariance with respect to improper rotations (with $\det \boldsymbol{\Omega} = -1$) as well as proper rotations (with $\det \boldsymbol{\Omega} = 1$).

The above properties translate, respectively, to the following conditions on the Fourier expression of the operator U:

$$\hat{\mathrm{U}}(\boldsymbol{\Omega}\boldsymbol{\omega}) = \boldsymbol{\Omega}\hat{\mathrm{U}}(\boldsymbol{\omega})\boldsymbol{\Omega}^{\mathsf{T}} \qquad \text{(rotation invariance)}; \tag{2.5}$$

$$\hat{\mathrm{U}}(\sigma\boldsymbol{\omega}) = \sigma^{2\gamma}\hat{\mathrm{U}}(\boldsymbol{\omega}) \qquad \text{(homogeneity)}. \tag{2.6}$$

The following theorem was proved by Arigovindan for $d = 2, 3$ [1]. It can be shown more generally to hold in any number of dimensions.

THEOREM 2.1 (Arigovindan [1]). *A vector convolution operator satisfying properties* (2.3) *and* (2.4) *has a Fourier expression of the form*

$$\hat{\boldsymbol{\Phi}}^{\gamma}_{\boldsymbol{\xi}}(\boldsymbol{\omega}) = \|\boldsymbol{\omega}\|^{2\gamma} \left[\mathrm{e}^{\xi_{\mathrm{irr}}} \frac{\boldsymbol{\omega}\boldsymbol{\omega}^{\mathsf{T}}}{\|\boldsymbol{\omega}\|^2} + \mathrm{e}^{\xi_{\mathrm{sol}}} \left(\mathbf{I} - \frac{\boldsymbol{\omega}\boldsymbol{\omega}^{\mathsf{T}}}{\|\boldsymbol{\omega}\|^2} \right) \right], \tag{2.7}$$

with $\boldsymbol{\xi} = (\xi_{\mathrm{irr}}, \xi_{\mathrm{sol}}) \in \mathbb{C}^2$.

It is easy to verify that the converse of the theorem is also true for arbitrary dimension d. Since we shall be considering real operators, in what follows we shall implicitly assume $\mathrm{e}^{\xi_{\mathrm{irr}}}, \mathrm{e}^{\xi_{\mathrm{sol}}} \in \mathbb{R}$ without further mention.

[2] We shall have to consider a right inverse of U that—unlike U—is not shift invariant and does not correspond to a convolution; hence the solution $\boldsymbol{B}_H$ will not be stationary (more on this later).

The operators introduced in Theorem 2.1 generalize vector Laplacians in two senses (fractional orders and reweighting of solenoidal and irrotational components). We shall therefore refer to them as fractional (vector) Laplacians and use the symbol $(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}$ to denote the operator with Fourier expression $\hat{\boldsymbol{\Phi}}^{\gamma}_{\boldsymbol{\xi}}$. To gain a better understanding of the action of fractional vector Laplacians, it is instructive at this point to recall the Fourier expressions (in standard Cartesian coordinates) of some related vector differential operators:[3]

$$
\begin{array}{rcl}
\mathbf{grad} & \overset{\mathcal{F}}{\longleftrightarrow} & \mathrm{j}\boldsymbol{\omega}; \\
\mathrm{div} & \overset{\mathcal{F}}{\longleftrightarrow} & (\mathrm{j}\boldsymbol{\omega})^{\mathsf{T}}; \\
\mathbf{curl} & \overset{\mathcal{F}}{\longleftrightarrow} & \begin{bmatrix} 0 & -\mathrm{j}\omega_3 & \mathrm{j}\omega_2 \\ \mathrm{j}\omega_3 & 0 & -\mathrm{j}\omega_1 \\ -\mathrm{j}\omega_2 & \mathrm{j}\omega_1 & 0 \end{bmatrix}; \\
\mathbf{grad}\,\mathrm{div} & \overset{\mathcal{F}}{\longleftrightarrow} & -\boldsymbol{\omega}\boldsymbol{\omega}^{\mathsf{T}}; \\
\mathbf{curl}\,\mathbf{curl} & \overset{\mathcal{F}}{\longleftrightarrow} & \|\boldsymbol{\omega}\|^2\mathbf{I} - \boldsymbol{\omega}\boldsymbol{\omega}^{\mathsf{T}}; \\
\boldsymbol{\Delta} & \overset{\mathcal{F}}{\longleftrightarrow} & -\|\boldsymbol{\omega}\|^2\mathbf{I}; \\
\mathrm{E} & \overset{\mathcal{F}}{\longleftrightarrow} & \frac{\boldsymbol{\omega}\boldsymbol{\omega}^{\mathsf{T}}}{\|\boldsymbol{\omega}\|^2}; \\
(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}} & \overset{\mathcal{F}}{\longleftrightarrow} & \hat{\boldsymbol{\Phi}}^{\gamma}_{\boldsymbol{\xi}}(\boldsymbol{\omega}) = \|\boldsymbol{\omega}\|^{2\gamma}\left[\mathrm{e}^{\xi_{\mathrm{irr}}}\frac{\boldsymbol{\omega}\boldsymbol{\omega}^{\mathsf{T}}}{\|\boldsymbol{\omega}\|^2} + \mathrm{e}^{\xi_{\mathrm{sol}}}\left(\mathbf{I} - \frac{\boldsymbol{\omega}\boldsymbol{\omega}^{\mathsf{T}}}{\|\boldsymbol{\omega}\|^2}\right)\right].
\end{array}
\tag{2.8}
$$

The penultimate operator (E) and its complement (Id − E) project a vector field onto its curl-free and divergence-free components, respectively. In other words, together they afford a Helmholtz decomposition of the vector field on which they act (these operators appear prominently in fluid dynamics literature [8, 9, 10, 46]). This is because

$$\mathrm{div}\,(\mathrm{Id} - \mathrm{E}) = 0 \quad \text{and} \quad \mathbf{curl}\ \mathrm{E} = 0.$$

In addition, one has

$$\mathrm{E}\ \mathbf{grad} = \mathbf{grad} \quad \text{and} \quad \mathrm{E}\ \mathbf{curl} = 0.$$

(Id − E) is known as the Leray projector in turbulence literature.

Our notation for the fractional vector Laplacian $(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}$ is motivated by the observation that it can be factorized as

$$(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}} = (-\boldsymbol{\Delta})^{\gamma}_{\mathbf{0}}\left[\mathrm{e}^{\xi_{\mathrm{irr}}}\mathrm{E} + \mathrm{e}^{\xi_{\mathrm{sol}}}(\mathrm{Id} - \mathrm{E})\right].$$

In view of the properties of the operator E, this factorization means that the operator $(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}$ combines a coordinatewise fractional Laplacian with a reweighting of the curl- and divergence-free components of the operand.

2.2. Some properties of $\hat{\boldsymbol{\Phi}}^{\gamma}_{\boldsymbol{\xi}}$. Let us now take a closer look at the family of matrix-valued functions $\hat{\boldsymbol{\Phi}}^{\gamma}_{\boldsymbol{\xi}}$, γ, $\mathrm{e}^{\xi_{\mathrm{irr}}}, \mathrm{e}^{\xi_{\mathrm{sol}}} \in \mathbb{R}$. They, of course, satisfy the required invariances:

$$\hat{\boldsymbol{\Phi}}^{\gamma}_{\boldsymbol{\xi}}(\boldsymbol{\Omega}\boldsymbol{\omega}) = \boldsymbol{\Omega}\hat{\boldsymbol{\Phi}}^{\gamma}_{\boldsymbol{\xi}}(\boldsymbol{\omega})\boldsymbol{\Omega}^{\mathsf{T}};$$
$$\hat{\boldsymbol{\Phi}}^{\gamma}_{\boldsymbol{\xi}}(\sigma\boldsymbol{\omega}) = \sigma^{2\gamma}\hat{\boldsymbol{\Phi}}^{\gamma}_{\boldsymbol{\xi}}(\boldsymbol{\omega}).$$

[3]Note that, while the curl operator is classically defined in three dimensions, the equivalents of $\mathbf{grad}\,\mathrm{div}$, $\mathbf{curl}\,\mathbf{curl}$, and $\boldsymbol{\Delta} = \mathbf{grad}\,\mathrm{div} - \mathbf{curl}\,\mathbf{curl}$ can be defined in any number of dimensions, for instance by their Fourier symbols. In fact, for arbitrary d, the equivalents of $-\mathbf{curl}\,\mathbf{curl}$ and $-\mathbf{grad}\,\mathrm{div}$ that appear in the definition of the vector Laplacian correspond, respectively, to the product of d-dimensional curl and divergence with their adjoints.

But they exhibit, in addition, the following properties.

($\hat{\mathbf{\Phi}}1$) *Closedness under multiplication.* We have

$$\hat{\mathbf{\Phi}}_{\boldsymbol{\xi}_1}^{\gamma_1}(\boldsymbol{\omega})\hat{\mathbf{\Phi}}_{\boldsymbol{\xi}_2}^{\gamma_2}(\boldsymbol{\omega}) = \hat{\mathbf{\Phi}}_{\boldsymbol{\xi}_1+\boldsymbol{\xi}_2}^{\gamma_1+\gamma_2}(\boldsymbol{\omega}),$$

which belongs to the same family. To see this, note that the matrix

$$\hat{\mathrm{E}}(\boldsymbol{\omega}) := \frac{\boldsymbol{\omega}\boldsymbol{\omega}^{\mathsf{T}}}{\|\boldsymbol{\omega}\|^2}$$

which appears in the definition of $\hat{\mathbf{\Phi}}_{\boldsymbol{\xi}}^{\gamma}$ is a projection, and therefore

$$\hat{\mathrm{E}}^2 \equiv \hat{\mathrm{E}} \quad \text{and} \quad \hat{\mathrm{E}}(\mathbf{I} - \hat{\mathrm{E}}) \equiv \mathbf{0}.$$

($\hat{\mathbf{\Phi}}2$) *Closedness under matrix inversion.* The inverse of the matrix $\hat{\mathbf{\Phi}}_{\boldsymbol{\xi}}^{\gamma}(\boldsymbol{\omega})$, for $\boldsymbol{\omega} \neq \mathbf{0}$, is equal to $\hat{\mathbf{\Phi}}_{-\boldsymbol{\xi}}^{-\gamma}(\boldsymbol{\omega})$, which is again in the same family. This follows from the previous property and the observation that $\hat{\mathbf{\Phi}}_{\mathbf{0}}^{0}(\boldsymbol{\omega})$ is the identity matrix for $\boldsymbol{\omega} \neq \mathbf{0}$.

($\hat{\mathbf{\Phi}}3$) *Closedness under Fourier transformation.* The family is closed under elementwise Fourier transforms in the particular fashion indicated by the following lemma.

LEMMA 2.2. *Let*

$$\hat{\mathbf{\Psi}}_{\boldsymbol{\xi}}^{\gamma} := \frac{2^{-\gamma}}{\Gamma(\gamma + \frac{d}{2})}\hat{\mathbf{\Phi}}_{\boldsymbol{\xi}}^{\gamma},$$

where Γ *denotes the Gamma function. The elementwise inverse Fourier transform of* $\hat{\mathbf{\Psi}}_{\boldsymbol{\xi}}^{\gamma}$ *(in the sense of generalized functions* [17]*) is given by the formula*

$$\mathcal{F}^{-1}\{\hat{\mathbf{\Psi}}_{\boldsymbol{\xi}}^{\gamma}\} = (2\pi)^{-\frac{d}{2}}\hat{\mathbf{\Psi}}_{\boldsymbol{\zeta}}^{-\gamma-\frac{d}{2}},$$

where $\boldsymbol{\zeta} = (\zeta_{\mathrm{irr}}, \zeta_{\mathrm{sol}})$ *is related to* $\boldsymbol{\xi} = (\xi_{\mathrm{irr}}, \xi_{\mathrm{sol}})$ *by*

$$\mathrm{e}^{\zeta_{\mathrm{irr}}} = \tfrac{2\gamma+d-1}{2\gamma}\mathrm{e}^{\xi_{\mathrm{irr}}} - \tfrac{d-1}{2\gamma}\mathrm{e}^{\xi_{\mathrm{sol}}} \quad \textit{and} \quad \mathrm{e}^{\zeta_{\mathrm{sol}}} = -\tfrac{1}{2\gamma}\mathrm{e}^{\xi_{\mathrm{irr}}} + \tfrac{2\gamma+1}{2\gamma}\mathrm{e}^{\xi_{\mathrm{sol}}}.$$

A proof can be found in Appendix A.

In particular, observe that if $\mathrm{e}^{\xi_{\mathrm{irr}}} = \mathrm{e}^{\xi_{\mathrm{sol}}}$, then $\mathrm{e}^{\zeta_{\mathrm{irr}}} = \mathrm{e}^{\zeta_{\mathrm{sol}}} = \mathrm{e}^{\xi_{\mathrm{irr}}} = \mathrm{e}^{\xi_{\mathrm{sol}}}$.

2.3. Inverse fractional Laplacians. The purpose of inverting the fractional Laplacian operator introduced in the previous subsection is to allow us to solve an equation of the form

$$(-\boldsymbol{\Delta})_{\boldsymbol{\xi}}^{\gamma}\boldsymbol{g} = \boldsymbol{h}. \tag{2.9}$$

This equation is understood in the sense of the identity

$$\langle(-\boldsymbol{\Delta})_{\boldsymbol{\xi}}^{\gamma}\boldsymbol{g}, \boldsymbol{f}\rangle = \langle\boldsymbol{h}, \boldsymbol{f}\rangle,$$

which must hold for all test functions $\boldsymbol{f}$ in some appropriate space.[4] In other words, the sides of the former equation are viewed as generalized functions belonging to the dual of the space of test functions $\boldsymbol{f}$.

[4]More precisely, the action of $(-\boldsymbol{\Delta})_{\boldsymbol{\xi}}^{\gamma}$ on $\boldsymbol{g}$ itself is defined by the duality relation

$$\langle(-\boldsymbol{\Delta})_{\boldsymbol{\xi}}^{\gamma}\boldsymbol{g}, \boldsymbol{f}\rangle = \langle\boldsymbol{g}, (-\boldsymbol{\Delta})_{\boldsymbol{\xi}}^{\gamma^*}\boldsymbol{f}\rangle,$$

where $(-\boldsymbol{\Delta})_{\boldsymbol{\xi}}^{\gamma^*}$ is the adjoint of the fractional Laplacian. With some abuse of notation, we shall denote $(-\boldsymbol{\Delta})_{\boldsymbol{\xi}}^{\gamma^*}$ by $(-\boldsymbol{\Delta})_{\boldsymbol{\xi}}^{\gamma}$, as the two operators share the same Fourier expression (they are, however, defined on different spaces). Also note that $\boldsymbol{g}$ and $\boldsymbol{h}$ need not belong to the same space, which in turn means that the test functions applied to them may come from different function spaces.

The solution to (2.9) is sought in some generalized function space (e.g., a subspace of $(\mathcal{S}')^d$). Solving (2.9) for general $\boldsymbol{h}$ is made possible by finding a *right inverse* of $(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}$ which injects $\boldsymbol{h}$ into the said space of solutions. Let us denote the desired right inverse by $(-\acute{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}}$. We shall define its action on $\boldsymbol{h}$ by the action of its adjoint $(-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\bar{\boldsymbol{\xi}}}$ on test functions:

$$\boldsymbol{g} = (-\acute{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}}\boldsymbol{h} \quad \Leftrightarrow \quad \langle \boldsymbol{g}, \boldsymbol{f}\rangle = \langle \boldsymbol{h}, (-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\bar{\boldsymbol{\xi}}}\boldsymbol{f}\rangle. \tag{2.10}$$

The adjoint operator is a *left inverse* of the dual Laplacian over the test function space since

$$\langle \boldsymbol{h}, (-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\bar{\boldsymbol{\xi}}}(-\boldsymbol{\Delta})^{\gamma}_{\bar{\boldsymbol{\xi}}}\boldsymbol{f}\rangle = \langle (-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}(-\acute{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}}\boldsymbol{h}, \boldsymbol{f}\rangle = \langle \boldsymbol{h}, \boldsymbol{f}\rangle.$$

We hinted previously that in order to use (2.2) with $\mathrm{U} = (-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}$, $\gamma > 0$, to define self-similar and isotropic random fields we would be seeking a particular right inverse of $(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}$ that retains its properties of homogeneity and rotation invariance. Equivalently, the adjoint (i.e., the left inverse) must be homogeneous and rotation invariant. Furthermore, it will be found necessary for our characterization that the range of the left inverse be a subspace of $(\mathrm{L}^2)^d$ (cf. subsection 3.1).

In connection with the fractional Laplacian we make the following observation. Let us first consider test functions belonging to the subspace $\mathcal{S}^d_0$ of $\mathcal{S}^d$ consisting of Schwartz functions with vanishing moments (i.e., zero derivatives of all orders at the origin of the Fourier space). $(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}$ is a bijection on this space, and hence also on its dual $(\mathcal{S}^d_0)'$, which can be identified with the quotient space of tempered distributions modulo polynomials, denoted by $(\mathcal{S}^d)'/\Pi$. The left and right inverses of $(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}$ therefore coincide on $\mathcal{S}^d_0$ and on $(\mathcal{S}^d_0)'$. On either space, they are both given by the integral

$$(2\pi)^{-d}\int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{j}\langle \boldsymbol{x},\boldsymbol{\omega}\rangle}\hat{\boldsymbol{\Phi}}^{-\gamma}_{-\boldsymbol{\xi}}(\boldsymbol{\omega})\boldsymbol{f}(\boldsymbol{\omega})\,\mathrm{d}\boldsymbol{\omega}. \tag{2.11}$$

However, from the identification of $(\mathcal{S}^d_0)'$ with $(\mathcal{S}^d)'/\Pi$ one sees that the extension of the right inverse to $(\mathcal{S}^d)'$ is not unique. This is precisely because $(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}$ has a nontrivial null space in $(\mathcal{S}^d)'$ due to the zero of its symbol at $\boldsymbol{\omega} = \boldsymbol{0}$. Correspondingly, the action of the left inverse on an arbitrary test function in $\mathcal{S}^d$ is not a priori well-defined, as its Fourier expression $(\hat{\boldsymbol{\Phi}}^{-\gamma}_{-\bar{\boldsymbol{\xi}}})$ is singular at $\boldsymbol{\omega} = \boldsymbol{0}$.

The problem of finding inverse operators that satisfy the desired properties (invariances and L^2-boundedness of the left inverse) can therefore be reformulated as that of choosing a particular *regularization* of the singular Fourier integral of (2.11) consistent with the said requirements. This will be the subject of the remainder of this subsection.

