

A theoretical study of steerable homogeneous operators, and applications

Research Internship Report – Master MVA

Lilian Besson*

“Élève normalien” in the Department of Mathematics
École Normale Supérieure de Cachan (France)

lilian.besson@ens-cachan.fr

Internship with Prof.Dr. **Michael Unser** and **Julien Fageot**
BIG team at EPFL, Lausanne (Switzerland)

April – August 2016

* If needed, see on-line at <http://lbo.k.vu/epfl2016> for a PDF version of this report, and additional resources (slides, code, figures, complete bibliography etc), open-sourced under the MIT License. Note that this document is my internship report, but also my Master thesis.

Abstract

This report presents the research I did between April and August 2016, in theoretical function analysis and operator theory, at the BIG team at EPFL.

We focus mainly on two aspects: convolution operators in any dimension; and steerable homogeneous convolutions in two dimensions (for images), and their possible applications. A lot of our results are valid regardless of the dimensions, so we stay in a general setting as long as we can, and then we restrict to operators on 2D images for the study of steerable operators.

We start by recalling the common notations used in signal processing and functional analysis, and then by recalling the main properties of a fundamental tool for this domain, the Fourier transform \mathcal{F} . We expose the main goals of our research, in order to motivate this theoretical study of convolution operators also from a practical point of view.

We chose to follow a very didactic approach, and so we almost redefine “from scratch” the theory of functional operators, along with proofs of its most important results. Our operators can have structural and geometric properties, namely linearity or continuity, and translation-, scaling-, rotation-invariance, unity – all these properties being already well-known – or steerability. We study extensively the links between all these properties, and we present many characterizations. This document is an attempt to summarize all these results, most of them are already well-known, but some of our latest characterizations appeared to be new results.

After a very broad section on operators, we focus on steerable convolutions G , mainly in 2D, as they appear to be the natural framework for image analysis tasks. Our main results consist in characterizations of steerable convolutions, first written as a sum of modulated and iterated real Riesz transforms. Then adding the γ -scale-invariance gives a nicer form, as a composition of a fractional Laplacian $(-\Delta)^{\gamma/2}$, some directional derivatives D_{α_i} and an invertible 0-scale-invariant factor G_0 ; and another form as a composition of elementary blocks $G_{\lambda,\alpha}$. This last form is very appealing for implementation, as it is enough to implement the elementary block, and to compose it to obtain every 2D steerable γ -SI convolution, and it has a strong theoretical interpretation: a G N -SI and steerable of order n_G gets decomposed as a product of N elementary blocks, all 1-SI steerable of order 1 or 2. And a simple universality theorem on $G_{\lambda,\alpha}$ reinforces this result: the elementary blocks are exactly the 1-SI steerable convolutions of order 1 or 2.

We conclude by explaining how to implement our operators and their inverse, and by presenting the results of some experiments on 2D stochastic processes, in order to illustrate the effects of our elementary blocks as well as more complicated operators. We highlight some properties on the examples, like the trade-of between the directionality of D_{α_i} and the isotropy of $(-\Delta)^{\gamma/2}$. Our operators could also be used to develop new splines (for new sampling schemes), and new Green’s functions (for new denoising and data recovery algorithms), but we did not have the time to fully study this.

We also studied unitary operators, but it is only included in an appendix as it was less connected to our main problem.

Outline

1	Introduction	5
1.1	Quick overview of our work and its contributions	5
1.2	Notations and symbols	6
1.3	Classic results on the Fourier transform	8
1.3.1	The Fourier transform \mathcal{F} and its inverse \mathcal{F}^{-1}	8
1.3.2	Rotation-invariance of the Fourier transform	9
1.3.3	The convolution theorem	9
1.4	Operators, “real world vs theory”	10
1.5	Short explanations on Green’s functions	10
2	Operators, and their geometric invariance properties	12
2.1	Defining functional operators	12
2.1.1	Test functions, and some tricks to simplify their domains	12
2.1.2	Operators on test functions	13
2.2	Classic properties of operators	13
2.2.1	Linear operators (L)	13
2.2.2	Continuous operators (C)	14
2.3	Geometric properties: translation-, scaling- and rotation-invariance	16
2.3.1	Additional hypothesis on the function space E	16
2.3.2	In fact, we need <i>topological</i> spaces	16
2.3.3	Translation-invariant operators (TI)	17
2.3.4	Scaling-invariant operators (0-SI)	19
2.3.5	γ -scaling-invariant operators (γ-SI)	20
2.3.6	Rotation-invariant operators (RI)	22
3	Classic results for operators, important examples, and characterizations for geometric invariance	26
3.1	Quick overview of some classic examples of operators	26
3.2	Schwartz’s kernel theorems	27
3.2.1	Characterization for LC – First Schwartz’s kernel theorem	27
3.2.2	Impulse response g of a LC TI operator G – Second Schwartz’s kernel theorem	29
3.2.3	Fourier multiplier \hat{g} of a LC TI operator G	30
3.2.4	Real operators and Hermitian symmetry	32
3.3	Important examples of convolutions: derivatives and Laplacians	32
3.3.1	Partial and directional derivatives	33
3.3.2	Laplacian and fractional Laplacian	34
3.4	Characterizations of geometric properties for integral and convolution operators	36
3.4.1	Characterization for LC γ-SI operators	37
3.4.2	Characterization for LC TI γ-SI operators	37
3.4.3	Characterization for LC RI operators	38
3.4.4	Characterization for LC TI RI operators	39
3.4.5	LC TI RI γ-SI operators are exactly fractional Laplacians	39
4	Steerable convolution operators in dimension 2	41
4.1	Defining steerability, in any dimension	41
4.1.1	Steerable functions and distributions	41
4.1.2	Steerable operators (St)	42
4.2	Special cases and examples of steerable operators	48
4.2.1	Steerability of order 0 or 1 ?	48
4.2.2	Some simple examples of steerable operators	48
4.3	2D steerable convolution operators	50
4.3.1	Representer theorems for rotation-invariance sub-spaces	50

4.3.2	Sum decomposition for steerable convolutions	52
4.3.3	A simpler form for this characterization of steerability?	55
4.4	Steerability and other geometric properties	55
4.4.1	Rotation-invariance for steerable convolution operators	56
4.4.2	γ -scale-invariance for 2D steerable convolution operators	56
4.5	Composing fractional Laplacians and directional derivatives	57
4.6	Main characterization for 2D steerable γ -SI convolution operators	61
4.6.1	First product decomposition	61
4.6.2	Rotating a 2D steerable γ -SI convolution operators	66
5	Implementing and inverting 2D steerable convolutions	68
5.1	Product decompositions <i>vs</i> sum decomposition	68
5.1.1	Factorizing G_0	68
5.1.2	Elementary blocks decomposition for 2D LC TI γ -SI St operators	69
5.1.3	A universality result for the elementary block $G_{\lambda,\alpha}$	70
5.1.4	A natural interpretation for this last characterization?	71
5.2	Implementation of 2D steerable γ -SI convolutions for images	71
5.3	Inverse steerable convolutions applied to Sparse Stochastic Processes	72
5.3.1	Details about the setting of our experiments	72
5.3.2	List of experiments and their figures	73
6	Conclusion	81
6.1	Other questions, possible future works	82
A	Appendix – On unitary operators (U)	83
A.1	Unitary operators (U), definitions and examples	83
A.2	Characterizations for unitary convolution operators	83
A.3	An extended example: the fractional-directional Hilbert transforms fdHT	85
A.3.1	Classic Hilbert transform and earlier generalizations in dimension 1 and 2	85
A.3.2	Definition of the fdHT	86
A.3.3	Elementary properties of the fdHT	87
A.3.4	Implementation of the fdHT	88
A.3.5	Non-universality but density of the fdHT among unitary 0-SI convolutions	89
A.3.6	Inverting a fdHT	94
A.4	Attempts of characterizing unitary 2D steerable convolutions	94
B	Appendix – Additional proofs	97
B.1	A composition of m directional derivatives has $n_G = m + 1$	97
B.2	Viète formulas to develop $\cos(k\theta)$ and $\sin(k\theta)$	99
B.3	The Fourier coefficients of $c(\theta)$ are mutually orthogonal projections	100
C	Acknowledgments	101
D	List of references	102

Internship advisors: Julien Fageot and Prof.Dr. Michael Unser

Location: LIB/BIG team, EPFL (Lausanne)

Master program: “Mathématiques, Vision, Apprentissage” (MVA)
in 2015–2016, at ENS de Cachan (France).

Grade: I got 20/20 for my internship.

1 Introduction

This first section is a short introduction, presenting the context and main goals and contributions of our research, and then the notations and symbols used in this document. We could not include reminders on every notions and topics that this work is built on, of course, but we remind the main properties of the Fourier transform and its inverse, \mathcal{F} and \mathcal{F}^{-1} , and short explanations on Green's functions of operators.

Remark 1.1 (Remarks for the MVA jury).

– Slides: This report comes along the slides used for the oral presentation I did in the BIG research team at the end of my internship, which covers a similar work, but in a simpler and more pedagogical way, so you might also be interested in reading them. You can find them on-line: goo.gl/vm8WPF or here (click)..

– Article version: This report is long, as I preferred to adopt a very didactic approach and make it as clear, complete and detailed as possible, by giving almost every proofs, a lot of examples, some illustrative figures, and numerical experiments. But we are working on a shorter version, to publish soon the most interesting and innovative results we obtained (mainly from sections 4, 5.1, 5.3).

1.1 Quick overview of our work and its contributions

Our research started as a study of *convolution operators*, as the team I joined was interested in them for some time [VSU15, FAU14, UFW16], and more specifically to *unitary* convolutions [CU10]. After some time on this topic, we obtained interesting results on a certain family of unitary convolutions, the *fractional-directional Hilbert transform*, generalizing in any dimension the previous generalizations of the 1D Hilbert transform [CU09]. Except in dimension 1 or 2, our results were quite limited, because of the intrinsic complexity of the sphere of higher dimensions. This part on unitary operators is only included in the appendix A, as we preferred to focus on our other direction of research.

We were also interested in *rotation-invariant* operators, as it is an simple property. We found a known strong characterization for rotation-invariant homogeneous convolutions, which turn out to be restricted to the Laplacian and its generalization (γ -fractional Laplacians) [UT14]. So we wanted to study a generalization of the rotation-invariance, called the *steerability* [VSU15, WCU13]. This notion is known and used in the community of filters for image analysis [Püs16], and classification¹, but we studied it from a theoretical point-of-view. In a nutshell, steerability means that the vector space spanned by the rotated versions of an operator G has only a finite dimension n_G , that we called the steerability order. An order 1 characterizes rotation-invariance, but it can be larger (e.g., 2D derivatives have order 2).

This class of steerable operator is more restricted, but the class of steerable convolutions was still too rich in our perspective of having a *simple parametrization for operators*. That is why our main interest relies on steerable operator that are also *homogeneous* (γ -scale-invariant), as they are a meaningful framework for “physically plausible” filters.

We obtained interesting theorems for this class of *homogeneous steerable convolutions* G in 2D, showing that once the order of steerability is fixed, they can all be parametrized

¹ For example, textures classification with steerable filters and deep learning, also studied in the BIG team at EPFL. Or for some historical use of steerability, cf. the steerable wavelets by Daubechies or [WCU13, UC13, WU14].

with a finite number of parameters, in a *sum decomposition*. We also have a decomposition for G in terms of a *product decomposition*, *i.e.*, a composition of elementary blocks $G_{\lambda,\alpha}$, that can be implemented very easily, and composed to give any homogeneous steerable operators. These parameters have an interesting physical interpretation, α is a direction, and λ is a measure of “partial directionality”.

This way, we can design, apply or implement a broad class of operators by simply choosing their order, and their parameters. And steerability also eases the computation of a rotated version of a steerable operator (simple linear algebra on its coefficients), and this is one of the most important practical use of steerability in computer vision [Püs16].

We tried to find an easy and parametric way to inverse our operators, as it would have open the way to more possible applications (*e.g.*, splines), but in practice it appeared that a closed formula for the inverse (or the Green’s function) of an partly-directional elementary block ($\lambda \neq 0, 1$) was hard to obtain, while it was already known for the special cases of purely isotropic ($\lambda = 0$) or purely directional ($\lambda = 1$) blocks [FAU14]. We were not able to find a formula but did not have the time nor the skills to disprove it.

However, we have been able to implement our operators and *their inverses*, and apply them numerically to Gaussian and compound Gaussian-Poisson *sparse processes* in 2D, giving interesting illustrations and insights on their behaviors.

1.2 Notations and symbols

We collect here the general notations and terminology² used in this report.

The set of non-negative integers, integers and real and complex numbers are denoted \mathbb{N} , \mathbb{Z} , \mathbb{R} and \mathbb{C} , respectively. \mathbb{N}^* , \mathbb{Z}^* , \mathbb{R}^* , \mathbb{C}^* stands for the set without 0, and \mathbb{R}_+ is the set of non-negative real numbers. $\#A$ denotes the number of elements in A (*i.e.*, its size).

We use bold letters to distinguish vectors from scalars. For example, $f(x)$ denotes a function of the real-valued variable x (with a domain of dimension $d = 1$), and $f(\mathbf{x})$ denotes a function of the multivariate variable $\mathbf{x} = (x_1, \dots, x_d)$ which lives in the considered domain of dimension d , either \mathbb{R}^d or \mathbb{C}^d (d will always stands for the dimension of the considered domain, instead of n).

For functions, we use the “.” symbol to denote the “missing variable”, for example $f(\cdot - \mathbf{s})$ denotes the function $\mathbf{x} \mapsto f(\mathbf{x} - \mathbf{s})$, *i.e.*, the *translation* of $f(\mathbf{x})$ by \mathbf{s} .

We use round brackets (\cdot) for continuous variables, and square brackets $[\cdot]$ for discrete variables. For example, $f[n, m]$ denotes a sequence defined on two integers (\mathbb{N}^2 , \mathbb{Z}^2 or $\llbracket 0, N \rrbracket \times \llbracket 0, M \rrbracket$), while for instance $f(x, y)$ will be defined on \mathbb{R}^2 .

We use bold capital letters to represent matrices, while their entries will be denoted with the subscripted lower case letters, *e.g.*, $\mathbf{A} = (a_{i,j})$. For a vector \mathbf{x} or a matrix \mathbf{A} , we denote its transpose by \mathbf{x}^T or \mathbf{A}^T , and its Hermitian transpose by \mathbf{x}^H or \mathbf{A}^H (conjugated and transposed).

The usual scalar product for vectors in \mathbb{R}^d will be denoted $\mathbf{x}^T \mathbf{y}$ or $\langle \mathbf{x}, \mathbf{y} \rangle$ and it is defined by $\sum_{i=1}^d x_i y_i$. For complex-valued vectors, the Hermitian scalar product is denoted

$$\mathbf{z}^T \mathbf{z}' = \langle \mathbf{z}, \mathbf{z}' \rangle \stackrel{\text{def}}{=} \sum_{i=1}^d \bar{z}_i \cdot z'_i = \sum_{i=1}^d \bar{z}_i z'_i.$$

As most of the references in our domain of interest are from engineers more than pure mathematicians, we prefer to use j (and **not** i) to denote the imaginary unit num-

² We try to use the more standard notations, and this paragraph is inspired from [Cha11, p9].

ber (“ $\sqrt{-1}$ ”). The conjugate of a complex number z is denoted \bar{z} (and **not** z^*), and the real and imaginary components of z are respectively denoted $\Re(z)$ and $\Im(z)$, while the modulus and argument (*i.e.*, phase) are denoted by $|z|$ and $\arg(z)$.

We use $\widehat{f}(\omega)$ to denote the (usual) Fourier transform of a function³ $f(\mathbf{x})$ defined on \mathbb{R}^d . When it has a meaning, it is defined on (true) functions by:

$$\widehat{f}(\omega) \stackrel{\text{def}}{=} \int_{\mathbb{R}^d} f(\mathbf{x}) \exp(-j\omega^T \mathbf{x}) d\mathbf{x} \quad (\text{for } \omega \in \mathbb{R}^d),$$

where \mathbf{x} is the *space* variable and ω is the *frequency* variable. As usual, we omit the domain of the integral whenever it is obvious from the context. And for a distribution ϕ , its Fourier transform $\mathcal{F}\{\phi\}$ is defined by duality bracket: $\langle \mathcal{F}\{\phi\}, f \rangle = \langle \phi, \mathcal{F}\{f\} \rangle$.

When it has a meaning, the *convolution* of two functions $f(\mathbf{x})$ and $g(\mathbf{x})$ is given by

$$(f * g)(\mathbf{x}) \stackrel{\text{def}}{=} \int f(\mathbf{s})g(\mathbf{x} - \mathbf{s})d\mathbf{s}.$$

We denote the first derivative of $f(x)$ by: $f'(x)$, $Df(x)$, or $df(x)/dx$. The k -th derivative is denoted by $f^{(k)}(x)$. We say that f is k -times differentiable if all its derivatives (up to order k) exists *and are continuous*. For a multivariate function $f(\mathbf{x}) = f(x_1, \dots, x_n)$, the partial derivative along x_i is represented by $D_i f(\mathbf{x}) = \frac{\partial f(\mathbf{x})}{\partial x_i}$ (cf. section 3.3.1).

We use $G_1 \circ G_2$ (often simplified as $G_1 G_2$) to denote the composition of two transformations (or operators) G_1 and G_2 , that is, $(G_1 \circ G_2)\{f\} \stackrel{\text{def}}{=} G_1(G_2\{f\})$ for every element f in the domain of G_2 (such that $G_2\{f\}$ is in the domain of G_1). Beware that the *order* of composition is usually important (\circ is **not** commutative); in general $G_1 G_2 \neq G_2 G_1$.

The Dirac delta distribution⁴ is denoted by $\delta(x)$. The Kronecker delta function is denoted by $\delta[n]$. The **modified sign** function is denoted by $\text{sign}(x)$, equals to 1 if $x > 0$ **or** $x = 0$, and -1 if $x < 0$.

We use standard notations for norms⁵, $\|\mathbf{x}\| = \|\mathbf{x}\|_2$ being the usual Euclidean norm.

For $0 < p < +\infty$, the p -norm⁶ on functions is defined by:

$$\|f\|_p \stackrel{\text{def}}{=} \left[\int |f(\mathbf{x})|^p d\mathbf{x} \right]^{1/p}.$$

The space of functions⁷ having finite p -norm is denoted $\mathbf{L}^p(\mathbb{R}^d, \mathbb{R}) = \mathbf{L}^p(\mathbb{R}^d)$ or simply by \mathbf{L}^p . An operator G which maps an \mathbf{L}^p function into an \mathbf{L}^q function is said to be bounded if there is a constant $C \geq 0$ such that $\|Gf\|_q \leq C\|f\|_p$.

We work in $\mathbf{L}^2(\mathbb{R}^d, \mathbb{R}) = \mathbf{L}^2(\mathbb{R}^d)$, the Hilbert space of *finite-energy* (or *square-integrable* functions), endowed with the inner product⁸ [Rud91]

$$\langle f, g \rangle = \int f(\mathbf{x})g(\mathbf{x})d\mathbf{x}.$$

³ With the usual abuse of notation between the function f and its value $f(\mathbf{x})$.

⁴ We will not give more details about distributions, if the reader is unfamiliar with this theory, we suggest to read the classic reference [Rud91] or this article on Wikipedia.

⁵ For a reminder on norms, see for instance this book [FR13, part. A.1] or this Wikipedia article.

⁶ $\|\cdot\|_p$ is a norm only for $p \geq 1$.

⁷ We assume the reader is familiar with the construction of this space as the quotient space of $\mathbf{L}^p(\mathbb{R}^d)$ for the equivalent relation “being equal almost everywhere”.

⁸ If the function are complex-valued, we always implicitly assume that the inner product is the Hermitian inner product, *i.e.*, that \bar{f} is used instead of f in the integral: $\langle f, g \rangle = \int_{\mathbb{R}^d} \overline{f(\mathbf{x})}g(\mathbf{x})d\mathbf{x}$, [Rud91].

The notation $\mathbf{H}^{\kappa,2}(\mathbb{R}^d, \mathbb{R})$ denotes the (real-valued) fractional Sobolev space of order $\kappa \in \mathbb{R}$, [Rud91], namely, the functions in $\mathbf{L}^2(\mathbb{R}^d, \mathbb{R})$, for which

$$\int (1 + \|\boldsymbol{\omega}\|^2)^\kappa |\widehat{f}(\boldsymbol{\omega})|^2 d\boldsymbol{\omega} < \infty. \quad (1)$$

For integer orders $\kappa \in \mathbb{N}$, this is the set of functions having (weak) partial derivatives up to order κ , such that the function *and* all its derivatives are in $\mathbf{L}^2(\mathbb{R}^d, \mathbb{R})$.

The class of infinitely differentiable functions that are rapidly decreasing at infinity, along with all their partial derivatives, is denoted by $\mathbf{S} = \mathbf{S}(\mathbb{R}, \mathbb{R})$ and $\mathbf{S}(\mathbb{R}^d, \mathbb{R}) = \mathbf{S}(\mathbb{R}^d)$ (the Schwartz class, [Rud91] and [UT14, p30]). Its (topological) dual, the set of continuous linear functional on \mathbf{S} , is denoted \mathbf{S}' , and is called the set of tempered distributions, generalized functions, or simply distributions.

For functions, the \propto symbol (“is proportional to”) has to be understood on the function not on a specific value: $f(x) \propto x^2$ does **not** mean that for any x , there is a constant c_x such that $f(x) = c_x x^2$ (this would be true for almost any function $f!$), it rather means that there is a constant c such that $f(\cdot) = c \cdot^2$, i.e., $\forall x, f(x) = c x^2$.

Of course, more specific notations will be introduced when needed.

1.3 Classic results on the Fourier transform

This short first part focuses on the classic results on the Fourier transform (section 1.3.1), especially on the convolution theorem (section 1.3.3), which is a fundamental tool for the study of convolution operator.

All these results can be found in many books, a good reference being the famous book by *Stein & Weiss* [SW71], and another one is [Hör83]. We quickly include them here for the sake of completeness, but proofs are not given. An interested reader is invited to consult a more complete reference.

1.3.1 The Fourier transform \mathcal{F} and its inverse \mathcal{F}^{-1}

As defined above in section 1.2, the Fourier transform \mathcal{F} is a functional, i.e., an operator on functions. One always have to be cautious on which space are these functions taken from, but in this report we will either consider true functions on $\mathbf{L}^2(\mathbb{R}^d, \mathbb{R})$ or $\mathbf{S}(\mathbb{R}^d, \mathbb{R})$, or distribution on $\mathbf{S}'(\mathbb{R}^d, \mathbb{R})$, on which the Fourier transform is well defined and invertible.

Properties 1.2 (Summary of properties of the Fourier transform).

The Fourier transform \mathcal{F} is a linear and continuous operator, operating on functions on $\mathbf{L}^2(\mathbb{R}^d, \mathbb{R})$ and $\mathbf{S}(\mathbb{R}^d, \mathbb{R})$, and on distribution on $\mathbf{S}'(\mathbb{R}^d, \mathbb{R})$.

In other words, \mathcal{F} is well defined on: $\mathbf{L}^2(\mathbb{R}^d, \mathbb{R}) \rightarrow \mathbf{L}^2(\mathbb{R}^d, \mathbb{R})$, $\mathbf{S}(\mathbb{R}^d, \mathbb{R}) \rightarrow \mathbf{S}(\mathbb{R}^d, \mathbb{R})$, and $\mathbf{S}'(\mathbb{R}^d, \mathbb{R}) \rightarrow \mathbf{S}'(\mathbb{R}^d, \mathbb{R})$.

Additionally, it is unitary (i.e., it conserves inner product) and an isometry (i.e., it conserves norms) on $\mathbf{L}^2(\mathbb{R}^d, \mathbb{R})$, and it has an inverse \mathcal{F}^{-1} (on $\mathbf{S}(\mathbb{R}^d, \mathbb{R})$ and on $\mathbf{L}^2(\mathbb{R}^d, \mathbb{R})$), satisfying $\mathcal{F}^{-1} \circ \mathcal{F} = \text{Id}$ and $\mathcal{F} \circ \mathcal{F}^{-1} = \text{Id}$.

Proof. ⁹ We do not prove these classic results here, but the interested reader can refer to reference books, e.g., [SW71] and [Rud91, Sec.7], or to more recent references for a short reminder with proofs, e.g., [UT14] or [Cha11]. □

⁹ Some proofs that are not new are included, for the sake of completeness, but they are included in small font like this one. A proof with a normal-sized font indicates that the proof concerns an innovative result or that the proof itself is somehow original.

What constant in front of the integral of the Fourier transform?

The exact (real-valued) constant used in the formula of the Fourier transform and its inverse is only relevant if we want $\mathcal{F}^{-1} \circ \mathcal{F} = \text{Id}$, and if we want \mathcal{F} to be an isometry (the wrong constant yields $\langle \mathcal{F}\{f_1\}, \mathcal{F}\{f_2\} \rangle = C^2 \langle f_1, f_2 \rangle$ for some $C \neq 1$). \mathcal{F} is usually defined with a leading $1/(2\pi)^{(d/2)}$ factor [D'A13, Def.3.2], as this choice gives the same constant for \mathcal{F}^{-1} , but in this case \mathcal{F} is not exactly an isometry (but it preserves orthogonality).

Warning: we will usually **not** explicit it, in order to keep the computations and formulas as simple as possible, but the reader should keep in mind that a constant, only depending on the dimension d , is always present when we manipulate \mathcal{F} and \mathcal{F}^{-1} .

1.3.2 Rotation-invariance of the Fourier transform

Before introducing properly the rotation operators r_R (see later for its definition, 2.41), we can give a semi-formal statement of the rotation-invariance of the Fourier transform.

Property 1.3 (Rotation-invariance of \mathcal{F} and \mathcal{F}^{-1}).

Both \mathcal{F} and its inverse \mathcal{F}^{-1} are rotation-invariant: for any rotation R , $r_R \circ \mathcal{F} = \mathcal{F} \circ r_R$, and $r_R \circ \mathcal{F}^{-1} = \mathcal{F}^{-1} \circ r_R$ (on $\mathbf{S}(\mathbb{R}^d, \mathbb{R})$ and on $\mathbf{L}^2(\mathbb{R}^d, \mathbb{R})$).

In simpler words, for any rotation $R \in \text{SO}(d)$, and any test function f ,

$$\forall \omega \in \mathbb{R}^d, \quad \mathcal{F}\{f\}(R\omega) = \mathcal{F}\{f(R\cdot)\}(\omega). \quad (2)$$

A short proof is given below, as a good example of manipulation of the Fourier transform as a multi-dimensional integral (on \mathbb{R}^d).

Proof. It is quite simple to obtain, as the (bijective) change of variable¹⁰ $\mathbf{x} \mapsto \mathbf{x}' = R^{-1}\mathbf{x}$ has a Jacobian equals to 1, so we can apply it in the integral defining the Fourier transform (or its inverse), without changing its value. For instance for \mathcal{F} , for any $\omega \in \mathbb{R}^d$, we have,

$$\begin{aligned} (r_R \circ \mathcal{F})\{f\}(\omega) &= \widehat{f}(R\omega) \stackrel{\text{def}}{=} \int_{\mathbb{R}^d} f(\mathbf{x}) \exp(-j(R\omega)^T \mathbf{x}) d\mathbf{x} \\ (\text{With } \mathbf{x}' = R^{-1}\mathbf{x}) &= \int_{\mathbb{R}^d} f(R\mathbf{x}') \exp(-j(R\omega)^T (R\mathbf{x}')) d\mathbf{x}' \\ (R \text{ conserves inner products}) &= \int_{\mathbb{R}^d} f(R\mathbf{x}') \exp(-j\omega^T \mathbf{x}') d\mathbf{x}' \\ &= \int_{\mathbb{R}^d} (r_R\{f\})(\mathbf{x}') \exp(-j\omega^T \mathbf{x}') d\mathbf{x}' \\ &= \widehat{r_R\{f\}}(\omega) = (\mathcal{F} \circ r_R)\{f\}(\omega). \end{aligned}$$

And the same technique works for \mathcal{F}^{-1} . □

1.3.3 The convolution theorem

The last important property of the Fourier transform we need to introduce is the so-called *convolution theorem*. One of the main interest of this operator \mathcal{F} , is its ability of mapping a convolution (on the *space domain*, i.e., \mathbf{x}) to a point-wise product (on the *frequency domain*, i.e., ω).