By a regularization of the singular Fourier integral operator of (2.11) we mean the following. Assume $\boldsymbol{f} \in \mathcal{S}^d_0$ to be a function with vanishing moments; $\boldsymbol{f}$ then satisfies $\partial_{\boldsymbol{k}}\hat{\boldsymbol{f}}(\boldsymbol{0}) = \boldsymbol{0}$ for all nonnegative multi-integers $\boldsymbol{k}$. As was already noted, the above Fourier integral converges for such $\boldsymbol{f}$. Consequently, the restriction of $(-\boldsymbol{\Delta})^{-\gamma}_{-\boldsymbol{\xi}}$ to this subspace of $\mathcal{S}^d$ is well-defined and inverts $(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}$ (and, by duality, the adjoint inverse can be applied to the dual of the image of this subspace). A *regularization* of $(-\boldsymbol{\Delta})^{-\gamma}_{-\boldsymbol{\xi}}$ is an extension of it to a larger class of functions, in our case $\mathcal{S}^d$.

There exists a *canonical* regularization of the above singular integral, which is homogeneous and rotation and shift invariant. It can be shown that the canonical regularization, which is the one considered by Gel'fand and Shilov [17] and Hörmander [20], corresponds to a convolution with a homogeneous generalized function. Unfortunately, this regularization fails the third of our requirements, namely L^2-boundedness. We shall therefore have to consider a different regularization of (2.11).

In what follows, we shall always limit our consideration to values of γ such that

$$2\gamma - \tfrac{d}{2} \notin \mathbb{Z}. \tag{2.12}$$

It will be seen later that this condition is equivalent to requiring that the Hurst exponent $H \notin \mathbb{Z}$ in the definition of fBm (see the discussion following Theorem 3.2).

To extend the definition of the (left) inverse from functions with vanishing moments to arbitrary test functions $\boldsymbol{f} \in \mathcal{S}^d$ let us introduce the *regularization operator*

$$\mathrm{R}^\gamma : \boldsymbol{f}(\cdot) \mapsto \boldsymbol{f}(\cdot) - \sum_{|\boldsymbol{k}| \leq \lfloor 2\gamma - \frac{d}{2} \rfloor} \mathbf{T}_{\boldsymbol{k}}[\boldsymbol{f}](\cdot)^{\boldsymbol{k}}, \tag{2.13}$$

where $\mathbf{T}_{\boldsymbol{k}}[\boldsymbol{f}]$ denotes the (vector) coefficient of $(\cdot)^{\boldsymbol{k}}$ in the Taylor series expansion of $\boldsymbol{f}(\cdot)$ around $\mathbf{0}$ (we use multiindex notation). Next, consider the operator

$$(-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}} : \boldsymbol{f} \mapsto (2\pi)^{-d} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{j}\langle \boldsymbol{x}, \boldsymbol{\omega} \rangle} \hat{\boldsymbol{\Phi}}^{-\gamma}_{-\boldsymbol{\xi}}(\boldsymbol{\omega}) [\mathrm{R}^\gamma \hat{\boldsymbol{f}}](\boldsymbol{\omega}) \, \mathrm{d}\boldsymbol{\omega} \tag{2.14}$$

(defined in the sense of the L^2 Fourier transform). This operator essentially removes sufficiently many terms from the Taylor expansion of $\hat{\boldsymbol{f}}(\boldsymbol{\omega})$ at $\boldsymbol{\omega} = \mathbf{0}$ so as to make the singularity of $\hat{\boldsymbol{\Phi}}^{-\gamma}_{\boldsymbol{\xi}}(\boldsymbol{\omega})$ square integrable. Of key importance is the fact that $(-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}}$ maps Schwartz test functions in $\mathcal{S}^d$ to square-integrable functions (assuming, as we already stated, that $2\gamma - \frac{d}{2} \notin \mathbb{Z}$).

PROPOSITION 2.3. *The operator* $(-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}}$ *maps* $\mathcal{S}^d$ *into* $(\mathrm{L}^2)^d$ *on the condition that* $2\gamma - d/2 \notin \mathbb{Z}$.

Proof. By Parseval's identity,

$$\begin{aligned}
\|(-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}} \boldsymbol{f}\|^2 &= (2\pi)^{-d} \|\hat{\boldsymbol{\Phi}}^{-\gamma}_{-\boldsymbol{\xi}} [\mathrm{R}^\gamma \hat{\boldsymbol{f}}](\boldsymbol{\omega})\|^2 \\
&= (2\pi)^{-d} \int_{\mathbb{R}^d} [\mathrm{R}^\gamma \hat{\boldsymbol{f}}]^{\mathsf{H}}(\boldsymbol{\omega}) \hat{\boldsymbol{\Phi}}^{-2\gamma}_{-2\mathrm{Re}\, \boldsymbol{\xi}}(\boldsymbol{\omega}) [\mathrm{R}^\gamma \hat{\boldsymbol{f}}](\boldsymbol{\omega}) \, \mathrm{d}\boldsymbol{\omega} \\
&= (2\pi)^{-d} \int_{\mathbb{R}^d} \sum_{1 \leq m,n \leq d} \overline{\mathrm{R}^\gamma \hat{f}_m(\boldsymbol{\omega})} \big[\hat{\boldsymbol{\Phi}}^{-2\gamma}_{-2\mathrm{Re}\, \boldsymbol{\xi}}(\boldsymbol{\omega})\big]_{mn} \mathrm{R}^\gamma \hat{f}_n(\boldsymbol{\omega}) \, \mathrm{d}\boldsymbol{\omega}.
\end{aligned}$$

We may consider the behavior of the integrand separately about $\boldsymbol{\omega} = \mathbf{0}$ and at infinity. First, note that $\overline{\mathrm{R}^\gamma \hat{f}_m(\boldsymbol{\omega})} \mathrm{R}^\gamma \hat{f}_n(\boldsymbol{\omega})$ has a zero of order at least $2\lfloor 2\gamma - d/2 \rfloor + 2$ at $\boldsymbol{\omega} = \mathbf{0}$ (cf. the definition of R^γ in (2.13)). Since the singularity of $\big[\hat{\boldsymbol{\Phi}}^{-2\gamma}_{-2\mathrm{Re}\, \boldsymbol{\xi}}(\boldsymbol{\omega})\big]_{mn}$ at $\boldsymbol{\omega} = \mathbf{0}$ is of order -4γ and

$$2\lfloor 2\gamma - d/2 \rfloor + 2 - 4\gamma > -d$$

for $2\gamma - d/2 \notin \mathbb{Z}$, the integral converges about $\boldsymbol{\omega} = \mathbf{0}$.

At infinity, $\overline{\mathrm{R}^\gamma \hat{f}_m(\boldsymbol{\omega})} \mathrm{R}^\gamma \hat{f}_n(\boldsymbol{\omega})$ is dominated by the polynomial term and grows at most like $\|\boldsymbol{\omega}\|^{2\lfloor 2\gamma - d/2 \rfloor}$, while $\big[\hat{\boldsymbol{\Phi}}^{-2\gamma}_{-2\mathrm{Re}\, \boldsymbol{\xi}}(\boldsymbol{\omega})\big]_{mn}$ decays like $\|\boldsymbol{\omega}\|^{-4\gamma}$. We have

$$2\lfloor 2\gamma - d/2 \rfloor - 4\gamma < -d,$$

from where it follows that the integral also converges at infinity. The $(\mathrm{L}^2)^d$ norm of $(-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}}\boldsymbol{f}$ is therefore bounded. □

The Hermitian adjoint[5] of $(-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}}$ is the operator

$$(-\acute{\boldsymbol{\Delta}})^{-\gamma}_{-\overline{\boldsymbol{\xi}}} : \boldsymbol{f} \mapsto (2\pi)^{-d} \int_{\mathbb{R}^d} \left[\mathrm{e}^{\mathrm{j}\langle \boldsymbol{x},\boldsymbol{\omega}\rangle} - \sum_{|\boldsymbol{k}|\le \lfloor 2\gamma - \frac{d}{2} \rfloor} \frac{\mathrm{j}^{|\boldsymbol{k}|}\boldsymbol{x}^{\boldsymbol{k}}\boldsymbol{\omega}^{\boldsymbol{k}}}{\boldsymbol{k}!} \right] \hat{\boldsymbol{\Phi}}^{-\gamma}_{-\overline{\boldsymbol{\xi}}}(\boldsymbol{\omega}) \hat{\boldsymbol{f}}(\boldsymbol{\omega})\, \mathrm{d}\boldsymbol{\omega} \tag{2.15}$$

(see Appendix B).

As was suggested, $(-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}}$ and $(-\acute{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}}$ are named, respectively, the *left* and *right* inverses of $(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}$. They satisfy

$$(-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}}(-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}} = \mathrm{Id} \quad \text{and} \quad (-\boldsymbol{\Delta})^{\gamma}_{\boldsymbol{\xi}}(-\acute{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}} = \mathrm{Id} \tag{2.16}$$

over $\mathcal{S}^d$. We may further extend the domain of $(-\acute{\boldsymbol{\Delta}})^{-\gamma}_{-\overline{\boldsymbol{\xi}}}$ to a subset of generalized functions (distributions)[6] on $\mathcal{S}^d$, using as definition the duality relation

$$\langle (-\acute{\boldsymbol{\Delta}})^{-\gamma}_{-\overline{\boldsymbol{\xi}}}\boldsymbol{g}, \boldsymbol{f}\rangle := \langle \boldsymbol{g}, (-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}}\boldsymbol{f}\rangle$$

wherever the right-hand side (r.h.s.) is meaningful and continuous for all $\boldsymbol{f} \in \mathcal{S}^d$.

It is easily verified that $(-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}}$ and, by duality, $(-\acute{\boldsymbol{\Delta}})^{-\gamma}_{-\overline{\boldsymbol{\xi}}}$ are rotation invariant and homogeneous. This fact is captured in our next proposition, which we shall prove with the aid of the following lemma.

LEMMA 2.4. $\mathrm{R}^{\gamma}[\boldsymbol{f}(\mathbf{M}^{-1}\cdot)](\boldsymbol{x}) = [\mathrm{R}^{\gamma}\boldsymbol{f}(\cdot)](\mathbf{M}^{-1}\boldsymbol{x})$.

Proof. By the uniqueness of the Taylor series expansion,

$$\text{r.h.s.} = \boldsymbol{f}(\mathbf{M}^{-1}\boldsymbol{x}) - \sum_{|\boldsymbol{k}|\le \lfloor 2\gamma - \frac{d}{2} \rfloor} \mathbf{T}_{\boldsymbol{k}}[\boldsymbol{f}](\mathbf{M}^{-1}\boldsymbol{x})^{\boldsymbol{k}} = \text{l.h.s.} \qquad \square$$

PROPOSITION 2.5. *The operators* $(-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}}$ *and* $(-\acute{\boldsymbol{\Delta}})^{-\gamma}_{-\overline{\boldsymbol{\xi}}}$ *are rotation invariant and homogeneous in the sense of* (2.3) *and* (2.4).

Proof. For a nonsingular real matrix $\mathbf{M}$,

$$\begin{aligned}(-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}}\boldsymbol{f}_{\mathbf{M}}(\boldsymbol{x}) &= (2\pi)^{-d} \int_{\mathbb{R}^d} |\det \mathbf{M}|\, \mathrm{e}^{\mathrm{j}\langle \boldsymbol{x},\boldsymbol{\omega}\rangle} \hat{\boldsymbol{\Phi}}^{-\gamma}_{-\boldsymbol{\xi}}(\boldsymbol{\omega})[\mathrm{R}^{\gamma}\mathbf{M}\hat{\boldsymbol{f}}](\mathbf{M}^{\mathsf{T}}\boldsymbol{\omega})\, \mathrm{d}\boldsymbol{\omega} \\ &= (2\pi)^{-d} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{j}\langle \mathbf{M}^{-1}\boldsymbol{x},\boldsymbol{\rho}\rangle} \hat{\boldsymbol{\Phi}}^{-\gamma}_{-\boldsymbol{\xi}}(\mathbf{M}^{-\mathsf{T}}\boldsymbol{\rho})\mathbf{M}[\mathrm{R}^{\gamma}\hat{\boldsymbol{f}}](\boldsymbol{\rho})\, \mathrm{d}\boldsymbol{\rho}\end{aligned}$$

by Lemma 2.4 and with the change of variables $\boldsymbol{\rho} = \mathbf{M}^{\mathsf{T}}\boldsymbol{\omega}$. Equations (2.5) and (2.6) can now be used to verify the rotation invariance and homogeneity of $(-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}}$ (and, by duality, of $(-\acute{\boldsymbol{\Delta}})^{-\gamma}_{-\overline{\boldsymbol{\xi}}}$). □

[5] This adjoint is with respect to the $\mathcal{S}^d$ scalar product

$$\langle \boldsymbol{f}, \boldsymbol{g}\rangle := \int_{\mathbb{R}^d} \boldsymbol{f}^{\mathsf{H}}(\boldsymbol{x})\boldsymbol{g}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \sum_{1\le i\le d} \langle f_i, g_i\rangle.$$

[6] By these we mean members of the dual $(\mathcal{S}^d)'$ of $\mathcal{S}^d$. As a matter of fact, $(\mathcal{S}^d)'$ can be identified with $(\mathcal{S}')^d$.

Finally, we note that the scalar counterparts of the vector left and right inverses were defined in our previous paper [51] (as generalizations of the one-dimensional definitions of Blu and Unser [4]) as follows:

$$
\begin{aligned}
(-\grave{\Delta})^{-\gamma} &: f \mapsto (2\pi)^{-d} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{j}\langle \boldsymbol{x}, \boldsymbol{\omega}\rangle} \|\boldsymbol{\omega}\|^{-2\gamma} \left[\hat{f}(\boldsymbol{\omega}) - \sum_{|\boldsymbol{k}| \leq \lfloor 2\gamma - \frac{d}{2} \rfloor} \frac{f^{(\boldsymbol{k})}(\mathbf{0}) \boldsymbol{\omega}^{\boldsymbol{k}}}{\boldsymbol{k}!} \right] \mathrm{d}\boldsymbol{\omega}; \\
(-\acute{\Delta})^{-\gamma} &: f \mapsto (2\pi)^{-d} \int_{\mathbb{R}^d} \left[\mathrm{e}^{\mathrm{j}\langle \boldsymbol{x}, \boldsymbol{\omega}\rangle} - \sum_{|\boldsymbol{k}| \leq \lfloor 2\gamma - \frac{d}{2} \rfloor} \frac{\mathrm{j}^{|\boldsymbol{k}|} \boldsymbol{x}^{\boldsymbol{k}} \boldsymbol{\omega}^{\boldsymbol{k}}}{\boldsymbol{k}!} \right] \|\boldsymbol{\omega}\|^{-2\gamma} \hat{f}(\boldsymbol{\omega}) \, \mathrm{d}\boldsymbol{\omega}.
\end{aligned}
\tag{2.17}
$$

They share the conjugacy and inversion properties of the vector inverses (cf. (2.10), (2.16)). Notice that the operand, f, is now scalar valued. Also, the vectorial parameters $\xi_{\mathrm{irr}}, \xi_{\mathrm{sol}}$ have no equivalent in the scalar case.

DIGRESSION 2.6. *The reader may be wondering why we should bother at all about singular integrals and distinct left and right inverses when we could have—as indicated in the introduction to this subsection—conveniently characterized the solution as an element of the space $(\mathcal{S}^d)'/\Pi$ of Schwartz distributions modulo polynomials, on which space the fractional Laplacian is bijective and therefore uniquely invertible. This would indeed be possible, since the space $\mathcal{S}_0^d$, being a subspace of a nuclear space, is again nuclear* [41, *Chapter* 5]*; therefore the theorems of Minlos (see* [26, 36]*) which we shall use in subsection* 3.1 *apply to it. By following this approach, one can characterize fractional Brownian vector fields as random elements of $(\mathcal{S}^d)'/\Pi$ (i.e., as random equivalence classes of tempered distributions modulo polynomials). One could in fact do even better by considering test functions with a finite number of vanishing moments and their dual spaces (Schwartz distributions modulo polynomials of some finite order), as was done by Dobrushin* [11] *in the scalar setting. (On a related note, the reader might also wish to consult the work of Vedel on the wavelet analysis of the Mumford process* [59]*; see also Bourdaud* [5]*.)*

However, the latter approach—although more straightforward from a technical point of view—does not provide us, at least immediately, with as complete a characterization of the stochastic solutions to (2.2) *as the one we shall see in the following sections.*

As far as the spaces of solutions are concerned, another possibility would be to use fractional Sobolev spaces, as proposed by Ruiz-Medina, Anh, and Angulo [45] *and Kelbert, Leonenko, and Ruiz-Medina* [23]*. It appears that this approach would work especially well when considering Gaussian self-similar vector fields. Working with spaces of generalized functions, on the other hand, allows us to use the method of characteristic functionals* [26, 36, 43]*, which shows its versatility when extending the work to the study of non-Gaussian random models.*

3. Vector fBm. A classical definition of the *scalar* isotropic fractional Brownian motion field with Hurst exponent H (denoted B_H, with $0 < H < 1$) goes as follows: B_H is a zero-mean Gaussian field satisfying $B_H(\mathbf{0}) = 0$, with stationary (Gaussian) increments whose variance depends on the step size as per

$$
\mathbf{E}\{|B_H(\boldsymbol{x}) - B_H(\boldsymbol{y})|^2\} = 2\alpha \|\boldsymbol{x} - \boldsymbol{y}\|^{2H}. \tag{3.1}
$$

This is a generalization of Lévy's characterization of multiparameter Brownian motion [30], to which it reduces for $H = \frac{1}{2}$. The above expectation, as a function of $\boldsymbol{x}$ and $\boldsymbol{y}$, is also known as the *variogram* of the field B_H (denoted here by $\mathrm{Vario}[B_H](\boldsymbol{x}, \boldsymbol{y})$).

A straightforward extension of the above definition to the multicomponent (vector) setting is obtained by requiring the vector-valued field $\boldsymbol{B}_H$ to satisfy

$$\mathbf{E}\{\|\boldsymbol{B}_H(\boldsymbol{x}) - \boldsymbol{B}_H(\boldsymbol{y})\|^2\} = 2\alpha'\|\boldsymbol{x} - \boldsymbol{y}\|^{2H}, \tag{3.2}$$

where one now considers the Euclidean norm of the increments instead of their absolute values [22]. This definition leaves the cross-correlation structure of the components of $\boldsymbol{B}_H$ unspecified; these are typically assumed to be independent, in which case the components become scalar fBm's of exponent H; i.e., the generalization is trivial. More generally, for a vector-valued random field, one may define a variogram matrix

$$\mathrm{Vario}[\boldsymbol{B}_H](\boldsymbol{x}, \boldsymbol{y}) := \mathbf{E}\left\{[\boldsymbol{B}_H(\boldsymbol{x}) - \boldsymbol{B}_H(\boldsymbol{y})][\boldsymbol{B}_H(\boldsymbol{x}) - \boldsymbol{B}_H(\boldsymbol{y})]^{\mathsf{H}}\right\}.$$

The scalar function given in (3.2) corresponds to the trace of this matrix.

A different approach to defining fBm consists in characterizing it as a linear transformation (essentially a fractional integral) of white noise. In this approach, one starts with a white noise measure on some suitable space and proceeds to derive the probabilistic law (probability measure on a certain space[7]) of fBm from there, showing that it is consistent with the definition given in (3.1) and (3.2). One advantage of this approach is that it is not, in its essence, limited to second-order statistical analysis; this means that one is in principle free to consider non-Gaussian white noises within the same framework.