¹⁰ This change of variable $\mathbf{x} \mapsto \mathbf{x}' = R^{-1}\mathbf{x}$ or $\mathbf{x} \mapsto \mathbf{x}' = R\mathbf{x}$ is used several times afterward.

Theorem 1.4. For any functions f, g in $\mathbf{L}^2(\mathbb{R}^d, \mathbb{R})$, the Fourier transform \mathcal{F} maps their convolution $(f * g)$ to the (point-wise) product of their Fourier transforms, i.e.,

$$\mathcal{F}(f * g) = (\mathcal{F}f) \cdot (\mathcal{F}g). \quad (3)$$

It is also true for functions in $\mathbf{S}(\mathbb{R}^d, \mathbb{R})$, and for distributions in $\mathbf{S}'(\mathbb{R}^d, \mathbb{R})$.

Easy Proof. Not proved here, cf. [SW71] or [Rud91, Th.7.2.(c)]. □

What we reminded here is very important for the rest of this report.

1.4 Operators, “real world vs theory”

In theory, there is two kinds of functions, as they can have either a continuous domain (\mathbb{R} or \mathbb{R}^2 , or \mathbb{R}^d in any dimension), or a discrete domains (\mathbb{Z} or $\llbracket 1, M \rrbracket$, \mathbb{Z}^2 or $\llbracket 1, N \rrbracket \times \llbracket 1, M \rrbracket$ for images, or \mathbb{Z}^d in any dimension). Discrete functions and their operators have been studied extensively, cf. for instance [Uns16], and in this work we only focus on continuous domain functions and their operators.

When we apply an operator to a “real-world” image, we need to switch from the continuous setting to the discrete setting, and this usually raises issues like sampling and approximation errors. We do not deal with these aspects here, but they should not be forgotten if one wants to implements our operators and use them in practice.

Operators for the discrete world include almost any transform one can think of one 1D signals or images. The possible applications of operators are countless, and this domain of research is very rich. They are also called “filters”, “image analysis operators” and other names, but they all are different ways of referring to an operator, which can also be seen as a black-box: feed it an input image, it gives an output image. For examples, we can consider a low-pass filter to reduce the noise on an audio signals, a Wiener filter to denoise an image, a contour detection algorithm for images, or a shape detector to locate molecules or cells in medical 2D or 3D images. The general public would be more familiar with filters and algorithms related to photos, the best example being Instagram filters!

1.5 Short explanations on Green’s functions

In simple words, a Green’s function ρ_G of a convolution G is the convolution kernel of its (left) inverse: $G^{-1}\{f\} = \rho_G * f$, satisfying $GG^{-1}\{f\} = f$.

In the world of Fourier multipliers, symbolically we have $\widehat{\rho_G}(\omega) = 1/\widehat{g}(\omega)$ (when it has a meaning) and so $\rho_G = \mathcal{F}^{-1}\{1/\widehat{g}\}$. All these notions of kernels and Fourier multipliers are introduced below in section 3.2.

Green’s functions are very useful, mainly for two aspects of signal processing:

- To develop new denoising algorithms, for compressive sensing (cf. [FR13]). A well-known example is the total-variation denoising (TV- L^1) in dimension 1, which is based on the derivative operator D and its inverse computed with its Green’s function ρ_D , cf. [Pey11]¹¹.

¹¹ Cf. this MATLAB notebook on “Total Variation Regularization for Denoising” www.numerical-tours.com/matlab/denoisingsimp_4_denoiseregul

- To generate splines, which are used to develop new sampling schemes. For instance, [UFW16] is a recent research article using Green's functions for generalized TV minimization, and the splines are used to express the solution of the minimization problem, which basically uses some well-chosen samples of the data (sampled at the splines' knots). A good reference is [UT14, Sec.6], more specifically Sec.6.2.2. For example, the derivative operator D in dimension 1 gives box-splines. The figure 6.3 in [UT14] shows the link between four objects, which are different ways of seeing the same thing: a Green's function (*i.e.*, an inverse operator), a spline, an interpolator (*i.e.*, a sampling scheme), and a wavelet (basis functions).

Green's functions, or inverse operators, can also be used for stochastic processes and sparse stochastic processes, cf. [UT14, Ch.5] for the whole picture.

We will not detail more all these notions, as it would go beyond the reach of this report. An interested reader could consult [UB07, BU07] for a good introduction on splines and Green's functions. We do talk much about Green's function, so the reader should not worry if she/he is not familiar with them.

We ran out of time to study more this aspect and to conclude the computations of a Green's function for our elementary blocks $G_{\lambda,\alpha}$.

Conclusion

This first section started to explain the goals of our work, as well as the outline of the next sections. We then detailed the notations and symbols used in this document, to be perfectly clear and to remind the user of some classic definitions and constructions in harmonic and Fourier analysis. Some important results on the Fourier transform were given, along with one proof, but the interested reader should consult a reference book if needed ([SW71, Hör83]). We concluded with a short reminder on Green's functions, and some explanations on their possible applications: in compressed sensing, they can be used to develop new algorithms as well as new splines.

The next section 2 studies functional operators and their geometric invariance properties, in a broad setting (dimension $d \in \mathbb{N}^*$). All the characterizations in section 3 are valid in any dimension, but for steerability we quickly restrict to 2D operators in section 4.

2 Operators, and their geometric invariance properties

This section first defines test functions and operators in 2.1, with an emphasis on the domain that can be considered, then the classic properties of linearity and continuity for operators are introduced in 2.2, and finally some geometric properties of operators are presented in 2.3. These properties of translation-, scaling- and rotation-invariance share the advantage of being easily understandable from a “physical” point of view, and can easily be explained in words and not only with a complicated formula. We try to prove every single result in this section, but proofs can be skipped if the reader is familiar with this theory.

2.1 Defining functional operators

We introduce here the notion of operator, and explain all the different properties that an operator can have, from the classic ones (linearity, continuity), to some less usual ones. Along definitions, examples are given regularly, but a detailed list of operators is given in the next section 3.1.

2.1.1 Test functions, and some tricks to simplify their domains

Let $d \in \mathbb{N}^*$, and $\Omega \subset \mathbb{R}^d$ be a domain for real-valued “test” functions $f : \Omega \rightarrow \mathbb{R}$. The following remarks justify the generality of this setting.

- We can restrict to $\Omega \subset \mathbb{R}^d$ instead of $\Omega \subset \mathbb{C}^d$, thanks to the usual isomorphism $\mathbb{C} \simeq \mathbb{R}^2$. It is always simpler to study functions of the real variable.
- Similarly, we can restrict to study real-valued functions, because if $f : \Omega \rightarrow \mathbb{C}$, we can study its real and imaginary parts *separately*.
- And with the same trick, we can also study multivariate-valued functions, by studying their components *separately*.

To avoid the (sometime delicate) manipulation of this sub-domain $\Omega \subset \mathbb{R}^d$, we will **always assume** that **the functions are defined on the whole domain $\Omega = \mathbb{R}^d$** (as it is usually done in this domain, [UB07, BU07]). We do not lose in generality, because if $f : \Omega \rightarrow \mathbb{R}$ is defined on a smaller domain¹², we can extend it to another $f^\# : \mathbb{R}^d \rightarrow \mathbb{R}$:

$$\forall \mathbf{x} \in \mathbb{R}^d, \quad f^\#(\mathbf{x}) \stackrel{\text{def}}{=} \begin{cases} f(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega = \text{Dom}(f) \\ 0 & \text{otherwise.} \end{cases}$$

Using this trick does not change the values of any integrals involving f on Ω :

$$\int_{\mathbb{R}^d} f^\#(\mathbf{x})\phi(\mathbf{x})d\mathbf{x} = \int_{\Omega} f(\mathbf{x})\phi(\mathbf{x})d\mathbf{x}$$

and therefore $f^\#$ and f will have the same Fourier transform, the same moments etc. We only consider functions spaces such that this integral $\int_{\mathbb{R}^d} f(\mathbf{x})\phi(\mathbf{x})d\mathbf{x}$ has a meaning, for instance $f, \phi \in \mathbf{L}^2(\mathbb{R}^d)$ or $f \in \mathbf{S}(\mathbb{R}^d), \phi \in \mathbf{S}'(\mathbb{R}^d)$.

¹² This is the case of images, for instance, as they are defined on a (small) compact set in 2D.

Note that we could lose some properties of smoothness (continuity, C^k etc) on the frontier of the domain Ω , but in this framework this is not an issue, as such properties are not studied point-wise but only globally (e.g., for the test functions in $\mathbf{S}(\mathbb{R}^d)$).

To sum up, in all this report, we only consider test functions f from \mathbb{R}^d to \mathbb{R} .

2.1.2 Operators on test functions

The main concept we studied is operators on such function spaces. An operator G is simply a function of higher order, taking a $f : \mathbb{R}^d \rightarrow \mathbb{R}$ and producing another function, denoted $G(f) = G\{f\} : \mathbb{R}^d \rightarrow \mathbb{R}$. We give below a more formal definition, along with examples. We use the letter G (or G_1, G_2, G', \dots) to denote operators, and f (or f_1, f_2, \dots) for test functions.

Let E be the function space we consider, which can be $\mathbf{L}^2(\mathbb{R}^d)$ or $\mathbf{S}(\mathbb{R}^d)$ for instance, and let F be another function space (can be equal to E or not).

Definition 2.1 (Functional operator). G is a functional operator, or simply an operator, from E to F when $G : E \rightarrow F$, i.e., $\forall f \in E, G(f) = G\{f\} \in F$, where $G(f) : \mathbf{x} \mapsto (G\{f\})(\mathbf{x})$, and **not** $G(f(\mathbf{x}))$.

Definition 2.2. G is said to operate on E when $G : E \rightarrow E$ (i.e., $G\{f\}$ has to be in E).

Example 2.3. To start with dummy examples, we can quote Id the *identity operator*: $\text{Id}\{f\} = f$, and the *null operator* $\mathbf{0}_{E \rightarrow E}$: $\mathbf{0}_{E \rightarrow E}(f) = \mathbf{0}_{\mathbb{R}^d \rightarrow \mathbb{R}} \in E$, both operating on E .

More interesting examples will be given regularly in this report.

Remark 2.4 (“Common domain and command dimension hypothesis”).

Without specifying it each time, it will always be assumed that the test functions in E and their images $G\{f\}$ in F are functions (or distributions) **on the same space domain**, $\Omega = \mathbb{R}^d$. The dimension d is both the “input” and “output” space dimensions. This simplifies greatly the notations, as for instance translation-invariance for operator (see later, definition 2.21) would have no meaning if the function $G\{f\}$ had a different domain (or domain dimension) than the function f .

2.2 Classic properties of operators

This short part recall the general definition of linearity and continuity for operators, and gives the first interesting examples of operators. The algebraic results presented here are very classic, but we chose to prove (almost all of) them cleanly, in order to present slowly how operators work and can be manipulated. An interested reader could consult a reference book for more results (any book on linear algebra); but *a reader familiar with all this could skip this part and jump directly to section 2.3.*

2.2.1 Linear operators (L)

We now require E and F to be vector spaces¹³.

¹³ The vector spaces considered in this report are usually on the field \mathbb{R} , unless it is clearly specified for \mathbb{C} . No other fields \mathbb{K} are used.

Definition 2.5 (Linear operator). G is a linear operator from E to F when

$$\forall \lambda_1, \lambda_2 \in \mathbb{R}, \forall f_1, f_2 \in E, \quad G\{\lambda_1 f_1 + \lambda_2 f_2\} = \lambda_1 G\{f_1\} + \lambda_2 G\{f_2\} \in F.$$

Definition 2.6 (Vector space of linear operators). The set of (real/complex) linear operators (from E to F) is (real/complex) a vector space (i.e., on the field \mathbb{R} or \mathbb{C}), denoted $\mathcal{L}(E, F) = \mathcal{L}_{\mathbb{R}}(E, F)$ or $\mathcal{L}_{\mathbb{C}}(E, F)$ (or simply $\mathcal{L}(E)$ if $F = E$).

Easy Proof. We prove two things in order to make this a definition:

- $\mathbf{0}_{E \rightarrow F}$ is linear so in $\mathcal{L}(E, F)$,
- Let $\mu_1, \mu_2 \in \mathbb{R}, G_1, G_2 \in \mathcal{L}(E, F)$ Then $\mu_1 G_1 + \mu_2 G_2 : f \mapsto (\mu_1 G_1 + \mu_2 G_2)\{f\} = \mu_1(G_1\{f\}) + \mu_2(G_2\{f\})$ which is still linear on $f \in E$, to F (because F is a vector space), so $\mu_1 G_1 + \mu_2 G_2 \in \mathcal{L}(E, F)$ too. \square

Remark 2.7. Except in very special cases, E, F and $\mathcal{L}(E, F)$ are infinitely dimensional, but they can be finitely dimensional.

For instance, for real function ($d = 1$), the function space $E = \text{Span}_{\mathbb{R}}\{\sin, \cos\}$ (of dimension 2) has a space of linear operators $\mathcal{L}(E)$ of finite dimension 4. Note that if E, F are finitely dimensional, then so is $\mathcal{L}(E, F)$.

Proposition 2.8. When it is well defined, composition preserves the linearity of operators, i.e., $\mathcal{L}(E)$ is also stable by composition.

Easy Proof. It is straightforward, by using the definition: if G_1, G_2 are linear, respectively in $\mathcal{L}(F, H)$ and $\mathcal{L}(E, F)$, and if we can write $G_1 \circ G_2$ (i.e., $\text{Im}(G_2) \subset \text{Dom}(G_1) = F$, e.g., G_1, G_2 both are operating on $\mathbf{L}^2(\mathbb{R}^d)$), then $G \stackrel{\text{def}}{=} G_1 \circ G_2$ is in $\mathcal{L}(E, H)$. Indeed, for $f_1, f_2 \in E, \lambda_1, \lambda_2 \in \mathbb{R}$,

$$\begin{aligned} G\{\lambda_1 f_1 + \lambda_2 f_2\} &= (G_1 \circ G_2)\{\lambda_1 f_1 + \lambda_2 f_2\} \\ &= G_1\{G_2\{\lambda_1 f_1 + \lambda_2 f_2\}\} \\ (G_2 \text{ is linear}) &= G_1\{\lambda_1(G_2\{f_1\}) + \lambda_2(G_2\{f_2\})\} \\ (G_1 \text{ is linear}) &= \lambda_1 G_1\{(G_2\{f_1\})\} + \lambda_2 G_1\{(G_2\{f_2\})\} \\ &= \lambda_1 G\{f_1\} + \lambda_2 G\{f_2\}. \end{aligned} \quad \square$$

Without any further hypothesis on the linear operators G , we cannot really say anything more than these definitions, in the general setting (E being unspecified).

2.2.2 Continuous operators (C)

We can ask our operators $G : E \rightarrow F$ to be *continuous*, in the sense of this definition:

Definition 2.9 (Continuity for operators). For $G \in \mathcal{L}(E, F)$, with E and F both having a norm¹⁴ denoted $\|\cdot\|_E$ and $\|\cdot\|_F$, we say that G is continuous when it preserves limits on test functions¹⁵:

$$\forall (f_n)_{n \in \mathbb{N}} \in E^{\mathbb{N}}, \forall f \in E, \quad \|f_n - f\|_E \xrightarrow{n \rightarrow +\infty} 0 \implies \|G\{f_n\} - G\{f\}\|_F \xrightarrow{n \rightarrow +\infty} 0.$$

A simpler way of writing it is $f_n \rightarrow f \implies G\{f_n\} \rightarrow G\{f\}$, for all $(f_n)_{n \in \mathbb{N}}, f$.

For example, on $E = \mathbf{L}^2(\mathbb{R}^d)$, the norm can be the usual 2-norm $\|\cdot\| = \|\cdot\|_2$.

Remark 2.10 (Concise notation for our class of operators). From now on, when there is no ambiguity, we will denote an operator (or the class of operators) satisfying a certain property with its corresponding bold capital letter (e.g., **C** for continuity). And if an operator satisfies more than one property, we will simply concatenate the bold capital letters, for instance a linear (**L**) and continuous (**C**) operator will be shorten to a **LC** operator, and so on.

This notation will be very convenient when we start to consider complex classes, such as linear (**L**), continuous (**C**), translation-invariant (**TI**), γ -scale-invariant (γ -**SI**), rotation-invariant (**RI**) and steerable (**St**) operators: instead of giving it a complicated name or always repeating this long list of properties, it is simply denoted **LC TI γ -SI St**. The same notation can either be interpreted as a shortcut for a list of adjective (when applied to one operator), or the set (the class) of operators having these properties. When the actual domain E (and image F) of the operators is important, it is specified before using this shortcut notation.

We will denote **C** the class of continuous operators on E (not studied by itself), and **LC** the class of linear and continuous operators (on E).

Proposition 2.11. *If E, F are linear spaces, this class **LC** is a linear sub-space of the space of operators from E to F , stable by composition if $F = E$.*

Easy Proof. It is well known that continuity of functions or operators is preserved by linear combination, by triangle inequality. And the null operator is obviously continuous.

Stability by composition is easy to obtain with the characterization of continuity as being bounded (for linear operators). Indeed, if G_1, G_2 are bounded, then $G = G_1 \circ G_2$ is still bounded: $\|G\{f\}\| = \|(G_1 \circ G_2)\{f\}\| = \|G_1\{G_2\{f\}\}\| \leq C_1 \|G_2\{f\}\| \leq C_1 C_2 \|f\| = C \|f\|$ with $C = C_1 C_2$. \square

Definition 2.12. *A linear continuous (**LC**) operator is called an integral operator¹⁶.*

Remark 2.13. It is well known that a linear operator G is continuous iff¹⁷ it is continuous on $\mathbf{0}_E$ iff it is bounded iff it is Lipschitz-continuous. Note that we will not prove continuity like this, because Schwartz's first kernel theorem 3.1 gives a simpler way to prove that an operator is **LC**, by directly providing it's Schwartz kernel k_G .

Remark 2.14. In the next paragraphs, we define some additional properties that an operator G can have. The properties **TI**, **0-SI** and **RI** are defined by requiring G to commute with a special family of operators (e.g., translation, dilatation etc). It might seem to be a pointless observation, but it proves directly that the classes of operators having one or more of these properties (e.g., **TI** or **TI + RI**¹⁸) are all stable under linear combinations, and under compositions (if $E = F$).

¹⁵ Or a (countable or not countable) family of semi-norms, e.g., the Sobolev norms $\|\cdot\|_{p,\alpha}$ on $E = \mathbf{S}(\mathbb{R}^d)$.

¹⁵ Or when it preserves limits on test functions for *all* the semi-norms, in the case of semi-norms.

¹⁶ This will be justified by the first Schwartz's kernel theorem, as $G\{f\} = \int k_G(\cdot, \mathbf{y})f(\mathbf{y})d\mathbf{y}$ is exactly the form of an integral operator.

¹⁷ We use the standard abbreviation "iff" for "if and only if", i.e., a logical equivalence $\dots \Leftrightarrow \dots$.

2.3 Geometric properties: translation-, scaling- and rotation-invariance

The properties introduced here are natural hypothesis for our operators. Such operators can for example be some sort of filters on images, and it is very meaningful to ask at least translation-invariance and scaling-invariance (or γ -scaling-invariance) for filters, as we will explain below. Having at least these two properties (**TI**, γ -**SI**) is a natural framework to study interesting and “physically meaningful” operators in 2D, and that is why our study of steerable operators is focused on **LC TI** γ -**SI** operators at the end (cf. section 4.4.2).

2.3.1 Additional hypothesis on the function space E

Before going into more details on these geometric transforms, we need to ask three properties to the function spaces E in order to clarify three points.

Remark 2.15 (Vector space). From now on, we **always** ask E and F to be vector spaces.

Remark 2.16 (Stability hypothesis for E). From now on, we **always** ask E, F to be stable by translations, scaling, γ -scaling and rotations, *i.e.*, E **should be stable by geometric transforms**.

It is the case for $E = \mathbf{L}^2(\mathbb{R}^d, \mathbb{R})$, $E = \mathbf{S}(\mathbb{R}^d, \mathbb{R})$ and $E = \mathbf{S}'(\mathbb{R}^d, \mathbb{R})$.

Remark 2.17 (Between \mathbf{S} and \mathbf{S}'). From now on, we **always** consider the setting where

$$\mathbf{S}(\mathbb{R}^d, \mathbb{R}) \subset E \subset \mathbf{S}'(\mathbb{R}^d, \mathbb{R}). \quad (4)$$

It is also the case for $E = \mathbf{L}^2(\mathbb{R}^d, \mathbb{R})$, $E = \mathbf{S}(\mathbb{R}^d, \mathbb{R})$ and $E = \mathbf{S}'(\mathbb{R}^d, \mathbb{R})$.

A consequence is that we can always assume at least that $G\{f\} \in \mathbf{S}'$.

2.3.2 In fact, we need *topological* spaces

Remark 2.18 (Topological Vector Spaces). We also need to precise that from now on, we always consider *topological vector spaces*, and not only vector spaces.

In a nutshell, it means that the vector spaces (E and F for functions, and the class of operators, *e.g.*, **LC** or **LC TI**) are endowed with a *topology*, compatible with their structure of vector spaces. Note that they are only asked to be topological, not normed.

And so whenever we say that E_1 is a linear sub-space of E , it means in fact that E_1 is a topological vector space too, included in E , *sharing the vector space structure* (*i.e.*, same $\mathbf{0}_E$, same sum $+$ and multiplication by a scalar \cdot), and *sharing the topological structure* (*i.e.*, all the open sets of E_1 are open sets of E).

In other words, *the topology is preserved whenever the linearity is preserved*. But for the sake of conciseness and simplicity, we preferred to give the proofs only for the linearity part, but all the results in the rest of this section are not only valid for vector spaces and linear sub-space, but *are also valid for topological vector spaces and topological linear sub-spaces*.

For more details on topological vector spaces, we let the reader consult a reference book, two good historical references being [Gro73] and [Bou81].

2.3.3 Translation-invariant operators (TI)

In a short sentence: *An operator is translation-invariant when its output on a translated input signal is the translated output on the non-translated input signal.*

For instance in 2D, it means that the operation on the input image does not depend on the exact location of patterns in the image, but on *relative* location. For example, if we are implementing a face detector, it should detect a face regardless of its location.

As soon as we consider translation operators, we see that it was a good idea to only consider functions of domain $\text{Dom}(f) = \Omega = \mathbb{R}^d$, as we need to be able to consider $f(\cdot - \mathbf{x}_0)$ for any $\mathbf{x}_0 \in \mathbb{R}^d$.

Definition 2.19 (Translation operator on E). For a fixed $\mathbf{x}_0 \in \mathbb{R}^d$, let $\tau_{\mathbf{x}_0}$ be defined by:

$$\tau_{\mathbf{x}_0} : E \rightarrow E, f \mapsto \tau_{\mathbf{x}_0}\{f\} \stackrel{\text{def}}{=} f(\cdot - \mathbf{x}_0) : \mathbb{R}^d \rightarrow \mathbb{R}, \mathbf{x} \mapsto f(\mathbf{x} - \mathbf{x}_0).$$

$\tau_{\mathbf{x}_0}$ is called the translation operator of offset \mathbf{x}_0 , or simply translation by \mathbf{x}_0 .

Lemma 2.20. For any $\mathbf{x}_0 \in \mathbb{R}^d$, $\tau_{\mathbf{x}_0}$ is a linear and continuous (LC) operator on E

Easy Proof. • Its linearity is obvious by the distributivity of function evaluation: $\tau_{\mathbf{x}_0}\{\lambda_1 f_1 + \lambda_2 f_2\}(\mathbf{x}) = (\lambda_1 f_1 + \lambda_2 f_2)(\mathbf{x} - \mathbf{x}_0) = \lambda_1 f_1(\mathbf{x} - \mathbf{x}_0) + \lambda_2 f_2(\mathbf{x} - \mathbf{x}_0) = \lambda_1 \tau_{\mathbf{x}_0}\{f_1\} + \lambda_2 \tau_{\mathbf{x}_0}\{f_2\}$.

- Its continuity comes from the fact that all the norms considered for function spaces (on \mathbb{R}^d) are integrals on the whole domain \mathbb{R}^d , and so the (bijective) change of variable $\mathbf{x} \mapsto \mathbf{x} - \mathbf{x}_0$ (of Jacobian 1) can always be applied, without changing the value of the norm. For instance, for classic p -norms (on \mathbb{R}^d), $\|\tau_{\mathbf{x}_0} f\|_p = \|f\|_p$. More complicated constants could pop out for semi-norms or norms based on derivatives, but there will always be a $C_{\mathbf{x}_0} \in \mathbb{R}_+$ such that $\|\tau_{\mathbf{x}_0} f\| = C_{\mathbf{x}_0} \|f\|$, hence $\tau_{\mathbf{x}_0}$ is continuous¹⁹.

Note: the same scheme of proof will be used for other operators. \square

We can now ask our operators $G : E \rightarrow F$ to be *translation-invariant* (also called *shift-invariant*), in the sense of the following definition:

Definition 2.21 (Translation invariance for operators). For $G \in \mathcal{L}(E, F)$, G is translation-invariant (TI) when it commutes with any translation operator $\tau_{\mathbf{x}_0}$, i.e., when

$$\forall \mathbf{x}_0 \in \mathbb{R}^d, \forall f \in E, \forall \mathbf{x} \in \mathbb{R}^d, (G\{f\})(\mathbf{x} - \mathbf{x}_0) = (G\{f(\cdot - \mathbf{x}_0)\})(\mathbf{x})$$

$$\text{ie. } = ((\tau_{\mathbf{x}_0} G)\{f\})(\mathbf{x}) = ((G\tau_{\mathbf{x}_0})\{f\})(\mathbf{x}).$$

A simpler way of writing it²⁰ is simply $\tau_{\mathbf{x}_0} \circ G = G \circ \tau_{\mathbf{x}_0}$, for all $\mathbf{x}_0 \in \mathbb{R}^d$.

We will denote **TI** the class of translation-invariant operators (not studied by itself), and by **LC TI** the class of linear, continuous and translation-invariant operators (on E).

An interesting general result is the following: for an operator H , the set of operators G commuting with it is always a linear sub-space.

¹⁹ And $\tau_{\mathbf{x}_0}$ is usually an isometry, i.e., $C_{\mathbf{x}_0} = 1$ except maybe for unusual function spaces E .

²⁰ When there is no risk of ambiguity, we tend to omit the composition symbol \circ , as usual.

Lemma 2.22 (Commutant²¹ of an operator). *Let $H \in \mathcal{L}(E)$ and let $\text{Comm}(H) \stackrel{\text{def}}{=} \{G \in \mathcal{L}(E), H \circ G = G \circ H\}$ be the commutant of H . It is a linear sub-space of $\mathcal{L}(E)$.*

Proof. First, the null-operator obviously commutes with H ($\mathbf{0}_{E \rightarrow E} \circ H = \mathbf{0}_{E \rightarrow E} = H \circ \mathbf{0}_{E \rightarrow E}$). Then, if $G_1, G_2 \in \text{Comm}(H)$, $\lambda \in \mathbb{R}$, then by linearity of H , $H \circ (G_1 + \lambda G_2) = H \circ G_1 + H \circ (\lambda G_2) = G_1 \circ H + \lambda(H \circ G_2) = G_1 \circ H + \lambda(G_2 \circ H) = G_1 \circ H + (\lambda G_2) \circ H = (G_1 + \lambda G_2) \circ H$. $\text{Comm}(H)$ is also a linear sub-space as the kernel of a linear meta-operator, $G \mapsto H \circ G - G \circ H$. \square

Remark 2.23. For invertible H , $\text{Comm}(H)$ can also be written as $\{G \in \mathcal{L}(E), H \circ G \circ H^{-1} = G\}$, so it can be seen as the set of fix point by the action $G \mapsto H \circ G \circ H^{-1}$. So if H is in a certain group A (e.g., group of translation or rotation operators), the set $\bigcap_{H \in A} \text{Comm}(H)$, of operators G commuting with every H in A (e.g., translation- or rotation-invariant operators), can also be interpreted in terms of fix points under a group action, named the *conjugation action* ($H \cdot G \stackrel{\text{def}}{=} H \circ G \circ H^{-1}$).

Some readers could be familiar with these concepts, but we did not study them because, despite our limited knowledge of this domain, we think that nothing from this theory of group actions could be useful for our questions.

Proposition 2.24. *If E, F are linear spaces, this class **LTI** is a linear sub-space of the space of operators from E to F , stable by composition if $F = E$.*

Easy Proof. In particular, for one translation operator $\tau_{\mathbf{x}_0}$, $\text{Comm}(\tau_{\mathbf{x}_0})$ is a linear sub-space. And by considering the intersection of all such linear sub-spaces, indexed²² by $\mathbf{x}_0 \in \mathbb{R}^d$, it implies that the set of operators commuting will all translations (i.e., the set of linear and translation-invariant operators, **LTI**) is a linear sub-space. \square

Corollary 2.25. *The class **LC TI** is a linear sub-space of $\mathcal{L}(E, F)$, stable by composition if $F = E$.*

Easy Proof. By intersection. \square

Warning 2.26. This class is usually denoted “**LSI**” in this community (e.g., [UT14]) for “**Linear Shift-Invariant**”, but the notation “**SI**” will instead be used for γ -**SI**, denoting the γ -scale-invariance and **not** the shift-invariance (see later).

Definition 2.27. *A linear continuous translation-invariant (**LC TI**) operator is called a convolution operator, or simply a convolution²³.*

²¹ This object has most surely been already studied but we did not find any common name for it, so “commutant” is our own way of calling it.

²² Even an infinite or a non-countable intersection of linear sub-spaces is a still a linear sub-space.

²³ This will be justified by the second Schwartz’s kernel theorem, as $G\{f\} = \int g(\cdot - \mathbf{y})f(\mathbf{y})d\mathbf{y} = g * f$ is exactly the form of a convolution operator.

2.3.4 Scaling-invariant operators (0-SI)

In a short sentence: *An operator is scaling-invariant when its output on a zoomed input signal is the zoomed output on this non-zoomed input signal.*

For instance in 2D, it means that the operation on the input image does not depend on the absolute scale of patterns in the image, but on relative scales. With the same example as above, a face detector should detect a face regardless of its size on the image (it should detect small faces in the background as well as large faces in the foreground).

Definition 2.28 (Scaling operator on E). For a fixed $a \in \mathbb{R}_+^*$ (i.e., $a \in \mathbb{R}$, $a > 0$), let d_a the operator on E defined by:

$$d_a : E \rightarrow E, f \mapsto d_a\{f\} \stackrel{\text{def}}{=} f(a \cdot) : \mathbb{R}^d \rightarrow \mathbb{R}, \mathbf{x} \mapsto f(a\mathbf{x}).$$

d_a is called the scaling operator (or dilatation operator) of dilatation factor²⁴ $a > 0$.

Some books use the inverse definition with $(d_a\{f\})(x) = f(x/a)$, but we preferred the simpler notation.

Lemma 2.29. d_a is a linear and continuous (LC) operator on E (for any linear space E). And d_a is **not**²⁵ translation-invariant (TI).

Easy Proof. • Its linearity is obvious: $d_a\{\lambda_1 f_1 + \lambda_2 f_2\}(\mathbf{x}) = (\lambda_1 f_1 + \lambda_2 f_2)(a\mathbf{x}) = \lambda_1 f_1(a\mathbf{x}) + \lambda_2 f_2(a\mathbf{x}) = \lambda_1 d_a\{f_1\} + \lambda_2 d_a\{f_2\}$.

- Its continuity comes from the fact that all the norms considered for function spaces (on \mathbb{R}^d) are integrals on the whole domain \mathbb{R}^d , and so the (bijective) change of variable $\mathbf{x} \mapsto a\mathbf{x}$ (of Jacobian a^d) can always be applied, without changing the value of the norm. For instance, for classic p -norms, $\|d_a f\|_p = a^d \|f\|_p$. More complicated constants can pop out for semi-norms or norms based on derivatives, but there will always be a $C_a \in \mathbb{R}_+$ such that $\|d_a f\| = C_a \|f\|$, hence d_a is continuous²⁶.

- And d_a is not **TI** in general:

$$(\tau_{\mathbf{x}_0} d_a)\{f\}(\mathbf{x}) = f(a\mathbf{x} - \mathbf{x}_0) \neq f(a(\mathbf{x} - \mathbf{x}_0)) = (d_a \tau_{\mathbf{x}_0})\{f\}(\mathbf{x}). \quad \square$$

We can now ask our operators $G : E \rightarrow F$ to be *scaling-invariant* (also sometimes called *dilatation-invariant*, or *homogeneous*), in the sense of the following definition:

Definition 2.30 (Scaling invariance for operators). For $G \in \mathcal{L}(E, F)$, we say that G is scaling-invariant (0-SI) when it commutes with any scaling operator d_a , i.e., when

$$\begin{aligned} \forall a \in \mathbb{R}_+^*, \forall f \in E, \forall \mathbf{x} \in \mathbb{R}^d, \quad (G\{f\})(a\mathbf{x}) &= (G\{f(a \cdot)\})(\mathbf{x}) \\ \text{ie.} \quad &= ((d_a G)\{f\})(\mathbf{x}) = ((G d_a)\{f\})(\mathbf{x}). \end{aligned}$$

A simpler way of writing it is $d_a \circ G = G \circ d_a$ for all $a \in \mathbb{R}_+^*$.

²⁴ Negative dilatation factor $a < 0$ are **never** considered, neither in the literature nor in this report, one reason being that they would mess up with the time-reversal property of the Fourier transform, another reason being that a simple derivative would not be scale-invariant if negative a were considered.

²⁵ Conversely, this gives directly that any translation operator $\tau_{\mathbf{x}_0}$ is **not** scale-invariant.

²⁶ And d_a is usually an isometry, up to a constant – and it always preserves orthogonality.

As for the translation-invariance, this definition also characterizes G by its commutativity with a certain family of operators (the dilatations here).

Remark 2.31. The definition has a very similar meaning for functions and operators: a function f is scale-invariant when applying a dilatation to its input *does not change* its value; while an operator G is scale-invariant when it *commutes* with any dilatation operator. But the correct way to “apply” a dilatation to an operator is to zoom-out the input and zoom-in the output, *i.e.*, to consider $G_a = d_a \circ G \circ d_{1/a}$. And so G is scale-invariant when $G_a = G$ for all $a > 0$, which means that applying a dilatation does not change G .

We will denote **0-SI** the class of scaling-invariant operators²⁷ (not studied by itself), and by **LC 0-SI** the class of linear, continuous and scaling-invariant operators (on E).

Proposition 2.32. *If E, F are linear spaces, this class **0-SI** is a linear sub-space of the space of operators from E to F , stable by composition if $F = E$.*

Easy Proof. The linearity is immediate as **0-SI** is the intersection of all $\text{Comm}(d_a)$ for $a \in \mathbb{R}^d$. And the stability by composition is simple to obtain with the definition: if G_1, G_2 are **0-SI**, let $G = G_1 \circ G_2$, then for any $a \in \mathbb{R}_+^*$, one have, by associativity of the composition,

$$\begin{aligned} d_a \circ G &= d_a \circ (G_1 \circ G_2) = (d_a G_1) G_2 = (G_1 d_a) G_2 \\ &= G_1 (d_a G_2) = G_1 (G_2 d_a) = (G_1 G_2) d_a = G \circ d_a. \end{aligned} \quad \square$$

Corollary 2.33. *LC **0-SI** operators also form a linear sub-space of $\mathcal{L}(E, F)$, stable by composition if $F = E$.*

Easy Proof. By intersection. □

2.3.5 γ -scaling-invariant operators (γ -SI)

In a short sentence: *An operator is γ -scaling-invariant when its output on a zoomed input signal is the zoomed output on this non-zoomed input signal, with a (possibly) different zoom factor, dependent geometrically on the input zoom factor.*

For instance in 2D, it means that the operation on the input image does not depend on the absolute scale of patterns in the image, but on relative scales (up to a geometric change of variable). For the same example, a face detector should detect a face regardless of its size on the image, but it could “react” stronger to larger faces.

Let $\gamma \in \mathbb{R}_+$ (*i.e.*, $\gamma \geq 0$). We can also ask our operators $G : E \rightarrow F$ to be γ -scaling-invariant (also sometimes called γ -homogeneous), in the sense of the following definition:

Definition 2.34. *For $G \in \mathcal{L}(E, F)$, we say that G is γ -scaling-invariant (γ -SI) when*

$$\begin{aligned} \forall a \in \mathbb{R}_+^*, \forall f \in E, \forall \mathbf{x} \in \mathbb{R}^d, \quad & (G\{f\})(\mathbf{x}/a) = a^\gamma (G\{f(\cdot/a)\})(\mathbf{x}) \\ \textit{i.e.} \quad & ((d_{1/a} \circ G)\{f\})(\mathbf{x}) = (d_{1/a}\{G\{f\}\})(\mathbf{x}) \\ & = a^\gamma (G\{d_{1/a}\{f\}\})(\mathbf{x}) = a^\gamma ((G \circ d_{1/a})\{f\})(\mathbf{x}). \end{aligned}$$

A simpler²⁸ way of writing it is $d_a \circ G \circ d_{1/a} = 1/a^\gamma G$, for all $a \in \mathbb{R}_+^$. This constant γ is called “the” order of homogeneity.*

²⁷ We use the notation **0-SI** instead of simply **SI** for “scale-invariant” to highlight the fact that it is a special case of γ -scale-invariance, with $\gamma = 0$, see the next subsection.

We can state and easily prove this simple result, justifying the definition of the order of homogeneity.

Lemma 2.35 (Uniqueness of the order of homogeneity for non-trivial operators).

If G is γ -SI and γ' -scale-invariant for two different $\gamma \neq \gamma'$, then it is the null operator. In other words, non-trivial operators have a unique order of homogeneity γ .

Proof. Such G will satisfy $d_{1/a} \circ G = a^\gamma (G \circ d_{1/a}) = a^{\gamma'} (G \circ d_{1/a})$, so if G is not the null operator, when applied to a function $d_a\{f\}$, for $f \notin \ker(G)$, then the equality gives $a^\gamma G\{f\} = a^{\gamma'} G\{f\}$, with $G\{f\} \neq 0$ (because $(G \circ d_{1/a})(d_a\{f\}) = G\{f\} \neq 0$ as $f \notin \ker(G)$), so for $a \neq 0, 1$, it gives $\gamma = \gamma'$. So the converse sense of this implication is the desired result: $\gamma \neq \gamma' \implies G = \mathbf{0}_{E \rightarrow E}$. \square

Remember that for functions, γ -homogeneity means that $f(a\mathbf{x}) = a^\gamma f(\mathbf{x})$. For example, in dimension $d = 1$, the γ -homogeneous functions on \mathbb{R}_+ are exactly the function²⁹ $r \mapsto r^\gamma$ (up to a constant), as proved below.

Lemma 2.36 (Uniqueness of γ -homogeneous functions on \mathbb{R}_+). *For $\gamma \in \mathbb{R}_+$, let $f : \mathbb{R}_+ \rightarrow \mathbb{R}$, then f is γ -homogeneous iff there is $C \in \mathbb{R}$, such that $\forall r \in \mathbb{R}_+$, $f(r) = Cr^\gamma$.*

Proof. Let $C \stackrel{\text{def}}{=} f(1)$, then f is γ -homogeneous iff $\forall r \in \mathbb{R}_+$, $f(r) = f(r.1) = r^\gamma f(1)$, so $f(r) = Cr^\gamma$. \square

This small result will be usefully applied afterward to the radial dependency ($\omega \mapsto \rho(\|\omega\|)$) of the Fourier multiplier of γ -SI operators (see later, section 3.4.5).

We will denote γ -SI the class of γ -scaling-invariant operators (not studied by itself), and by LC γ -SI the class of linear, continuous and γ -scaling-invariant operators (on E).

Proposition 2.37. *If E, F are linear spaces, this class γ -SI is a linear sub-space of the space of operators from E to F .*

But the sum of two non-trivial γ -SI operator and γ' -SI operator is scale-invariant iff $\gamma = \gamma'$. Additionally, the class γ -SI is stable under composition iff $\gamma = 0$.

Easy Proof. γ -SI cannot be written with commutants, but it is easy to prove its stability by linear combinations. The null operator is γ -SI, and if G_1, G_2 are γ -SI and $\lambda_1, \lambda_2 \in \mathbb{R}$, then:

$$\begin{aligned} d_{1/a} \circ (\lambda_1 G_1 + \lambda_2 G_2) &= \lambda_1 d_{1/a} \circ G_1 + \lambda_2 (d_{1/a} \circ G_2) \\ &= \lambda_1 a^\gamma (G_1 \circ d_{1/a}) + \lambda_2 a^\gamma (G_2 \circ d_{1/a}) \\ &= a^\gamma (\lambda_1 G_1 + \lambda_2 G_2) \circ d_{1/a}. \end{aligned}$$

A simple example for the second point is to consider, naively, in dimension 1, $r \mapsto r$ and $r \mapsto r^2$, whose sum $r \mapsto r + r^2$ is clearly not γ -SI (for any γ). The same example can be easily extended to higher dimensions.

For the last point, the indirect sense is obvious, as we already proved that the 0-SI class is stable by composition. The direct sense is a consequence of the next proposition 2.39, and the fact that $2\gamma = \gamma$ iff $\gamma = 0$. \square

²⁸ Beware: the order is important here.

²⁹ The positive variable is denoted r and not x here, because this result will be used for the radial part of functions in higher dimensions later on, e.g., $f(r) = f(\|\omega\|)$.

Corollary 2.38. *LC γ -SI operators also form a linear sub-space of $\mathcal{L}(E, F)$.*

Easy Proof. By intersection. □

We can also say something of the composition of two γ -SI operator and γ' -SI operator, to generalize the fact that 0-scale-invariance is preserved by operator composition.

Proposition 2.39. *Let $\gamma, \gamma' \in \mathbb{R}_+$, and G, G' a γ -SI operator and a γ' -SI operator, respectively on F to H and E to F . Then $G'' = G \circ G'$ is also scale-invariant, of order $\gamma'' = \gamma + \gamma'$, on E to H .*

Easy Proof. We simply apply the definition:

$$\begin{aligned} \forall a \in \mathbb{R}_+^*, \forall f \in E, \forall \mathbf{x} \in \mathbb{R}^d, & (G''\{f\})(\mathbf{x}/a) \\ &= ((G \circ G')\{f\})(\mathbf{x}/a) = (G\{G'\{f\}\})(\mathbf{x}/a) \\ \text{(as } G \text{ is } \gamma\text{-SI)} &= a^\gamma G\{G'\{f\}(\cdot/a)\}(\mathbf{x}) \\ \text{(as } G' \text{ is } \gamma'\text{-SI)} &= a^\gamma a^{\gamma'} G\{G'\{f(\cdot/a)\}\}(\mathbf{x}) \\ &= a^{\gamma+\gamma'} (G \circ G'\{f(\cdot/a)\})(\mathbf{x}) \\ &= a^{\gamma''} (G''\{f(\cdot/a)\})(\mathbf{x}). \end{aligned}$$

And so³⁰ $G'' = G \circ G'$ is indeed $\gamma'' = \gamma + \gamma'$ scale-invariant. □

2.3.6 Rotation-invariant operators (RI)

In a short sentence: *An operator is rotation-invariant when its output on a rotated input signal is the rotated output on this non-rotated input signal.*

For instance in 2D, it means that the operation on the input image does not depend on the absolute orientation of patterns in the image, but on relative orientation. With the same example as above, a contour detector should detect a contour regardless of its orientation on the image.

Definition 2.40 (Rotations on \mathbb{R}^d). *A d -dimensional rotation is a square matrix, in $\mathbb{M}_{d,d}(\mathbb{R})$, which is orthogonal: $R^T R = R R^T = \mathbb{I}_d$. Such matrix R has a determinant $\det(R) = \pm 1$, and it corresponds to a linear operator on \mathbb{R}^d .*

The group of (proper and improper) rotations on \mathbb{R}^d is denoted $O(d)$, but is not used.

*We only consider **proper rotations** in this report³¹, i.e., when $\det(R) = +1$, and the group of proper rotations on \mathbb{R}^d is denoted $SO(d)$.*

In small dimensions, rotations are well known:

- The case $d = 1$ is empty: there is only one (proper) rotation, the identity.

³⁰ Such result will be obvious for **LC TI** operators, by using their Fourier multipliers (see later).

³¹ Improper rotations in dimension 2 simply are axial symmetries, and including them in the definition of rotation-invariance was reducing too much the range of interesting **RI** operators. Note that some books do not clarify whether they are considering improper and proper rotations, e.g., [UT14, Def.5.3, p90].

- The case $d = 2$ is simple but quite rich. (Proper) rotations on \mathbb{R}^2 are parametrized by a single parameter, $\theta \in [0, 2\pi)$: $R_\theta = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$, which operates on vectors of \mathbb{R}^2 . They have an interesting³² composition property: $R_{\theta_1} \circ R_{\theta_2} = R_{\theta_1 + \theta_2}$.
- Higher dimensional rotation-invariant operators were not studied by themselves.

From a rotation matrix R , we can associate a functional operator r_R .

Definition 2.41 (Rotation operator on E). For a fixed $R \in \text{SO}(d)$ (i.e., a rotation, operating on vectors), let r_R the corresponding operator on functions (on E) defined by:

$$r_R : E \rightarrow E, f \mapsto r_R\{f\} \stackrel{\text{def}}{=} f(R \cdot) : \mathbb{R}^d \rightarrow \mathbb{R}, \mathbf{x} \mapsto f(R\mathbf{x}).$$

r_R is called the rotation operator of rotation matrix³³ R .

Lemma 2.42. r_R is a linear and continuous (LC) operator on E (for any E).
And r_R is **not**³⁴ translation-invariant (TI) but is 0-scale-invariant (0-SI).

Easy Proof. • Its linearity is obvious: $r_R\{\lambda_1 f_1 + \lambda_2 f_2\}(\mathbf{x}) = (\lambda_1 f_1 + \lambda_2 f_2)(R\mathbf{x}) = \lambda_1 f_1(R\mathbf{x}) + \lambda_2 f_2(R\mathbf{x}) = (\lambda_1 r_R\{f_1\} + \lambda_2 r_R\{f_2\})(\mathbf{x})$.

- Its continuity comes from the fact that all the norms considered for function spaces (on \mathbb{R}^d) are integrals on the whole domain \mathbb{R}^d , and so the (bijective) change of variable $\mathbf{x} \mapsto R\mathbf{x}$ (of Jacobian 1) can always be applied, without changing the value of the norm. Hence, for classic p -norms, $\|R \circ f\|_p = \|f \circ R\|_p$. More complicated constants can pop out for semi-norms or norms based on derivatives, but there will always be a $C_R \in \mathbb{R}_+$ such that $\|r_R f\| = C_R \|f\|$, hence r_R is continuous³⁵.
- And r_R is not TI and but is 0-SI: $(\tau_{\mathbf{x}_0} r_R)\{f\}(\mathbf{x}) = f(R\mathbf{x} - \mathbf{x}_0) \neq f(R(\mathbf{x} - \mathbf{x}_0)) = (r_R \tau_{\mathbf{x}_0})\{f\}(\mathbf{x})$, and $(d_a r_R)\{f\}(\mathbf{x}) = f(a(R\mathbf{x})) = f(R(a\mathbf{x})) = (r_R d_a)\{f\}(\mathbf{x})$. □

This other lemma is not required for the next results in this section, but it is an interesting result on rotation and convolution, that will be used afterward in some proofs about steerability, and it concerns the rotation operators.

Lemma 2.43 (Rotation of a convolution). For $R \in \text{SO}(d)$, and f, g in E , we have $r_R\{g * f\} = (r_R\{g\}) * (r_R\{f\})$.

³² So there is an isomorphism between the group $\text{SO}(2)$ and $\mathbb{R}/2\pi\mathbb{Z} = \mathbb{T}$ the 2π -periodic torus.

³³ R and r_R have different domain so we preferred to use two different symbols. Note that some references make no difference, but it can be confusing: R operates on d -dimensional vectors, while r_R operates on functions!

³⁴ Conversely, this gives directly that any translation operator $\tau_{\mathbf{x}_0}$ is **not** rotation-invariant, and that any dilatation operator d_a is rotation-invariant.

³⁵ And r_R is an isometry, up to a constant – at least it preserves orthogonality. It is usually an isometry (i.e., $C_R = 1$) except maybe for unusual function space E .

Easy Proof. It is quite simple to obtain, as the (bijective) change of variable $\mathbf{y} \mapsto \mathbf{y}' = R^{-1}\mathbf{y}$ has a Jacobian³⁶ equals to 1, so it can be applied in the integral defining the convolution without changing its value.

Let $\mathbf{x} \in \mathbb{R}^d$, then

$$\begin{aligned} r_R\{(g * f)\}(\mathbf{x}) &= (g * f)(R\mathbf{x}) = \\ &= \int_{\mathbf{y} \in \mathbb{R}^d} g(R\mathbf{x} - \mathbf{y})f(\mathbf{y})d\mathbf{y} \\ (\text{With } \mathbf{y}' = R^{-1}\mathbf{y}) &= \int_{\mathbf{y}' \in \mathbb{R}^d} g(R(\mathbf{x} - \mathbf{y}'))f(R\mathbf{y}')d\mathbf{y}' \\ &= \int_{\mathbf{y} \in \mathbb{R}^d} (r_R\{g\})(\mathbf{x} - \mathbf{y})(r_R\{f\})(\mathbf{y})d\mathbf{y} \\ &= ((r_R\{g\}) * (r_R\{f\}))(\mathbf{x}) \end{aligned}$$

For functions, this computation is clear. For distribution, one needs to be more cautious, but the results is still true. \square

We can now ask our operators $G : E \rightarrow F$ to be *rotation-invariant*, in the sense of the following definition:

Definition 2.44 (Rotation invariance for operators). For $G \in \mathcal{L}(E, F)$, we say that G is rotation-invariant (**RI**) when it commutes with any rotation operator r_R , i.e., when

$$\begin{aligned} \forall R \in \text{SO}(d), \forall f \in E, \forall \mathbf{x} \in \mathbb{R}^d, \quad &(G\{f\})(R\mathbf{x}) = (G\{f(R\cdot)\})(\mathbf{x}) \\ \text{ie.} &= ((r_R G)\{f\})(\mathbf{x}) = ((G r_R)\{f\})(\mathbf{x}). \end{aligned}$$

A simpler way of writing it is $r_R \circ G = G \circ r_R$, for all $R \in \text{SO}(d)$.

Remark 2.45. As before, we only consider proper rotations for the definition of rotation-invariance (and for steerability, see later definition 4.1). All these properties are studied because they have a “physical” interpretation: rotation-invariance for a function means isotropy of its values (e.g., for a 2D image, it means that we observe circles of different lightness, centered at $(0, 0)$). And for an operator (a filter, as called in signal processing), rotation-invariance means that if the filter is applied to a rotated version of the input image, it will give the same output as if we rotate its output on the image.

For instance, steerability have been studied for wavelets, especially in the BIG team during the last few years (for example cf. Szuszanna Püspöki’s PhD thesis [Püs16]), and applied to contour and shape detection with great success, where rotation-invariance is an appealing property (a contour should be detected correctly no matter its orientation).

Remark 2.46. The definition has a very similar meaning for functions and operators: a function f is rotation-invariant when applying a rotation to its input *does not change*

³⁶ This change of variable will be used for other proofs as well. For a quick reminder on the change of variable theorem and the notion of Jacobian, cf. this Wikipedia article.

its value; while an operator G is rotation-invariant when it *commutes* with any rotation operator. But the correct way to “apply” a rotation to an operator is to anti-rotate the input and rotate the output, *i.e.*, to consider $G_R = r_R \circ G \circ r_{R^{-1}}$ (cf. remark 4.6). And so G is rotation-invariant when *iff* $G_R = G$ for all $R \in \text{SO}(d)$, which exactly means that applying a rotation on G does not change it.

We will denote **RI** the class of rotation-invariant operators (not studied by itself), and by **LC RI** the class of linear, continuous and rotation-invariant operators (on E).

Proposition 2.47. *This class **RI** is a linear sub-space of $\mathcal{L}(E, F)$.*

Easy Proof. As before for **TI**, it is immediate when **RI** is written as the intersection of all the $\text{Comm}(r_R)$ for $R \in \text{SO}(d)$. \square

Corollary 2.48. *LC RI operators also form a linear sub-space of $\mathcal{L}(E, F)$.*

Easy Proof. By intersection. \square

Remark 2.49. As it will be proved afterward in section 3.4.4, rotation-invariance is a very strong property, and usually it reduces greatly the freedom on the operator. In our study of rotation-invariant operators, the Laplacian will in fact be the only rotation-invariant and scale-invariant **LC TI** operator (along with all the fractional Laplacians³⁷).

Conclusion

This first section introduced the notion of test functions and operators, and presented some algebraic and geometric properties that operators can have. We chose to focus on the three main geometric invariance properties, translation-, scaling- and rotation-invariance. Many results were stated about the structure of the some interesting class of operators, and what is important is the fact that all the main class of operators are topological vector spaces (cf. remark 2.18).

Another geometric property?

Remark 2.50 (Unitary operators). We chose to include the definition and study of unitary operators in appendix A, along with the study of the Hilbert transform and its generalizations, because they are less relevant³⁸ to the rest of our work, but interesting enough to be included.

³⁷ Linguistic remark: I was not sure what is the correct English rule to follow here, between laplacian, laplacians, Laplacian or Laplacians. It seemed clearer to put a “s” to indicate plurals, and a capital letter as usual for the Laplacian. Forgive me if you disagree with my choice: “one Laplacian, two Laplacians”.

³⁸ An interested reader could consult this appendix, but should keep in mind that it is less coherent with the rest of our research.

3 Classic results for operators, important examples, and characterizations for geometric invariance

This section will give general results about operators, mainly characterizations for each of the algebraic and geometric properties introduced in the previous section 2.3.

Along these general considerations, some examples of operators are presented, first in a small overview of the classic operators, and then by studying in more details two important examples, both coming from the study of differential equations: derivatives (and directional derivatives) and Laplacian (and fractional Laplacian).