In the scalar setting, it has been indicated previously in one way or another that the linear transformation of white noise which produces fBm corresponds, in effect, to the right inverse of the scalar fractional Laplacian introduced in (2.17) (see, for instance, Tafti, Van De Ville, and Unser [51], Benassi, Jaffard, and Roux [3], Leonenko [28], and Kelbert, Leonenko, and Ruiz-Medina [23] for the multidimensional case and Samorodnitsky and Taqqu [47] and Blu and Unser [4] for the unidimensional one). One may therefore say that the scalar fractional Brownian field B_H *solves*—we shall elaborate on this—the fractional Poisson equation

$$(-\Delta)^{H/2+d/4} B_H = \epsilon_H W \tag{3.3}$$

subject to boundary conditions imposed by the right inverse (zero at the origin); i.e.,

$$B_H = \epsilon_H (-\acute{\Delta})^{-H/2-d/4} W.$$

In the above formula W denotes a Gaussian white noise field (defined in subsection 3.1) and ϵ_H is a special constant related to α in (3.1) by

$$\epsilon_H = \sqrt{(2\pi)^{\frac{d}{2}} \frac{2^{2H+d/2}\Gamma(H + \frac{d}{2})}{|\Gamma(-H)|}\alpha}.$$

Given that the only essential limitation on H values in the above characterization is the exclusion of integer H (as a consequence of (2.12)), it can also serve as a natural generalization of the definition of fBm to $H > 1$ [4, 51].

So far in this section we have identified two approaches towards defining scalar fBm's: first by means of the variogram and then through a transformation of white

[7]The space of tempered distributions is standard [19], although other choices are also possible (cf. the monographs by Vakhania [58] and Talagrand [52] and the papers by Ruiz-Medina, Anh, and Angulo [45] and Kelbert, Leonenko, and Ruiz-Medina [23]).

noise. Early on in this section we also highlighted what may be considered a fundamental property of any reasonable vector generalization of fBm (see (3.2)). We noted that a relatively trivial random vector field with the said property could be constructed by grouping together d independent scalar fBm's.

Next, we shall propose a more general definition of vector fBm consistent with the trace structure of (3.2). We shall not, however, approach the problem by imposing this requirement directly. Instead, following the line of reasoning sketched in the previous paragraph, our characterization relies on solving a stochastic fractional partial differential equation similar to (3.3). From there, we shall then proceed to derive the variogram of the model in section 4 and show that it has the desired trace property.

3.1. The whitening model. As hinted above, we shall take generalized vector fBm to be the solution of the fractional Poisson equation

$$(-\boldsymbol{\Delta})^{H/2+d/4}_{\boldsymbol{\xi}} \boldsymbol{B}_{H,\boldsymbol{\xi}} = \epsilon_H \boldsymbol{W} \tag{3.4}$$

defined using the right inverse as per

$$\boldsymbol{B}_{H,\boldsymbol{\xi}} := \epsilon_H (-\acute{\boldsymbol{\Delta}})^{-H/2-d/4}_{-\boldsymbol{\xi}} \boldsymbol{W}, \tag{3.5}$$

where $\boldsymbol{W}$ is a white noise vector field, to be defined shortly. The first identity is known as a *whitening* equation in signal processing parlance (although there it is applied only to stationary processes). We shall limit our consideration to *real* random fields.

Equations (3.4) and (3.5) may be understood as equivalences in law in a sense we shall now describe. The main reference for the underlying theory of generalized random fields is Gel'fand and Vilenkin [18].

$\boldsymbol{B}_{H,\boldsymbol{\xi}}$ and $\boldsymbol{W}$ are taken to be generalized random fields, i.e., random elements of the continuous duals of certain spaces of test functions. Let us use $\boldsymbol{X}$ to denote one such random element. Under some reasonable consistency conditions, by a generalization of Kolmogorov's extension theorem [36], the stochastic law (infinite-dimensional σ-additive probability measure) of $\boldsymbol{X}$ is fully specified—in the sense of a σ-additive measure on the σ-algebra of Borel cylinder sets—by way of indicating all finite joint distributions of its "scalar products" with test functions. These products are classical random variables denoted as $\langle \boldsymbol{X}, \boldsymbol{f} \rangle$, with $\boldsymbol{f}$ belonging to the desired test function space.

By Minlos's infinite-dimensional generalization of Bochner's theorem [26, 36], it is also possible to uniquely specify the stochastic law of a real random field $\boldsymbol{X}$ by its *characteristic functional*, defined as the expectation

$$L_{\boldsymbol{X}}(\boldsymbol{f}) := \mathbf{E}\{\mathrm{e}^{\mathrm{j}\langle \boldsymbol{X}, \boldsymbol{f} \rangle}\},$$

provided the test functions belong to a nuclear space. More precisely, a probability measure on a dual nuclear space gives rise to a positive-definite and continuous characteristic functional, and, conversely, any positive-definite and continuous functional on a nuclear space that evaluates to 1 at $\boldsymbol{f} \equiv \boldsymbol{0}$ uniquely determines a probability measure on the dual space.

The characteristic functional serves as an infinite-dimensional equivalent of the characteristic function of a random variable. In particular, for any finite number of test functions $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_N$ (N arbitrary), the N-variable function

$$\varphi(\omega_1, \ldots, \omega_N) := L_{\boldsymbol{X}} \left(\sum_{1 \le i \le N} \omega_i \boldsymbol{f}_i \right)$$

is the joint characteristic function of the random variables $\langle \boldsymbol{X}, \boldsymbol{f}_i \rangle$, $1 \le i \le N$, which is in one-to-one correspondence with their finite-dimensional joint measure by the finite-dimensional version of Bochner's theorem. Characteristic functionals have an entire theory of their own on which we shall not elaborate here, referring instead to Gel'fand and Vilenkin [18] and the survey article by Prohorov [43].

Another useful functional that one may consider is the *correlation form* of $\boldsymbol{X}$, defined as

$$\langle\!\langle \boldsymbol{f}, \boldsymbol{g} \rangle\!\rangle_{\boldsymbol{X}} := \mathbf{E}\left\{ \overline{\langle \boldsymbol{X}, \boldsymbol{f} \rangle} \langle \boldsymbol{X}, \boldsymbol{g} \rangle \right\} \quad \text{for} \quad \boldsymbol{f}, \boldsymbol{g} \in \mathcal{S}^d.$$

For real Gaussian fields, it can be shown that the correlation form and characteristic functional are related by

$$L_{\boldsymbol{X}}(\boldsymbol{f}) = \exp[-\tfrac{1}{2} \langle\!\langle \boldsymbol{f}, \boldsymbol{f} \rangle\!\rangle_{\boldsymbol{X}}], \tag{3.6}$$

which is consistent with the understanding that a Gaussian field is completely specified by its second-order statistics.

A reasonable definition of scalar *white noise* can be given as a random field W that has independent values at every point in the sense that for any two test functions f, g with disjoint supports $\langle W, f \rangle$ and $\langle W, g \rangle$ are independent. With the additional assumption that the field has Gaussian statistics, one is led to the standard definition of scalar white Gaussian noise as the field with characteristic functional

$$L_W(f) = \exp[-\tfrac{1}{2} \|f\|_2^2]$$

($\|\cdot\|_2$ denotes the L^2 norm). This random field exists as a random element of $\mathcal{S}'$ (i.e., it corresponds to a unique probability measure on $\mathcal{S}'$), as Minlos [36] has shown. The above characteristic functional also defines a cylinder probability measure on subspaces of L^2.

We shall define the standard white Gaussian noise *vector* $\boldsymbol{W}$ as the field with characteristic functional

$$L_{\boldsymbol{W}}(\boldsymbol{f}) = \exp[-\tfrac{1}{2} \|\boldsymbol{f}\|_2^2] = \exp\left[-\tfrac{1}{2} \sum_{1 \le k \le d} \|f_k\|_2^2 \right].$$

It is clear that $\boldsymbol{W}$ corresponds to a vector of independent scalar white noise fields. Its correlation form is given by the relation

$$\langle\!\langle \boldsymbol{f}, \boldsymbol{g} \rangle\!\rangle_{\boldsymbol{W}} = \langle \boldsymbol{f}, \boldsymbol{g} \rangle = \sum_{1 \le k \le d} \langle f_k, g_k \rangle. \tag{3.7}$$

DIGRESSION 3.1. *The general form of the characteristic functional of a (not necessarily Gaussian) one-dimensional white noise process can be found in Gel'fand and Vilenkin* [18]. *In the multivariate setting, we note here in particular the characteristic functional of a Poisson white noise field P consisting of Dirac impulses with independent and identically distributed amplitudes with probability measure P_a and a spatial Poisson distribution with parameter λ:*

$$L_P(f) = \exp\left[\lambda \int \int_{\mathbb{R}^d} (\mathrm{e}^{\mathrm{j} a f(\boldsymbol{x})} - 1) \, \mathrm{d}\boldsymbol{x} \, P_a(\mathrm{d}a) \right].$$

Poisson white noise can be used to define non-Gaussian stochastic fractals that agree with fractional Brownian models in their second-order statistics [57].

With our framework set as it is, all we now need is a means to derive the law of $\boldsymbol{B}_{H,\boldsymbol{\xi}}$ from that of $\boldsymbol{W}$, i.e., to give probabilistic meaning to (3.5). This we shall do as follows.

By definition, the action of an operator on a generalized function (random or deterministic) is described by the action of its adjoint on test functions. In particular, we have (cf. (2.10))

$$\boldsymbol{B}_{H,\boldsymbol{\xi}} = \epsilon_H(-\boldsymbol{\grave{\Delta}})^{-H/2-d/4}_{-\boldsymbol{\xi}}\boldsymbol{W} \quad \Leftrightarrow \quad \langle \boldsymbol{B}_{H,\boldsymbol{\xi}}, \boldsymbol{f}\rangle = \langle \boldsymbol{W}, \overline{\epsilon_H}(-\boldsymbol{\grave{\Delta}})^{-H/2-d/4}_{-\overline{\boldsymbol{\xi}}}\boldsymbol{f}\rangle \tag{3.8}$$

for all test functions $\boldsymbol{f} \in \mathcal{S}^d$. One may interpret the right-hand equality as an equivalence in joint law for all finite collections of test functions $\boldsymbol{f}$.

We shall now make use of (3.8), (3.7), and (2.14) to find the correlation form of vector fBm:

$$\begin{aligned}
\langle\!\langle \boldsymbol{f}, \boldsymbol{g}\rangle\!\rangle_{\boldsymbol{B}_{H,\boldsymbol{\xi}}} &= \mathbf{E}\left\{\overline{\langle \boldsymbol{B}_{H,\boldsymbol{\xi}}, \boldsymbol{f}\rangle}\langle \boldsymbol{B}_{H,\boldsymbol{\xi}}, \boldsymbol{g}\rangle\right\} \\
&= \mathbf{E}\left\{\overline{\langle \epsilon_H(-\boldsymbol{\grave{\Delta}})^{-H/2-d/4}_{-\boldsymbol{\xi}}\boldsymbol{W}, \boldsymbol{f}\rangle}\langle \epsilon_H(-\boldsymbol{\grave{\Delta}})^{-H/2-d/4}_{-\boldsymbol{\xi}}\boldsymbol{W}, \boldsymbol{g}\rangle\right\} \\
&= |\epsilon_H|^2\mathbf{E}\left\{\overline{\langle \boldsymbol{W}, (-\boldsymbol{\grave{\Delta}})^{-H/2-d/4}_{-\overline{\boldsymbol{\xi}}}\boldsymbol{f}\rangle}\langle \boldsymbol{W}, (-\boldsymbol{\grave{\Delta}})^{-H/2-d/4}_{-\overline{\boldsymbol{\xi}}}\boldsymbol{g}\rangle\right\} \\
&= |\epsilon_H|^2\langle\!\langle (-\boldsymbol{\grave{\Delta}})^{-H/2-d/4}_{-\overline{\boldsymbol{\xi}}}\boldsymbol{f}, (-\boldsymbol{\grave{\Delta}})^{-H/2-d/4}_{-\overline{\boldsymbol{\xi}}}\boldsymbol{g}\rangle\!\rangle_{\boldsymbol{W}} \\
&= |\epsilon_H|^2\langle (-\boldsymbol{\grave{\Delta}})^{-H/2-d/4}_{-\overline{\boldsymbol{\xi}}}\boldsymbol{f}, (-\boldsymbol{\grave{\Delta}})^{-H/2-d/4}_{-\overline{\boldsymbol{\xi}}}\boldsymbol{g}\rangle \\
&= \frac{|\epsilon_H|^2}{(2\pi)^d}\int_{\mathbb{R}^d}[\mathrm{R}^{H/2+d/4}\hat{\boldsymbol{f}}]^{\mathsf{H}}(\boldsymbol{\omega})\hat{\boldsymbol{\Phi}}^{-H-\frac{d}{2}}_{-2\mathrm{Re}\,\boldsymbol{\xi}}(\boldsymbol{\omega})[\mathrm{R}^{H/2+d/4}\hat{\boldsymbol{g}}](\boldsymbol{\omega})\,\mathrm{d}\boldsymbol{\omega}.
\end{aligned}$$

In view of the above identity and (3.6) we have the following theorem.

THEOREM 3.2. *The characteristic functional of the vector fBm $\boldsymbol{B}_{H,\boldsymbol{\xi}}$ is given by*

$$L_{\boldsymbol{B}_{H,\boldsymbol{\xi}}}(\boldsymbol{f}) = \exp\left[-\frac{|\epsilon_H|^2}{2(2\pi)^d}\int_{\mathbb{R}^d}[\mathrm{R}^{\frac{2H+d}{4}}\hat{\boldsymbol{f}}]^{\mathsf{H}}(\boldsymbol{\omega})\hat{\boldsymbol{\Phi}}^{-H-\frac{d}{2}}_{-2\mathrm{Re}\,\boldsymbol{\xi}}(\boldsymbol{\omega})[\mathrm{R}^{\frac{2H+d}{4}}\hat{\boldsymbol{f}}](\boldsymbol{\omega})\,\mathrm{d}\boldsymbol{\omega}\right], \tag{3.9}$$

with $\hat{\boldsymbol{\Phi}}^{-H-\frac{d}{2}}_{-2\mathrm{Re}\,\boldsymbol{\xi}}(\boldsymbol{\omega})$ and the regularization operator $\mathrm{R}^{\frac{2H+d}{4}}$ defined as in (2.7) *and* (2.13)*, respectively.*

We remark that the positive-definiteness and continuity of $L_{\boldsymbol{B}_{H,\boldsymbol{\xi}}}$ follow from the positive-definiteness of $L_{\boldsymbol{W}}$ and the continuity of $L_{\boldsymbol{W}}$ and $(-\boldsymbol{\grave{\Delta}})^{-H/2-d/4}_{-\boldsymbol{\xi}}$. These imply the existence of a probability measure corresponding to the given characteristic form, also for $H > 1$, thus extending the definition of fBm outside the usual range of 0 to 1 (however, by (2.12), integer Hurst exponents are once again excluded).

4. Some properties of vector fBm. In this section we shall establish some of the main properties of the random fields defined in the previous section.

4.1. Self-similarity. Vector fBm fields are statistically self-similar (fractal) in the sense that the random field $\boldsymbol{B}_{H,\boldsymbol{\xi}}(\sigma\cdot)$ has the same statistical character as the

field $\sigma^H \boldsymbol{B}_{H,\boldsymbol{\xi}}$. This can be shown as follows:

$$\begin{aligned}
\langle \boldsymbol{B}_{H,\boldsymbol{\xi}}(\sigma\cdot), \boldsymbol{f}(\cdot)\rangle &= \langle \epsilon_H[(-\acute{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\boldsymbol{\xi}} \boldsymbol{W}](\sigma\cdot), \boldsymbol{f}(\cdot)\rangle && \text{by (3.4)},\\
&= \langle \sigma^{H+\frac{d}{2}} \epsilon_H (-\acute{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\boldsymbol{\xi}} \{\boldsymbol{W}(\sigma\cdot)\}, \boldsymbol{f}(\cdot)\rangle && \text{by (2.4)},\\
&= \langle \sigma^{H+\frac{d}{2}} \epsilon_H \boldsymbol{W}(\sigma\cdot), (-\grave{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\bar{\boldsymbol{\xi}}} \boldsymbol{f}(\cdot)\rangle && \text{by duality},\\
&= \langle \sigma^{H+\frac{d}{2}} \sigma^{-\frac{d}{2}} \epsilon_H \boldsymbol{W}(\cdot), (-\grave{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\bar{\boldsymbol{\xi}}} \boldsymbol{f}(\cdot)\rangle && \text{by the homogeneity of } \boldsymbol{W},\\
&= \langle \sigma^{H} \epsilon_H (-\acute{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\boldsymbol{\xi}} \{\boldsymbol{W}(\cdot)\}, \boldsymbol{f}(\cdot)\rangle && \text{by duality},\\
&= \langle \sigma^{H} \boldsymbol{B}_{H,\boldsymbol{\xi}}(\cdot), \boldsymbol{f}(\cdot)\rangle && \text{by (3.4)}.
\end{aligned}$$

4.2. Rotation invariance. For any orthogonal transformation matrix $\boldsymbol{\Omega}$, the random fields $\boldsymbol{B}_{H,\boldsymbol{\xi}}$ and $\boldsymbol{\Omega}\boldsymbol{B}_{H,\boldsymbol{\xi}}(\boldsymbol{\Omega}^{\mathsf{T}}\cdot)$ follow the same stochastic law. The demonstration is similar to the previous one.

4.3. Nonstationarity. Vector fBm is nonstationary. The operator $(-\grave{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\boldsymbol{\xi}}$ is not translation invariant, and consequently the random variables

$$\langle \boldsymbol{B}_{H,\boldsymbol{\xi}}, \boldsymbol{f}(\cdot)\rangle = \langle \epsilon_H \boldsymbol{W}, (-\grave{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\bar{\boldsymbol{\xi}}} \{\boldsymbol{f}(\cdot)\}\rangle$$

and

$$\langle \boldsymbol{B}_{H,\boldsymbol{\xi}}, \boldsymbol{f}(\cdot+\boldsymbol{h})\rangle = \langle \epsilon_H \boldsymbol{W}, (-\grave{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\bar{\boldsymbol{\xi}}} \{\boldsymbol{f}(\cdot+\boldsymbol{h})\}\rangle$$

are not identically distributed in general.