The two main theorems are very well known, theorems 3.1 and 3.5 are the so called Schwartz's kernel theorems, and we use them to give interesting characterization of being **RI**, **0-SI** and/or γ -**SI** for **LC** and **LC TI** operators. The proofs of every results, except the first Schwartz theorem, are included here.

Extensive study ? We tried to conduct an exhaustive study of operators with certain properties, where every combination of properties has been studied. For the sake of conciseness and because a lot of these combinations do not yield easy or interesting results, we only included here the more interesting ones. It is not surprising to observe that not much can be said about operators having some geometric properties if they are not convolution. Therefore, operators are assumed to be linear, continuous and translation-invariant (*i.e.*, convolution) from now on.

3.1 Quick overview of some classic examples of operators

We start by giving a small list of linear operators, that will be used a lot in this section and the next ones. It can be seen as a short "catalog", and nothing here is new.

The following simple operators have already been defined and studied above, because they are required to define the translation-, scaling- and rotation-invariance:

- The identity Id , and the constant operators $C\text{Id}$ ($C \in \mathbb{R}$),
- The translations $\tau_{\mathbf{x}_0}$ were introduced in definition 2.19, and studied in lemma 2.20,
- The dilatations d_a were introduced in definition 2.28, and studied in lemma 2.29,
- The rotations r_R were introduced in definition 2.41, and studied in lemma 2.42.

Two other interesting families of operators are *differential* operators, which can be either directional or isotropic, and they are studied later in details:

- The partial derivatives D_i and directional derivatives $D_{\vec{u}}$, are introduced in definitions 3.18 and 3.21, and studied in proposition 3.23,
- The Laplacian Δ , and fractional Laplacians $(-\Delta)^{\gamma/2}$, are introduced in definitions 3.25 and 3.27, and studied in proposition 3.30.

Finally, other operators will also be presented afterward in this report, but they are not studied by themselves and not used for our main questions:

- The Hilbert transform H , and its generalizations, the directional Hilbert transform $H_{\vec{u}}$ or directional Hilbert transform H_{τ} , and our generalization the fractional-directional Hilbert transform (**fdHT**) $H_{\tau, \vec{u}}$, are introduced in definition A.9, and studied in proposition A.11. We chose to include the study of the Hilbert transform and its generalizations in appendix A.3, along with the definition and study of unitary operators.
- The Riesz transform \mathcal{R} , and the iterated Riesz transforms \mathcal{R}^n , are not studied in this report (even if we spent some time working on them). They are simply quoted and their form is reminded in example 3.11, in order to recognize a sum of iterated real Riesz transform in an important decomposition studied later in section 4.3.2.

This small subsection gave a list of some well-known operators. They are all linear continuous (**LC**), and most of them are translation-invariant (**TI**). The first examples ($\tau_{\mathbf{x}_0}$, d_a , r_R) are used to define the geometric invariance properties (**TI**, **0-SI** or γ -**SI**, **RI**), and the differential operators will be used in the next section 4.6 as elementary blocks in some decompositions of operators as a composition of simpler terms.

Before going into these innovative results, we need to recall the well-known Schwartz's kernel theorem (3.1), and study its consequences, first for convolution operators (**LC TI**) in section 3.2.2, and then for operators with other properties in section 3.4. Derivatives and Laplacian are also studied in the light of the second Schwartz kernel, in section 3.3.

3.2 Schwartz's kernel theorems

For the following characterizations, we consider the general setting with $d \in \mathbb{N}^*$ (as before), and $E = \mathbf{S}(\mathbb{R}^d)$ (very restrictive on the test functions f) and $F = \mathbf{S}'(\mathbb{R}^d)$ (very **not** restrictive on the transformed functions $G\{f\}$, as they can be distribution). Most of the results can also be applied for operators in $\mathbf{L}^2(\mathbb{R}^d)$, without losing in generality.

Note: All the operators considered from now on are **at least linear and continuous (LC)**.

3.2.1 Characterization for LC – First Schwartz's kernel theorem

The first characterization we can give is a well known result, characterizing the continuity of linear operators G by the existence of their Schwartz kernel k_G .

Such operator is called a *kernel operator*, or an *integral operator*, and they have been extensively studied. We will not focus on this class **LC** per se, but on some sub-classes.

Theorem 3.1 (Schwartz's kernel theorem). $G \in \mathcal{L}(\mathbf{S}(\mathbb{R}^d), \mathbf{S}'(\mathbb{R}^d))$ is continuous iff there exists a **distribution**³⁹ k_G on $\mathbf{S}'(\mathbb{R}^d \times \mathbb{R}^d)$, called the Schwartz kernel of G , such that

$$\forall f \in \mathbf{S}(\mathbb{R}^d), \forall \mathbf{x} \in \mathbb{R}^d, G\{f\}(\mathbf{x}) = \int_{\mathbb{R}^d} k_G(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y}. \quad (5)$$

Some interesting explanations can be found in [Uns16, Th.11.18], but the proof is another story.

³⁹ Some operator G has a kernel k_G which is a (true) function, but we **have to** consider generalized functions, i.e., distributions in \mathbf{S}' : the simplest possible operator, $G = \text{Id}$, has the 0-Dirac δ_0 for its kernel.

Hard Proof. Any proof of this strong result is technical and long (about 2 pages), making intensive usage of distributions, duality brackets and partial derivatives; and it goes beyond the reach of this report, so we advise the interested reader to consult a specialized book for the proof. It also requires a strong knowledge of distribution theory. See for instance in [Hör83, Th.5.2.1, Part.5.2], where the author employs a slightly more theoretical formulation. \square

In the sense of the distributions and the duality bracket notation⁴⁰ (on \mathbf{S} and \mathbf{S}'), and a tensor product ($f \otimes \phi$), then k_G is determined uniquely by this relationship

$$\forall f \in \mathbf{S}(\mathbb{R}^d), \forall \phi \in \mathbf{S}'(\mathbb{R}^d), \langle G\{f\}, \phi \rangle = \langle k_G, f \otimes \phi \rangle. \quad (6)$$

Examples 3.2. We can start by giving a few examples. We do not compute the kernels, we simply give it and let the reader check that it works.

- For the identity $G = \text{Id}$, its kernel is $k_G(\mathbf{x}, \mathbf{y}) = \delta_{\mathbf{y}}(\mathbf{x})$,
- For a translation $G = \tau_{\mathbf{x}_0}$, its kernel is $k_G(\mathbf{x}, \mathbf{y}) = \delta_{\mathbf{y}}(\mathbf{x} - \mathbf{x}_0)$,
- For a dilatation, $G = d_a$, its kernel is $k_G(\mathbf{x}, \mathbf{y}) = \delta_{\mathbf{y}}(a\mathbf{x})$.
- For a rotation, $G = r_R$, its kernel is $k_G(\mathbf{x}, \mathbf{y}) = \delta_{R^{-1}\mathbf{y}}(\mathbf{x}) = \delta_{\mathbf{y}}(R\mathbf{x})$.

Computing the Schwartz kernel k_G for more complicated operators can be hard, for instance the case of a Laplacian is not as simple as these examples,

Proposition 3.3 (Explicit form for $k_G(\mathbf{x}, \mathbf{y})$). *For G LC, $k_G(\mathbf{x}, \mathbf{y})$ has an explicit form given by the value in \mathbf{x} of the image by G of the Dirac delta distribution (in \mathbf{y}):*

$$k_G(\mathbf{x}, \mathbf{y}) = G\{\delta_{\mathbf{y}}\}(\mathbf{x}). \quad (7)$$

Proof. We simply take the formula, $G\{f\}(\mathbf{x}) = \int_{\mathbb{R}^d} k_G(\mathbf{x}, \mathbf{y}') f(\mathbf{y}') d\mathbf{y}'$, and apply it to $f = \delta_{\mathbf{y}}$ (for a fixed $\mathbf{y} \in \mathbb{R}^d$): $G\{\delta_{\mathbf{y}}\}(\mathbf{x}) = \int_{\mathbb{R}^d} k_G(\mathbf{x}, \mathbf{y}') \delta_{\mathbf{y}}(\mathbf{y}') d\mathbf{y}' = k_G(\mathbf{x}, \mathbf{y})$. \square

Note that in general, this explicit form is pretty but can be hard to compute concretely.

Examples 3.4. However, for simple operators, we can check that the examples of kernels given above all satisfy this formula, and compute directly their kernels k_G :

- For the identity $G = \text{Id}$, $k_G(\mathbf{x}, \mathbf{y}) = \delta_{\mathbf{y}}(\mathbf{x}) = (\text{Id}\{\delta_{\mathbf{y}}\})(\mathbf{x})$ trivially,
- For a translation $G = \tau_{\mathbf{x}_0}$, $k_G(\mathbf{x}, \mathbf{y}) = \delta_{\mathbf{y}}(\mathbf{x} - \mathbf{x}_0) = (\tau_{\mathbf{x}_0}\{\delta_{\mathbf{y}}\})(\mathbf{x})$,
- For a dilatation $G = d_a$, $k_G(\mathbf{x}, \mathbf{y}) = \delta_{\mathbf{y}}(a\mathbf{x}) = (d_a\{\delta_{\mathbf{y}}\})(\mathbf{x})$.
- For a rotation, $G = r_R$, $k_G(\mathbf{x}, \mathbf{y}) = \delta_{\mathbf{y}}(R\mathbf{x}) = (r_R\{\delta_{\mathbf{y}}\})(\mathbf{x})$.

⁴⁰ Or in the sense of the inner product on \mathbf{L}^2 . We chose to not dive too much into the details of the theory of distributions, and it is the only mention of the notion of duality bracket on a tensor product.

3.2.2 Impulse response g of a LC TI operator G – Second Schwartz's kernel theorem

This second theorem is also a well known result, characterizing the translation-invariance of integral operators G by a property on their Schwartz kernel k_G .

The second Schwartz's kernel theorem is very easy to obtain from the first one, and the explicit form of the kernel $k_G(\mathbf{x}, \mathbf{y})$ given by proposition 3.3. It is often called the Schwartz's convolution theorem, because this specific form for the kernel k_G , depending only on $\mathbf{x} - \mathbf{y}$, allows to express $G\{f\}$ as a *convolution* between g and f .

Theorem 3.5 (Schwartz's convolution theorem). *A continuous $G \in \mathcal{L}(\mathbf{S}(\mathbb{R}^d), \mathbf{S}'(\mathbb{R}^d))$ is translation-invariant iff there exists a **distribution** g , on $\mathbf{S}'(\mathbb{R}^d, \mathbb{R})$, called the impulse response of G , such that $k_G(\mathbf{x}, \mathbf{y}) = g(\mathbf{x} - \mathbf{y})$, i.e.,*

$$\forall f \in \mathbf{S}(\mathbb{R}^d), \forall \mathbf{x} \in \mathbb{R}^d, \quad G\{f\}(\mathbf{x}) = \int_{\mathbb{R}^d} g(\mathbf{x} - \mathbf{y})f(\mathbf{y})d\mathbf{y} = (g * f)(\mathbf{x}). \quad (8)$$

Proof. This second theorem is easy to obtain, thanks to the explicit form for $k_G(\mathbf{x}, \mathbf{y})$ proved in proposition 3.3. Indeed, as soon as G is translation-invariant, we have

$$\begin{aligned} k_G(\mathbf{x}, \mathbf{y}) &= (G\{\delta_{\mathbf{y}}\})(\mathbf{x}) = (G\{\tau_{\mathbf{y}}\delta_{\mathbf{0}}\})(\mathbf{x}) \\ &= ((G \circ \tau_{\mathbf{y}})\{\delta_{\mathbf{0}}\})(\mathbf{x}) \\ (G \text{ is translation-invariant}) &= ((\tau_{\mathbf{y}} \circ G)\{\delta_{\mathbf{0}}\})(\mathbf{x}) \\ &= (\tau_{\mathbf{y}}\{G\{\delta_{\mathbf{0}}\}\})(\mathbf{x}) = (\tau_{\mathbf{y}}\{k_G(\cdot, \mathbf{0})\})(\mathbf{x}) \\ &= (k_G(\cdot, \mathbf{0}))(\mathbf{x} - \mathbf{y}) = k_G(\mathbf{x} - \mathbf{y}, \mathbf{0}). \end{aligned}$$

And so we can define $g \stackrel{\text{def}}{=} k_G(\cdot, \mathbf{0})$, and rewrite $\int_{\mathbb{R}^d} k_G(\mathbf{x}, \mathbf{y})f(\mathbf{y})d\mathbf{y}$ as $\int_{\mathbb{R}^d} k_G(\mathbf{x} - \mathbf{y}, \mathbf{0})f(\mathbf{y})d\mathbf{y} = \int_{\mathbb{R}^d} g(\mathbf{x} - \mathbf{y})f(\mathbf{y})d\mathbf{y} = (g * f)(\mathbf{x})$, as claimed.

The converse sense is easy, and is left as an exercise for the reader. \square

Examples 3.6. For the same three basic examples as above:

- The identity, and any translation $\tau_{\mathbf{x}_0}$, are convolutions, as $k_{\tau_{\mathbf{x}_0}}(\mathbf{x} - \mathbf{y}) = \delta_0(\mathbf{x} - \mathbf{y} - \mathbf{x}_0) = \delta_{\mathbf{x}_0}(\mathbf{x} - \mathbf{y}) = g(\mathbf{x} - \mathbf{y})$ with $g = \delta_{\mathbf{x}_0}$.
- A non-trivial dilatation d_a (i.e., for $a \neq 1$) is **not** a convolution. Indeed, $k_{d_a}(\mathbf{x}, \mathbf{y}) = \delta_0(a\mathbf{x} - \mathbf{y})$ cannot be written as function of $\mathbf{x} - \mathbf{y}$ for the simple reason that if $a \neq 1$, $\mathbf{y} = \mathbf{x} \implies g(\mathbf{x} - \mathbf{y}) = g(0) \neq \delta_0((a - 1)\mathbf{x})$ in general (the right hand-side being non-constant, for $a \neq 1$).
- Similarly, a non-trivial rotation r_R is **not** a convolution ($k_{r_R}(\mathbf{x}, \mathbf{y}) = \delta_0(R\mathbf{x} - \mathbf{y})$ is not a function of $\mathbf{x} - \mathbf{y}$, for $R \neq \mathbb{I}_d$).

Proposition 3.7 (Explicit form for $g(\mathbf{x})$). *For G LC TI, $g(\mathbf{x})$ has an explicit form:*

$$\forall \mathbf{x} \in \mathbb{R}^d, \quad g(\mathbf{x}) = G\{\delta_{\mathbf{0}}\}(\mathbf{x}). \quad (9)$$

Easy Proof. It is an obvious consequence of the fact that $k_G(\mathbf{x}, \mathbf{y}) = g(\mathbf{x} - \mathbf{y})$, so $g(\mathbf{x}) = k_G(\mathbf{x}, \mathbf{0})$, and proposition 3.3 giving the explicit form for $k_G(\mathbf{x}, \mathbf{y})$ for LC operators. \square

3.2.3 Fourier multiplier \hat{g} of a LC TI operator G

This form of $(g * f)(\mathbf{x})$ will be extremely convenient in the Fourier domain, as convolutions become products. In full generality, $F \subset \mathbf{S}'$ (cf. remark 2.17), and so thanks to properties 1.2, $G\{f\}$ is at least in \mathbf{S}' . So we can always consider the Fourier transform $\mathcal{F}\{G\{f\}\}$ of the function or distribution $G\{f\}$, whatever the spaces E, F (as long as they satisfy (4)).

$$\forall \omega \in \mathbb{R}^d, \mathcal{F}\{G\{f\}\}(\omega) = \mathcal{F}\{(g * f)\}(\omega)$$

So thanks to the Fourier convolution theorem 1.4.

$$\begin{aligned} \mathcal{F}\{G\{f\}\}(\omega) &= (\mathcal{F}\{g\} \cdot \mathcal{F}\{f\})(\omega) \\ \text{So } \mathcal{F}\{G\{f\}\}(\omega) &= \hat{g}(\omega) \cdot \hat{f}(\omega) \end{aligned} \quad (10)$$

Definition 3.8. This \hat{g} is called the Fourier multiplier of the operator G .

In full generality, it is a distribution (i.e., on $\mathbf{S}'(\mathbb{R}^d, \mathbb{R})$). In practice however, \hat{g} is often a complex-valued function, defined on \mathbb{R}^d : $\hat{g} : \mathbb{R}^d \rightarrow \mathbb{C}$.

When $G\{f\}$ stays⁴¹ a function, an importance consequence of (10) is the simplicity of applying an operator given by its Fourier multiplier: a LC TI operator is applied on an input function f by a simple multiplication, up-to a (possibly complicated) computation of \mathcal{F} and \mathcal{F}^{-1} :

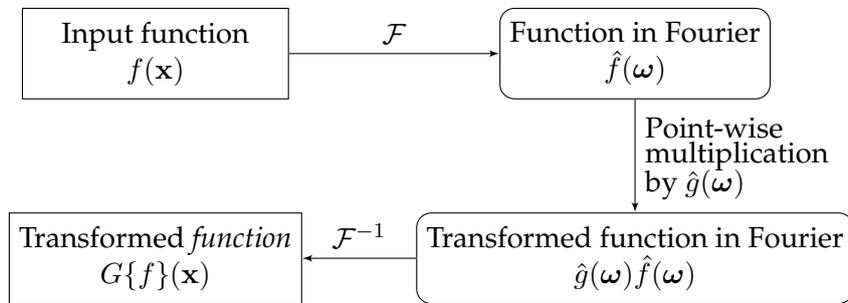


Figure 1: Applying G , given by its Fourier multiplier \hat{g} , on an input function f .

Remark 3.9. In the discrete setting, applying \mathcal{F} and \mathcal{F}^{-1} will never be an issue, because we work with Fast Fourier Transforms (`fft`, `fft2`) and Inverse FFT (`ifft`, `ifft2`) on signals and images, as we explain later in section 5.2, and figure 3.

Examples 3.10. We can give a few examples of very simple Fourier multipliers:

- For $G = \text{Id}$, $\hat{g}(\omega) = 1$ is the simplest Fourier multiplier.
- For a translation $G = \tau_{\mathbf{x}_0}$, $\hat{g}(\omega) = e^{-j\mathbf{x}_0^T \omega}$, thanks to the translation / time-shifting property of the Fourier transform [Rud91, Th.7.2.(a)].

⁴¹ It is always the case in practice.

Example 3.11 (Riesz transform in dimension $d = 2$). The (complex) Riesz transform is the **LC TI** operators on $\mathbf{L}(\mathbb{R}^2, \mathbb{C})$ defined by its (complex, not Hermitian symmetric) Fourier multiplier (in polar coordinate):

$$\mathcal{R}\{f\}(\mathbf{x}) \xleftrightarrow{\mathcal{F}} \widehat{\mathcal{R}}(\boldsymbol{\omega}) = \widehat{\mathcal{R}}(r, \theta) = e^{j\theta}.$$

The *real Riesz transform* is given by $\widehat{\mathcal{R}}_{\text{real}}(r, \theta) = \cos(\theta)$.

The complex and real Riesz transform and their iterates are **LC TI** and scale-invariant. For more results and details on the Riesz transform, cf. [WCU13] and [UC13].

After these few examples, a general result can be stated about the Fourier multipliers of **LC TI** operators.

Proposition 3.12 (Measurability of Fourier multipliers, and order of decay).

Usually in practice, \widehat{g} is a true function, at least measurable on \mathbb{R}^d .

In particular, there is no general result stating that \widehat{g} is continuous, or bounded, or even in any $\mathbf{L}^p(\mathbb{R}^d, \mathbb{R})$ space.

However, we *assume*⁴² that \widehat{g} always has an order of decay $\kappa \in \mathbb{R}$ such that $\widehat{g}(\boldsymbol{\omega})(1 + \|\boldsymbol{\omega}\|^2)^\kappa$ is in $\mathbf{L}^2(\mathbb{R}^d, \mathbb{R})$.

Proof. Its measurability comes directly from the fact that it is the Fourier transform of g , and classic results of Fourier analysis, [SW71].

For a more specific result, cf. [UT14, Th.3.5, p41] where they characterize the fact that G operates on $\mathbf{L}^p(\mathbb{R}^d, \mathbb{R})$ with their \widehat{g} .

The existence of the *order of decay* $\kappa \in \mathbb{R}$, does not come from Fourier analysis, it is an additional assumption. Note that it does *not* imply that \widehat{g} is in the Sobolev space $\mathbf{H}^{\kappa,2}(\mathbb{R}^d, \mathbb{R})$ (cf. (1)), as \widehat{g} may be not in $\mathbf{L}^2(\mathbb{R}^d, \mathbb{R})$. \square

Another very useful property of **LC TI** operators operating on and to functions (*i.e.*, $G\{f\}$ stays a function) is the surprising fact that they *all* commute with each other, as long as they operator on an to the same domain, *i.e.*, $E = F$.

Theorem 3.13 (Commutativity of **LC TI** operators). If $E = F$, and G_1, G_2 two **LC TI** operators on E . Then $G_1 \circ G_2 = G_2 \circ G_1$.

Easy Proof. With the use of their Fourier multipliers \widehat{g}_1 and \widehat{g}_2 , it is almost instantaneous. Let $G = G_1 \circ G_2$ and $G' = G_2 \circ G_1$, and let us prove directly their equality.

$$\begin{aligned} \forall \boldsymbol{\omega} \in \mathbb{R}^d, \mathcal{F}\{G\{f\}\}(\boldsymbol{\omega}) &= \mathcal{F}\{G_1\{G_2\{f\}\}\}(\boldsymbol{\omega}) \\ &= (\mathcal{F}\{g_1\} \cdot \mathcal{F}\{G_2\{f\}\})(\boldsymbol{\omega}) \\ &= (\mathcal{F}\{g_1\} \cdot (\mathcal{F}\{g_2\} \cdot \mathcal{F}\{f\}))(\boldsymbol{\omega}) \\ &= \widehat{g}_1(\boldsymbol{\omega}) \cdot \widehat{g}_2(\boldsymbol{\omega}) \cdot \widehat{f}(\boldsymbol{\omega}) \end{aligned}$$

But this product is for point-wise values of the *functions* (and not distributions), in the Fourier domain, *i.e.*, product of simple complex values, and so it commutes,

$$\begin{aligned} &= \widehat{g}_2(\boldsymbol{\omega}) \cdot \widehat{g}_1(\boldsymbol{\omega}) \cdot \widehat{f}(\boldsymbol{\omega}) \\ &= \mathcal{F}\{G_2\{G_1\{f\}\}\}(\boldsymbol{\omega}) = \mathcal{F}\{G'\{f\}\}(\boldsymbol{\omega}). \end{aligned} \quad \square$$

⁴² It is false to say in general that \widehat{g} has an order of decay κ just because it is measurable.

3.2.4 Real operators and Hermitian symmetry

It was an implicit hypothesis from the beginning, but let us define now the real operators.

Definition 3.14 (Real operators). *An operator G is real when it maps real-valued functions f to real-valued functions $G\{f\}$.*

The notion of Hermitian symmetry for **complex**-valued functions is useful to study real operators.

Definition 3.15. $h : \mathbb{R}^d \rightarrow \mathbb{C}$ is Hermitian symmetric when $\forall \omega \in \mathbb{R}^d$, $h(-\omega) = \overline{h(\omega)}$.

Lemma 3.16. *Hermitian symmetry is preserved by **real** linear combinations, and by point-wise function product.*

Easy Proof. It is very easy to prove, with basic algebraic manipulations. \square

This notion is used for theorem 3.17, a very useful result characterizing the operators G that conserves the fact that $f(\mathbf{x}) \in \mathbb{R}$: a real-valued signal stays real. It will be used a lot in our study of steerable operator.

As for the geometric invariance properties, G is real *iff* g is, but not \hat{g} .

Theorem 3.17 (Characterization for real **LC TI** operators). *A **LC TI** operator G is real *iff* g is real (if g is a function) *iff* \hat{g} is Hermitian symmetric (if \hat{g} is a function).*

Proof. For all f , $G\{f\}$ is real-valued *iff* for all f , $\mathcal{F}(G\{f\})$ is Hermitian symmetric, thanks to a well known property of the Fourier transform *iff* for all f , $\omega \mapsto \hat{g}(\omega)\hat{f}(\omega)$ is Hermitian symmetric, but f is real-valued so \hat{f} is already Hermitian symmetric, hence *iff* $\omega \mapsto \hat{g}(\omega)$ is Hermitian symmetric *iff* g is real. \square

In fact, this hypothesis is required (maybe implicitly) for almost every operator studied in this report. One could check that all the previously given examples of operators have indeed an Hermitian symmetric Fourier multiplier \hat{g} . For instance, the **fdHT** operators defined later in appendix A.3.2 are real operators, because their Fourier multiplier are Hermitian symmetric.

This part introduced the second Schwartz theorem, and we now have two new points of view to work with G : either its impulse response g or its Fourier multiplier \hat{g} . Before giving other general results, two families of important operators are presented in the next subsection.

3.3 Important examples of convolutions: derivatives and Laplacians

We present here two important families of convolutions, that can be defined easily as operators, or by giving their Fourier multipliers.

In the next section, some of our most important theorems on 2D steerable operators will highlight the importance of the two “elementary blocks” that are directional derivatives and Laplacians (cf. proposition 4.46).

3.3.1 Partial and directional derivatives

A very interesting family of operators on differentiable functions are the derivatives. We first recall the definition of the d partial derivatives D_1, \dots, D_d , and then we define a directional derivative.

Definition 3.18 (Partial derivatives). For $i \in \{1, \dots, d\}$, D_i the i -th partial derivative operator, defined on differentiable functions (e.g., on $\mathbf{S}(\mathbb{R}^d, \mathbb{R})$), is

$$D_i f : \mathbf{x} \mapsto \frac{\partial f}{\partial x_i}(\mathbf{x}) \stackrel{\text{def}}{=} \lim_{h \rightarrow 0, h \neq 0} \frac{f(\mathbf{x} + h\mathbf{e}_i) - f(\mathbf{x})}{h}. \quad (11)$$

where \mathbf{e}_i denotes the i -th vector of the canonical orthonormal basis for \mathbb{R}^d .

Examples 3.19. • For $d = 1$, there is only one derivative operator: $D : f \rightarrow f'$.

• For $d = 2$, the two derivatives are: $D_1 = D_x : f \rightarrow \frac{\partial f}{\partial x}$ and $D_2 = D_y : f \rightarrow \frac{\partial f}{\partial y}$.

Proposition 3.20. A partial derivative is a linear, continuous, translation-invariant and 1-scale-invariant operator (**LC TI 1-SI**), on the Schwartz class of infinitely differentiable functions of rapid decay⁴³ $\mathbf{S}(\mathbb{R}^d, \mathbb{R})$ and on distribution $\mathbf{S}'(\mathbb{R}^d, \mathbb{R})$.

Easy Proof. • Linearity is obvious by definition and by the linearity of a limit.