4.4. Stationary nth-order increments. We shall now show that the increments of order $\lfloor H\rfloor + 1$ of the field $\boldsymbol{B}_{H,\boldsymbol{\xi}}$ are stationary. In particular, for $0 < H < 1$, $\boldsymbol{B}_{H,\boldsymbol{\xi}}$ has stationary first-order increments, as is the case for standard fBm [53]. To show this, let us first define the nth-order symmetric difference operator $\mathrm{D}_{\boldsymbol{h}_1,\ldots,\boldsymbol{h}_n}$ recursively by the relations

$$\begin{gathered}
\mathrm{D}_{\boldsymbol{h}_1} : \boldsymbol{f}(\cdot) \mapsto \boldsymbol{f}(\cdot + \tfrac{\boldsymbol{h}_1}{2}) - \boldsymbol{f}(\cdot - \tfrac{\boldsymbol{h}_1}{2}),\\
\mathrm{D}_{\boldsymbol{h}_1,\ldots,\boldsymbol{h}_n} := \mathrm{D}_{\boldsymbol{h}_n} \mathrm{D}_{\boldsymbol{h}_1,\ldots,\boldsymbol{h}_{n-1}},
\end{gathered}$$

with $\boldsymbol{h}_1,\ldots,\boldsymbol{h}_n \in \mathbb{R}^d\backslash\{\boldsymbol{0}\}$. The above operator is represented in the Fourier domain by the expression

$$\prod_{1\le i\le n} 2\sin\tfrac{\langle \boldsymbol{h}_i,\boldsymbol{\omega}\rangle}{2}.$$

We have the following theorem.

THEOREM 4.1. *The vector fBm field $\boldsymbol{B}_{H,\boldsymbol{\xi}}$ has stationary increments of order $\lfloor H\rfloor + 1$; that is, the random field*

$$\mathrm{D}_{\boldsymbol{h}_1,\ldots,\boldsymbol{h}_n} \boldsymbol{B}_{H,\boldsymbol{\xi}}$$

with $n = \lfloor H\rfloor + 1$ is stationary, irrespective of the lengths and directions of the steps $\boldsymbol{h}_i$, $1 \le i \le n$.

Proof. We proceed to show that the characteristic functional of the increment field $\boldsymbol{I}_n := \mathrm{D}_{\boldsymbol{h}_1,\dots,\boldsymbol{h}_n}\boldsymbol{B}_{H,\boldsymbol{\xi}}$ is shift invariant, i.e.,

$$L_{\boldsymbol{I}_n}(\boldsymbol{f}(\cdot-\boldsymbol{h})) = L_{\boldsymbol{I}_n}(\boldsymbol{f}),$$

which then directly implies the stationarity of $\boldsymbol{I}_n$.

Indeed, one may write

$$\begin{aligned}
L_{\boldsymbol{I}_n}(\boldsymbol{f}(\cdot-\boldsymbol{h})) &= \mathbf{E}\big\{\exp[\mathrm{j}\langle \boldsymbol{I}_n, \boldsymbol{f}(\cdot-\boldsymbol{h})\rangle]\big\}\\
&= \mathbf{E}\big\{\exp[\mathrm{j}\langle \mathrm{D}_{\boldsymbol{h}_1,\dots,\boldsymbol{h}_n}\boldsymbol{B}_{H,\boldsymbol{\xi}}, \boldsymbol{f}(\cdot-\boldsymbol{h})\rangle]\big\}\\
&= \mathbf{E}\big\{\exp[\mathrm{j}\langle \boldsymbol{B}_{H,\boldsymbol{\xi}}, \mathrm{D}_{-h_n,\dots,-h_1}\boldsymbol{f}(\cdot-\boldsymbol{h})\rangle]\big\}\\
&= L_{\boldsymbol{B}_{H,\boldsymbol{\xi}}}(\mathrm{D}_{-h_n,\dots,-h_1}\{\boldsymbol{f}(\cdot-\boldsymbol{h})\})\\
&= \exp\left[-\frac{|\epsilon_H|^2}{2(2\pi)^d}\int_{\mathbb{R}^d}\left[\mathrm{R}^{\frac{2H+d}{4}}\left\{\hat{\boldsymbol{f}}(\boldsymbol{\omega})\mathrm{e}^{-\mathrm{j}\langle \boldsymbol{h},\boldsymbol{\omega}\rangle}\prod_{1\le i\le \lfloor H\rfloor+1} 2\sin\tfrac{\langle \boldsymbol{h}_i,\boldsymbol{\omega}\rangle}{2}\right\}\right]^{\mathsf{H}}\right.\\
&\qquad\left.\cdot\,\hat{\boldsymbol{\Phi}}_{-2\mathrm{Re}\,\boldsymbol{\xi}}^{-H-\frac{d}{2}}(\boldsymbol{\omega})\left[\mathrm{R}^{\frac{2H+d}{4}}\left\{\hat{\boldsymbol{f}}(\boldsymbol{\omega})\mathrm{e}^{-\mathrm{j}\langle \boldsymbol{h},\boldsymbol{\omega}\rangle}\prod_{1\le i\le \lfloor H\rfloor+1} 2\sin\tfrac{\langle \boldsymbol{h}_i,\boldsymbol{\omega}\rangle}{2}\right\}\right]\mathrm{d}\boldsymbol{\omega}\right].
\end{aligned}$$

Next, note that the partial derivatives at $\boldsymbol{\omega}=\boldsymbol{0}$ of the function

$$\hat{\boldsymbol{f}}(\boldsymbol{\omega})\mathrm{e}^{-\mathrm{j}\langle \boldsymbol{h},\boldsymbol{\omega}\rangle}\prod_{1\le i\le \lfloor H\rfloor+1} 2\sin\tfrac{\langle \boldsymbol{h}_i,\boldsymbol{\omega}\rangle}{2},\quad \boldsymbol{\omega}\in\mathbb{R}^d,$$

which appears as the argument of the regularization operator $\mathrm{R}^{\frac{2H+d}{4}}$ in the integral, all vanish up to order $\lfloor H\rfloor+1$ at least; this means that the first $\lfloor H\rfloor+1$ terms of its Taylor expansion around the origin are all zero. As a result, the said function is a fixed point of the regularization operator $\mathrm{R}^{\frac{2H+d}{4}}$ (cf. (2.13)).

All this means that we have

$$\begin{aligned}
L_{\boldsymbol{I}_n}(\boldsymbol{f}(\cdot-\boldsymbol{h})) &= \exp\left[-\frac{|\epsilon_H|^2}{2(2\pi)^d}\int_{\mathbb{R}^d}\left[\hat{\boldsymbol{f}}(\boldsymbol{\omega})\mathrm{e}^{-\mathrm{j}\langle \boldsymbol{h},\boldsymbol{\omega}\rangle}\prod_{1\le i\le \lfloor H\rfloor+1} 2\sin\tfrac{\langle \boldsymbol{h}_i,\boldsymbol{\omega}\rangle}{2}\right]^{\mathsf{H}}\right.\\
&\qquad\left.\cdot\,\hat{\boldsymbol{\Phi}}_{-2\mathrm{Re}\,\boldsymbol{\xi}}^{-H-\frac{d}{2}}(\boldsymbol{\omega})\left[\hat{\boldsymbol{f}}(\boldsymbol{\omega})\mathrm{e}^{-\mathrm{j}\langle \boldsymbol{h},\boldsymbol{\omega}\rangle}\prod_{1\le i\le \lfloor H\rfloor+1} 2\sin\tfrac{\langle \boldsymbol{h}_i,\boldsymbol{\omega}\rangle}{2}\right]\mathrm{d}\boldsymbol{\omega}\right]\\
&= \exp\left[-\frac{|\epsilon_H|^2}{2(2\pi)^d}\int_{\mathbb{R}^d}[\hat{\boldsymbol{f}}(\boldsymbol{\omega})]^{\mathsf{H}}\hat{\boldsymbol{\Phi}}_{-2\mathrm{Re}\,\boldsymbol{\xi}}^{-H-\frac{d}{2}}(\boldsymbol{\omega})[\hat{\boldsymbol{f}}(\boldsymbol{\omega})]\prod_{1\le i\le \lfloor H\rfloor+1} 4\sin^2\tfrac{\langle \boldsymbol{h}_i,\boldsymbol{\omega}\rangle}{2}\,\mathrm{d}\boldsymbol{\omega}\right] \qquad (4.1)\\
&= L_{\boldsymbol{I}_n}(\boldsymbol{f}),
\end{aligned}$$

which is what we set out to prove. ∎

4.5. The variogram and correlation form of vector fBm. As was seen in the previous paragraph, for $0<H<1$ the random field $\boldsymbol{B}_{H,\boldsymbol{\xi}}$ has stationary first-order increments. In this case we may define its variogram (or second-order structure function) as the correlation matrix of the stationary increment $\boldsymbol{B}_{H,\boldsymbol{\xi}}(\boldsymbol{x})-\boldsymbol{B}_{H,\boldsymbol{\xi}}(\boldsymbol{y}) = \mathrm{D}_{\boldsymbol{x}-\boldsymbol{y}}\boldsymbol{B}_{H,\boldsymbol{\xi}}(\frac{\boldsymbol{x}+\boldsymbol{y}}{2})$. Formally, this is to say

$$\mathrm{Vario}[\boldsymbol{B}_{H,\boldsymbol{\xi}}](\boldsymbol{x},\boldsymbol{y}) := \mathbf{E}\big\{\boldsymbol{I}(\boldsymbol{0})[\boldsymbol{I}(\boldsymbol{0})]^{\mathsf{H}}\big\},$$

with $\boldsymbol{I} := \mathrm{D}_{\boldsymbol{x}-\boldsymbol{y}}\boldsymbol{B}_{H,\boldsymbol{\xi}}(\cdot - \frac{\boldsymbol{x}+\boldsymbol{y}}{2})$.

We shall proceed as follows to evaluate the above expression. First, we shall find the cross correlation of the (stationary) scalar random fields $[\boldsymbol{I}]_i$ and $[\boldsymbol{I}]_j$ that constitute the components of the vector $\boldsymbol{I}$. The ijth element of the variogram matrix then corresponds to the value of the said correlation function at $\mathbf{0}$.

We first obtain the correlation form of $\boldsymbol{I}$ from its characteristic functional (derived from (4.1) by setting $n = 1$) by identification (cf. (3.6)):

$$\langle\!\langle \boldsymbol{f}, \boldsymbol{g} \rangle\!\rangle_{\boldsymbol{I}} = \frac{|\epsilon_H|^2}{(2\pi)^d} \int_{\mathbb{R}^d} 4\sin^2\langle \tfrac{\boldsymbol{x}-\boldsymbol{y}}{2}, \boldsymbol{\omega}\rangle [\hat{\boldsymbol{f}}(\omega)]^{\mathsf{H}} \hat{\boldsymbol{\Phi}}^{-H-\frac{d}{2}}_{-2\mathrm{Re}\,\boldsymbol{\xi}}(\boldsymbol{\omega})[\hat{\boldsymbol{g}}(\omega)]\, \mathrm{d}\boldsymbol{\omega}.$$

Next, let $\boldsymbol{f} = \hat{\boldsymbol{e}}_i\phi$ and $\boldsymbol{g} = \hat{\boldsymbol{e}}_j\psi$, where $\hat{\boldsymbol{e}}_i$ and $\hat{\boldsymbol{e}}_j$ denote standard unit vectors in $\mathbb{R}^d$ and ϕ and ψ are scalar test functions. We have

$$\begin{aligned}\mathbf{E}\big\{\overline{\langle [\boldsymbol{I}]_i, \phi\rangle}\langle [\boldsymbol{I}]_j, \psi\rangle\big\} &= \langle\!\langle \boldsymbol{f}, \boldsymbol{g} \rangle\!\rangle_{\boldsymbol{I}} \\ &= (2\pi)^{-d} \int_{\mathbb{R}^d} 4\sin^2\langle \tfrac{\boldsymbol{x}-\boldsymbol{y}}{2}, \boldsymbol{\omega}\rangle \big[|\epsilon_H|^2 \hat{\boldsymbol{\Phi}}^{-H-\frac{d}{2}}_{-2\mathrm{Re}\,\boldsymbol{\xi}}(\boldsymbol{\omega})\big]_{ij} \overline{\hat{\phi}(\omega)}\hat{\psi}(\omega)]\, \mathrm{d}\boldsymbol{\omega}.\end{aligned}$$

By the kernel theorem, this last expression can be written in the spatial domain as

$$(2\pi)^{-d} \int_{\mathbb{R}^d} c(\boldsymbol{t} - \boldsymbol{\tau})\phi(\boldsymbol{t})\psi(\boldsymbol{\tau})\, \mathrm{d}\boldsymbol{t}\mathrm{d}\boldsymbol{\tau},$$

where $c(\boldsymbol{t})$ is the generalized cross-correlation function of the random fields $[\boldsymbol{I}]_i$ and $[\boldsymbol{I}]_j$. $c(\boldsymbol{t})$ is given by the inverse Fourier transform of

$$4\sin^2\langle \tfrac{\boldsymbol{x}-\boldsymbol{y}}{2}, \boldsymbol{\omega}\rangle \big[|\epsilon_H|^2 \hat{\boldsymbol{\Phi}}^{-H-\frac{d}{2}}_{-2\mathrm{Re}\,\boldsymbol{\xi}}(\boldsymbol{\omega})\big]_{ij} = \big(\mathrm{e}^{\mathrm{j}\langle \boldsymbol{x}-\boldsymbol{y}, \boldsymbol{\omega}\rangle} - 2 + \mathrm{e}^{\mathrm{j}\langle \boldsymbol{y}-\boldsymbol{x}, \boldsymbol{\omega}\rangle}\big)\big[|\epsilon_H|^2 \hat{\boldsymbol{\Phi}}^{-H-\frac{d}{2}}_{-2\mathrm{Re}\,\boldsymbol{\xi}}(\boldsymbol{\omega})\big]_{ij},$$

which, by Lemma 2.2, is equal to

$$\alpha\big[\hat{\boldsymbol{\Phi}}^{H}_{\boldsymbol{\eta}}(\boldsymbol{t} + \boldsymbol{x} - \boldsymbol{y})\big]_{ij} - 2\alpha\big[\hat{\boldsymbol{\Phi}}^{H}_{\boldsymbol{\eta}}(\boldsymbol{t})\big]_{ij} + \alpha\big[\hat{\boldsymbol{\Phi}}^{H}_{\boldsymbol{\eta}}(\boldsymbol{t} + \boldsymbol{y} - \boldsymbol{x})\big]_{ij} \tag{4.2}$$

with $\boldsymbol{\eta} = (\eta_{\mathrm{irr}}, \eta_{\mathrm{sol}})$ given by

$$\begin{aligned}\mathrm{e}^{\eta_{\mathrm{irr}}} &= \tfrac{2H+1}{2H+d}\mathrm{e}^{-2\mathrm{Re}\,\xi_{\mathrm{irr}}} + \tfrac{d-1}{2H+d}\mathrm{e}^{-2\mathrm{Re}\,\xi_{\mathrm{sol}}}; \\ \mathrm{e}^{\eta_{\mathrm{sol}}} &= \tfrac{1}{2H+d}\mathrm{e}^{-2\mathrm{Re}\,\xi_{\mathrm{irr}}} + \tfrac{2H+d-1}{2H+d}\mathrm{e}^{-2\mathrm{Re}\,\xi_{\mathrm{sol}}}.\end{aligned} \tag{4.3}$$

In particular, we find the ijth element of the variogram matrix by evaluating (4.2) at $\boldsymbol{t} = \mathbf{0}$. This, along with the even symmetry of $\hat{\boldsymbol{\Phi}}^{H}_{\boldsymbol{\eta}}$, yields

$$2\alpha\big[\hat{\boldsymbol{\Phi}}^{H}_{\boldsymbol{\eta}}(\boldsymbol{x} - \boldsymbol{y})\big]_{ij}$$

as the ijth element of the variogram. We have thus proved the following theorem.

THEOREM 4.2. *The variogram of a normalized vector fBm with parameters $H \in (0,1)$ and $\boldsymbol{\xi} = (\xi_{\mathrm{irr}}, \xi_{\mathrm{sol}})$ is*

$$\mathrm{Vario}[\boldsymbol{B}_{H,\boldsymbol{\xi}}](\boldsymbol{x}, \boldsymbol{y}) = 2\alpha\hat{\boldsymbol{\Phi}}^{H}_{(\eta_{\mathrm{irr}}, \eta_{\mathrm{sol}})}(\boldsymbol{x} - \boldsymbol{y}), \tag{4.4}$$

where the dependence of $(\eta_{\mathrm{irr}}, \eta_{\mathrm{sol}})$ on $(\xi_{\mathrm{irr}}, \xi_{\mathrm{sol}})$, H, and d is dictated by (4.3).

COROLLARY 4.3. *For* $0 < H < 1$ *we have*

$$\mathbf{E}\{\|\boldsymbol{B}_{H,\boldsymbol{\xi}}(\boldsymbol{x}) - \boldsymbol{B}_{H,\boldsymbol{\xi}}(\boldsymbol{y})\|^2\} = \alpha'\|\boldsymbol{x}-\boldsymbol{y}\|^{2H}$$

with

$$\alpha' = [\mathrm{e}^{-2\mathrm{Re}\,\xi_{\mathrm{irr}}} + (d-1)\mathrm{e}^{-2\mathrm{Re}\,\xi_{\mathrm{sol}}}]\alpha.$$

The proof is immediate, once it is observed that the above expectation is nothing but the trace of (4.4). We have thus shown that the new definition of fBm is consistent with (3.2).

We further remark that, by (4.3), $\mathrm{e}^{\xi_{\mathrm{irr}}} = \mathrm{e}^{\xi_{\mathrm{sol}}}$ implies $\mathrm{e}^{\eta_{\mathrm{irr}}} = \mathrm{e}^{\eta_{\mathrm{sol}}}$ (and vice versa). Consequently, in the case of classical fBm (where $\xi_{\mathrm{irr}} = \xi_{\mathrm{sol}}$) the variogram matrix is diagonal and the vector components are uncorrelated (and hence independent, due to Gaussianity).

4.6. Wavelet analysis and stationarity. The utility of wavelet analysis in studying fractal processes and turbulent flow has been noted frequently since the early days of wavelet theory, and the stationarizing effect of wavelet transforms on fBm has been widely documented [13, 16, 31, 34, 35, 60]. In this connection, an interesting observation can be made with regard to the scalar products of $\boldsymbol{B}_{H,\boldsymbol{\xi}}$ with test functions that have sufficiently many vanishing moments and, in particular, with respect to the representation of $\boldsymbol{B}_{H,\boldsymbol{\xi}}$ in a biorthogonal wavelet system.[8]

Let $\psi_{n,\boldsymbol{k}} \in \mathrm{L}^2$ and $\tilde{\psi}_{n,\boldsymbol{k}} \in \mathrm{L}^2$ symbolize the primal and dual basis functions of a biorthogonal wavelet system, with n denoting the resolution and $\boldsymbol{k}$ indicating position on a refinable lattice in $\mathbb{R}^d$.