- The continuity of D_i is true on rapid decaying Schwartz functions, as $f_n \rightarrow f$ means by definition that all the semi-norms $\|f_n - f\|_{p,\alpha} \rightarrow 0$, and some of the semi-norms concern the i -th partial derivative of f_n and f , so $f_n \rightarrow f \implies D_i f_n \rightarrow D_i f$, in the sense of the Schwartz class.
- By the definition of D_i as a limit, we can compute the effect of a translation $\tau_{\mathbf{x}_0}$:

$$\begin{aligned} ((\tau_{\mathbf{x}_0} \circ D_i)\{f\})(\mathbf{x}) &= \lim_{h \rightarrow 0, h \neq 0} \frac{f(\mathbf{x} - \mathbf{x}_0 + h\mathbf{e}_i) - f(\mathbf{x} - \mathbf{x}_0)}{h} \\ &= \lim_{h \rightarrow 0, h \neq 0} \frac{(\tau_{\mathbf{x}_0}\{f\})(\mathbf{x} + h\mathbf{e}_i) - (\tau_{\mathbf{x}_0}\{f\})(\mathbf{x})}{h} \\ &= ((D_i \circ \tau_{\mathbf{x}_0})\{f\})(\mathbf{x}), \end{aligned}$$

so any D_i is translation-invariant.

- The 1-scaling-invariance comes directly from the definition, with the (bijective) change of variable $h \mapsto h/a$ (of Jacobian $1/a^d$ in \mathbb{R}^d).
- And finally, D_i operates on $\mathbf{S}(\mathbb{R}^d, \mathbb{R})$ because a partial derivative of a Schwartz function is still a Schwartz function, by definition (cf. [UT14, p30] if needed).
- For a distribution ϕ on $\mathbf{S}(\mathbb{R}^d, \mathbb{R})$ (i.e., $\phi \in \mathbf{S}'(\mathbb{R}^d, \mathbb{R})$), D_i is defined as usual by a duality bracket: $\langle D_i \phi, f \rangle = \langle \phi, D_i f \rangle$, and it is a classical result from [Rud91] that D_i also operates on $\mathbf{S}'(\mathbb{R}^d, \mathbb{R})$. \square

A first generalization of the partial derivative can easily be obtained with directional derivatives, and a second generalization will be composed directional derivatives.

⁴³ Considering only the once-differentiable functions is not enough, as all the family of semi-norms are required to prove continuity.

Definition 3.21 (Directional derivatives). For a unitary vector $\vec{u} \in \mathbb{S}^{n-1}$, the directional derivative of direction \vec{u} is defined as $D_{\vec{u}} = u_1 D_1 + \dots + u_d D_d = \langle \vec{u}, \nabla \cdot \rangle$.

Examples 3.22. We start by giving two simple examples in dimensions 1 and 2 :

- In dimension $d = 1$, the only directional derivatives are $D : f \rightarrow f'$ and $-D$.
- In dimension $d = 2$, a directional derivative is parametrized by a unitary vector \vec{u} on the sphere \mathbb{S}^1 , so it can be parametrized by a unique angle $\alpha \in \mathbb{R}$, $\vec{u} = \vec{u}_\alpha \stackrel{\text{def}}{=} [\cos(\alpha) \ \sin(\alpha)]^T$. We write D_α instead of $D_{\vec{u}_\alpha}$, it is simply $D_\alpha : f \mapsto \cos(\alpha) \frac{\partial f}{\partial x} + \sin(\alpha) \frac{\partial f}{\partial y}$. We will prove later that it has a Fourier multiplier given by $\widehat{D_\alpha}(\omega) = \widehat{D_\alpha}(r, \theta) = \cos(\alpha)(jr \cos(\theta)) + \sin(\alpha)(jr \sin(\theta)) = jr \cos(\theta - \alpha)$.

This formula has the advantage of highlighting the “directionality” of D_α : it differentiates along the line of angle α , and this can also be seen as first changing θ to $\theta - \alpha$ (i.e., applies a rotation of angle $-\alpha$), and then differentiating along the x -axis with D_1 (Fourier multiplier $jr \cos(\theta)$).

This example is very important, as directional derivatives are used a lot in our study of steerable operators, cf. section 4.6 later.

Proposition 3.23. A directional derivative is also **LC TI 1-SI** on $\mathbf{S}(\mathbb{R}^d, \mathbb{R})$ and $\mathbf{S}'(\mathbb{R}^d, \mathbb{R})$.

Easy Proof. It is obvious, as the class **LC TI 1-SI** is a linear sub-space of $\mathcal{L}(E, F)$, and a directional derivative is nothing but a linear combination of partial derivatives, which are **LC TI 1-SI** from proposition 3.20 above. \square

Directional derivatives can be composed to obtain *higher order directional derivatives*, e.g., $G = D_{\vec{u}_1} \circ D_{\vec{u}_2} \circ \dots \circ D_{\vec{u}_k}$ of **order** $k \in \mathbb{N}$.

Lemma 3.24. Any composition of $k \in \mathbb{N}$ directional derivatives $D_{\vec{u}_1} \circ D_{\vec{u}_2} \circ \dots \circ D_{\vec{u}_k}$ is **LC TI** and **k-SI**, and their order do not matter⁴⁴.

Easy Proof. Obvious with the theorems of stability by composition, and composition of order of homogeneity, see later in corollary 2.25 and in proposition 2.39. $k = 0$ gives the identity, which is indeed **LC TI 0-SI**. \square

We presented here the derivatives and directional derivatives in the general setting ($d \in \mathbb{N}^*$), but we will study their steerability later in section 4.2.2.

3.3.2 Laplacian and fractional Laplacian

On twice-differentiable functions, the classic Laplacian operator is well known and has been studied and used for a long time. It has a generalization called the fractional Laplacian, which can be hard to express on functions, but it has a simple Fourier multiplier.

⁴⁴ This remark on the commutativity of the derivatives will be obvious because they are convolutions and convolutions all commute with each other (theorem 3.13), but it is interesting, as for instance it gives the Schwarz theorem: $\frac{\partial^2 f(x,y)}{\partial x \partial y} = \frac{\partial^2 f(x,y)}{\partial y \partial x}$. However, in the framework of distribution, it is trivial.

Definition 3.25 (Laplacian). The Laplacian operator (Δ) is defined on $E = \mathbf{S}(\mathbb{R}^d, \mathbb{R})$ by

$$(\Delta\{f\})(\mathbf{x}) = \Delta f(\mathbf{x}) \stackrel{\text{def}}{=} \sum_{i=1}^d \frac{\partial^2 f}{\partial x_i^2}(\mathbf{x}). \quad (12)$$

And it is defined on $\mathbf{S}'(\mathbb{R}^d, \mathbb{R})$ by duality bracket, $\langle \Delta\phi, f \rangle = \langle \phi, \Delta f \rangle$.

Theorem 3.26. The Laplacian is a linear, continuous, translation-invariant, rotation-invariant, and 2-scale-invariant operator (**LC TI RI 2-SI**), on $E = \mathbf{S}(\mathbb{R}^d, \mathbb{R})$ and $\mathbf{S}'(\mathbb{R}^d, \mathbb{R})$.

Easy Proof. As a special case of fractional Laplacian, it is a consequence of the more general theorem given below. But proving quickly a few things manually is interesting:

- Its linearity and translation-invariant both come from the linearity and translation-invariant of any iterated partial derivative.
- And similarly for its continuity, on $E = \mathbf{S}(\mathbb{R}^d, \mathbb{R})$ or $E = \mathbf{S}'(\mathbb{R}^d, \mathbb{R})$.
- The Laplacian is also rotation-invariant, but this is harder to prove manually without the use of its Fourier multiplier ($\widehat{\Delta}(\omega) = -\|\omega\|^2$, see below in definition 3.27).
- Any partial derivative $\frac{\partial^2}{\partial x_i^2}$ is 2-scale-invariant, so the Laplacian is also 2-0-SI. \square

We could define a generalized Laplacian on functions, but it is much easier to just give its Fourier multiplier.

Definition 3.27 (Fractional Laplacian). For $\gamma \in \mathbb{R}_+$, let the fractional Laplacian operator, denoted $(-\Delta)^{\gamma/2}$, be defined on $E = \mathbf{S}(\mathbb{R}^d, \mathbb{R})$ and on $\mathbf{S}'(\mathbb{R}^d, \mathbb{R})$, by its Fourier multiplier $\widehat{(-\Delta)^{\gamma/2}}(\omega) \stackrel{\text{def}}{=} \|\omega\|^\gamma$:

$$(-\Delta)^{\gamma/2} f(\mathbf{x}) \xleftrightarrow{\mathcal{F}} \|\omega\|^\gamma \mathcal{F}\{f\}(\omega). \quad (13)$$

Remark 3.28. This notation $(-\Delta)^{\gamma/2}$ is used to have a simple form of the Fourier multiplier (without a $-$ symbol), but its weakness is that for $\gamma = 2$, $(-\Delta)^{2/2} = -\Delta$ and **not** Δ . Still, the 2-fractional Laplacian is the classic Laplacian (up to a change of sign). And for any $n \in \mathbb{N}^*$, $(-\Delta)^{2n/2} = (-1)^n \Delta^n$, so the fractional Laplacian of order $2n$ is the n -fold iterated Laplacian (again, up to a change of sign). Also, $(-\Delta)^{0/2}$ is indeed the identity operator (Fourier multiplier = 1), so this power notation is well justified. In fact, $(-\Delta)$ is the good object to be “fractionalized”, as $(-\Delta)^{\gamma/2}$ appeared as the only way to rationally generalize the Laplacian [UT14].

We already know that **LC TI** operators commute, so fractional Laplacians obviously commute, but we can say more about a composition of two fractional Laplacians $(-\Delta)^{\gamma_1} \circ (-\Delta)^{\gamma_2}$: it is still a fractional Laplacian, of order $\gamma_1 + \gamma_2$.

Lemma 3.29. For any $\gamma_1, \gamma_2 \in \mathbb{R}_+$, $(-\Delta)^{\gamma_1} \circ (-\Delta)^{\gamma_2} = (-\Delta)^{\gamma_1 + \gamma_2} = (-\Delta)^{\gamma_2} \circ (-\Delta)^{\gamma_1}$. In other words, there is a monoid morphism between $(\mathbb{R}_+, +)$ and $(\{(-\Delta)^\gamma, \gamma \in \mathbb{R}_+\}, \circ)$.

Easy Proof. From the definition with their Fourier multipliers, it is instantaneous.

And it is only a monoid morphism and not a group nor semi-group because the scale-invariance order γ can only be non-negative (there is no inverse). \square

A fractional Laplacian has a very simple Fourier multiplier $\|\omega\|^\gamma$, so it is not surprising that it yields an operator with nice properties: it has all the properties defined in section 2.3!

Proposition 3.30. *The γ -fractional Laplacian is a linear, continuous, translation-invariant, rotation-invariant, and γ -scale-invariant operator (LC TI RI γ -SI), on $E = \mathbf{S}(\mathbb{R}^d, \mathbb{R})$ but **not**⁴⁵ on $\mathbf{S}'(\mathbb{R}^d, \mathbb{R})$.*

Easy Proof. It will be obvious with the general characterization on the Fourier multipliers of LC TI γ -SI and LC TI RI operators, see later theorems 3.35 and 3.39. Note that rotation-invariant implies steerability (of order $n_{(-\Delta)^{\gamma/2}} = 1$), as shown in section 4.4.1. \square

What is even stronger is the converse theorem, stating that $(-\Delta)^{\gamma/2}$ is the **only** operator having all these properties (up to a constant).

Theorem 3.31. *Let G be a LC TI RI γ -SI operator on $E = \mathbf{S}(\mathbb{R}^d, \mathbb{R})$ or on $\mathbf{S}'(\mathbb{R}^d, \mathbb{R})$, then there is $C \in \mathbb{R}$ such that $G = C(-\Delta)^{\gamma/2}$.*

Easy Proof. We will prove it later, see below theorem 3.40. \square

Extend the differential operators to L^2 functions?

By definition, the operators D and Δ are only valid for once- and twice-differentiable functions so one might wonder if they could be extended on $L^2(\mathbb{R}^d, \mathbb{R})$. The simple answer is no, from the functions point-of-view, but from the Fourier transform point-of-view we can. And from the *distribution* point-of-view, we can generalize D_α and $(-\Delta)^{\gamma/2}$ to $\mathbf{S}'(\mathbb{R}^d, \mathbb{R})$.

Remark 3.32 (Differentiable operators on images?). It might be scary to consider a derivative of an image, as it is never smooth if we see it as a function. But for implementation, we will use the discrete derivatives and the discrete Laplacians.

And any other operator G can be *approximatively* applied to an image by going into the Fourier domain, multiplying point-wise by its Fourier multiplier \hat{g} , and coming back in the space domain. It can only be an approximation, but it works.

3.4 Characterizations of geometric properties for integral and convolution operators

In this section, we give some general results characterizing the linear and convolution operators that have some geometric properties (0-SI or γ -SI, and RI, or both of them).

All these results are valid for linear operators on $\mathbf{S}(\mathbb{R}^d, \mathbb{R})$, in *any* dimension $d \in \mathbb{N}^*$.

⁴⁵ But the γ -fractional Laplacian can also be defined on a rich class of distribution, smaller than $\mathbf{S}'(\mathbb{R}^d, \mathbb{R})$, [UT14]. The details of this result goes beyond the reach of this report.

3.4.1 Characterization for LC γ -SI operators

We do not start by studying the special case 0-SI but instead we directly study the general setting γ -SI. This theorem is not so intuitive but it has a simple form and an easy proof.

Theorem 3.33. *G LC is γ -SI iff $k_G(\mathbf{x}, \mathbf{y}) = a^{\gamma+d} k_G(a\mathbf{x}, a\mathbf{y})$ for all $a \in \mathbb{R}_+^*$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$.*

Remark 3.34. At first sight, this form seems weird, and the next result is more natural. The $+d$ in the exponent for a simply comes from the Jacobian of the (bijective) change of variable $\mathbf{x} \mapsto a\mathbf{x}$ (on \mathbb{R}^d).

Computational Proof. ⁴⁶ As G is LC, we can work with its Schwartz kernel k_G .

G is γ -SI iff $a^\gamma G\{f\}(a\mathbf{x}) = G\{f(a\cdot)\}(\mathbf{x})$ for any f, \mathbf{x}, a , but $G\{f\}(\mathbf{x}) = \int_{\mathbb{R}^d} k_G(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y}$, so $a^\gamma G\{f\}(a\mathbf{x}) = a^\gamma \int_{\mathbb{R}^d} k_G(a\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y}$ and $G\{f(a\cdot)\}(\mathbf{x}) = \int_{\mathbb{R}^d} k_G(\mathbf{x}, \mathbf{y}) f(a\mathbf{y}) d\mathbf{y}$.

They are equal iff $\int_{\mathbb{R}^d} a^\gamma k_G(a\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y} = \int_{\mathbb{R}^d} k_G(\mathbf{x}, \mathbf{y}) f(a\mathbf{y}) d\mathbf{y}$ so the (bijective) change of variable $\mathbf{y}' = a\mathbf{y}$ of Jacobian a^d on \mathbb{R}^d gives (for non-zero a) $\int_{\mathbb{R}^d} a^\gamma k_G(a\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y} = \int_{\mathbb{R}^d} k_G(\mathbf{x}, \mathbf{y}'/a) f(\mathbf{y}') a^{-d} d\mathbf{y}'$.

And so G is γ -SI iff $\int_{\mathbb{R}^d} (a^{\gamma+d} k_G(a\mathbf{x}, \mathbf{y}) - k_G(\mathbf{x}, \mathbf{y}/a)) f(\mathbf{y}) d\mathbf{y} = 0$ for any \mathbf{x} , and any f , and so iff $a^{\gamma+d} k_G(a\mathbf{x}, \mathbf{y}) = k_G(\mathbf{x}, \mathbf{y}/a)$ for all \mathbf{x}, \mathbf{y} (as they are orthogonal to any f – if we work in \mathbf{L}^2 , or as they have the same image on any test function f – if we work in \mathbf{S}).

Finally, the same substitution $\mathbf{y}' = a\mathbf{y}$ gives, without the Jacobian this time (it is not a substitution on an integral but on the free variable \mathbf{y}): G is γ -SI iff $a^{\gamma+d} k_G(a\mathbf{x}, a\mathbf{y}') = k_G(\mathbf{x}, \mathbf{y}')$ for all \mathbf{x}, \mathbf{y}' , i.e., iff $a^{\gamma+d} k_G(a\cdot, a\cdot) = k_G(\cdot, \cdot)$ as claimed. \square

3.4.2 Characterization for LC TI γ -SI operators

For LC TI operators (convolution), γ -scale-invariance can be characterized on their Fourier multipliers instead of their impulse response. This theorem is more intuitive than the previous one, and it has an even easier proof.

Theorem 3.35. *G LC TI is γ -SI iff $\hat{g}(a\boldsymbol{\omega}) = a^\gamma \hat{g}(\boldsymbol{\omega})$ for all $a \in \mathbb{R}_+^*$, $\boldsymbol{\omega} \in \mathbb{R}^d$, i.e., \hat{g} is γ -SI.*

Computational Proof. As G is LC TI, we can indeed work with its impulse response g , and the previous theorem directly gives that G is γ -SI iff $\forall a > 0, \forall \mathbf{x}, \mathbf{y}, k_G(\mathbf{x}, \mathbf{y}) = a^{\gamma+d} k_G(a\mathbf{x}, a\mathbf{y})$ iff $\forall a > 0, \forall \mathbf{x}, \mathbf{y}, g(\mathbf{x} - \mathbf{y}) = a^{\gamma+d} g(a(\mathbf{x} - \mathbf{y}))$, and applying the Fourier transform and its time-scaling property (in dimension d) gives $\mathcal{F}\{d_a g\} = a^{-d} d_{1/a} \mathcal{F}\{g\}$ (for non-zero a), so $\forall a > 0, \forall \mathbf{x}, \mathbf{y}, g(\mathbf{x} - \mathbf{y}) = a^{\gamma+d} g(a(\mathbf{x} - \mathbf{y}))$ iff $\forall a > 0, \forall \mathbf{x}, g(\mathbf{x}) = a^{\gamma+d} g(a\mathbf{x})$, iff $\forall a > 0, \forall \boldsymbol{\omega}', \hat{g}(\boldsymbol{\omega}') = a^\gamma \hat{g}(\boldsymbol{\omega}'/a)$ iff $\forall a > 0, \forall \boldsymbol{\omega}, \hat{g}(a\boldsymbol{\omega}) = a^\gamma \hat{g}(\boldsymbol{\omega})$ (with $\boldsymbol{\omega} = \boldsymbol{\omega}'/a$), which is exactly the result we announced. \square

This theorem makes it clear that a LC TI G is γ -SI (as an operator) iff \hat{g} is γ -SI (as a function). The characterization on k_G for G to be LC TI seemed unnatural (with this $+d$ in the exponent), so this theorem is reassuring: for LC TI G , G is indeed γ -SI iff some underlying function is γ -homogeneous.

⁴⁶ Some proofs are harder or longer than other, and we indicate them like here, to warn the reader.

Isolating the γ -scale-invariance part as a fractional Laplacian

What is even stronger for **LC TI** γ -**SI** operators is the next result, which allows to *isolate* the γ -**SI** part as a (fractional) Laplacian, composed with a second term which turns out to be **LC TI** 0-**SI**. It really shows that having a $\|\omega\|^\gamma$ multiplicative term in its Fourier multiplier is *the only way* to have γ -scale-invariance for a convolution.

Theorem 3.36 (Decomposition of a **LC TI** γ -**SI** operator). *Let $\gamma \in \mathbb{R}_+$, and $G = G_\gamma$ be an operator from E to F . G_γ is **LC TI** γ -**SI** iff there exists a **LC TI** 0-**SI** G_0 such that $G_\gamma = G_0 \circ (-\Delta)^{\gamma/2}$.*

Proof. • Converse sense:

If $G_\gamma = G_0 \circ (-\Delta)^{\gamma/2}$ for such G_0 , then by composition G_γ is also linear, continuous, translation-invariant (**LC TI**). And because G_0 is 0-scale-invariant, and the fractional Laplacian is γ -scale-invariant, their composition is γ -**SI** (thanks to proposition 2.39). Hence G_γ is indeed **LC TI** γ -**SI**.

• Direct sense:

In the Fourier domain, let $A(\omega) = \mathcal{F}\{G_\gamma\}(\omega)/\mathcal{F}\{(-\Delta)^{\gamma/2}\}(\omega)$ for $\omega \neq \mathbf{0}_{\mathbb{R}^d}$. Because G_γ and $(-\Delta)^{\gamma/2}$ are γ -**SI** operators, A is 0-scale-invariant:

$$\begin{aligned} \forall \omega \neq \mathbf{0}_{\mathbb{R}^d}, \forall a \in \mathbb{R}_+^*, \quad A(a\omega) &= \frac{\mathcal{F}\{G_\gamma\}(a\omega)}{\mathcal{F}\{(-\Delta)^{\gamma/2}\}(a\omega)} \\ &= \frac{a^\gamma \mathcal{F}\{G_\gamma\}(\omega)}{a^\gamma \mathcal{F}\{(-\Delta)^{\gamma/2}\}(\omega)} \\ &= \frac{\mathcal{F}\{G_\gamma\}(\omega)}{\mathcal{F}\{(-\Delta)^{\gamma/2}\}(\omega)} = A(\omega). \end{aligned}$$

We also need to justify why A can be seen as a Fourier multiplier. By definition, A is measurable, and we can consider its inverse Fourier transform a . So there is a valid impulse response a such that $A(\omega) = \hat{a}(\omega)$ and it indeed defines an operator G_0 which is **LC TI** (as a convolution), and 0-**SI** (because A is 0-scale-invariant). \square

Remark 3.37. Note that there is no reason for G_0 to be unitary (cf. appendix A). At first, this was the result we were hoping to get, based on some examples, but it cannot be achieved. A simple counter example is a directional derivative, e.g., D_x in dimension $d = 2$, which can be decomposed as a 1-scale-invariant part (i.e., half Laplacian $(-\Delta)^{1/2}$, of Fourier multiplier $\propto \|\omega\|$) and a 0-scale-invariant part (of Fourier multiplier $\propto \omega = (r, \theta) \mapsto j \cos(\theta)$) which is clearly **not** unitary.

3.4.3 Characterization for **LC RI** operators

This first characterization of rotation-invariance has a nice formulation as some sort⁴⁷ of “partial” rotation-invariance of the kernel k_G :

⁴⁷ Rotation-invariance for a function (or distribution) of two variables should be defined as $k_G(\mathbf{x}, \mathbf{y}) = k_G(R\mathbf{x}, R'\mathbf{y})$ for any **two** rotations R, R' . Here we only have the special case when $R = R'$.

Theorem 3.38. G LC is RI iff $k_G(\mathbf{x}, \mathbf{y}) = k_G(R\mathbf{x}, R\mathbf{y})$ for all $R \in \text{SO}(d)$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$.

Computational Proof. As G is LC, we can indeed work with its Schwartz kernel k_G .

G is RI iff $G\{f\}(R\mathbf{x}) = G\{f(R\cdot)\}(\mathbf{x})$ for any f, \mathbf{x}, R , but $G\{f\}(\mathbf{x}) = \int_{\mathbb{R}^d} k_G(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y}$, so $G\{f\}(R\mathbf{x}) = \int_{\mathbb{R}^d} k_G(R\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y}$ and $G\{f(R\cdot)\}(\mathbf{x}) = \int_{\mathbb{R}^d} k_G(\mathbf{x}, \mathbf{y}) f(R\mathbf{y}) d\mathbf{y}$.

They are equal iff $\int_{\mathbb{R}^d} k_G(R\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y} = \int_{\mathbb{R}^d} k_G(\mathbf{x}, \mathbf{y}) f(R\mathbf{y}) d\mathbf{y}$ so the (bijective) change of variable $\mathbf{y}' = R\mathbf{y}$ gives $\int_{\mathbb{R}^d} k_G(R\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y} = \int_{\mathbb{R}^d} k_G(\mathbf{x}, R^{-1}\mathbf{y}') f(\mathbf{y}') d\mathbf{y}'$.

And so G is RI iff $\int_{\mathbb{R}^d} (k_G(R\mathbf{x}, \mathbf{y}) - k_G(\mathbf{x}, R^{-1}\mathbf{y})) f(\mathbf{y}) d\mathbf{y} = 0$ for any \mathbf{x}, f , and so iff $k_G(R\mathbf{x}, \mathbf{y}) = k_G(\mathbf{x}, R^{-1}\mathbf{y})$ for all \mathbf{x}, \mathbf{y} (as they are orthogonal to any f if we work in L^2 , as they have the same image on any test function if we work in \mathbf{S}).

Finally, the same change of variable $\mathbf{y}' = R\mathbf{y}$ gives: G is RI iff $k_G(R\mathbf{x}, R\mathbf{y}') = k_G(\mathbf{x}, \mathbf{y}')$ for all \mathbf{x}, \mathbf{y}' , i.e., iff $k_G(R\cdot, R\cdot) = k_G(\cdot, \cdot)$ as claimed. \square

3.4.4 Characterization for LC TI RI operators

For LC TI operators, rotation-invariant can be characterized on their Fourier multipliers, which has to be rotation-invariant functions:

Theorem 3.39. G LC TI is RI iff $\hat{g}(R\boldsymbol{\omega}) = \hat{g}(\boldsymbol{\omega})$ for all $R \in \text{SO}(d)$, $\boldsymbol{\omega} \in \mathbb{R}^d$, i.e., \hat{g} is RI.

Proof. It is very simply to obtain, as G is LC TI, we can work with its impulse response g , and the previous theorem directly gives that G is RI iff $\forall R, \forall \mathbf{x}, \mathbf{y}, k_G(\mathbf{x}, \mathbf{y}) = k_G(R\mathbf{x}, R\mathbf{y})$ iff $\forall R, \forall \mathbf{x}, \mathbf{y}, g(\mathbf{x} - \mathbf{y}) = g(R(\mathbf{x} - \mathbf{y}))$, iff $\forall R, g = r_R\{g\}$, and as \mathcal{F} commutes with r_R , $\hat{g} = \mathcal{F}\{g\} = \mathcal{F}\{r_R\{g\}\} = r_R\{\mathcal{F}\{g\}\} = r_R\hat{g}$, so G is RI iff $\forall R, r_R\hat{g} = \hat{g}$. \square

This theorem makes it clear that a LC TI G is RI (as an operator) iff \hat{g} is RI (as a function). A RI function on \mathbb{R}^d is *purely radial*, so such \hat{g} will only depend on $\|\boldsymbol{\omega}\|$ and not $\boldsymbol{\omega}$. Note that it narrows greatly the richness of the Fourier multiplier.

3.4.5 LC TI RI γ -SI operators are exactly fractional Laplacians

This last result is very strong, we announced it above when we worked on fractional Laplacians without proving it, but let us rephrase it and prove it here.

Theorem 3.40 (Universality of the fractional Laplacian). G is LC TI RI γ -SI on $\mathbf{S}(\mathbb{R}^d, \mathbb{R})$ iff \hat{g} is γ -SI and rotation-invariant⁴⁸, iff $\hat{g}(\boldsymbol{\omega}) \propto \|\boldsymbol{\omega}\|^\gamma$, i.e., there exists a constant $C \in \mathbb{R}$ such that $\hat{g}(\boldsymbol{\omega}) = C\|\boldsymbol{\omega}\|^\gamma$, for all $\boldsymbol{\omega} \in \mathbb{R}^d$, iff $G \propto (-\Delta)^{\gamma/2}$.

Proof. It is simply the consequence of the two previous theorems 3.35 and 3.39, characterizing LC TI γ -SI and LC TI RI operators.

- Direct sense: For a LC TI RI γ -SI operator G , \hat{g} have to be γ -SI and RI, so it has a simple dependency on $\|\boldsymbol{\omega}\|$ (as it is γ -SI). For non-zero $\boldsymbol{\omega}$, $\hat{g}(\boldsymbol{\omega}) = \hat{g}(\|\boldsymbol{\omega}\| \cdot \boldsymbol{\omega} / \|\boldsymbol{\omega}\|) = \|\boldsymbol{\omega}\|^\gamma \hat{g}(\boldsymbol{\omega} / \|\boldsymbol{\omega}\|) = C\|\boldsymbol{\omega}\|^\gamma$, with $C \stackrel{\text{def}}{=} \hat{g}(\boldsymbol{\omega} / \|\boldsymbol{\omega}\|)$ being constant, as \hat{g} is RI.