By construction, all wavelets at a given resolution n are lattice shifts of one another ($\boldsymbol{k} \in \mathbb{Z}^d$ indexes the refinable lattice $\mathbf{Q}\mathbf{D}^{-n}\mathbb{Z}^d$ with dilation matrix $\mathbf{D} \in \mathbb{Z}^{d\times d}$, $|\det \mathbf{D}| > 1$):

$$\psi_{n,\boldsymbol{k}}(\boldsymbol{x}) = \psi_{n,\mathbf{0}}(\boldsymbol{x} - \mathbf{Q}\mathbf{D}^{-n}\boldsymbol{k}).$$

Consider the discrete random field $w_{n,i}$ defined by

$$w_{n,i}[\boldsymbol{k}] := \langle \boldsymbol{B}_{H,\boldsymbol{\xi}}, \hat{\boldsymbol{e}}_i\tilde{\psi}_{\boldsymbol{k}}\rangle.$$

Then

$$w_{n,i}[\boldsymbol{k}] = \langle \epsilon_H(-\acute{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\boldsymbol{\xi}}\boldsymbol{W}, \hat{\boldsymbol{e}}_i\tilde{\psi}_{n,\boldsymbol{k}}\rangle = \langle \epsilon_H\boldsymbol{W}, (-\grave{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\bar{\boldsymbol{\xi}}}\hat{\boldsymbol{e}}_i\tilde{\psi}_{n,\boldsymbol{k}}\rangle. \tag{4.5}$$

Assuming that $\tilde{\psi}_{n,\boldsymbol{k}}$ has vanishing moments (Fourier-domain zeros at $\boldsymbol{\omega} = 0$) up to degree $\lfloor H \rfloor$ so that $\mathrm{R}^{\frac{2H+d}{4}}\hat{\boldsymbol{e}}_i\tilde{\psi}_{n,\boldsymbol{k}} = \hat{\boldsymbol{e}}_i\tilde{\psi}_{n,\boldsymbol{k}}$, we have

$$(-\grave{\boldsymbol{\Delta}})^{-\frac{2H+d}{4}}_{-\bar{\boldsymbol{\xi}}}\hat{\boldsymbol{e}}_i\tilde{\psi}_{n,\boldsymbol{k}} \;=\; (2\pi)^{-d}\int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{j}\langle \boldsymbol{x},\boldsymbol{\omega}\rangle}\big[\hat{\boldsymbol{\Phi}}^{-\frac{2H+d}{4}}_{-\boldsymbol{\xi}}(\boldsymbol{\omega})\big]^{\mathsf{H}}_i\hat{\tilde{\psi}}_{n,\boldsymbol{k}}(\boldsymbol{\omega})\,\mathrm{d}\boldsymbol{\omega} \;\in\; \mathrm{L}^2;$$

[8] See Mallat [32] for detailed definitions and properties of wavelet systems. For an account of fractional-order splines and wavelets that are derived from fractional derivative operators see Unser and Blu [55]. A fundamental link between splines and fBm processes was studied in two papers by the same authors [4, 56] and extended to the multiparameter setting by Tafti, Van De Ville, and Unser [51]. The last reference also provides a detailed account of polyharmonic cardinal fractional splines in arbitrary dimensions and their connection with the wavelet analysis of scalar fBm fields.

and it becomes clear that the left inverse is shift invariant over this particular subspace of functions (in contrast to the general case of functions with nonvanishing moments for which it is not, due to the space-dependent operation of $\mathrm{R}^{\frac{2H+d}{4}}$). But then, since $\boldsymbol{W}$ is stationary, by (4.5) we may conclude that in this case the discrete random process $w_{n,i}[\cdot]$ is stationary.

In other words, a wavelet analysis of vector fBm with wavelets whose moments of degrees up to $\lfloor H \rfloor$ vanish yields stationary coefficients at each resolution.

For an overview of how matrix-valued wavelets can be used to estimate the parameters of vector fBm we refer the reader to Tafti et al. [49] and Tafti and Unser [50].

4.7. Link with the Helmholtz decomposition of vector fields. It is possible to study the divergence and curl of vector fBm (the latter for $d = 3$, where it is defined) using adjoint operators.

Taking an arbitrary scalar test function ϕ, for the divergence we have

$$\begin{aligned}
\langle \operatorname{div} \boldsymbol{B}_{H,\boldsymbol{\xi}}, \phi \rangle &= -\langle \boldsymbol{B}_{H,\boldsymbol{\xi}}, \mathbf{grad}\, \phi \rangle \\
&= -\langle \boldsymbol{W}, (-\grave{\boldsymbol{\Delta}})_{-\boldsymbol{\xi}}^{-\frac{H}{2}-\frac{d}{4}} \mathbf{grad}\, \phi \rangle \\
&= -\langle \boldsymbol{W}, \mathrm{e}^{-\overline{\xi_{\mathrm{irr}}}} (-\grave{\boldsymbol{\Delta}})_{\mathbf{0}}^{-\frac{H}{2}-\frac{d}{4}} \mathbf{grad}\, \phi \rangle \\
&= -\mathrm{e}^{-\overline{\xi_{\mathrm{irr}}}} \langle \boldsymbol{W}, (-\grave{\boldsymbol{\Delta}})_{\mathbf{0}}^{-\frac{H}{2}-\frac{d}{4}} \mathbf{grad}\, \phi \rangle, \qquad (4.6)
\end{aligned}$$

where the penultimate step can be verified easily in the Fourier domain as follows:

$$\begin{aligned}
(-\grave{\boldsymbol{\Delta}})_{-\boldsymbol{\xi}}^{-\frac{H}{2}-\frac{d}{4}} \mathbf{grad}\, \phi(\boldsymbol{x}) &= (2\pi)^{-d} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{j}\langle \boldsymbol{x}, \boldsymbol{\omega} \rangle} \|\boldsymbol{\omega}\|^{-H-\frac{d}{2}} \\
&\qquad \cdot \left[\mathrm{e}^{-\overline{\xi_{\mathrm{irr}}}} \frac{\boldsymbol{\omega}\boldsymbol{\omega}^{\mathsf{T}}}{\|\boldsymbol{\omega}\|^2} + \mathrm{e}^{-\overline{\xi_{\mathrm{sol}}}} \left(\mathbf{I} - \frac{\boldsymbol{\omega}\boldsymbol{\omega}^{\mathsf{T}}}{\|\boldsymbol{\omega}\|^2} \right) \right] (\mathrm{j}\boldsymbol{\omega}) \hat{\phi}(\boldsymbol{\omega})\, \mathrm{d}\boldsymbol{\omega} \\
&= (2\pi)^{-d} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{j}\langle \boldsymbol{x}, \boldsymbol{\omega} \rangle} (\mathrm{j}\boldsymbol{\omega}) \|\boldsymbol{\omega}\|^{-H-\frac{d}{2}} \left[\mathrm{e}^{-\overline{\xi_{\mathrm{irr}}}} + 0 \right] \mathrm{d}\boldsymbol{\omega} \\
&= \mathrm{e}^{-\overline{\xi_{\mathrm{irr}}}} (-\grave{\boldsymbol{\Delta}})_{\mathbf{0}}^{-\frac{H}{2}-\frac{d}{4}} \mathbf{grad}\, \phi.
\end{aligned}$$

Similarly, one can take an arbitrary vector test function $\boldsymbol{f}$ and write the following with regard to the curl:

$$\begin{aligned}
\langle \mathbf{curl}\, \boldsymbol{B}_{H,\boldsymbol{\xi}}, \boldsymbol{f} \rangle &= \langle \boldsymbol{B}_{H,\boldsymbol{\xi}}, \mathbf{curl}\, \boldsymbol{f} \rangle \\
&= \langle \boldsymbol{W}, (-\grave{\boldsymbol{\Delta}})_{-\boldsymbol{\xi}}^{-\frac{H}{2}-\frac{d}{4}} \mathbf{curl}\, \boldsymbol{f} \rangle \\
&= \langle \boldsymbol{W}, \mathrm{e}^{-\overline{\xi_{\mathrm{sol}}}} (-\grave{\boldsymbol{\Delta}})_{\mathbf{0}}^{-\frac{H}{2}-\frac{d}{4}} \mathbf{curl}\, \boldsymbol{f} \rangle \\
&= \mathrm{e}^{-\overline{\xi_{\mathrm{sol}}}} \langle \boldsymbol{W}, (-\grave{\boldsymbol{\Delta}})_{\mathbf{0}}^{-\frac{H}{2}-\frac{d}{4}} \mathbf{curl}\, \boldsymbol{f} \rangle. \qquad (4.7)
\end{aligned}$$

The derivation is comparable to that of the previous result, with the difference that one needs to use the Fourier matrix of the curl operator (cf. (2.8)).

We may then deduce from (4.6) that as $|\exp(-\xi_{\mathrm{irr}})| \to 0$, $\boldsymbol{B}_{H,\boldsymbol{\xi}}$ assumes a divergence-free nature. It follows likewise from (4.7) that as $|\exp(-\xi_{\mathrm{sol}})| \to 0$, $\boldsymbol{B}_{H,\boldsymbol{\xi}}$ becomes curl-free.

5. Simulation. The random vector fields that we have described can be simulated on a digital computer in several ways. A simple approach is to take the definition (3.4) in conjunction with (2.15) and apply the operator to simulated white Gaussian noise in the Fourier domain.

A more complex scheme can be set up by considering the scalar products of the vector field with *measurement* test functions and deriving the joint probability distributions of the resulting Gaussian samples, which can then be simulated using standard techniques. For example, a localized test function of the form $\hat{\boldsymbol{e}}_i\psi$ would measure the ith component of the field about a certain location. In implementing this scheme, one may for instance take these measurement functions to be wavelets and simulate the field in keeping with subsection 4.6, taking advantage of the fact that wavelet transform coefficients of vector fBm are stationary (cf. subsection 4.6). (In this connection, see also Elliott and Majda [12]. A Fourier-based technique for simulating processes with power-law spectra was presented in Viecelli and Canfield [61].)

The reader can find examples of simulated two-dimensional vector fBm in Figures 1 and 2. These figures were generated from a single 512×512 pseudorandom noise sequence with different values for the parameters H, ξ_{irr}, and ξ_{sol}; they are available in color only in the online version.

In each instance, we have provided two complementary visualizations. Images on the left were produced by a visualization technique known as line integral convolution (LIC), which consists of local directional smoothing of a noise image in the direction of flow [6] (we used Mathematica's implementation). In these images, more neutral tones indicate larger magnitudes. Arrows are superimposed in white.

In the images on the right, the hue angle encodes local direction, while the local amplitude of the field is indicated by the saturation level (smaller amplitudes are washed out).

The change in smoothness with increasing H is visible in these images, as is the clear effect of the parameters ξ_{irr} and ξ_{sol} on the directional behavior of the field, exhibiting nearly divergence-free and nearly rotation-free extremes as well as the middle ground.

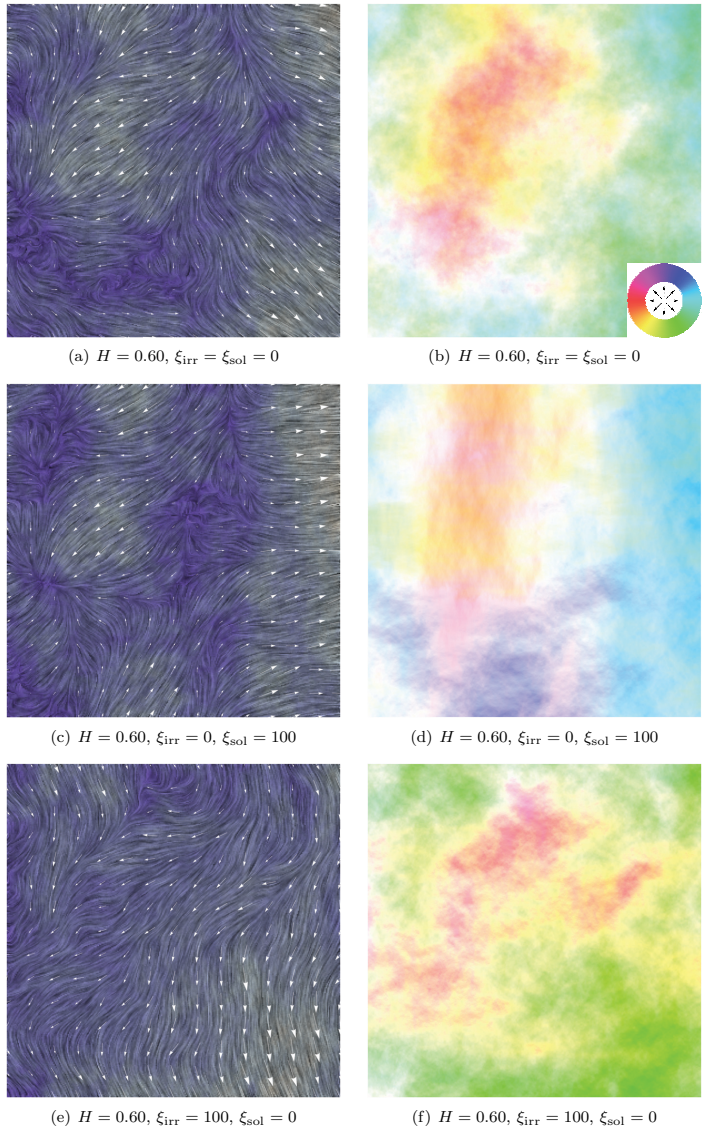
More examples can be found online at http://bigwww.epfl.ch/tafti/gal/vfBm/.

6. Conclusion. In this paper we introduced a family of random vector fields that extend fBm models by providing a means of correlating vector coordinates, which are independent in classical vector fBm models.

The first step in our investigation was to identify vector operators that are invariant under rotations and scalings of the coordinate system and can therefore be used to define random fields that are self-similar (fractal) and rotation invariant. The specific formulation of rotation invariance considered in the present work was inspired by the way physical vector fields transform under changes of coordinates. The operators identified in this step turned out to be generalizations of the vector Laplacian.

Our study of the said operators was aimed at characterizing random vector field models with the desired invariances as solutions of a whitening equation with the said generalized fractional Laplacians acting as whitening operators. To this end, we next addressed the problem of inverting the fractional Laplacian operators. This required us to introduce a new way of regularizing singular integrals, in order to define continuous inverse fractional Laplacian operators that are homogeneous and rotation invariant.

Once these inverse operators were identified, we were able to set the problem of characterizing the random models in the framework of Gel'fand and Vilenkin's theory of stochastic analysis. Specifically, we used the method of characteristic functionals to

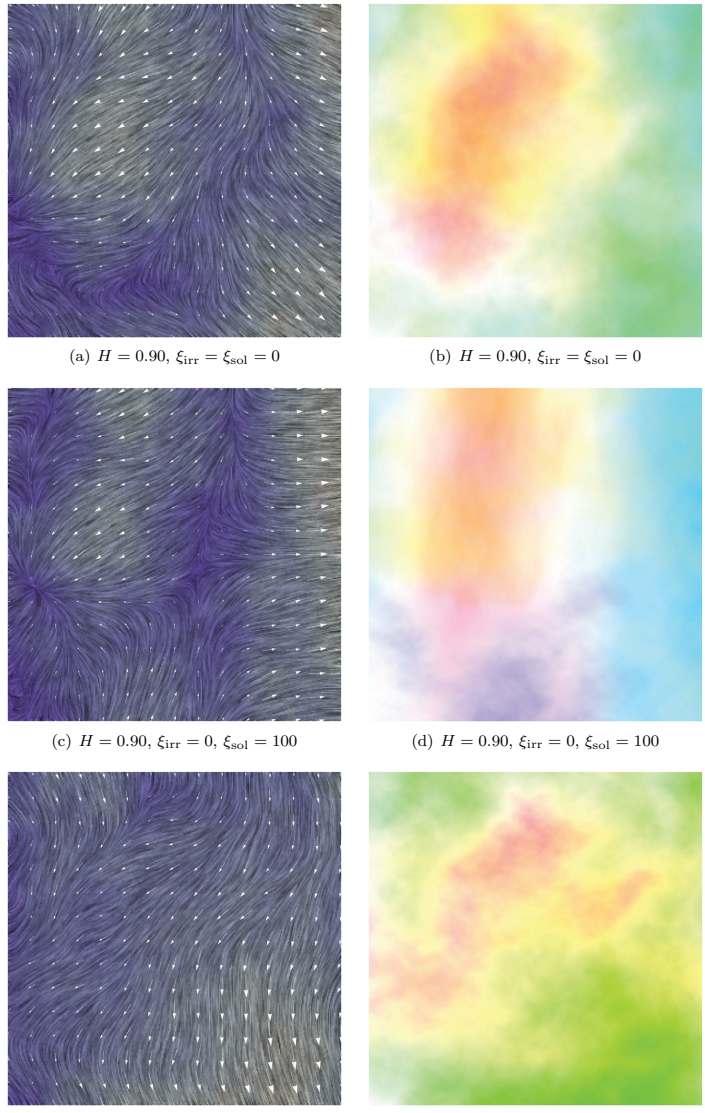


(a) $H = 0.60$, $\xi_{\text{irr}} = \xi_{\text{sol}} = 0$ (b) $H = 0.60$, $\xi_{\text{irr}} = \xi_{\text{sol}} = 0$

(c) $H = 0.60$, $\xi_{\text{irr}} = 0$, $\xi_{\text{sol}} = 100$ (d) $H = 0.60$, $\xi_{\text{irr}} = 0$, $\xi_{\text{sol}} = 100$

(e) $H = 0.60$, $\xi_{\text{irr}} = 100$, $\xi_{\text{sol}} = 0$ (f) $H = 0.60$, $\xi_{\text{irr}} = 100$, $\xi_{\text{sol}} = 0$

FIG. 1. *Simulated vector fBm with* $H = 0.6$ *and varying* ξ_{irr} *and* ξ_{sol}. *Left column: LIC visualization with arrows superimposed in white. Right column: directional behavior with local direction coded by the hue angle (see inset) and local amplitude represented by color saturation level (smaller amplitudes are bleached out).*



(a) $H = 0.90$, $\xi_{\text{irr}} = \xi_{\text{sol}} = 0$ (b) $H = 0.90$, $\xi_{\text{irr}} = \xi_{\text{sol}} = 0$

(c) $H = 0.90$, $\xi_{\text{irr}} = 0$, $\xi_{\text{sol}} = 100$ (d) $H = 0.90$, $\xi_{\text{irr}} = 0$, $\xi_{\text{sol}} = 100$

(e) $H = 0.90$, $\xi_{\text{irr}} = 100$, $\xi_{\text{sol}} = 0$ (f) $H = 0.90$, $\xi_{\text{irr}} = 100$, $\xi_{\text{sol}} = 0$

FIG. 2. *Simulated vector fBm with* $H = 0.9$ *and varying* ξ_{irr} *and* ξ_{sol}. *See the caption of Figure* 1 *for a description of the different visualizations.*

provide a complete probabilistic characterization of the new random vector fields. Using this methodology, we were also able to extend the definition of fractional Brownian fields to Hurst exponents beyond the usual range of $0 < H < 1$.

Similar to classical fBm, the fractional Brownian vector fields introduced in the present work are nonstationary but have stationary nth-order increments and can also be stationarized by means of wavelet analysis. In addition, in accordance with classical fBm models, these random fields exhibit statistical self-similarity (fractality) and rotation invariance, which are in fact properties they inherit from inverse fractional Laplacian operators. On the other hand, the directional properties of these new models have no scalar counterpart. Significantly, these models can exhibit a range of vectorial behavior, from completely irrotational (curl-free) to fully solenoidal (divergence-free).