⁴⁸ With the vocabulary of functions in real analysis, \hat{g} is γ -homogeneous and isotropic.

- Converse sense: It is obvious that $\hat{g}(\omega) = C\|\omega\|^\gamma$ yields a γ -**SI** and **RI** Fourier multiplier, *i.e.*, a **LC TI RI** γ -**SI** operator, thanks to the converse sense of both theorems 3.35 and 3.39. \square

This theorem makes it clear that a **LC TI** G is **RI** and γ -**SI** (as an operator) *iff* \hat{g} is **RI** and γ -**SI** (as a function). And the only operators in this class are the fractional Laplacian (up to a constant).

A corollary of this theorem is another (naive) result showing that imposing too many properties on G reduces the freedom on \hat{g} (and on G) so much that operators proportional to the identity are the only possible one.

Corollary 3.41. For $\gamma = 0$ (*i.e.*, for “true” scale-invariant operators), G is **LC TI RI** 0-**SI** *iff* \hat{g} is constant *iff* G is proportional to the identity $G \propto \text{Id}$ (*i.e.*, $\exists C \in \mathbb{R}$ such that $G = C \text{Id}$, $G : f \mapsto Cf$).

Easy Proof. It is a very intuitive result, so we will give an informal proof. Asking G to be **RI** and 0-**SI** is equivalent to ask \hat{g} to be both scale-invariant and isotropic, *i.e.*, not depending on the norm $\|\omega\|$ neither on its direction $\frac{\omega}{\|\omega\|}$, so not depending on ω , *i.e.*, \hat{g} has to be constant. It is also the consequence of the previous theorem, with $\gamma = 0$, as $G = C(-\Delta)^{\gamma/2} = C \text{Id}$. \square

Conclusion

In this section, we studied linear and continuous operators (**LC**) with the help of their Schwartz kernel k_G , and convolution operators (**LC TI**) with the help of their impulse response g or Fourier multiplier \hat{g} . Some important examples were given, including translations, dilatations and rotations, and differential operators (derivatives and directional derivatives, Laplacians and fractional Laplacians). Thanks to the two Schwartz theorems, we gave some characterizations of scale-invariance, γ -scale-invariance, and/or rotation-invariance, for integral and convolution operators. Combining these characterizations gave an interesting result, showing the universality of the fractional Laplacian among the **LC TI RI** γ -**SI** operators.

4 Steerable convolution operators in dimension 2

We give here the general definition of a steerable operator on test functions of domain \mathbb{R}^d , first with a definition on the operator, and then with two equivalent definitions (either of g or \hat{g}). Then we start by giving a few general results valid for any d , and then we present a more complete study for the specific setting of $d = 2$. Our main contributions are in this setting, especially the characterizations 4.31 and 4.49.

In all this section, G denotes a linear and continuous operator (LC), with no additional hypothesis if nothing is specified.

4.1 Defining steerability, in any dimension

This property is the last one we introduce for our operators G in $\mathcal{L}(E, F)$. Note that this notion and some results presented in this section and the next one are new, to the best of our knowledge.

The goal of steerability is to generalize the rotation-invariance, as we explained above how restrictive rotation-invariance can be: indeed it reduces the dimension of the domain of $\hat{g}(\omega)$ from d (for ω), to only 1 (for $\|\omega\| \in \mathbb{R}_+$).

4.1.1 Steerable functions and distributions

Steerability for function or distribution, as we define it here, is a natural generalization of rotation-invariance, and proposition 4.4 will soon back up this intuition.

Note that steerable functions have already been studied, for instance [HOT98, Def.1] studies a more general notion with a generic Lie transformation group instead of simply the rotations. They just consider one example with rotations in 2D in page 4, but the rest of their work is more general (and much more complicated).

Definition 4.1 (Steerability). *For $f \in E$, we say that f is a steerable function (St) when V_f , the span of its rotated versions, has a finite dimension i.e., when $V_f \stackrel{\text{def}}{=} \text{Span}_{\mathbb{R}}\{r_R\{f\}, R \in \text{SO}(d)\}$ has a finite dimension (as a linear sub-space of E).*

Definition 4.2 (Order of steerability). *If f is steerable, we write n_f the dimension of V_f : $n_f \stackrel{\text{def}}{=} \dim_{\mathbb{R}} V_f \in \mathbb{N}$, and it is called its order of steerability.*

We can give a few examples of simple functions that are steerable:

Examples 4.3. • The null function $\mathbf{0}_E$ of E is steerable, of order $n_{\mathbf{0}_E} = 0$ (and the converse is clearly true).

- A rotation-invariant function f is steerable, of order 1, as $\forall R \in \text{SO}(d), r_R\{f\} = f$ (and it is an equivalence, cf. proposition 4.4).
- In dimension $d = 1$, the only (proper) rotation is the identity, so steerability is not interesting, as *any* function will be steerable in dimension $d = 1$.
- On \mathbb{R}^d , any linear form $f : \mathbf{x} \mapsto a_1x_1 + \dots + a_dx_d$ is steerable, of order $n_f \leq d$. Indeed, $V_f \subset \text{Vect}_{\mathbb{R}}\{\mathbf{x} \mapsto x_i, i \in [1, \dots, d]\}$. And we can even say that $n_f = d$ if f is non-zero, as V_f contains all the linear forms, and so it contains all the canonical projections, because any hyperplan of \mathbb{R}^d can be written as the rotated version of the hyperplan $\ker(f) = \mathbf{a}^\perp$ ($\mathbf{b}^\perp = R(R^{-1}\mathbf{a})^\perp$).

As wanted, this first notion of steerability generalizes rotation-invariance for $f \in E$.

Proposition 4.4 (RI = St of order 1). *Let f be a non-zero function, in $\mathbf{S}(\mathbb{R}^2, \mathbb{R})$ or in $\mathbf{L}^2(\mathbb{R}^2, \mathbb{R})$, or a non-null distribution in $\mathbf{S}'(\mathbb{R}^d, \mathbb{R})$.*

Then f is rotation-invariant (RI) iff f is steerable (St), with $n_f = 1$.

Hard Proof. • **Direct sense:** By definition, $\forall R \in \text{SO}(d), r_R\{f\} = f$, so $V_f = \text{Span}_{\mathbb{R}}\{f\} = \{\lambda f, \lambda \in \mathbb{R}\} = \text{“}\mathbb{R}f\text{”}$ has dimension $n_f = 1$, so f is steerable.

- **Converse sense for functions:** If V_f has dimension 1, f is non-zero, and it means that for a $R \in \text{SO}(d)$, there exists one $\lambda_R \in \mathbb{R}$ such that $r_R\{f\} = \lambda_R f$. It is unique as f is not the null function.

First, as R can be written⁴⁹ as $(R')^2$ for some $R' \in \text{SO}(d)$, and so $r_R\{f\} = r_{(R')^2}\{f\} = ((r_{R'})^2)\{f\} = (\lambda_{R'})^2 f = \lambda_R f$, so by uniqueness $\lambda_R = (\lambda_{R'})^2 \geq 0$. Now, let prove that $\lambda_R \in \{-1, +1\}$: function norms are always some sort of integrals over the domain \mathbb{R}^d and on f (or $\partial^\beta f$ for Sobolev semi-norms), and so there are all rotation-invariant, as the (bijective) change of variable $\mathbf{x} \mapsto R\mathbf{x}$ has a Jacobian = +1 and is a bijection on \mathbb{R}^d . So $\|f\| = \|r_R\{f\}\|$, which gives that $|\lambda_R| = 1$ by homogeneity of the norm $\|\cdot\|$.

These two points prove that $\lambda_R = 1$, for any R , and so $r_R\{f\} = \lambda_R f = f$, so f is indeed rotation-invariant.

- **Converse sense for distributions:** For distributions, λ_R still exists (on the duality bracket, $\langle r_R\{\phi\}, f \rangle = \lambda_R \langle \phi, f \rangle$ for $\phi \in \mathbf{S}'(\mathbb{R}^d, \mathbb{R})$ and $f \in \mathbf{S}(\mathbb{R}^d, \mathbb{R})$), and we can still easily prove that $\lambda_R \geq 0$. By duality bracket, $|\langle r_R\{\phi\}, f \rangle| = |\langle \phi, r_R\{f\} \rangle|$. For some class C of functions, this gives $\sup_{f \in C} |\langle r_R\{\phi\}, f \rangle| = \sup_{f \in C} |\langle \phi, r_R\{f\} \rangle|$, so if C is stable by rotations and such that $\sup_{f \in C} |\langle \phi, f \rangle| < +\infty$, thanks to $r_R\{\phi\} = \lambda_R \phi$ we have $|\lambda_R| \sup_{f \in C} |\langle \phi, f \rangle| = \sup_{f \in C} |\langle \phi, f \rangle|$, so $|\lambda_R| = 1$, yielding the same conclusion. \square

4.1.2 Steerable operators (St)

We can finally ask our linear operators $G : E \rightarrow F$ to be *steerable*. To the best of our knowledge, we are the first to try to define and study this notion of steerable operators *in this way*, and in the light of Schwartz’s kernel theorem and Fourier multipliers.

When we try to generalize the steerability of functions to operators, the first definition we could think of is the following definition 4.5 on the operator part, but other definitions could consider the impulse response g of G and on the Fourier multiplier \hat{g} of G (proposition 4.9). It turns out that the three definitions are equivalent.

Definition 4.5 (Steerability for operators). *For $G \in \mathcal{L}(E, F)$ a linear operator, we say that G is a steerable operator (St) when $V_G \stackrel{\text{def}}{=} \text{Span}_{\mathbb{R}}\{r_R \circ G \circ r_{R^{-1}}, R \in \text{SO}(d)\}$ has a finite dimension (as a linear sub-space of $\mathcal{L}(E, F)$).*

The order of steerability of G is also (uniquely) defined as $n_G \stackrel{\text{def}}{=} \dim_{\mathbb{R}} V_G$.

⁴⁹ For instance, for $d = 2$, take half the angle of R , $R' = R_{\theta_0/2}$ if $R = R_{\theta_0}$; and in higher dimension it is still true, only harder to visualize.

Remark 4.6 (Why $r_R \circ G \circ r_{R^{-1}}$?). The first thing that we thought of when we tried to generalize the steerability from functions to operators is to define V_G the same way we defined V_f : with $r_R\{G\} = r_R \circ G$. But this does not work at all: it would not be a generalization of rotation-invariance, and it is weirdly non-symmetric: if $G = G_1 \circ G_2$, the intuition says that turning G is equivalent to turning both G_1 and G_2 , but $r_R\{G\} = (r_R \circ G_1) \circ G_2$: only the left term G_1 gets “turned” by R with this incorrect definition.

With the correct definition, it works as expected: $r_R \circ G \circ r_{R^{-1}} = r_R \circ (G_1 \circ G_2) \circ r_{R^{-1}} = r_R \circ G_1 \circ (r_{R^{-1}} r_R) \circ G_2 \circ r_{R^{-1}} = (r_R \circ G_1 \circ r_{R^{-1}}) \circ (r_R \circ G_2 \circ r_{R^{-1}})$, which indeed appears as the composition of both G_1 and G_2 “turned” by R .

Example 4.7 ($r_R \circ G \circ r_{R^{-1}}$ or $r_{R^{-1}} \circ G \circ r_R$?). One could wonder if the order is the good one. A simple sanity check is to consider, in dimension 2, $G = D_x$ and $R = R_{+\pi/2}$. The intuition behind this notion of “turned” operator is that applying $R_{+\pi/2}$ to D_x should give D_y . We let the reader check that $r_{R_{+\pi/2}} \circ D_x \circ r_{R_{+\pi/2}^{-1}} = r_{R_{+\pi/2}} \circ D_x \circ r_{R_{-\pi/2}} = D_y$ as wanted, but $r_{R_{+\pi/2}^{-1}} \circ D_x \circ r_{R_{+\pi/2}} = r_{R_{-\pi/2}} \circ D_x \circ r_{R_{+\pi/2}} = r_{R_{-\pi}} D_y \circ r_{R_{\pi}} : f \mapsto \{D_y(f(\cdot))\}(\cdot) \neq D_y$.

Remark 4.8 (Conjugation action). This notation $r_R \circ G \circ r_{R^{-1}}$ will ring a bell to anyone who studied group action, as it is one of the most well-known and studied extensively in the context of group action. This is called a *conjugation action*, and here the group of rotation operators r_R (i.e., $\{r_R, R \in \text{SO}(d)\}$) acts on the group of convolutions G .

This first definition of steerability is the most general definition we can give (G was only asked to be linear). By restricting to **LC TI** operators, we can also use their impulse response g or Fourier multiplier \hat{g} , and express the steerability of G as the steerability of g or \hat{g} (as function or distribution, and not operators). Fortunately, these two definitions turn out to be equivalent to the first one.

Proposition 4.9 (Other definition of steerability for **LC TI** operators). *For $G : E \rightarrow F$ a **LC TI** operator, we say that G is a steerable operator (**St**) when g is steerable (as a distribution) or when \hat{g} is steerable (as a function), i.e., when either $V_g = \text{Span}_{\mathbb{R}}\{r_R\{g\}, R \in \text{SO}(d)\}$ or $V_{\hat{g}} = \text{Span}_{\mathbb{R}}\{r_R\{\hat{g}\}, R \in \text{SO}(d)\}$ has a finite dimension.*

The order of steerability of G is also $n_G = n_g = \dim_{\mathbb{R}} V_g = n_{\hat{g}} = \dim_{\mathbb{R}} V_{\hat{g}}$.

Long Proof. We need to prove the equivalence between the steerability of g and the steerability of \hat{g} . It comes from a property we highlighted before, the rotation-invariance of the Fourier transform (property 1.3), $r_R \circ \mathcal{F} = \mathcal{F} \circ r_R$. Let prove that g is steerable iff \hat{g} is steerable, [Püs16].

- **Direct sense:** If g is steerable, we denote by $\beta = (g_1, \dots, g_{n_g})$ a basis of V_g (which is finite dimensional by definition). The only thing we need⁵⁰ is that the we can take the Fourier transform of the g_i .

⁵⁰ But beware that we cannot chose a basis such that $g_i = r_{R_i} g$ for a certain rotation $R_i \in \text{SO}(d)$, but that is not an issue as this kind of proof works on the $(g_i)_{1 \leq i \leq n_g}$ without the need to have a specific form for the basis.

Let $R \in \text{SO}(d)$, then by definition of the basis β , there is $(\lambda_i)_{1 \leq i \leq n_g} \in \mathbb{R}^{n_g}$ such that $r_R\{g\} = \sum_{i=1}^{n_g} \lambda_i g_i$. And $\hat{g}(\omega) = \mathcal{F}\{g\}(\omega)$, and so for $\omega \in \mathbb{R}^d$,

$$\begin{aligned} (r_R\{\hat{g}\})(\omega) &= (r_R\{\mathcal{F}\{g\}\})(\omega) = ((r_R \circ \mathcal{F})\{g\})(\omega) \\ (\text{as } \mathcal{F} \text{ is RI}) &= ((\mathcal{F} \circ r_R)\{g\})(\omega) \\ &= (\mathcal{F}\{r_R\{g\}\})(\omega) = \left(\mathcal{F} \left\{ \sum_{i=1}^{n_g} \lambda_i g_i \right\} \right) (\omega) \\ (\text{as } \mathcal{F} \text{ is linear}) &= \sum_{i=1}^{n_g} \lambda_i (\mathcal{F}\{g_i\})(\omega) \end{aligned}$$

$$\text{So finally, } (r_R\{\hat{g}\})(\omega) = \sum_{i=1}^{n_g} \lambda_i \hat{g}_i(\omega).$$

And so, \hat{g} is steerable, because there is $(\hat{g}_i)_{1 \leq i \leq n_g}$ such that for every $R \in \text{SO}(d)$, there exists $(\lambda_i)_{1 \leq i \leq n_g} \in \mathbb{R}^{n_g}$ satisfying $(r_R\{\hat{g}\}) = \sum_{i=1}^{n_g} \lambda_i \hat{g}_i$.

It also proves that $n_{\hat{g}} \leq n_g$ as $(\hat{g}_i)_{1 \leq i \leq n_g}$ is (at least) a spanning family of $V_{\hat{g}}$.

- Converse sense: The same computation as for the direct sense can be done here, by using the fact that $g = \mathcal{F}^{-1}\{\hat{g}\}$, and the rotation-invariance of \mathcal{F}^{-1} (property 1.3), and its linearity. It also proves that $n_g \leq n_{\hat{g}}$ as $(g_i)_{1 \leq i \leq n_{\hat{g}}}$ is at least a spanning family of V_g .

Both sense gave an inequality on dimensions, so $n_g = n_{\hat{g}}$. \square

Finally, we can sum up the previous results in this proposition.

Proposition 4.10 (Equivalence between the three notions of steerability). *G is steerable iff g is steerable iff \hat{g} is steerable. And we have $n_G = n_{\hat{g}} = n_g$.*

Long Proof. We already proved the second equivalence, so we can focus for instance on G is **St** iff \hat{g} is **St**. This proof is based on the explicit formula linking g and G , proposition 3.7, $g(\mathbf{x}) = G\{\delta_0\}(\mathbf{x})$, and the fact that δ_0 is rotation-invariant (as a distribution, because $\mathbf{x} = 0$ iff $R\mathbf{x} = 0$).

- Direct sense: If G is steerable, i.e., if V_G has finite dimension, then there exists n_G operators $G_1, \dots, G_{n_G} \in \text{SO}(d)$, of impulse response g_i , such that for all $R \in \text{SO}(d)$, there is $\lambda_1, \dots, \lambda_{n_G} \in \mathbb{R}$ satisfying $r_R G r_{R^{-1}} = \sum_{i=1}^{n_G} \lambda_i G_i$. We have $(r_R g)(\mathbf{x}) = g(R\mathbf{x}) = (G\{\delta_0\})(R\mathbf{x}) = (r_R G)\{\delta_0\}(\mathbf{x})$, and $\delta_0 = r_{R^{-1}}\delta_0$ as it is **RI**, so $r_R g = (r_R G)\{r_{R^{-1}}\delta_0\} = (r_R G r_{R^{-1}})\{\delta_0\}$. Therefore, $r_R g = \sum_{i=1}^{n_G} \lambda_i (G_i)\{\delta_0\} = \sum_{i=1}^{n_G} \lambda_i g_i$, which proves that g is steerable, of order $n_g \leq n_G$ (because (g_1, \dots, g_{n_G}) spans V_g).
- Converse sense: The same computation works in a similar way, proving that $r_R G r_{R^{-1}}$ and $\sum_{i=1}^{n_G} \lambda_i G_i$ has the same impulse response ($r_R g$ and $\sum_{i=1}^{n_G} \lambda_i g_i$), and so they are equal (as $G\{f\} = g * f$). So if g is **St**, G is also **St**, of order $n_G \leq n_g$.

Both sense gave an inequality on dimensions, so $n_G = n_g (= n_{\hat{g}})$. \square

The following result is not new and has already been used earlier, but we prefer to summarize the computations in a lemma.

Lemma 4.11 (Three forms for a rotated operator).

Let G be a convolution, and for $R \in \text{SO}(d)$, let $G_R = r_R \circ G \circ r_{R^{-1}}$. Then his impulse response is $g_R = r_R\{g\}$, and his Fourier multiplier is $\widehat{g}_R = r_R\{\widehat{g}\}$.

Easy Proof. For the impulse response, the closed-form expression gives $g_R = G_R\{\delta_0\} = (r_R \circ G \circ r_{R^{-1}})\{\delta_0\} = (r_R \circ G)\{r_{R^{-1}}\{\delta_0\}\}$ and δ_0 is **RI** so $g_R = (r_R \circ G)\{\delta_0\} = r_R\{G\{\delta_0\}\} = r_R\{g\}$.

The same result can be obtained with the convolution formula $G\{f\} = (g * f)$ and lemma 2.43 about the rotation of a convolution: $G_R\{f\} = (r_R \circ G \circ r_{R^{-1}})\{f\} = r_R(G\{r_{R^{-1}}\{f\}\}) = r_R(g * (r_{R^{-1}}\{f\})) = (r_R\{g\}) * (r_R\{r_{R^{-1}}\{f\}\}) = (r_R\{g\}) * f = (g_R * f)$.

And for the Fourier multiplier, it comes directly from the rotation-invariance of \mathcal{F} (property 1.3): $\widehat{g}_R = \widehat{r_R\{g\}} = \mathcal{F}\{r_R\{g\}\} = r_R\{\mathcal{F}\{g\}\} = r_R\{\widehat{g}\}$. \square

Example 4.12. For example, in dimension $d = 2$, G is steerable when

$$\dim_{\mathbb{R}} V_{\widehat{g}} = \dim_{\mathbb{R}} \text{Span}_{\mathbb{R}}\{\widehat{g}(R_{\theta}\cdot), \theta \in [0, 2\pi)\} < +\infty$$

Note that such a parametric form (in θ) for $V_{\widehat{g}}$ is only valid in dimension 2.

Examples 4.13. These examples are the operator versions of the few examples of steerable functions given above:

- The null operator on E is steerable, of order $n_{0_{\mathcal{L}(E)}} = 0$ (and the conserve is true).
- A rotation-invariant operator (**RI**) is steerable, of order 1 (and it is an equivalence, see after proposition 4.22).
- On \mathbb{R}^2 , any directional derivative D_{α} is steerable, of order $n_{D_{\alpha}} = 2$. A composition of $m \in \mathbb{N}$ directional derivatives is also steerable, of order $m + 1$ (see after in lemma 4.41).

The steerability order $n_f, n_G \in \mathbb{N}$ indicates somehow the “directional variability” of a function f or an operator G . Being able to have $n_G > 1$ (e.g., for derivatives on \mathbb{R}^2), proves that steerability is really a richer notion than rotation-invariance, as proposition 4.4 already proved that it contains rotation-invariance (when $n_G = 1$).

We will denote **St** the class of steerable operators, and **LC TI St** the class of linear, continuous, translation-invariant and steerable operators (on E). We tend to prefer the longer version **LC TI St**, the class of **steerable convolution operators**, to always remember than steerability does **not** require translation-invariance. As for the previous classes, these two classes are stable under linear combinations, and by compositions

Proposition 4.14. *The class **LC TI St** is a linear sub-space of $\mathcal{L}(E, F)$, and*

$$n_{G_1+G_2} \leq n_{G_1} \times n_{G_2}. \quad (14)$$

Long Proof. For the stability by linear combinations, this time we prefer to separate between the (external) scalar product λG and the (internal) sum $G_1 + G_2$, instead of doing both at the same time with $\lambda_1 G_1 + \lambda_2 G_2$ as we did before.

- The null operator is steerable (of order $n_G = 0$), as shown above.

- If G is steerable, and $\lambda \in \mathbb{R}^*$, then obviously $\lambda G : f \mapsto (\lambda G)\{f\} = \lambda(G\{f\}) = \lambda G\{f\}$ (scalar product) is still steerable⁵¹, as $\text{Span}_{\mathbb{R}}\{r_R \circ (\lambda G), R \in \text{SO}(d)\} = \lambda \text{Span}_{\mathbb{R}}\{r_R G, R \in \text{SO}(d)\} = \lambda V_G = V_G$ (as $\lambda \neq 0$), which is finite dimensional by hypothesis on G (being steerable). So $V_{\lambda G}$ is V_G if $\lambda \neq 0$, but is the null singleton $\{\mathbf{0}_{E \rightarrow E}\}$ if $\lambda = 0$, and therefore the order of steerability of λG is given by

$$n_{\lambda G} = \begin{cases} n_G & \text{if } \lambda \neq 0, \\ 0 & \text{if } \lambda = 0. \end{cases}$$

- If G_1, G_2 are two steerable operators, then for $R \in \text{SO}(d)$ we simply have $r_R(G_1 + G_2) = r_R G_1 + r_R G_2 \in V_{G_1} + V_{G_2}$ as $r_R G_1 \in V_{G_1}$ and $r_R G_2 \in V_{G_2}$, and this sum⁵² is a linear space, therefore $V_{G_1+G_2} = \text{Span}\{r_R(G_1 + G_2), R \in \text{SO}(d)\} \subset V_{G_1} + V_{G_2}$, and so $V_{G_1+G_2}$ is again finitely dimensional, of dimension $n_{G_1+G_2} \leq n_{G_1} + n_{G_2}$, which is exactly (14). \square

Remark 4.15 (Optimal inequality). One could wonder if this inequality (14) could be improved, as it is sometimes an equality (e.g., $G_1 = 0$). In fact, we cannot improve it to $n_{G_1+G_2} = \max(n_{G_1}, n_{G_2})$, as for instance if $G_1 = \text{Id}, G_2 = -G_1$, then $n_{G_1+G_2} = n_{\mathbf{0}_E} = 0 < \max(n_{G_1}, n_{G_2}) = \max(1, 1) = 1$.

Proposition 4.16. *Composition preserves steerability, i.e., if G and G' are operators on E and are both steerable, then $G'' = G \circ G'$ is still steerable, and $n_{G''}$ satisfies:*

$$n_{G \circ G'} \leq n_G \times n_{G'}. \quad (15)$$

Proof. Let G, G' two steerable operators, of order $n = n_G$ and $n' = n_{G'}$, and $G'' = G \circ G'$. Thanks to proposition 4.10, we can work on either $g = g * g'$ or $\hat{g} = \hat{g} \hat{g}'$. As we already observed, $r_R\{g\} = (r_R\{g'\}) * (r_R\{g\})$ and $r_R\{\hat{g}\} = (r_R\{\hat{g}'\})(r_R\{\hat{g}\})$.

If we assume G and G' to be steerable of respective order n and n' , then by definition we can take $(G_i)_{1 \leq i \leq n}$ and $(G'_k)_{1 \leq k \leq n'}$ two bases of V_G and $V_{G'}$. No need to explicitly write G_i as a rotated version of G , it will be easier to just work with G_i and G'_k . And it is easier to work with their impulse responses, g and g' , (g_i) and (g'_k) .

Let $R \in \text{SO}(d)$, there is n numbers $\lambda_i \in \mathbb{R}$ and n' numbers $\lambda'_k \in \mathbb{R}$, such that $r_R g = \sum_{i=1}^n \lambda_i g_i$ and $r_R g' = \sum_{k=1}^{n'} \lambda'_k g'_k$. Therefore,

$$r_R g'' = \left(\sum_{i=1}^n \lambda_i g_i \right) * \left(\sum_{k=1}^{n'} \lambda'_k g'_k \right) = \sum_{i=1}^n \sum_{k=1}^{n'} \underbrace{(\lambda_i \lambda'_k)}_{\stackrel{\text{def}}{=} \lambda''_l} \underbrace{(g_i * g'_k)}_{\stackrel{\text{def}}{=} g''_l}$$

$$\text{So the change of variable } (i, k) \mapsto l \text{ gives } = \sum_{l=1}^{n \times n'} \lambda''_l g''_l.$$

And this proves that $V_{g''}$ has finite dimension, as this family $(g''_l)_{1 \leq l \leq n \times n'}$ is spanning $r_R g''$ and so $V_{g''}$. So g'' is steerable, of order $n'' \leq n \times n'$. Hence G'' is also steerable, and we have $n_{G \circ G'} \leq n_G \times n_{G'}$, giving (15). \square

⁵¹ In other words, the steerability property does not care about the non-zero constants.

⁵² The last sum has to be understood as a Minkowsky sum between two linear sub-spaces: $A + B \stackrel{\text{def}}{=} \{a + b, a \in A, b \in B\}$.

Remark 4.17 (Special cases?). This formula (15) can be quickly checked:

- If $G' = \mathbf{0}_E$, $n_G = 0$ and $G'' = \mathbf{0}_E$, so $n_G = 0$ is absorbing in (15).
- If $G' = \text{Id}_E$, $n_G = 1$ and $G'' = G'$, so $n_G = 1$ is neutral in (15).

Remark 4.18 (Optimal inequality). One could wonder if this inequality $n_{G_1 \circ G_2} \leq n_{G_1} \times n_{G_2}$ (15) can be improved as it is sometimes an equality (e.g., $G_1 = \text{Id}$). In fact, we cannot improve it to $n_{G_1 \circ G_2} = n_{G_1} \times n_{G_2}$, as for instance if $G_1 = G_2 = D_x$, then $n_{G_1 \circ G_2} = n_{D_x^2} = 3 < n_{G_1} \times n_{G_2} = 2 \times 2 = 4$.

Remark 4.19 (Easy to check in 2D). We will have a strong characterization for steerable convolutions, valid only in 2D (theorem 4.31), and one of its consequences is the fact that steerability is preserved by composition for 2D steerable convolutions, as given by proposition 4.16.

An obvious corollary is the following, stating that composing by a **RI** operator preserves steerability and gives the same steerability order.

Corollary 4.20. *If G_1 is rotation-invariant and G_2 steerable, then $G = G_1 \circ G_2$ is also steerable and $n_G \leq n_{G_2}$, with an equality iff $G_1 \neq \mathbf{0}_{\mathcal{L}(E)}$.*

Proof. We reprove it differently here, to see when $n_G \leq n_{G_2}$ can be strict.

We can work with their Fourier multipliers, $\hat{g} = \hat{g}_1 \cdot \hat{g}_2$, and for $R \in \text{SO}(d)$, we have $r_R \hat{g} = (r_R \hat{g}_1) \cdot (r_R \hat{g}_2)$ (point-wise multiplication), thanks to lemma 2.43. So if G_1 is **RI**, $r_R \hat{g}_1 = \hat{g}_1$ (theorem 3.39), and so:

$$\begin{aligned} V_{\hat{g}} &= \text{Span}\{r_R \hat{g}, R \in \text{SO}(d)\} \\ &= \text{Span}\{(r_R \hat{g}_1) \cdot (r_R \hat{g}_2), R \in \text{SO}(d)\} \\ &= \text{Span}\{\hat{g}_1 \cdot (r_R \hat{g}_2), R \in \text{SO}(d)\} \\ &= \{\hat{g}_1 \cdot h, h \in \text{Span}\{r_R \hat{g}_2, R \in \text{SO}(d)\}\} \\ &= \hat{g}_1 \cdot V_{\hat{g}_2}, \end{aligned}$$

(with a slight abuse of notation at the last line.) But it shows that $\dim_{\mathbb{R}} V_{\hat{g}} \leq \dim_{\mathbb{R}} V_{\hat{g}_2}$, and it is an equality iff multiplying by \hat{g}_1 does not reduce the dimension of $V_{\hat{g}_2}$. \square

In this first subsection about steerability, we defined steerable functions and steerable operators, in any dimension $d \in \mathbb{N}$, and we gave some general results about steerability (e.g., $n_f = 1$ iff f is **RI**) and the class of steerable convolution operators.

In the rest of this section, we study some examples and special cases of steerable operators in any dimension ($n_G = 0, 1$), and the links between steerability and the other geometric invariance properties. We then focus on 2D steerable convolutions, for which the Laplacians and directional derivatives are important construction blocks.

4.2 Special cases and examples of steerable operators

In this short subsection we show two very basic cases, for operators with a steerability order of 0 or 1, and then we study some simple examples of steerable operators.

4.2.1 Steerability of order 0 or 1 ?

Proposition 4.21. *G is the null operator (from E to F) iff $n_g = 0$.*

Easy Proof. $G = \mathbf{0}_{E \rightarrow F}$ iff $g = \mathbf{0}_E$ and all its rotated versions are $\mathbf{0}_E$ too iff $V_g = \{\mathbf{0}_E\}$ (by definition) iff $n_g = 0$ ⁵³. \square

Proposition 4.22. *If G is a non-zero LC operator, G is RI iff G is steerable (St) with $n_g = 1$.*

Easy Proof. G is steerable iff g is steerable, and so the result obtained for distributions (proposition 4.4) directly gives: G is RI iff g is RI iff g is St with $n_g = 1$. \square

Note that we did not find any similar results for higher values $n_g \geq 2$.

4.2.2 Some simple examples of steerable operators

In dimension 2, we use the polar notation $\omega = (r, \theta)$ (for $\omega \in \mathbb{R}^2$, $r \in \mathbb{R}_+$, $\theta \in [0, 2\pi)$ or $\theta \in \mathbb{R}$) almost everywhere, from now on.

The generalized Riesz transform is steerable

Example 4.23. For any $k \in \mathbb{Z}$, the Fourier multiplier \hat{g} in $\mathbf{L}^2(\mathbb{R}^2, \mathbb{C})$ defined by $\hat{g}(\omega) = \hat{g}(r, \theta) = e^{jk\theta}$ is steerable with $n_{\hat{g}} = 2$. It corresponds to a generalization of the (complex) Riesz transform.

Indeed, $R_{\theta_0}\{\hat{g}\}(\omega) = \hat{g}(R_{\theta_0}\omega) = \hat{g}(r, \theta + \theta_0) = \cos(k\theta_0)e^{jk\theta} + j \sin(k\theta_0)e^{jk\theta}$ and if they are seen as functions of (r, θ) , they are in $\text{Span}_{\mathbb{R}}\{e^{jk\theta}, e^{j(k\theta + \pi/2)}\}$, so $V_{\hat{g}}$ is finite dimensional, and has dimension $n_{\hat{g}} = n_G = 2$.

Note that none of the two functions spanning $V_{\hat{g}}$ are impulse responses of real operators, e.g., $(r, \theta) \mapsto e^{j(k\theta + \pi/2)} = je^{jk\theta}$ does **not** correspond to a real operator, as it is not Hermitian symmetric. And for $k \in \mathbb{N}^*$, it gives the (complex) Riesz transform \mathcal{R} for $k = 1$ and its iterate \mathcal{R}^k . For more details on the complex Riesz transform, see [UC13, WCU13].

This example can be generalized into the following lemma, giving a simple but expressive elementary block for steerable functions of $\mathbf{L}^2(\mathbb{R}^2, \mathbb{C})$, with a free radial part $\rho(r)$ (in $\mathbf{L}^2(\mathbb{R}_+)$), that will be used for the representer theorems below.

Lemma 4.24. *For any $k \in \mathbb{Z}$ and $\rho \in \mathbf{L}^2(\mathbb{R}_+)$ non-zero, the function u , defined by $u(r, \theta) = \rho(r)e^{jk\theta}$, is steerable, with $n_u = 2$.*

This result can be seen as a direct consequence of the more general corollary 4.20 proved above, but we also prove it manually to explain properly this idea of separating the radial and polar dependencies, in 2D.

⁵³ Remember that the only linear sub-space of dimension 0 is the null singleton – and not any singleton, as a linear sub-space has to contain the null vector!

Easy Proof. The previous example already proved that the polar part $v(r, \theta) = e^{jk\theta}$ is steerable, so let a basis v_1, v_2 of V_v . Let $R \in \text{SO}(d)$, and so there exists $\lambda_1, \lambda_2 \in \mathbb{R}$ such that $r_{R\theta_0}\{v\}(r, \theta) = e^{jk(\theta+\theta_0)} = \lambda_1 v_1(\theta) + \lambda_2 v_2(\theta)$.

The radial dependency does not change anything for steerability, as $r_{R\theta_0}\{u\}(r, \theta) = u(r, \theta + \theta_0) = \rho_k(r)e^{jk(\theta+\theta_0)}$, and so $r_{R\theta_0}\{u\}(r, \theta) = \rho_k(r)(\lambda_1 v_1(\theta) + \lambda_2 v_2(\theta)) = \lambda_1 u_1(r, \theta) + \lambda_2 u_2(r, \theta)$, proving that u is also steerable with $n_u \leq 2$ (as u_1, u_2 is spanning V_u).

For non-zero radial profile, $u_1(r, \cdot), u_2(r, \cdot)$ is free for some $r > 0$, so u_1, u_2 is free. \square

Real Riesz transform and iterated real Riesz transforms are steerable

Example 4.25. For any $k \in \mathbb{N}$, the Fourier multiplier \hat{g} in $\mathbf{L}^2(\mathbb{R}^2, \mathbb{C})$ defined by $\hat{g}(r, \theta) = j \cos(k\theta)$ yields a 2D steerable convolution.

Indeed, $R_{\theta_0}\{\hat{g}\}(\omega) = \hat{g}(R_{\theta_0}\omega) = \hat{g}(r, \theta + \theta_0) = \cos(\theta_0) \cos(k\theta) + j \sin(\theta_0) \cos(k\theta) \in \text{Span}_{\mathbb{R}}\{(r, \theta) \mapsto e^{j\beta\theta}, (r, \theta) \mapsto e^{j(\beta\theta+\pi/2)}\}$, so $V_{\hat{g}}$ is finite dimensional, and has dimension $n_{\hat{g}} = n_G = 2$. $\hat{g}(\omega) = \hat{g}(r, \theta) = j \cos(k\theta)$ is steerable with $n_{\hat{g}} = 2$. It corresponds to the iterated real Riesz transform. They are real operators as this time their Fourier multipliers are clearly Hermitian symmetric.

Directional derivatives are steerable

In dimension 2, we can easily study the steerability of directional derivatives, but it will be harder in higher dimensions (as their Fourier multiplier is not easy to compute).

Proposition 4.26. *In dimension $d = 2$, partial derivatives and directional derivatives are steerable, of steerability order $n_{D_\alpha} = 2$ (for any $\alpha \in \mathbb{R}$), and we have a simple formula for the rotation of \widehat{D}_α :*

$$\forall \theta_0 \in \mathbb{R}, \quad r_{R\theta_0}\{\widehat{D}_\alpha\} = \widehat{D}_{\alpha-\theta_0}. \quad (16)$$

Easy Proof. As a partial derivative D_i is nothing but a special case of directional derivative (with a unit vector $\vec{u} = \vec{u}_\alpha \in \mathbb{S}^1$ of angle $\alpha \in \{0, \frac{\pi}{2}\}$), we can study D_α as a general case.

D_α is steerable iff its Fourier multiplier is steerable (proposition 4.10). And we know that $\widehat{D}_\alpha(\omega) = \widehat{D}_\alpha(r, \theta) = jr \cos(\theta - \alpha)$ (see examples 3.22), so

$$\begin{aligned} \forall \omega = (r, \theta) \in \mathbb{R}^2, \quad r_{R\theta_0}\{\widehat{D}_\alpha\}(\omega) &= \widehat{D}_\alpha(R_{\theta_0}\omega) \\ (\text{In polar coordinates,}) &= \widehat{D}_\alpha(R_{\theta_0}(r, \theta)) \\ &= \widehat{D}_\alpha(r, \theta + \theta_0) = jr \cos((\theta + \theta_0) - \alpha) \\ &= jr \cos(\theta - (\alpha - \theta_0)) = \widehat{D}_{\alpha-\theta_0}(\omega). \end{aligned}$$

This identity on $r_{R\theta_0}\{\widehat{D}_\alpha\}$ is enough to have the steerability of \widehat{D}_α , as

$$\begin{aligned} V_{\widehat{D}_\alpha} &= \text{Span}_{\mathbb{R}}\{r_R\{\widehat{D}_\alpha\} : R \in \text{SO}(d)\} = \text{Span}_{\mathbb{R}}\{r_{R\theta_0}\{\widehat{D}_\alpha\} : \theta_0 \in \mathbb{R}\} \\ &= \text{Span}_{\mathbb{R}}\{\widehat{D}_{\alpha-\theta_0} : \theta_0 \in \mathbb{R}\} = \text{Span}_{\mathbb{R}}\{\widehat{D}_{\theta_0} : \theta_0 \in \mathbb{R}\} \\ &= \text{Span}_{\mathbb{R}}\{\widehat{D}_x, \widehat{D}_y\}. \end{aligned}$$

Therefore $V_{\widehat{D}_\alpha}$ has finite dimension, equals to $\dim_{\mathbb{R}} \text{Span}_{\mathbb{R}}\{\widehat{D}_x, \widehat{D}_y\} \leq 2$.

And we know that \widehat{D}_x and \widehat{D}_y are free, as functions on \mathbb{R}^2 , because $\widehat{D}_x : \omega \mapsto jw_1$ and $\widehat{D}_y : \omega \mapsto jw_2$, so $V_{\widehat{D}_\alpha}$ has dimension exactly 2. \square

Remark 4.27 (And for higher dimensions?). In higher dimension ($d > 2$), $\widehat{D_{\vec{u}}}$ is not as simple as in 2D, but we also have that $r_R \widehat{D_{\vec{u}}} = \widehat{D_{R^{-1}\vec{u}}}$, which spans exactly $\widehat{D_1}, \dots, \widehat{D_d}$ (for $R \in \text{SO}(d)$). So $\widehat{D_{\vec{u}}}$ is also steerable, of order d .

Fractional Laplacians are steerable

Laplacians and fractional Laplacians are rotation-invariant, as proved above in proposition 3.30, and so they are (trivially) steerable, of order $n_{\Delta} = 1$ and $n_{(-\Delta)^{\gamma/2}} = 1$.

4.3 2D steerable convolution operators

We now focus on steerable convolution operators in 2D, as our motivation mainly lies in possible applications to image analysis and stochastic processes for images. This subsection starts by giving two versions of a representer theorem for rotation-invariant subspaces of Fourier multiplier, and then applies them to prove a first characterization for steerable convolution, that can be written as a sum of (complex) Riesz transform, modulated by complex-valued radial factors (lemma 4.24). Special cases of this form are then studied, and the next subsection will give a simpler form for γ -SI operators.

We **only** consider operators on $E = \mathbf{L}^2(\mathbb{R}^2, \mathbb{R})$ from now on. The first theorem 4.31 is a very powerful characterization of steerable convolutions.

4.3.1 Representer theorems for rotation-invariance sub-spaces

The proof of theorem 4.31 is not too hard, providing that we have the following intermediate result. The first theorem comes from [VSU15, Th.2.3], and the next one is a simple rewriting, adapted to our setting. We refer to them as “representer” theorems because they exhibit the form of the elements of a basis, like usual representer theorems⁵⁴ do.

Theorem 4.28 (Representer theorem for rotation-invariance sub-spaces).

A q -dimensional linear sub-space of $\mathbf{L}^2(\mathbb{R}^2, \mathbb{C})$ is stable by rotation iff it is spanned by q functions $\phi_1(r, \theta), \dots, \phi_q(r, \theta)$ of $\mathbf{L}^2(\mathbb{R}^2, \mathbb{C})$, that have the following separable form in polar coordinates:

$$\phi_n(\omega) = \phi_n(r, \theta) = \rho_n(r) e^{jk_n\theta} \quad \text{where } k_n \in \mathbb{Z} \text{ (all different)}. \quad (17)$$

The radial dependencies $\rho_n(r)$ can be complex-valued, and are in $\mathbf{L}^2(\mathbb{R}_+, \mathbb{C})$.

Remark 4.29 (Short chronology). In [VSU15, Sec.2.2], the authors evoke a proof of a more general result based on Lie groups (cf. [HOT98]), but we chose to not spend any extra time studying these more complicated results (as it would have required some additional time to understand the underlying theory), and we checked their “simple” proof, based on Fourier series (see below, (21)). The converse sense was known from 1991, and the general result for the direct sense seems to be known from 1998, but the particular result in this setting only dates from last year (2015, [VSU15]).

⁵⁴ For instance, see this Wikipedia article for the representer theorem in statistical learning and in RKHS.

Long and Hard Proof. • Converse sense: The sufficient part is easy to check. If these q functions spans the sub-space A , then it is easy to check that it is stable by rotation: $r_{R_\alpha}\{\phi_n\}(r, \theta) = \phi_n(r, \theta + \alpha) = e^{jk_n\alpha}\phi_n(r, \theta) \in \mathbb{C}\phi_n(r, \theta) \in \text{Span}_{\mathbb{C}}\{\phi_1(r, \theta), \dots, \phi_q(r, \theta)\} = A$, for each $n \in \{1, \dots, q\}$, so $r_{R_\alpha}A \subset A$ as expected.

- Direct sense: We adapt here the proof given in [VSU15, Appendix A].

Let A be this subspace, and ϕ_1, \dots, ϕ_q a basis of A . Because we are in $\mathbf{L}^2(\mathbb{R}^d, \mathbb{C})$, an Hermitian space, we can ask this basis to be orthonormal (by using the Gram-Schmidt process if needed). Let ϕ be a vector-valued function defined by $\phi(r, \theta) = [\phi_1(r, \theta), \dots, \phi_q(r, \theta)]$ for $r, \theta \in \mathbb{R}^2$ ($\phi: \mathbb{R}^2 \rightarrow \mathbb{C}^q$). A is stable by rotation *iff* for all $\theta \in \mathbb{R}$, $r_{R_\theta}\{\phi\}$ has all its entries is in the span of ϕ *iff* there exists a $(q \times q)$ matrix-valued function \mathbf{c} of θ such that

$$\forall \theta \in \mathbb{R}, \quad r_{R_\theta}\{\phi\} = \mathbf{c}(\theta)\phi. \quad (18)$$

$\mathbf{c}(\theta)\phi$ is a matrix-vector product, giving a function of $\mathbb{R}^2 \rightarrow \mathbb{C}^q$. This matrix $\mathbf{c}(\theta) \in \mathcal{M}_{q \times q}(\mathbb{C})$ is uniquely defined because ϕ is a basis, and its entries are given by

$$\forall \theta \in \mathbb{R}, \quad \mathbf{c}_{i_1, i_2}(\theta) = \langle r_{R_\theta}\{\phi_{i_1}\}, \phi_{i_2} \rangle. \quad (19)$$

And now for $\theta_1, \theta_2 \in \mathbb{R}$, $r_{R_{\theta_1 + \theta_2}} = r_{R_{\theta_1}} \circ r_{R_{\theta_2}}$ and the linearity of $r_{R_{\theta_1}}$ gives

$$\mathbf{c}(\theta_1 + \theta_2) = \mathbf{c}(\theta_1)\mathbf{c}(\theta_2). \quad (20)$$

The form of the entries of $\mathbf{c}(\theta)$ (19) implies that they are 2π -periodic, and bounded, thanks to Cauchy-Schwarz inequality, both $r_{R_\theta}\{\phi_{i_1}\}$ and ϕ_{i_2} being in $\mathbf{L}^2(\mathbb{R}^2, \mathbb{C})$. Therefore, the Fourier coefficients of \mathbf{c} belong to $\mathbf{L}^2(\mathbb{T})$, \mathbb{T} being the 2π -periodic torus ($\mathbb{R}/2\pi\mathbb{Z}$), and so we can expand $\mathbf{c}(\theta)$ as a Fourier series, with $q \times q$ matrix-valued Fourier coefficients \mathbf{c}_k (for $k \in \mathbb{Z}$), given by

$$\forall k \in \mathbb{Z}, \quad \mathbf{c}_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{c}(\theta) e^{-jk\theta} d\theta. \quad (21)$$

Using (20), we obtain the following identities (seen as matrix product),

$$\begin{cases} \mathbf{c}_k^2 = \mathbf{c}_k & \text{for every } k \in \mathbb{Z}, \\ \mathbf{c}_{k_1} \mathbf{c}_{k_2} = \mathbf{0} & \text{for every } k_1 \neq k_2 \in \mathbb{Z}. \end{cases}$$

Their proof can be found in the appendix B.3. These identities imply that the (matrix-valued) Fourier coefficients \mathbf{c}_k are mutually orthogonal projection matrices. But they are $q \times q$ matrices, so there can be at most q orthonormal projection directions \mathbf{u}_n ($n = 1, \dots, p$ with $p \leq q$), and we define k_n the index of the corresponding projection matrices ($k_n \in \mathbb{Z}$, all distinct).

The matrix of projection of direction \mathbf{u}_n will in fact be written⁵⁵ $\mathbf{c}_{k_n} = \mathbf{u}_n \cdot \mathbf{u}_n^H$, and so the Fourier series of $\mathbf{c}(\theta)$ reduces to a finite number of terms (at most q), of the form $\mathbf{c}_{k_n} e^{jk_n\theta}$, *i.e.*, $(\mathbf{u}_n \cdot \mathbf{u}_n^H) e^{jk_n\theta}$.

⁵⁵ \mathbf{u}_n is a complex-valued vector of dimension q , and so the *outer product* $\mathbf{u}_n \cdot \mathbf{u}_n^H$ is a complex-valued matrix of dimension $q \times q$, as \mathbf{c}_{k_n} should be.

And (18) implies that $\phi(r, \theta) = \mathbf{c}(\theta)\phi(r, 0)$. This was enough to prove the separability between r and θ , but we need the form on \mathbf{c}_{k_n} to be able to define $\rho_n(r)$ as $\mathbf{u}_n^H \phi_n(r, 0)$, which proves that

$$\forall r, \theta, \quad \phi(r, \theta) = \sum_{n=1}^q \mathbf{u}_n \rho_n(r) e^{jk_n \theta}.$$

And we can then conclude that the components of ϕ , ϕ_1, \dots, ϕ_q , all depend linearly on functions of the form (17). \square

As we will see it and use it often later, Hermitian symmetry implies some symmetry conditions on the indexes k_n , and some complex symmetric on the radial profiles.

Theorem 4.30 (Repres. theorem for rotation-invariance Hermitian-sym. sub-space). *A q -dimensional linear sub-space of $\mathbf{L}^2(\mathbb{R}^2, \mathbb{C})$, of Hermitian-symmetric Fourier multipliers, is stable by rotation iff it is spanned by q functions $\phi_1(r, \theta), \dots, \phi_q(r, \theta)$ that have the following separable form in polar coordinates:*

$$\phi_n(\omega) = \phi_n(r, \theta) = \rho_n(r) e^{jk_n \theta} \quad \text{where } k_n \in \mathbb{Z}. \quad (22)$$

and $k_n \in \mathbb{Z}$ is present iff $-k_n \in \mathbb{Z}$ is present (for $k_{n'} = -k_n$ for some n'). The radial dependencies $\rho_n(r)$ are in $\mathbf{L}^2(\mathbb{R}_+, \mathbb{R})$, can be complex-valued, but they satisfies $\rho_{n'}(r) = (-1)^{k_n} \overline{\rho_n(r)}$, and in the linear combinations the coefficients has to satisfy $a_{n'} = (-1)^{k_n} \overline{a_n}$.

The Hermitian symmetry conditions are really not so clear here, and cannot be simplified, but they will be easier to understand for the characterization of steerable operators, see below proposition 4.35.

4.3.2 Sum decomposition for steerable convolutions

The first general characterization for steerable convolutions on $\mathbf{L}^2(\mathbb{R}^2, \mathbb{R})$ is the following. It will be a direct application of the previous theorem, with $q = n_G$ functions, and $N = \max_{\mathbb{N}}\{|k_n|, 1 \leq n \leq p\} = K_G$. We first consider complex steerable convolutions, and then use the Hermitian symmetry to give a characterization for real operators. This first characterization of *complex LC TI St* operators will write their Fourier multiplier as a **sum** of Riesz transforms, each term $e^{jk\theta}$ being modulated by a radial factor $\rho_k(r)$.

Theorem 4.31 (Characterization of 2D complex LC TI St operators).

G is a non-zero LC TI St operator on $\mathbf{L}^2(\mathbb{R}^2, \mathbb{C})$ iff there exists $K_G \in \mathbb{N}$, and $(\rho_k)_{-K_G \leq k \leq K_G} : \mathbb{R}_+ \rightarrow \mathbb{C}$ such that

$$\forall \omega = (r, \theta) \in \mathbb{R}^2, \quad \hat{g}(r, \theta) = \sum_{-K_G \leq k \leq K_G} \rho_k(r) e^{jk\theta}. \quad (\heartsuit)$$

Long and Hard Proof. • Converse sense: The sufficient sense is easy to check, as each function $(r, \theta) \mapsto \rho_k(r) e^{jk\theta}$ is LC TI and steerable (lemma 4.24), and a complex linear combination (that conserves Hermitian symmetry) conserves steerability. And the existence of the Fourier multiplier (implicit on the right side) ensures that G is LC TI, so it is indeed LC TI St on $\mathbf{L}^2(\mathbb{R}^2)$.

- Direct sense: The necessary sense is a consequence of the previous theorem 4.28. It is only valid for sub-spaces of $\mathbf{L}^2(\mathbb{R}^2, \mathbb{R})$, and as we said in proposition 3.12, in the general setting we do not have \hat{g} in $\mathbf{L}^2(\mathbb{R}^2, \mathbb{R})$.

However, by hypothesis⁵⁶ (cf. proposition 3.12), \hat{g} will always have a certain “order of decay” $\kappa \in \mathbb{R}$, so that $(1 + r^2)^\kappa \hat{g}(r, \theta)$ is in $\mathbf{L}^2(\mathbb{R}^2, \mathbb{R})$ (cf. proposition 3.12). And so, up to a purely-radial multiplicative term that will be included in all the radial functions $\rho_k(r)$, we can safely assume that \hat{g} is in $\mathbf{L}^2(\mathbb{R}^2)$.

The proof is simpler to understand if we do not keep this additional term $(1 + r^2)^\kappa$. Because we might modify the radial functions $\rho_k(r)$, we will not be able to affirm that they are in \mathbf{L}^2 , and that is why this theorem kept them unspecified⁵⁷ (in “ $\mathbb{R}_+ \rightarrow \mathbb{R}$ ”).

If G is steerable, by definition $n_G = n_{\hat{g}} < +\infty$, and so we can apply the theorem to the n_G -dimensional subspace of $\mathbf{L}^2(\mathbb{R}^2, \mathbb{R})$, $V_{\hat{g}}$, which is obviously stable by rotations (by definition, $V_{\hat{g}}$ contains all the rotated versions of \hat{g} , $r_R\{\hat{g}\}$). $V_{\hat{g}}$ is indeed in $\mathbf{L}^2(\mathbb{R}^2, \mathbb{R})$ thanks to the trick we exposed above (maybe it is $V_{(r, \theta) \mapsto \hat{g}(r, \theta)(1+r^2)^\kappa}$ instead).

Therefore, there is n_G functions $\phi_n(r, \theta)$ spanning $V_{\hat{g}}$, that are in $\mathbf{L}^2(\mathbb{R}^2)$, and have the form $\phi_n(r, \theta) = \rho_n(r)e^{jk_n\theta}$ for $k_n \in \mathbb{Z}$ (with the radial component $\rho_n(r)$ in $\mathbf{L}^2(\mathbb{R}_+)$). This family is spanning $V_{\hat{g}}$, therefore any $u \in V_{\hat{g}}$ can be written as $u = \sum_{n=1}^{n_G} a_n \rho_n(r)e^{jk_n\theta}$ for $(a_n)_{1 \leq n \leq n_G} \in \mathbb{C}^{n_G}$.