Considering the versatility of these stochastic vector field models, potential stochastic modeling applications can exist in different disciplines such as fluid mechanics and turbulence physics, field theory, and image processing.

Appendix A. Proof of Lemma 2.2. Let

$$f_\lambda(\boldsymbol{\omega}) := 2^{-\frac{\lambda}{2}} \frac{\|\boldsymbol{\omega}\|^\lambda}{\Gamma(\frac{\lambda+d}{2})}.$$

We note the following facts concerning the above function:

$$\mathcal{F}^{-1}\{f_\lambda(\boldsymbol{\omega})\} = (2\pi)^{-\frac{d}{2}} f_{-\lambda-d}(\boldsymbol{x});$$

$$\frac{f_\lambda(\boldsymbol{\omega})}{\|\boldsymbol{\omega}\|^2} = \tfrac{1}{\lambda+d-2} f_{\lambda-2}(\boldsymbol{\omega});$$

$$\partial_i\partial_j f_\lambda(\boldsymbol{\omega}) = \tfrac{\lambda}{\lambda+d-2} f_{\lambda-2}(\boldsymbol{\omega}) \left[\delta_{ij} + (\lambda-2)\frac{\omega_i\omega_j}{\|\boldsymbol{\omega}\|^2}\right],$$

where δ_{ij} is Kronecker's delta.

The ijth element of $\hat{\boldsymbol{\Psi}}^\gamma_{\boldsymbol{\xi}}$ is

$$\left[\hat{\boldsymbol{\Psi}}^\gamma_{\boldsymbol{\xi}}(\boldsymbol{\omega})\right]_{ij} = f_{2\gamma}(\boldsymbol{\omega}) \left[\mathrm{e}^{\xi_{\mathrm{sol}}}\delta_{ij} - (\mathrm{e}^{\xi_{\mathrm{sol}}} - \mathrm{e}^{\xi_{\mathrm{irr}}})\frac{\omega_i\omega_j}{\|\boldsymbol{\omega}\|^2}\right].$$

Using the cited properties of f_λ we can write

$$\begin{aligned}
\mathcal{F}^{-1}\left\{\left[\hat{\boldsymbol{\Psi}}^\gamma_{\boldsymbol{\xi}}(\boldsymbol{\omega})\right]_{ij}\right\} &= (2\pi)^{-\frac{d}{2}}\left[\mathrm{e}^{\xi_{\mathrm{sol}}}\delta_{ij} f_{-2\gamma-d}(\boldsymbol{x}) + \tfrac{\mathrm{e}^{\xi_{\mathrm{sol}}}-\mathrm{e}^{\xi_{\mathrm{irr}}}}{2\gamma+d-2}\partial_i\partial_j f_{-2\gamma-d+2}(\boldsymbol{x})\right] \\
&= (2\pi)^{-\frac{d}{2}} f_{-2\gamma-d}(\boldsymbol{x})\left[\mathrm{e}^{\xi_{\mathrm{sol}}}\delta_{ij} + \tfrac{\mathrm{e}^{\xi_{\mathrm{sol}}}-\mathrm{e}^{\xi_{\mathrm{irr}}}}{2\gamma}\left(\delta_{ij} - (2\gamma+d)\frac{x_i x_j}{\|\boldsymbol{x}\|^2}\right)\right] \\
&= (2\pi)^{-\frac{d}{2}} f_{-2\gamma-d}(\boldsymbol{x})\left[\mathrm{e}^{\zeta_{\mathrm{irr}}}\frac{x_i x_j}{\|\boldsymbol{x}\|^2} + \mathrm{e}^{\zeta_{\mathrm{sol}}}\left(\delta_{ij} - \frac{x_i x_j}{\|\boldsymbol{x}\|^2}\right)\right]
\end{aligned}$$

with

$$\mathrm{e}^{\zeta_{\mathrm{irr}}} = \tfrac{2\gamma+d-1}{2\gamma}\mathrm{e}^{\xi_{\mathrm{irr}}} - \tfrac{d-1}{2\gamma}\mathrm{e}^{\xi_{\mathrm{sol}}} \quad \text{and} \quad \mathrm{e}^{\zeta_{\mathrm{sol}}} = -\tfrac{1}{2\gamma}\mathrm{e}^{\xi_{\mathrm{irr}}} + \tfrac{2\gamma+1}{2\gamma}\mathrm{e}^{\xi_{\mathrm{sol}}}. \qquad \square$$

Appendix B. Conjugacy of $(-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}}$ and $(-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}}$. We proceed to show that for all test functions $\boldsymbol{f}$ and $\boldsymbol{g} \in \mathcal{S}^d$,

$$\langle(-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}}\boldsymbol{f}, \boldsymbol{g}\rangle = \langle \boldsymbol{f}, (-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}}\boldsymbol{g}\rangle.$$

Using Parseval's identity and the definition of $(-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}}$ in (2.14) we can write

$$\begin{aligned}\langle \boldsymbol{f}, (-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}} \boldsymbol{g}\rangle &= (2\pi)^{-d} \int_{\mathbb{R}^d} [\hat{\boldsymbol{f}}(\boldsymbol{\omega})]^{\mathsf{H}} \hat{\boldsymbol{\Phi}}^{-\gamma}_{-\boldsymbol{\xi}}(\boldsymbol{\omega}) [\mathrm{R}^{\gamma} \hat{\boldsymbol{g}}](\boldsymbol{\omega}) \, \mathrm{d}\boldsymbol{\omega} \\ &= (2\pi)^{-d} \int_{\mathbb{R}^d} [\hat{\boldsymbol{f}}(\boldsymbol{\omega})]^{\mathsf{H}} \hat{\boldsymbol{\Phi}}^{-\gamma}_{-\boldsymbol{\xi}}(\boldsymbol{\omega}) \left[\hat{\boldsymbol{g}}(\boldsymbol{\omega}) - \sum_{|\boldsymbol{k}| \leq \lfloor 2\gamma - \frac{d}{2} \rfloor} \mathbf{T}_{\boldsymbol{k}}[\hat{\boldsymbol{g}}] \boldsymbol{\omega}^{\boldsymbol{k}} \right] \mathrm{d}\boldsymbol{\omega}. \end{aligned} \tag{B.1}$$

Moreover,

$$\mathbf{T}_{\boldsymbol{k}}[\hat{\boldsymbol{g}}] = \frac{\hat{\boldsymbol{g}}^{(\boldsymbol{k})}(\mathbf{0})}{\boldsymbol{k}!} = \int_{\mathbb{R}^d} \frac{(-\mathrm{j})^{\boldsymbol{k}} \boldsymbol{x}^{\boldsymbol{k}}}{\boldsymbol{k}!} \boldsymbol{g}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}.$$

By combining this and (B.1) we get

$$\begin{aligned}\langle \boldsymbol{f}, (-\grave{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}} \boldsymbol{g}\rangle &= (2\pi)^{-d} \int_{\mathbb{R}^d} [\hat{\boldsymbol{f}}(\boldsymbol{\omega})]^{\mathsf{H}} \hat{\boldsymbol{\Phi}}^{-\gamma}_{-\boldsymbol{\xi}}(\boldsymbol{\omega}) \\ &\qquad \cdot \left[\int_{\mathbb{R}^d} \left(\mathrm{e}^{-\mathrm{j}\langle \boldsymbol{x}, \boldsymbol{\omega} \rangle} - \sum_{|\boldsymbol{k}| \leq \lfloor 2\gamma - \frac{d}{2} \rfloor} \frac{(-\mathrm{j})^{\boldsymbol{k}} \boldsymbol{x}^{\boldsymbol{k}} \boldsymbol{\omega}^{\boldsymbol{k}}}{\boldsymbol{k}!} \right) \boldsymbol{g}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \right] \mathrm{d}\boldsymbol{\omega} \\ &= \int_{\mathbb{R}^d} \left[(-\acute{\boldsymbol{\Delta}})^{-\gamma}_{-\boldsymbol{\xi}} \boldsymbol{f}(\boldsymbol{x}) \right]^{\mathsf{H}} \boldsymbol{g}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}, \end{aligned}$$

where the last step follows from exchanging the order of integration and using the definition of the right inverse given in (2.15), together with the identity $\left[\hat{\boldsymbol{\Phi}}^{-\gamma}_{-\boldsymbol{\xi}}(\boldsymbol{\omega})\right]^{\mathsf{H}} = \hat{\boldsymbol{\Phi}}^{-\gamma}_{-\bar{\boldsymbol{\xi}}}(\boldsymbol{\omega})$. The last integral is equal to the scalar product $\langle (-\acute{\boldsymbol{\Delta}})^{-\gamma}_{-\bar{\boldsymbol{\xi}}} \boldsymbol{f}, \boldsymbol{g} \rangle$ by definition. □

Acknowledgments. The authors thank the anonymous referees for their detailed and helpful comments, which have helped improve the paper. They also thank Dr. P. Thévenaz for help with the illustrations.

REFERENCES

[1] M. Arigovindan, *Variational Reconstruction of Vector and Scalar Images from Non-uniform Samples*, Ph.D. thesis, EPFL, Biomedical Imaging Group, Lausanne, Switzerland, http://bigwww.epfl.ch/publications/arigovindan0502.html (2005).

[2] M. Avellaneda and A. J. Majda, *Simple examples with features of renormalization for turbulent transport*, Philos. Trans. Roy. Soc. London Ser. A, 346 (1994), pp. 205–233.

[3] A. Benassi, S. Jaffard, and D. Roux, *Elliptic Gaussian random processes*, Rev. Mat. Iberoamericana, 13 (1997), pp. 19–90.

[4] T. Blu and M. Unser, *Self-similarity: Part* II*—Optimal estimation of fractal processes*, IEEE Trans. Signal Process., 55 (2007), pp. 1364–1378.

[5] G. Bourdaud, *Localisation et multiplicateurs des espaces de Sobolev homogenes*, Manuscripta Math., 60 (1988), pp. 93–130.

[6] B. Cabral and L. Leedom, *Imaging vector fields using line integral convolution*, in SIGGRAPH '93: Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques, ACM, New York, 1993, pp. 263–270.

[7] R. Carmona, *Renormalization theories for incompressible flows with Gaussian statistics*, in Monte Carlo Simulations in Oceanography: Proceedings of the 9th 'Aha Huliko'a Hawaiian Winter Workshop, University of Hawaii at Manoa, Honolulu, HI, 1997, pp. 81–89.

[8] A. Craya, *Contribution à l'analyse de la turbulence associée à des vitesses moyennes*, Publ. Sci. Tech. Ministère de l'Air (Paris) 345, SDIT, Paris, 1958.

[9] E. Deriaz, M. Farge, and K. Schneider, *Craya decomposition using compactly supported biorthogonal wavelets*, Appl. Comput. Harmon. Anal., 28 (2010), pp. 267–284.
[10] E. Deriaz and V. Perrier, *Orthogonal Helmholtz decomposition in arbitrary dimension using divergence-free and curl-free wavelets*, Appl. Comput. Harmon. Anal., 26 (2009), pp. 249–269.
[11] R. L. Dobrushin, *Gaussian and their subordinated self-similar random generalized fields*, Ann. Probab., 7 (1979), pp. 1–28.
[12] F. W. Elliott, Jr., and A. J. Majda, *A wavelet Monte Carlo method for turbulent diffusion with many spatial scales*, J. Comput. Phys., 113 (1994), pp. 82–111.
[13] M. Farge, *Wavelet transforms and their applications to turbulence*, in Annual Review of Fluid Mechanics 24, Annual Reviews, Palo Alto, CA, 1992, pp. 395–457.
[14] M. Farge, N. Kevlahan, V. Perrier, and É. Goirand, *Wavelets and turbulence*, Proc. IEEE, 84 (1996), pp. 639–669.
[15] D. J. Field, *Scale-invariance and self-similar "wavelet" transforms: An analysis of natural scenes and mammalian visual systems*, in Wavelets, Fractals, and Fourier Transforms, M. Farge, J. C. R. Hunt, and J. C. Vassilicos, eds., Clarendon Press, Oxford, UK, 1993.
[16] P. Flandrin, *Wavelet analysis and synthesis of fractional Brownian motion*, IEEE Trans. Inform. Theory, 38 (1992), pp. 910–916.
[17] I. M. Gel'fand and G. E. Shilov, *Generalized Functions. Volume* I: *Properties and Operations*, Academic Press, New York, 1964.
[18] I. M. Gel'fand and N. Ya. Vilenkin, *Generalized Functions. Volume* IV: *Applications of Harmonic Analysis*, Academic Press, New York, 1964.
[19] T. Hida, *Harmonic analysis on the space of generalized functions*, Theory Probab. Appl., 15 (1970), pp. 117–122.
[20] L. Hörmander, *The Analysis of Linear Partial Differential Operators*. I. *Distribution Theory and Fourier Analysis*, Grundlehren Math. Wiss. 256, Springer-Verlag, Berlin, 1983.
[21] H. E. Hurst, *Long-term storage capacity of reservoirs*, Trans. Amer. Soc. Civil Engineers, 116 (1951), pp. 770–808.
[22] J.-P. Kahane, *Some Random Series of Functions*, 2nd ed., Cambridge University Press, Cambridge, UK, 1985.
[23] M. Ya. Kelbert, N. N. Leonenko, and M. D. Ruiz-Medina, *Fractional random fields associated with stochastic fractional heat equations*, Adv. in Appl. Probab., 37 (2005), pp. 108–133.
[24] M. S. Keshner, $1/f$ *noise*, Proc. IEEE, 70 (1982), pp. 212–218.
[25] V. Klyatskin, W. Woyczynski, and D. Gurarie, *Diffusing passive tracers in random incompressible flows: Statistical topography aspects*, J. Statist. Phys., 84 (1996), pp. 797–836.
[26] A. N. Kolmogorov, *A note on the papers of R. A. Minlos and V. Sazonov*, Theory Probab. Appl., 4 (1959), pp. 221–223.
[27] W. E. Leland, W. Willinger, M. S. Taqqu, and D. V. Wilson, *On the self-similar nature of ethernet traffic*, SIGCOMM Comput. Commun. Rev., 25 (1995), pp. 202–213.
[28] N. Leonenko, *Limit Theorems for Random Fields with Singular Spectrum*, Math. Appl. 465, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999.
[29] P. Lévy, *Le mouvement brownien*, Mémor. Sci. Math. 126, Gauthier–Villars, Paris, 1954.
[30] P. Lévy, *Processus stochastiques et mouvement brownien*, 2ème ed., Gauthier–Villars Paris, 1965; reprinted, Éditions Jacques Gabay, Sceaux, France, 1992.
[31] S. Mallat, *A theory for multiresolution signal decomposition: The wavelet representation*, IEEE Trans. Pattern Anal. Mach. Intell., 11 (1989), pp. 674–693.
[32] S. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed., Academic Press, San Diego, CA, 1999.
[33] B. B. Mandelbrot, *Gaussian self-affinity and fractals*, in Globality, The Earth, $1/f$ Noise, and R/S, Springer-Verlag, New York, 2002.
[34] E. Masry, *The wavelet transform of stochastic processes with stationary increments and its application to fractional Brownian motion*, IEEE Trans. Inform. Theory, 39 (1993), pp. 260–264.
[35] Y. Meyer, F. Sellan, and M. Taqqu, *Wavelets, generalized white noise and fractional integration: The synthesis of fractional Brownian motion*, J. Fourier Anal. Appl., 5 (1999), pp. 465–494.
[36] R. A. Minlos, *Generalized random processes and their extension to a measure*, in Selected Transl. Math. Statist. and Prob., Vol. 3, AMS, Providence, RI, 1963, pp. 291–313.
[37] A. S. Monin and A. M. Yaglom, *Statistical Fluid Mechanics: Mechanics of Turbulence. Vol.* II, MIT Press, Cambridge, MA, 1975; reprinted, Dover, Mineola, NY, 2007.
[38] S. A. Orszag, *Numerical methods for the simulation of turbulence*, Phys. Fluids, 12 (1969),

Suppl. II, pp. 250–257.
[39] A. P. Pentland, *Fractal-based description of natural scenes*, IEEE Trans. Pattern Anal. Mach. Intell., 6 (1984), pp. 661–674.
[40] B. Pesquet-Popescu and J. Lévy-Véhel, *Stochastic fractal models for image processing*, IEEE Signal Process. Mag., 19 (2002), pp. 48–62.
[41] A. Pietsch, *Nuclear Locally Convex Spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete 66, Springer-Verlag, New York, 1972.
[42] M. S. Pinsker, *The theory of curves in Hilbert space with stationary nth increments*, Izv. Akad. Nauk SSSR. Ser. Mat., 19 (1955), pp. 319–344 (in Russian).
[43] Yu. V. Prohorov, *The method of characteristic functionals*, in Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 2, University of California Press, Berkeley, CA, 1961, pp. 403–419; also available online from http://projecteuclid.org/euclid.bsmsp/1200512613.
[44] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed., McGraw–Hill, New York, 1976.
[45] M. D. Ruiz-Medina, V. V. Anh, and J. M. Angulo, *Fractional generalized random fields of variable order*, Stochastic Anal. Appl., 22 (2004), pp. 775–799.
[46] P. Sagaut and C. Cambon, *Homogeneous Turbulence Dynamics*, Cambridge University Press, Cambridge, UK, 2008.
[47] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, Chapman and Hall, New York, 1994.
[48] N. Sarshar and X. Wu, *On rate-distortion models for natural images and wavelet coding performance*, IEEE Trans. Image Process., 16 (2007), pp. 1383–1394.
[49] P. D. Tafti, R. Delgado-Gonzalo, A. F. Stalder, and M. Unser, *Fractal modelling and analysis of flow-field images*, in Proceedings of the Seventh IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI 2010), Rotterdam, The Netherlands, 2010, pp. 49–52.
[50] P. D. Tafti and M. Unser, *Self-similar random vector fields and their wavelet analysis*, in Proceedings of the SPIE Conference on Mathematical Imaging: Wavelets XIII, M. Papadakis, V. K. Goyal, and D. Van De Ville, eds., Vol. 7446, San Diego, CA, 2009, pp. 74460Y–1–74460Y–8.
[51] P. D. Tafti, D. Van De Ville, and M. Unser, *Invariances, Laplacian-like wavelet bases, and the whitening of fractal processes*, IEEE Trans. Image Process., 18 (2009), pp. 689–702.
[52] M. Talagrand, *Pettis integral and measure theory*, Mem. Amer. Math. Soc., 51 (1984), no. 307.
[53] M. S. Taqqu, *Fractional Brownian motion and long-range dependence*, in Theory and Applications of Long-Range Dependence, P. Doukhan, G. Oppenheim, and M. S. Taqqu, eds., Birkhäuser Boston, Boston, MA, 2003, pp. 5–38.
[54] M. S. Taqqu, V. Teverovsky, and W. Willinger, *Is network traffic self-similar or multifractal?*, Fractals, 5 (1997), pp. 63–73.
[55] M. Unser and T. Blu, *Fractional splines and wavelets*, SIAM Rev., 42 (2000), pp. 43–67.
[56] M. Unser and T. Blu, *Self-similarity: Part I—Splines and operators*, IEEE Trans. Signal Process., 55 (2007), pp. 1352–1363.
[57] M. Unser and P. D. Tafti, *Stochastic Models for Sparse and Piecewise-Smooth Signals*, manuscript, in review.
[58] N. N. Vakhania, *Probability Distributions on Linear Spaces*, North–Holland Ser. Probab. Appl. Math., North–Holland, New York, 1981.
[59] B. Vedel, *Confinement of the infrared divergence for the Mumford process*, Appl. Comput. Harmon. Anal., 21 (2006), pp. 305–323.
[60] D. Veitch and P. Abry, *A wavelet-based joint estimator of the parameters of long-range dependence*, IEEE Trans. Inform. Theory, 45 (1999), pp. 878–897.
[61] J. A. Viecelli and E. H. Canfield, *Functional representation of power-law random fields and time series*, J. Comput. Phys., 95 (1991), pp. 29–39.
[62] A. M. Yaglom, *Correlation Theory of Stationary and Related Random Functions. Volume* I. *Basic Results*, Springer Ser. Statist., Springer-Verlag, New York, 1987.
[63] A. M. Yaglom and M. S. Pinsker, *Random processes with stationary increments of the nth order*, Doklady Akad. Nauk SSSR (N.S.), 90 (1953), pp. 731–734.

Partial List of Symbols

The following is a partial list of some of the more volatile and/or mysterious notations used in the thesis (the majority are from Chapter 2). To complement this list, we remind the reader that in the thesis we have made free use of the multi-index notation (see the entry for $\mathrm{U}_{c,n}^{-\lambda}$ below) and also employed a loose form of Einstein's tensor notation from time to time (see **2.bk**).

$\mathbb{1}_X$	Indicator function of the set X: $\mathbb{1}_X(x) = 1$ if $x \in X$ and $= 0$ otherwise.
X^Y	Product set/space: $X^Y = \prod_{y \in Y} X = \{f : Y \to X\}$.
$\pi_I : X^Y \to X^I$	Projection operator mapping $f : Y \to X$ to its restriction to $I \subset Y$.
$\mathscr{P}_F$	Finitely- or countably-additive probability law associated with some model/process/field F.
$\widehat{\mathscr{P}_F}$	Characteristic function(al) of the law $\mathscr{P}_F$.
$\xrightarrow{w}$	Convergence in probability law/measure.
$\mathbb{R}$	Reals.
$\mathbb{R}_+$	Positive reals.
$\mathbb{C}$	Complex numbers.
$\mathbb{N}$	Natural numbers: $\mathbb{N} = \{1, 2, 3, \ldots\}$.
$\mathbb{N}_0$	$\mathbb{N}_0 = \mathbb{N} \cup \{0\} = \{0, 1, 2, 3, \ldots\}$.
$\mathbb{Z}$	Integers.
d	Dimension of the domain.

$\dot{\mathbb{R}}^d$ — Punctured Euclidean domain (**2.b**): $\dot{\mathbb{R}}^d = \mathbb{R}^d \backslash \{0\}$.

T_τ, $\tau \in T = \mathbb{R}^d$ — Translation by τ (**2.n**):

$$\mathrm{T}_\tau : f \mapsto f(\cdot - \tau).$$

S_σ, $\sigma \in \Sigma = \mathbb{R}_+$ — Scaling by σ (**2.n**):

$$\mathrm{S}_\sigma : f \mapsto \sigma^{-\frac{d}{2}} f(\sigma^{-1}\cdot).$$

$\mathrm{R}_{\omega,s}$, $\omega \in \mathrm{O}(d)$ — Scalar rotation by the orthogonal matrix ω (**2.n**):

$$\mathrm{R}_{\omega,s} : f \mapsto f(\omega^{\mathrm{T}}\cdot).$$

$\mathrm{R}_{\omega,v}$, $\omega \in \mathrm{O}(d)$ — Vector rotation by the orthogonal matrix ω (**2.n**):

$$\mathrm{R}_{\omega,v} : f \mapsto \omega f(\omega^{\mathrm{T}}\cdot).$$

O — Inversion with respect to the origin in $\dot{\mathbb{R}}^d$ (**2.n**):

$$\mathrm{O} : f(x) \mapsto |x|^{-d} f\left(\frac{x}{|x|^2}\right).$$

$\mathrm{O}(d)$ — Real orthogonal group of degree n; identified with the group of orthogonal matrices in $\mathbb{R}^{d\times d}$.

$\widehat{\phi} = \phi^\wedge$ — Fourier transform of ϕ:

$$\widehat{\phi}(\xi) = \phi^\wedge(\xi) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}\langle x,\xi\rangle} \phi(x)\, \mathrm{d}x.$$

$\mathscr{D}(\mathbb{R}^d)$ — Space of compactly-supported smooth test functions with its standard topology.

$\mathscr{S}(\mathbb{R}^d)$ — Schwartz space of rapidly decaying smooth functions with its standard topology.

$\mathscr{E}$ — Place-holder for $\mathscr{D}$ and $\mathscr{S}$ (and occasionally also for other nuclear spaces).

$\mathscr{E}(\dot{\mathbb{R}}^d)$ — Space of functions $\phi \in \mathscr{E}(\mathbb{R}^d)$ such that $\mathrm{O}\phi \in \mathscr{E}(\mathbb{R}^d)$ as well (**2.o**).

$\widehat{\mathscr{E}}(\dot{\mathbb{R}}^d)$ — Space of functions whose Fourier transform belongs to $\mathscr{E}(\dot{\mathbb{R}}^d)$ (**2.o**).

δ — Dirac's delta distribution.

δ_{ij} — Kronecker's delta: $\delta_{ij} = 1$ if $i = j$ and $= 0$ otherwise.

Scalar homogeneous and rotation-invariant distributions

ρ_c^λ, $\lambda, c \in \mathbb{C}$ — Homogeneous isotropic distributions in $\mathscr{E}(\mathbb{R}^d)$ (**2.aa**):

$$\rho_c^\lambda(x) := c\frac{|x|^\lambda}{2^{\frac{\lambda}{2}}\Gamma(\frac{\lambda+d}{2})}.$$

For $\lambda = -d - 2m$, $m \in \mathbb{N}$, we have

$$\rho_c^{-d-2m} = c\frac{(2\pi)^{\frac{d}{2}}}{2^{m+1}\Gamma(m+\frac{d}{2})}(-\Delta)^m\delta.$$

In addition, the following identities hold (**2.ae**, **2.af**):

$$\mathrm{S}_\sigma\rho_c^\lambda = \sigma^{-\frac{d}{2}-\lambda}\rho_c^\lambda; \qquad \text{(homogeneity)}$$

$$\mathrm{R}_{\omega,s}\rho_c^\lambda = \rho_c^\lambda; \qquad \text{(scalar rotation-invariance)}$$

$$(\rho_c^\lambda)^\wedge = \rho_c^{-d-\lambda} = \rho_c^{\widehat{\lambda}}, \qquad \text{with } \widehat{\lambda} := -d-\lambda; \qquad \text{(Fourier transform)}$$

$$(\widehat{\lambda}+2)\partial_i\rho_c^\lambda(x) = -\lambda x_i\rho_c^{\lambda-2}(x); \qquad \text{(First-order derivatives)}$$

$$\Delta\,\rho_c^\lambda = \lambda\rho_c^{\lambda-2}; \qquad \text{(Laplacian)}$$

$$\delta_{ij}\rho_c^\lambda = -\frac{\widehat{\lambda}}{\lambda+2}\partial_{ij}\rho_c^{\lambda+2} + \frac{\lambda}{\widehat{\lambda}+2}x_ix_j\rho_c^{\lambda-2} = -\frac{\widehat{\lambda}}{\lambda+2}\partial_{ij}\rho_c^{\lambda+2} - \frac{\lambda}{\widehat{\lambda}+2}\left(\partial_{ij}\rho_c^{\widehat{\lambda}+2}\right)^\wedge. \qquad \text{(Second-order derivatives)}$$

Additional formulae for products and convolutions can be found in **2.ag**.

Tensor homogeneous and rotation-invariant distributions

$P^\lambda_{\underline{\bullet}}$, $\lambda \in \mathbb{C}$

Various parametrizations of homogeneous vector-rotation-invariant tensor distributions in $(\mathscr{E}')^{d\times d}(\mathbb{R}^d)$ (**2.av**), with $\underline{\bullet}$ a place-holder for one of three different parametrizations, by $\underline{s} = (s_1, s_2) \in \mathbb{C}^2$, $\underline{r} = (r_1, r_2)$, or $\underline{k} = (k_1, k_2)$:

$$\begin{aligned}
[P^\lambda_{\underline{s}}]_{ij} &:= \partial_{ij}|x|^2\rho^\lambda_{s_1} + x_i x_j \,\Delta\, \rho^\lambda_{s_2} \\
&= -\widehat{\lambda}\partial_{ij}\rho^{\lambda+2}_{s_1} + \lambda x_i x_j \rho^{\lambda-2}_{s_2} \\
&= -\widehat{\lambda}\partial_{ij}\rho^{\lambda+2}_{s_1} - \lambda\left(\partial_{ij}\rho^{\widehat{\lambda}+2}_{s_2}\right)^\wedge \\
&= -\widehat{\lambda}(\lambda+2)\partial_{ij}\,\Delta^{-1}\,\rho^\lambda_{s_1} - \lambda(\widehat{\lambda}+2)\frac{x_i x_j}{|x|^2}\rho^\lambda_{s_2} \\
[P^\lambda_{\underline{r}}]_{ij} &:= \partial_{ij}\,\Delta^{-1}\,\rho^\lambda_{r_1} + \left(\delta_{ij} - \partial_{ij}\,\Delta^{-1}\right)\rho^\lambda_{r_2} \\
&= \delta_{ij}\rho^\lambda_{r_2} + \frac{1}{\lambda+2}\partial_{ij}\rho^{\lambda+2}_{r_1-r_2}; \\
[P^\lambda_{\underline{k}}]_{ij} &:= \frac{x_i x_j}{|x|^2}\rho^\lambda_{k_1} + \left(\delta_{ij} - \frac{x_i x_j}{|x|^2}\right)\rho^\lambda_{k_2} \\
&= \delta_{ij}\rho^\lambda_{k_2} - \frac{1}{\widehat{\lambda}+2}x_i x_j \rho^{\lambda-2}_{k_1-k_2};
\end{aligned}$$

with $\widehat{\lambda} = -d - \lambda$ as before.
Different parametrizations of the same distribution are related thus:

$$\begin{aligned}
\begin{pmatrix} r_1 \\ r_2 \end{pmatrix} &= \begin{pmatrix} -\widehat{\lambda}(\lambda+2) & (\widehat{\lambda}+1)(\widehat{\lambda}+2) \\ 0 & \widehat{\lambda}+2 \end{pmatrix}\begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \\
&= \frac{1}{\widehat{\lambda}}\begin{pmatrix} -\lambda-1 & 1-d \\ -1 & \widehat{\lambda}+1 \end{pmatrix}\begin{pmatrix} k_1 \\ k_2 \end{pmatrix}; \\
\begin{pmatrix} k_1 \\ k_2 \end{pmatrix} &= \begin{pmatrix} (\lambda+1)(\lambda+2) & -\lambda(\widehat{\lambda}+2) \\ \lambda+2 & 0 \end{pmatrix}\begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \\
&= \frac{1}{\lambda}\begin{pmatrix} -\widehat{\lambda}-1 & 1-d \\ -1 & \lambda+1 \end{pmatrix}\begin{pmatrix} r_1 \\ r_2 \end{pmatrix}.
\end{aligned}$$

The following properties are fulfilled:

$$\begin{aligned}
&\mathrm{S}_\sigma P^\lambda_{\underline{\bullet}} = \sigma^{-\frac{d}{2}-\lambda}P^\lambda_{\underline{\bullet}}; && \text{(homogeneity)} \\
&\mathrm{R}_{\omega,v}P^\lambda_{\underline{\bullet}} = P^\lambda_{\underline{\bullet}}\omega; && \text{(vector rotation-invariance)} \\
&(P^\lambda_{(s_1,s_2)})^\wedge = P^{\widehat{\lambda}}_{(s_2,s_1)}. && \text{(Fourier transform)}
\end{aligned}$$

Formulae for the Fourier transform of the other parametrizations are given in **2.ay**. Product and convolution formulae can be found in **2.ba**. For results concerning the vector derivatives and Helmholtz decomposition of $P^\lambda_{\underline{\bullet}}$, see the corresponding properties of the associated convolution operator in **§2.4**.

Operators associated with homogeneous and rotation-invariant distributions

U^λ_c

Scalar homogeneous and translation- and rotation-invariant convolution operator associated with ρ^λ_c (**2.ac**):

$$\mathrm{U}^\lambda_c : \phi \;\mapsto\; \phi * \widehat{\rho^\lambda_c} = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\langle\cdot,\xi\rangle}\; \rho^\lambda_c(\xi)\; \widehat{\phi}(\xi)\; \mathrm{d}\xi.$$

$\mathrm{U}^{-\lambda}_{c,n}$

L_p-continuous homogeneous and rotation-invariant modification of the previous operator (**§2.2.2**):

$$\mathrm{U}^{-\lambda}_{c,n} = \mathrm{U}^{-\lambda}_c - \mathrm{Reg}^{-\lambda}_{c,n}$$

with $n = \lfloor \mathrm{Re}\,\lambda + \frac{d}{p} \rfloor - d$, $\mathrm{Re}\,\lambda + \frac{d}{p} \notin \mathbb{N}$, and with $\mathrm{Reg}^{-\lambda}_{c,n}$ defined as

$$\mathrm{Reg}^\lambda_{c,n} \;:\; \phi \;\mapsto\; \sum_{|k|\le n} \frac{\partial_k \widehat{\rho^\lambda_c}}{k!} \int_{\mathbb{R}^d} (-y)^k \phi(y)\; \mathrm{d}y,$$

where k is a multi-index in $\mathbb{N}^d_0$ with $|k| := \sum_i k_i$, $(-y)^k := \prod_i (-y_i)^{k_i}$, and $k! := \prod_i k_i!$, and the sum is zero when empty (**2.ai**).
The $n+1$st finite differences of $\mathrm{U}^{-\lambda}_{c,n}$ are the same as those of the unmodified operator $\underline{\mathrm{U}}^{-\lambda}_c$ (**2.ao**).

$\underline{\mathrm{U}}^\lambda_{\underline{\bullet}}$

Vector homogeneous and translation- and rotation-invariant convolution operator associated with $P^\lambda_{\underline{\bullet}}$ (**2.bb**):

$$\underline{\mathrm{U}}^\lambda_{\underline{\bullet}} : \phi \;\mapsto\; \phi * (P^\lambda_{\underline{\bullet}})^\wedge = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\langle\cdot,\xi\rangle}\; P^\lambda_{\underline{\bullet}}(\xi)\; \widehat{\phi}(\xi)\; \mathrm{d}\xi.$$

$\underline{U}^{-\lambda}_{\underline{\bullet},n}$ — L_p-continuous homogeneous and rotation-invariant modification of $\underline{U}^{-\lambda}_{\underline{\bullet}}$ (**§2.3.2**):

$$\underline{U}^{-\lambda}_{\underline{r},n} = \underline{U}^{-\lambda}_{\underline{r}} - \underline{\mathrm{Reg}}^{-\lambda}_{\underline{r}}$$

with n related to p as in the scalar case and $\underline{\mathrm{Reg}}^{-\lambda}_{\underline{r}}$ defined in **2.bg**.
The $n+1$st finite difference of $\underline{U}^{-\lambda}_{\underline{\bullet},n}$ (in any combination of directions) is the same as that of the unmodified operator $\underline{U}^{-\lambda}_{\underline{\bullet}}$ (**2.bi**).
Other properties of $\underline{U}^{-\lambda}_{\underline{\bullet},n}$ in connection with vector derivatives and the Helmholtz decomposition are discussed in **§2.4**.

References

[AB91] L. Amodei and M. N. Benbourhim, *A vector spline approximation*, J. Approx. Theory **67** (1991), no. 1, 51–79.

[AGL05] J. Ahrens, B. Geveci, and C. Law, *Paraview: An end-user tool for large-data visualization*, Visualization Handbook (C. D. Hansen and C. R. Johnson, eds.), Butterworth-Heinemann, Burlington, 2005, pp. 717–732.

[AGLV09] P. Abry, P. Gonçalves, and J. Lévy Véhel (eds.), *Scaling, fractals and wavelets*, Digital Signal and Image Processing Series, ISTE, London, 2009.

[And95] T. Ando, *Matrix Young inequalities*, Operator theory in function spaces and Banach lattices, Oper. Theory Adv. Appl., vol. 75, Birkhäuser, Basel, 1995, pp. 33–38.

[Ari05] M. Arigovindan, *Variational reconstruction of vector and scalar images from non-uniform samples*, Ph.D. thesis, École Polytechnique Fédérale de Lausanne, 2005.

[ASJ+07] M. Arigovindan, M. Sühling, C. Jansen, P. Hunziker, and M. Unser, *Full motion and flow field recovery from echo Doppler data*, IEEE Trans. Med. Imaging **26** (2007), no. 1, 31–45.

[BCI02] A. Benassi, S. Cohen, and J. Istas, *Identification and properties of real harmonizable fractional Lévy motions*, Bernoulli **8** (2002), no. 1, 97–115.

[BFB94] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, *Performance of optical flow techniques*, Int. J. Comput. Vis. **12** (1994), 43–77.

[BI02] A. Benassi and J. Istas, *Processus autosimilaires*, Lois d'échelle, fractales et ondelettes (P. Abry, P. Gon calvès, and J. Lévy Véhel, eds.), Traité IC2, vol. 2, Hermes Lavoisier, 2002.

[BJR97] A. Benassi, S. Jaffard, and D. Roux, *Elliptic gaussian random processes*, Rev. Mat. Ibero. **13** (1997), no. 1, 19–90.

[Bog07] V. I. Bogachev, *Measure theory. Vol. I, II*, Springer-Verlag, Berlin, 2007.

[Bro64] R. Brown, *Function spaces and product topologies*, Quart. J. Math. Oxford Ser. (2) **15** (1964), 238–250.

[BU07] T. Blu and M. Unser, *Self-similarity: Part II—Optimal estimation of fractal processes*, IEEE Trans. Signal Process. **55** (2007), no. 4, 1364–1378.

[CL93] B. Cabral and L. Leedom, *Imaging vector fields using line integral convolution*, Proc. SIGGRAPH 93 (New York, NY, USA), ACM, 1993, pp. 263–270.

[Dob79] R. L. Dobrushin, *Gaussian and their subordinated self-similar random generalized fields*, Ann. Probab. **7** (1979), no. 1, 1–28.

[DR04] F. Dodu and C. Rabut, *Irrotational or divergence-free interpolation*, Numerische Mathematik **98** (2004), no. 3, 477–498.

[Dys04] F. Dyson, *A meeting with Enrico Fermi*, Nature **427** (2004), no. 6972, 297.

[FAH+08] A. Frydrychowicz, R. Arnold, D. Hirtler, C. Schlensak, A. F. Stalder, J. Hennig, M. Langer, and M. Markl, *Multidirectional flow analysis by cardiovascular magnetic resonance in aneurysm development following repair of aortic coarctation*, J. Cardiovasc. Magn. Reson. **10** (2008), no. 1, 30.

[FDON06] M. A. T. Figueiredo, J. B. Dias, J. P. Oliveira, and R. D. Nowak, *On total variation denoising: A new majorization-minimization algorithm and an experimental comparison with wavalet denoising*, Proc. 2006 IEEE Int. Conf. Image Process., October 2006, pp. 2633–2636.

[Fla89] P. Flandrin, *On the spectrum of fractional Brownian motions*, IEEE Trans. Inf. Theory **35** (1989), no. 1, 197–199.

[Fla92] ———, *Wavelet analysis and synthesis of fractional Brownian motion*, IEEE Trans. Inf. Theory **38** (1992), no. 2, 910–916.

[Fol01] G. B. Folland, *How to integrate a polynomial over a sphere*, Am. Math. Mon. **108** (2001), no. 5, 446–448.

[Fre03] D. H. Fremlin, *Measure theory, Vol. 2: Broad foundations*, Torres Fremlin, Colchester, 2003.

[Fre04a] _____, *Measure theory, Vol. 1: The irreducible minimum*, Torres Fremlin, Colchester, 2004.

[Fre04b] _____, *Measure theory, Vol. 3: Measure algebras*, Torres Fremlin, Colchester, 2004.

[Fre06] _____, *Measure theory, Vol. 4: Topological measure spaces*, Torres Fremlin, Colchester, 2006.

[Fre08] _____, *Measure theory, Vol. 5: Set-theoretic measure theory*, Torres Fremlin, Colchester, 2008.

[Gir95] D. A. Girard, *The fast Monte-Carlo cross-validation and C_L procedures: comments, new results and application to image recovery problems*, Comput. Statist. **10** (1995), no. 3, 205–258.

[GR92] D. Geman and G. Reynolds, *Constrained restoration and the recovery of discontinuities*, IEEE Trans. Pattern Analysis Mach. Intell. **14** (1992), no. 3, 367–383.

[GS64] I. M. Gel'fand and G. E. Shilov, *Generalized functions, Vol. I: Properties and operations*, Academic Press, 1964, translated by E. Saletan.

[GV64] I. M. Gel'fand and N. Ya. Vilenkin, *Generalized functions, Vol. IV: Applications of harmonic analysis*, Academic Press, 1964, translated by A. Feinstein.

[Hau14] F. Hausdorff, *Grundzüge der Mengenlehre*, Veit, Leipzig, reprinted, Chelsea Publishing Company, New York, N.Y., 1914.

[Hör90] L. Hörmander, *The analysis of linear partial differential operators, Vol. I: Distribution theory and Fourier analysis*, 2nd ed., Springer-Verlag, Berlin, 1990.

[Hur51] H. E. Hurst, *Long-term storage capacity of reservoirs*, Trans. Am. Soc. Civ. Eng. **116** (1951), 770–808.

[Hur56] H. E. Hurst, *Methods of using long-term storage in reservoirs*, Proc. Inst. Civ. Eng. **5** (1956), no. 5, 519–543.

[Kol33] A. Kolmogoroff, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Julius Springer, Berlin, 1933.

[Kol35] _____, *La transformation de Laplace dans les espaces linéaires*, C. r. hebd. séances Acad. sci. Paris **200** (1935), 1717–1718.

[Kol40] A. N. Kolmogoroff, *Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum*, C. r. (dokl.) Acad. Sci. URSS **26** (1940), no. 2, 115–118.

[Kol56] A. N. Kolmogorov, *Foundations of the theory of probability*, 2nd English ed., Chelsea Publishing Company, 1956, translated by N. Morrison, with an added bibliography by A. T. Bharucha-Reid.

[Kol59] ———, *A note on the papers of R. A. Minlos and V. Sazonov*, Theory Probab. Appl. **4** (1959), 221–223.

[Kön97] H. König, *Measure and integration*, Springer-Verlag, Berlin, 1997.

[Kön06a] ———, *The Lebesgue decomposition theorem for arbitrary contents*, Positivity **10** (2006), no. 4, 779–793.

[Kön06b] ———, *New versions of the Radon-Nikodým theorem*, Arch. Math. (Basel) **86** (2006), no. 3, 251–260.

[Kön06c] ———, *Stochastic processes on the basis of new measure theory*, Positivity IV, Tech. Univ. Dresden, Dresden, 2006, pp. 79–92.

[Kön08] ———, *Measure and integral: new foundations after one hundred years*, Functional analysis and evolution equations, Birkhäuser, Basel, 2008, pp. 405–422.

[Lé54] P. Lévy, *Le mouvement brownien*, vol. Fascicule CXXVI, Imprimerie Gauthier–Villars, Paris, 1954, Publié sous le patronage de l'Académie des sciences de Paris.

[Lé65] ———, *Processus stochastiques et mouvement brownien*, 2nd ed., Gauthier–Villars & Co., 1965, reprinted in 1992 by Éditions Jacques Gabay.

[Mal89] S. Mallat, *A theory for multiresolution signal decomposition: The wavelet representation*, IEEE Trans. Pattern Analysis Mach. Intell. **11** (1989), no. 7, 674–693.

[Mas93] E. Masry, *The wavelet transform of stochastic processes with stationary increments and its application to fractional Brownian motion*, IEEE Trans. Inf. Theory **39** (1993), no. 1, 260–264.

[MHB+07] M. Markl, A. Harloff, T. A. Bley, M. Zaitsev, B. Jung, E. Weigang, M. Langer, J. Hennig, and A. Frydrychowicz,

Time-resolved 3D MR velocity mapping at 3T: Improved navigator-gated assessment of vascular anatomy and blood flow, J. Magn. Reson. Imaging **25** (2007), no. 4, 824–831.

[Min63] R. A. Minlos, *Generalized random processes and their extension to a measure*, Selected Transl. Math. Statist. and Prob., Vol. 3, Amer. Math. Soc., Providence, R.I., 1963, pp. 291–313.

[MP86] S. G. Mikhlin and S. Prössdorf, *Singular integral operators*, Springer-Verlag, Berlin, 1986, translated by A. Böttcher and R. Lehmann.

[MST99] Y. Meyer, F. Sellan, and M. Taqqu, *Wavelets, generalized white noise and fractional integration: The synthesis of fractional Brownian motion*, J. Fourier Analysis Appl. **5** (1999), no. 5, 465–494.

[MV68] B. B. Mandelbrot and J. W. Van Ness, *Fractional Brownian motions, fractional noises and applications*, SIAM Rev. **10** (1968), no. 4, 422–437.

[MX01] J. D. Mason and Y. Xiao, *Sample path properties of operator-self-similar Gaussian random fields*, Teor. Veroyatnost. i Primenen. **46** (2001), no. 1, 94–116.

[Pet51] B. J. Pettis, *On the extension of measures*, Ann. of Math. (2) **54** (1951), 186–197.

[Pie72] A. Pietsch, *Nuclear locally convex spaces*, Springer-Verlag, New York, 1972, translated by W. H. Ruckle.

[Pin55] M. S. Pinsker, *The theory of curves in Hilbert space with stationary nth increments*, Izv. Akad. Nauk SSSR. Ser. Mat. **19** (1955), 319–344.

[Pin64] ———, *Information and information stability of random variables and processes*, Holden-Day Inc., San Francisco, Calif., 1964, translated and edited by A. Feinstein.

[PLV94] R.-F. Peltier and J. Lévy Véhel, *A new method for estimating the parameter*, Research Report RR-2396, Institut national de recherche en informatique et en automatique (INRIA), 1994.

[PLV95] ———, *Multifractional Brownian motion: Definition and preliminary results*, Research Report RR-2645, Institut national de recherche en informatique et en automatique (INRIA), August 1995.

[PPL02] B. Pesquet-Popescu and J. Lévy Véhel, *Stochastic fractal models for image processing*, IEEE Signal Process. Mag. **19** (2002), no. 5, 48–62.

[Pro56] Yu. V. Prohorov, *Convergence of random processes and limit theorems in probability theory*, Teor. Veroyatnost. i Primenen. **1** (1956), 177–238.

[Pro61] ———, *The method of characteristic functionals*, Proc. 4th Berkeley Symp. Math. Statist. Prob., vol. 2, Univ. of Calif. Press, 1961, pp. 403–419.

[RBU08] S. Ramani, T. Blu, and M. Unser, *Monte-Carlo SURE: A black-box optimization of regularization parameters for general denoising algorithms*, IEEE Trans. Image Process. **17** (2008), no. 9, 1540–1554.

[ROF92] L. I. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Physica D: Nonlinear Phenom. **60** (1992), no. 1–4, 259–268.

[Rud76] W. Rudin, *Principles of mathematical analysis*, 3rd ed., McGraw-Hill, Inc., 1976.

[Saz58] V. Sazonov, *A remark on characteristic functionals*, Teor. Veroyatnost. i Primenen. **3** (1958), 201–205.

[SC00] D. Suter and F. Chen, *Left ventricular motion reconstruction based on elastic vector splines*, IEEE Trans. Med. Imaging **19** (2000), no. 4, 295–305.

[Sch66] L. Schwartz, *Théorie des distributions*, Publications de l'Institut de Mathématique de l'Université de Strasbourg, No. IX-X. Nouvelle édition, entiérement corrigée, refondue et augmentée, Hermann, Paris, 1966.

[Sch73] ———, *Radon measures on arbitrary topological spaces and cylindrical measures*, Published for the Tata Institute of Fundamental Research, Bombay by Oxford University Press, London, 1973.

[Sch81] ———, *Geometry and probability in Banach spaces*, Lecture Notes in Mathematics, vol. 852, Springer-Verlag, Berlin, 1981.

[SF76] O. G. Smolyanov and S. V. Fomin, *Measures on linear topological spaces*, Russ. Math. Surv. **31** (1976), no. 4, 1–53.

[ST94] G. Samorodnitsky and M. S. Taqqu, *Stable non-Gaussian random processes: Stochastic models with infinite variance*, Chapman and Hall, New York, 1994.

[Sta09] A. F. Stalder, *Quantitative analysis of blood flow and vessel wall parameters using 4D-flow sensitive MRI*, Ph.D. thesis, Universität Freiburg, Technische Fakultät (bisher: Fak. f. Angew. Wiss.), 2009.

[Ste70] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, vol. 30, Princeton University Press, 1970.

[SU] Q. Sun and M. Unser, *Left-inverses of fractional Laplacian and sparse stochastic processes*, Adv. Comput. Math., to appear.

[Sut94] D. Suter, *Motion estimation and vector splines*, Proc. 1994 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR94), June 1994, pp. 939–942.

[TDGSU10] P. D. Tafti, R. Delgado-Gonzalo, A. F. Stalder, and M. Unser, *Fractal modelling and analysis of flow-field images*, Proc. 7th IEEE Int. Symp. Biomed. Imaging (ISBI 2010) (Rotterdam, The Netherlands), June 2010.

[TDGSU11] ———, *Variational enhancement and denoising of flow field images*, Proc. 8th IEEE Int. Symp. Biomed. Imaging (ISBI 2011) (Chicago, US-IL), 2011.

[TU09] P. D. Tafti and M. Unser, *Self-similar random vector fields and their wavelet analysis*, Proc. SPIE Conf. Math. Imaging: Wavelets XIII (San Diego CA, USA) (V. K. Goyal, M. Papadakis, and D. Van De Ville, eds.), vol. 7446, 74460Y, August 2009, pp. 74460Y–1–74460Y–8.

[TU10a] ———, *Fractional Brownian models for vector field data*, Proc. 2010 IEEE Int. Symp. Inf. Theory (2010), 1738–1742.

[TU10b] ———, *Fractional Brownian vector fields*, SIAM Multiscale Mode. Simul. **8** (2010), no. 5, 1645–1670.

[TVU08] P. D. Tafti, D. Van De Ville, and M. Unser, *Innovation modelling and wavelet analysis of fractal processes in bio-imaging*, Proc. 5th IEEE Int. Symp. Biomed. Imaging (ISBI'08) (Paris, France), May 2008, pp. 1501–1504.

[TVU09] ———, *Invariances, Laplacian-like wavelet bases, and the whitening of fractal processes*, IEEE Trans. Image Process. **18** (2009), no. 4, 689–702.

[UT11] M. Unser and P. D. Tafti, *Stochastic models for sparse and piecewise-smooth signals*, IEEE Trans. Signal Process. **59** (2011), no. 3, 989–1006.

[VA99] D. Veitch and P. Abry, *A wavelet-based joint estimator of the parameters of long-range dependence*, IEEE Trans. Inf. Theory **45** (1999), no. 3, 878–897.

[Wie24] N. Wiener, *The average value of a functional*, Proc. London Math. Soc. **s2-22** (1924), no. 1, 454–467.

[Wir04] A. Wirgin, *The inverse crime*, ArXiv e-print (2004), arXiv:math-ph/0401050v1.

[WO92] G. W. Wornell and A. V. Oppenheim, *Estimation of fractal signals from noisy measurements using wavelets*, IEEE Trans. Signal Process. **40** (1992), no. 3, 611–623.

[Yag87] A. M. Yaglom, *Correlation theory of stationary and related random functions. Vol. I: Basic results*, Springer Series in Statistics, Springer-Verlag, New York, 1987.

[YP53] A. M. Yaglom and M. S. Pinsker, *Random processes with stationary increments of the nth order*, Dokl. Akad. Nauk SSSR (N.S.) **90** (1953), 731–734.

Curriculum Vitæ

Education

- PhD, Doctoral Programme in Computer, Communication, and Information Sciences, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, 2011.
- MASc (Distinction), Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada, 2006.
- BSc, Electrical Engineering, Sharif University of Technology, Tehran, Iran, 2003.
- High School Diploma, National Organization for Development of Exceptional Talents (NODET) Allameh Helli High School, Tehran, Iran, 1999.

Honours and Awards

In reverse chronological order:

- *Best Student Paper Award,* 8th IEEE International Symposium on Biomedical Imaging, Chicago, US-IL, 2011.
- *Certificate of Excellence* and award for *Outstanding Thesis Research,* Department of Electrical and Computer Engineering, McMaster University, Canada, 2005.
- Graduate Studies Scholarship and Departmental Bursary, McMaster University, Canada, 2003–2006.
- Ranked 48th among 250,000+ participants (top 0.02%) in the National University Entrance Examination (Concours), Iran, 1999.
- Studied in the highly selective gifted education programme supervised by the National Organization for Development of Exceptional Talents (NODET), Tehran, Iran, 1992–1999.

Professional Memberships

- The Institute of Electrical and Electronics Engineers (IEEE)
- The Society for Industrial and Applied Mathematics (SIAM)
- The American Mathematical Society (AMS)
- The Swiss Mathematical Society (SMS)

Peer-Reviewed Publications

Journal Articles:

- P. D. Tafti and M. Unser, *'On regularized reconstruction of vector fields',* IEEE Trans. Image Process., to appear.
- M. Unser and P. D. Tafti, *'Stochastic models for sparse and piecewise-smooth signals',* IEEE Trans. Signal Process., vol. 59, no. 3, pp. 989–1006, Mar. 2011.
- P. D. Tafti and M. Unser, *'Fractional Brownian vector fields',* SIAM Multiscale Model. Simul., vol. 8, no. 5, pp. 1645–70, 2010.
- P. D. Tafti, D. Van De Ville, and M. Unser, *'Invariances, Laplacian-like wavelet bases, and the whitening of fractal processes',* IEEE Trans. Image Process., vol. 18, no. 4, pp. 689-702, Apr. 2009.
- P. D. Tafti, S. Shirani, and X. Wu, *'On interpolation and resampling of discrete data',* IEEE Signal Process. Lett., vol. 13, no. 12, pp. 733–736, Dec. 2006.

Conference Papers:

- P. D. Tafti, R. Delgado-Gonzalo, A. F. Stalder, and M. Unser, *'Variational enhancement and denoising of flow field images',* Proc. 8th IEEE Int. Symp. Biomed. Imaging (ISBI 2011), pp. 1061–4, US-IL, Mar.–Apr. 2011 (winner of the Best Student Paper Award).
- P. D. Tafti and M. Unser, *'Fractional Brownian models for vector field data',* Proc. 2010 IEEE Int. Symp. Inf. Theory (ISIT 2010), Austin, US-TX, Jun. 2010.
- P. D. Tafti, R. Delgado-Gonzalo, A. F. Stalder, and M. Unser, *'Fractal modelling and analysis of flow-field images',* Proc. 7th IEEE Int. Symp. on Biomed. Imaging (ISBI 2010), pp. 49–52, Rotterdam, The Netherlands, Apr. 2010.
- R. Delgado-Gonzalo, P. D. Tafti, and M. Unser, *'Fractional Laplacian pyramids',* Proc. 2009 IEEE Int. Conf. Image Process. (ICIP 2009), pp. 3809–12, Cairo, Egypt, Nov. 2009.
- P. D. Tafti and M. Unser, *'Self-similar random vector fields and their wavelet analysis',* Proc. SPIE Wavelets XIII, vol. 7446, pp. 7446Y-1–8,

San Diego, US-CA, Aug. 2009.
- P. D. Tafti and X. Wu, *'Automatic tonal harmonization for multi-spectral mosaics',* Proc. 2008 IEEE Int. Conf. Image Process. (ICIP 2008), pp. 1876–9, San Diego, US-CA, Oct. 2008.
- P. D. Tafti, D. Van De Ville, and M. Unser, *'Innovation modelling and wavelet analysis of fractal processes in bio-imaging',* Proc. 5th IEEE Int. Symp. Biomed. Imaging (ISBI 2008), pp. 1501–4, Paris, France, 2008.
- P. D. Tafti, S. Shirani, and X. Wu, *'Multi-dimensional average-interpolating refinement on arbitrary lattices',* Proc. 2005 IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP 2005), vol. 4, pp. IV-597–600, Philadelphia, US-PA, 2005.

Selected Software Projects

- `SPARSEPROC`, software for innovation modelling of sparse processes (in `MATLAB`), 2011.
- `VREG`, software for variational reconstruction of vector fields (in `MATLAB`), 2010–2011.
- Software for wavelet-based fractal analysis of images, volume data, and vector fields (in `MATLAB`), 2006–2009.
- `GEOCOL`, software for spectral harmonization of multi-spectral satellite images using L_1 minimization (in `ANSI C`), 2006.
- `COLEX`, a collection of colorimetry tools (in the Plan 9 dialect of `ANSI C`, using Bell Labs-style multi-threading), 2005.

Computer Skills

- Programming: `ANSI C`, Unix shell scripting, `MATLAB`, *&c.*
- Operating Systems: NetBSD, Linux, Mac OS X, Windows
- Publishing and Presentation: TEX/LATEX, HTML, CSS, Adobe Photoshop, Adobe InDesign, *&c.*

Language Skills

- Persian, English: fluent; French: CEFR level B1/B2

Hobbies and Non-Technical Interests

- Classical and traditional music, reading, cartooning