Instead of parameterizing the $k_n \in \mathbb{Z}$, we can introduce more coefficients a_k for $k \in \mathbb{Z}$ and simply chose them equals to 0 if k is not one of the $(k_n)_{1 \leq n \leq n_G}$ (and we also define $\phi_k(r) = 1$ in this case). If we define $K_G = \max_{\mathbb{N}}\{|k_n|, 1 \leq n \leq p\}$, any $u \in V_{\hat{g}}$ can be also written as $u = \sum_{-K_G \leq k \leq K_G} a_k \phi_k(r)e^{jk\theta}$. This second form has the advantage of being parametric in only one integer $K_G \in \mathbb{N}$, coefficients a_k and the radial functions ϕ_k .

Finally, as $\hat{g} \in V_{\hat{g}}$, it can be written as wanted

$$\hat{g}(r, \theta) = \sum_{-K_G \leq k \leq K_G} a_k \rho_k(r)e^{jk\theta}.$$

We simply rename $\rho_k(r) := a_k \rho_k(r)$ to give it the form of the theorem. □

We can give some examples of **LC TI** steerable operators written in this form (♡).

Examples 4.32. • If $K_G = 0$, then $\hat{g}(r, \theta) = \rho_0(r)$ is purely radial (but with an ad-hock radial profile), hence it is rotation-invariant ($n_G = 1$) but not necessarily scale-invariant.

- For example, a fractional Laplacian $G = (-\Delta)^{\gamma/2}$ has the wanted form, with $S_G = \{0\}$ (so $K_G = 0$ and $n_G = 1$, cf. proposition 4.33) and $\rho_0(r) = r^\gamma$.

⁵⁶ This could be justified better by considering weighted Sobolev spaces, but we preferred to had the existence of the order of decay κ in the hypothesis on the “usually encountered” \hat{g} .

⁵⁷ But as soon as we work with γ -scale-invariant operators, they will have $\rho_k(r) = r^\gamma$, and $\kappa = -(\gamma/2+1)$ will be enough to have $(1 + r^2)^\kappa \rho_k(r)$ in $\mathbf{L}^2(\mathbb{R}_+, \mathbb{R})$.

- A directional derivative D_α (for $\alpha \in \mathbb{R}$) can be written under this form: $\widehat{D}_\alpha(r, \theta) = jr \cos(\theta - \alpha) = jr (\cos(\alpha) \cos(\theta) + \sin(\alpha) \sin(\theta))$.

And Euler's identity gives $j \cos(\theta) = \frac{j}{2} (e^{j\theta} + e^{-j\theta})$ and $j \sin(\theta) = \frac{j}{2} (e^{j\theta} - e^{-j\theta})$ so $\widehat{D}_\alpha(r, \theta) = a_1 r e^{j\theta} + a_{-1} r e^{-j\theta}$ is indeed in the form of the theorem, with $a_1 = \frac{j}{2} \cos(\alpha) - \frac{1}{2} \sin(\alpha) \in \mathbb{C}$ and $a_{-1} = \frac{j}{2} \cos(\alpha) + \frac{1}{2} \sin(\alpha) \in \mathbb{C}$. It has $S_G = \{-1, +1\}$, $K_G = 1$ and $n_G = 2 = K_G + 1$ (as proved above in proposition 4.42). We observe that $a_{-1} r = -\overline{a_1} r = -\overline{a_1} r = (-1)^1 \overline{a_1} r$, which will be proved in proposition 4.35.

- Another example, for $G = D_x D_y$, $\widehat{g}(\omega) = (j\omega_1)(j\omega_2) = -r^2 \cos(\theta) \sin(\theta)$ so $\widehat{g}(\omega) = -r^2 \frac{1}{2} \sin(2\theta) = r^2 \left(\frac{j}{4} e^{j2\theta} - \frac{j}{4} e^{-j2\theta} \right) = a_2 r^2 e^{j2\theta} - a_{-2} r^2 e^{-j2\theta}$. It has $S_G = \{-2, +2\}$, $K_G = 2$ and $n_G = 3 = K_G + 1$ (as proved below in proposition 4.42). This time, we observe that $a_{-2} r^2 = \overline{a_2} r^2 = \overline{a_2} r^2 = (-1)^2 \overline{a_2} r^2$.

This form (♥) is very general, as nothing is known about the radial functions ρ_k , except that they have to be measurable (as \widehat{g}). The polar components (*i.e.*, dependency in θ) are a weighted sum of iterated (complex) Riesz transforms (Fourier multiplier $e^{jk\theta}$, $k > 0$) and iterated anti-Riesz transforms ($e^{-jk\theta}$, $k > 0$), but we will rewrite it in terms of real Riesz transforms later on.

Before that, it is interesting to note that we can bound the steerability order of a steerable operator, if it is written in this form.

Proposition 4.33. *For a non-zero 2D steerable γ -SI convolution operator G written in this form (♥), there is a link between its steerability order n_G and K_G the higher "frequency" intervening in $\widehat{g}(r, \theta)$: $n_G \leq K_G + 1$. We recall that K_G is*

$$K_G = \max |S_G| = \max_{\mathbb{N}} \{|k|, k \in \mathbb{Z}, a_k \neq 0\} = \text{"max}\{|k|, k \in \mathbb{Z}, e^{jk\theta} \text{ appears in } \widehat{g}(r, \theta)\}\text{"}$$

Easy Proof. Note that for $G = \mathbf{0}_E$, $n_G = 0$ but $K_G = -\infty$ (as a maximum of an empty set, by convention), thus the formula can only be valid for non-zero operators.

In fact, we do not need to prove it now, because for *non-zero* 2D steerable convolution, adding the scale-invariant will give the decomposition (34), and so proposition 4.42 directly gives this result on n_G and K_G . \square

Remark 4.34 (Can we have an equality?). A simple counter-example can be $\widehat{g}(r, \theta) = \cos(N\theta)$ for $N \in \mathbb{N}$, $N \geq 4$. It has $K_G = N \geq 4$, but $r_{R_\alpha} \{\widehat{g}\}(r, \theta) = \cos(N(\theta + \alpha)) = \cos(N\theta) \cos(N\alpha) - \sin(N\theta) \sin(N\alpha)$. So $V_{\widehat{g}} \subset \text{Span}_{\mathbb{R}} \{ \cos(N\theta), \sin(N\theta) \}$ (as functions of (r, θ)). So \widehat{g} and G are indeed steerable with $n_G \leq 2$ (obvious from theorem 4.31). In fact $n_G = 2$, as G is not **RI**. So this G has $n_G = 2$ and $K_G = N \geq 4$, *i.e.*, $n_G < K_G$ and so the inequality $n_G \leq K_G + 1$ cannot be improved.

As said above, the Hermitian symmetry of its Fourier multiplier can be characterized easily on the (complex) radial part $(\rho_k(r))_{-K_G \leq k \leq K_G}$, for a steerable operator written on the form of the first theorem. We have eliminated the indexes k_n and simply used $k \in \{-K_G \leq k \leq K_G\}$, so this result is much simpler to write.

Proposition 4.35 (Characterization of real **LC TI St** operators). *For a steerable operator G written in this form (♡), G maps real-valued functions to real-valued functions (i.e., G is on $L(\mathbb{R}^2, \mathbb{R})$) iff*

$$\forall r \in \mathbb{R}_+, \forall k \in \mathbb{Z}, \rho_{-k}(r) = (-1)^k \overline{\rho_k(r)}. \quad (23)$$

Computational Proof. Thanks to theorem 3.17, G is a real operator iff \hat{g} is Hermitian symmetric, so when \hat{g} is written in this form (♡), \hat{g} is Hermitian symmetric

$$\begin{aligned} \Leftrightarrow \forall r, \theta, \hat{g}(r, \theta + \pi) &= \overline{\hat{g}(r, \theta)} \\ \Leftrightarrow \forall r, \theta, \sum_{k \in \mathbb{Z}} \rho_k(r) (-1)^k e^{jk\theta} &= \sum_{k \in \mathbb{Z}} \overline{\rho_k(r)} e^{-jk\theta} \end{aligned}$$

We prefer to write the sum for $k \in \mathbb{Z}$ and keep in mind that it is always a finite sum on $k \in S_G$, as we define $a_k = 0$ if $k \notin S_G$

$$\Leftrightarrow \forall r, \theta, \sum_{k \in \mathbb{Z}} \left(\rho_k(r) (-1)^k - \overline{\rho_{-k}(r)} \right) e^{jk\theta} = 0$$

But the functions $\theta \mapsto e^{jk\theta}$ are all linearly independent (for $k \in \mathbb{Z}$), so having this equality for all values of θ forces all the coefficients to be zeros (and it is an equivalence as the converse is true)

$$\Leftrightarrow \forall r, \forall k \in \mathbb{Z}, \rho_k(r) (-1)^k = \overline{\rho_{-k}(r)}. \quad \square$$

4.3.3 A simpler form for this characterization of steerability?

Remark 4.36. In particular, if the radial dependencies are real-only, the condition $\rho_{-k}(r) = \overline{\rho_k(r)}$ simply means $\rho_{-k}(r) = \rho_k(r)$, and so each term for k and $-k$ in the sum can be simplified into $\rho_k(r) (e^{jk\theta} + (-1)^k e^{-jk\theta})$. This will be the beginning of the proof of the simpler characterization for **LC TI St** operators that are also γ -**SI**.

Common radial part?

This simpler form could be interesting to study, but we obviously lose in generality:

$$\forall \omega = (r, \theta) \in \mathbb{R}^2, \hat{g}(r, \theta) = \rho(r) \sum_{-K_G \leq k \leq K_G} a_k e^{jk\theta}.$$

A simple counter example can be $\hat{g}(r, \theta) = r + r^2 j \cos(\theta) = r + r^2 j \frac{1}{2} (e^{j\theta} + e^{-j\theta})$, which gives a steerable and real convolution G , but the radial dependency of $\hat{g}(r, \theta)$ cannot be factorized into one function $\rho(r)$.

This form was intermediate between the general form, with (possibly) different $\rho_k(r)$, and the form for γ -**SI** operators, with a shared radial dependency in the form of r^γ .

We tried to find a more subtle characterization for **LC TI St** operators having a *common* radial dependency $\rho(r)$ in the form (♡), but nothing interesting came out of this direction, unfortunately.

4.4 Steerability and other geometric properties

From the beginning, we mainly studied steerable convolutions and not simply steerable operators, i.e., steerable operators that are also linear, continuous and translation-

invariant; so only two⁵⁸ geometric properties are left to be studied: rotation-invariance (**RI**), and γ -scale-invariance (γ -**SI**).

4.4.1 Rotation-invariance for steerable convolution operators

We already observed in proposition 4.22 the link between steerability and rotation-invariance ($n_G = 1$), but a more interesting geometric property to study is the scaling-invariance (or the γ -scale-invariance to not lose in generality).

4.4.2 γ -scale-invariance for 2D steerable convolution operators

Quite surprisingly, it is easy to obtain a much stronger characterization for **LC TI St** operators that are also γ -**SI**. There is a sufficient and necessary condition on *all* the radial functions ρ_k appearing in the characterization (\heartsuit), which all have to be γ -**SI** (i.e., being proportional to $r \mapsto r^\gamma$, thanks to lemma 2.36). The proof is quite elegant as it turns out that we do not need to work on each ρ_k separately.

Theorem 4.37 (Characterization of 2D **LC TI γ -SI St** operators). *G is a **LC TI γ -SI St** operator on $\mathbf{L}^2(\mathbb{R}^2, \mathbb{R})$ iff there exists $K_G \in \mathbb{N}$, and $(a_k)_{-K_G \leq k \leq K_G} \in \mathbb{C}$ such that*

$$\forall \omega = (r, \theta) \in \mathbb{R}^2, \quad \widehat{g}(r, \theta) = r^\gamma \sum_{-K_G \leq k \leq K_G} a_k e^{jk\theta}. \quad (24)$$

Computational Proof. • **Direct sense:** If G is **LC TI St**, \widehat{g} has the form of the first theorem. Instead of trying to prove manually that each ϕ_k is γ -**SI**, we simply need to observe that if G is γ -**SI**, then:

$$\forall r > 0, \quad \widehat{g}(r, \theta) = r^\gamma \widehat{g}(1, \theta).$$

And so, by the first theorem (\heartsuit),

$$\forall r > 0, \quad \widehat{g}(r, \theta) = r^\gamma \sum_{-K_G \leq k \leq K_G} \underbrace{\rho_k(1)}_{=a_k} e^{jk\theta}.$$

Thus, defining a_k as $\phi_k(1) \in \mathbb{C}$ directly gives the desired form.

- **Converse sense:** It is almost obvious, as the existence of the Fourier multiplier (implicit on the right side) ensures that G is **LC TI**. A Fourier multiplier of this form is γ -**SI**, Hermitian symmetric, so G is γ -**SI** too and operates on $\mathbf{L}^2(\mathbb{R}^2)$ (i.e., $G\{f\}$ is still real-valued). And the last point to check is the steerability: every term of the sum is steerable and the preservation of the steerability by (finite) sums (proposition 4.14). \square

Remark 4.38. Note that the coefficients a'_k are *not necessarily* the same as the a_k of the first theorem 4.31, but the K_G integer stays the same.

⁵⁸ As we said above, we did not study the unity property as extensively as the others. However, there is two small results on unitary steerable convolutions in the appendix A.4.

In the simplified form (24), we can characterize the Hermitian symmetry of \hat{g} directly on its (complex) coefficients $(a_k)_{-K_G \leq k \leq K_G}$.

Proposition 4.39. *For a γ -SI steerable operator G written in this form (24), G maps real-valued functions to real-valued functions iff $\forall k, a_{-k} = (-1)^k \overline{a_k}$.*

Proof. It is simply the consequence of the same result (proposition 4.35) for steerable operators (non-necessarily γ -SI), applied to the special case of a common radial profile, $\phi_k(r) = a_k r^\gamma, \forall k \in \mathbb{Z}$. \square

Corollary 4.40. *If $S_G \stackrel{\text{def}}{=} \{k \in \mathbb{Z}, a_k \neq 0\}$, then the previous result implies that S_G has to be “even”, i.e., symmetric regarding 0: $\forall k \in \mathbb{Z}, k \in S_G \Leftrightarrow -k \in S_G$.*

For instance, for a single directional derivative, $S_G = \{-1, +1\}$, as seen above.

Factorizing a Laplacian ?

This form $\hat{g}(r, \theta) = r^\gamma \sum_{-K_G \leq k \leq K_G} a_k e^{jk\theta}$ for the Fourier multiplier of a 2D steerable γ -SI convolution can be seen as a special case of theorem 3.36, where we explained that a LC TI γ -SI operator G can be decomposed as $G = (-\Delta)^{\gamma/2} \circ G_0$ for G_0 LC TI 0-SI. This result showed that the γ -SI part can be isolated into a Laplacian, and this is exactly what is done here: G is decomposed as $G = (-\Delta)^{\gamma/2} \circ G_0$ for G_0 LC TI 0-SI, with G_0 of Fourier multiplier $g_0(\omega) = \sum_{-K_G \leq k \leq K_G} a_k e^{jk\theta}$, written in the form (\heartsuit), in the special case of a steerable convolution not depending on $r = |\omega|$ (i.e., 0-SI).

Trigonometric polynomial

We can observe that $g_0(\omega) = g_0(e^{j\theta}) = \sum_{-K_G \leq k \leq K_G} a_k e^{jk\theta}$ is exactly a *trigonometric polynomial*, that is Hermitian symmetric like $g(\omega)$. We will study the roots of this trigonometric polynomial (on the circle), in the last characterization of steerable convolutions (see below in section 4.6).

4.5 Composing fractional Laplacians and directional derivatives

We already proved that any fractional Laplacian and any directional derivative are steerable, and a their composition is also steerable. Moreover, we can say a little bit more:

Lemma 4.41 (Composition of directional derivatives). *Let $G = D_{\alpha_1} \dots D_{\alpha_m}$, for $m \in \mathbb{N}$, then G is steerable, of order $n_G = m + 1$, and of max frequency $K_G = \max |S_G| = m$.*

Proof. Steerability is preserved by composition, and for (partial) derivatives, we can check quite easily that $V_G \subset \text{Span}_{\mathbb{R}} \{D_x^m, D_x^{m-1} D_y, D_x^{m-2} D_y^2, \dots, D_x^2 D_y^{m-2}, D_x D_y^{m-1}, D_y^m\}$. And so by definition G is steerable of order $n_G \leq m + 1$.

And it is an equality, $n_G = m + 1$, as proven in appendix B.1. The proof is long but technical and not so interesting, as it consists mainly in trigonometric manipulations, so we preferred to move it to the appendix.

The result for max frequency is clear when we simply develop $\hat{g}(r, \theta)$. It is written $\prod_{i=1}^m j r \cos(\theta - \alpha_i) = r^m j^m \frac{1}{2^m} \prod_{i=1}^m (e^{j(\theta - \alpha_i)} - e^{-j(\theta - \alpha_i)})$ and so when

we develop the product we get a (Hermitian symmetry) complex linear combination of terms of the form $e^{jk\theta - j(\alpha_{i_1} \pm \dots \pm \alpha_{i_k})}$ for $-m \leq k \leq m$ and $\{i_1, \dots, i_k\} \subset \{1, \dots, m\}$ (of size exactly k , i.e., all i_1, \dots, i_k are different). \square

We can state the following result to sum-up the properties of a composition of a fractional Laplacian and m directional derivatives, and the rest of this subsection will study the converse of this proposition.

Proposition 4.42. *Let $\gamma \in \mathbb{R}_+$, and $m \in \mathbb{N}$, $C \in \mathbb{R}$, and $(\alpha_i)_{1 \leq i \leq m} \in \mathbb{R}^m$, and let the operator $G = C(-\Delta)^{\gamma/2} D_{\alpha_1} \dots D_{\alpha_m}$. Then G is **LC TI**, $(\gamma + m)$ scale-invariant, and steerable, with $K_G = \max |S_G| = m$, and $n_G = m + 1$, on $E = \mathbf{S}(\mathbb{R}^2, \mathbb{R})$.*

Proof. • The linearity, continuity, $\gamma + m$ scale-invariance and steerability comes from the previous theorems 2.39 and 4.14, of stability by compositions and additivity of the orders of scale-invariance.

• And the previous lemma 4.41 gives the announced values for K_G and n_G , as the fractional Laplacian is rotation-invariant (and so composing by it does not change the order of steerability neither the max frequency K_G). \square

Example 4.43 ($D_x + D_y$). For example, the steerable convolution $G = D_x + D_y = D_1 + D_2 = D_{\alpha=0} + D_{\alpha=\pi/2}$ can be written in the form of this proposition, with no Laplacian and just one directional derivative (as $n_G = 2 = K_G - 1$, $m = K_G = 1$ is enough):

$$G = D_x + D_y = \sqrt{2}D_{\pi/4}.$$

Easy Proof. Let $G = D_x + D_y$, of Fourier multiplier $\hat{g}(r, \theta) = jr \cos(\theta) + jr \sin(\theta)$, which can be written as $jr(\cos(\theta) + \sin(\theta)) = jr\sqrt{2} \cos(\theta - \frac{\pi}{4})$, that we recognized as the Fourier multiplier of the fractional derivative of angle $+\frac{\pi}{4}$, so $G = \sqrt{2}D_{\pi/4}$. \square

Remark 4.44 (Only one Laplacian?). In the form given to G of the previous result, one could wonder why there is only one Laplacian $((-\Delta)^{\gamma/2})$ but (possibly) more than one directional derivatives ($m \geq 0$).

- First, with $\gamma = 0$ the Laplacian will be the identity, so this form includes the possibility of not composing a Laplacian.
- And then, thanks to proposition 2.39, and the commutativity of **LC TI** operators (including Laplacians and derivatives), if there is more than one fractional Laplacian, let say $(-\Delta)^{\gamma_1/2} \circ \dots \circ (-\Delta)^{\gamma_{m'}/2}$, it gets reduced into one Laplacian $(-\Delta)^{(\gamma_1 + \dots + \gamma_{m'})/2}$.

Therefore, this form with exactly one fractional Laplacian in the composition is as general as the form with $m' \in \mathbb{N}$ fractional Laplacians.

The same result is clearly wrong for directional derivatives, as two derivatives cannot be factorized into only one (e.g., $D_x D_y$ cannot be written as one D_α).

We tried to get a converse result of the previous proposition, that would be a form of universality of this family composed of Laplacians and directional derivatives, but clearly we had no serious hope for it. The converse result, stated like this, is wrong, as the following example will show.

Example 4.45 (Counter-example). Without a more subtle analysis of the properties of operators on this form $G \stackrel{\text{def}}{=} C(-\Delta)^{\gamma/2} D_{\alpha_1} \dots D_{\alpha_m}$, the converse result of proposition 4.42 is wrong, as there exists an operator G **LC TI**, m -scale-invariant, steerable, with $n_G = m + 1$ and $K_G = m$, that cannot be written in this form.

Easy Proof. Let $G_1 = D_{\alpha_1} \dots D_{\alpha_m}$ and $G_2 = (-\Delta)^{(m-1)/2} D_{\alpha_{m+1}}$, which are both **LC TI**, m -scale-invariant and steerable, and they have $n_{G_1} = m + 1$ (cf. lemma 4.41) and $n_{G_2} = 2$ (cf. proposition 4.42) and they have both $K_{G_1} = K_{G_2} = m$. But their sum cannot be written as $C(-\Delta)^{\gamma/2} D_{\alpha'_1} \dots D_{\alpha'_{m'}}$ (we can check this on their Fourier multiplier). \square

The more general result we got about the converse sense is the following, where G is not written as a simple composition of a fractional Laplacian and some directional derivatives, but as a **sum** of such products.

Proposition 4.46 (Weak converse of proposition 4.42). *Let G is **LC TI**, $\gamma + m$ scale-invariant (for $\gamma \in \mathbb{R}_+$, $m \in \mathbb{N}$), and steerable, with $K_G = \max |S_G| = m$, on $E = \mathbf{S}(\mathbb{R}^2, \mathbb{R})$ (i.e., G has the form of (24)),*

$$\widehat{g}(r, \theta) = Cr^{\gamma+m} \sum_{-m \leq k \leq m} a_k e^{jk\theta}, \quad (25)$$

then G can be decomposed as a **sum** of terms of this form

$$C(-\Delta)^{\gamma_i/2} D_{\alpha_{i,1}} \dots D_{\alpha_{i,m_i}}. \quad (26)$$

all satisfying $\gamma_i + m_i = \gamma + m$. In this form, G has $n_G = 1 + \max_i \{m_i\}$.

It is only a “weak” converse because there is a **sum** of terms, and not just one, as it was the case for the direct sense (proposition 4.42).

Long and Hard Proof. The converse sense is clearly true, as each term is obviously **LC TI** and steerable, and $\gamma_i + m_i = \gamma + m$ implies that the i -th term is also $\gamma + m$ scale-invariant, and so is their sum.

Proving the direct sense is just a matter of a long and technical trigonometric computation, in order to rewrite $\widehat{g}(r, \theta) = r^{\gamma+m} \sum_{-m \leq k \leq m} a_k e^{jk\theta}$ as a **sum** of products like $\prod_{i=1}^K r_j \cos(\theta - \alpha_i)$, which corresponds to the Fourier multiplier of $\bigcirc_{i=1}^K D_{\alpha_i}$. It consists in some trigonometric manipulations to convert the angular dependency $\sum_{-m \leq k \leq m} a_k e^{jk\theta}$ (dependency on θ).

This proof makes an intensive use of the fact that G is a real operator, to have $a_{-k} = (-1)^k \overline{a_k}$, $\forall k \in \{-m, \dots, m\}$ (proposition 4.39). Instead of keeping complex coefficients a_k we write them as $a_k = b_k + jc_k$ for $b_k, c_k \in \mathbb{R}$. The symmetry reads $b_{-k} = (-1)^k b_k$ for real parts, and $c_{-k} = (-1)^{k+1} c_k$ for imaginary parts.

Let us work on $\sum_{-m \leq k \leq m, k \neq 0} a_k e^{jk\theta}$, as the term for $k = 0$ gives a simple Laplacian term (which is already in the form (26)). So are dealing with this sum: $\sum_{1=k}^m a_k e^{jk\theta} + a_{-k} e^{-jk\theta}$. Each term can first be written and studied independently:

$$\begin{aligned} a_k e^{jk\theta} + a_{-k} e^{-jk\theta} &= (b_k + jc_k) e^{jk\theta} + ((-1)^k b_k + j(-1)^{k+1} c_k) e^{-jk\theta} \\ &= b_k \underbrace{(e^{jk\theta} + (-1)^k e^{-jk\theta})}_{\substack{=2 \cos(k\theta) \text{ if } k \in 2\mathbb{N}, \\ =2j \sin(k\theta) \text{ if } k \in 2\mathbb{N}+1}} + jc_k \underbrace{(e^{jk\theta} + (-1)^{k+1} e^{-jk\theta})}_{\substack{=2j \sin(k\theta) \text{ if } k \in 2\mathbb{N}, \\ =2 \cos(k\theta) \text{ if } k \in 2\mathbb{N}+1}} \end{aligned}$$

And so by regrouping the real terms and the imaginary terms together:

$$\begin{aligned} \sum_{k=1}^m a_k e^{jk\theta} + a_{-k} e^{-jk\theta} &= 2 \sum_{\substack{1 \leq k \leq m, \\ k \in 2\mathbb{N}}} \underbrace{(b_k \cos(k\theta) - c_k \sin(k\theta))}_{\in \mathbb{R}} \\ &\quad + 2j \sum_{\substack{1 \leq k \leq m, \\ k \in 2\mathbb{N}+1}} \underbrace{(b_k \sin(k\theta) + c_k \cos(k\theta))}_{\in \mathbb{R}} \end{aligned}$$

We can check that we did not lose the Hermitian symmetry, as the real part only contains the even indexes k so is even (as a function of θ), and the imaginary part only contains the odd indexes k so is odd (as a function of θ).

For both parts, we will develop each terms $\cos(k\theta)$ and $\sin(k\theta) = \cos(k(\theta - \frac{\pi}{2k}))$, thanks to the Viète formula (lemma B.1, given and proved in appendix B.2). We only work with cosines, because we aim at recognizing the Fourier multipliers of compositions of directional derivatives (of the form $j^r \cos(\theta - \alpha)$).

For the real part,

$$\begin{aligned} &2 \sum_{\substack{1 \leq k \leq m, \\ k \in 2\mathbb{N}}} (b_k \cos(k\theta) - c_k \sin(k\theta)) \\ &= 2 \sum_{\substack{1 \leq k \leq m, \\ k \in 2\mathbb{N}}} \left(b_k \left[\sum_{l=0}^{\lfloor k/2 \rfloor} \binom{k}{2l} (-1)^l (\cos(\theta))^{k-2l} (\cos(\theta - \frac{\pi}{2}))^{2l} \right] \right. \\ &\quad \left. - c_k \left[\sum_{l=0}^{\lfloor k/2 \rfloor} \binom{k}{2l} (-1)^l (\cos(\theta - \frac{\pi}{2k}))^{k-2l} (\cos(\theta - (\frac{\pi}{2} + \frac{\pi}{2k})))^{2l} \right] \right). \end{aligned}$$

And for the imaginary part,

$$\begin{aligned} &2 \sum_{\substack{1 \leq k \leq m, \\ k \in 2\mathbb{N}+1}} (b_k \sin(k\theta) + c_k \cos(k\theta)) \\ &= 2 \sum_{\substack{1 \leq k \leq m, \\ k \in 2\mathbb{N}+1}} \left(b_k \left[\sum_{l=0}^{\lfloor k/2 \rfloor} \binom{k}{2l} (-1)^l (\cos(\theta - \frac{\pi}{2k}))^{k-2l} (\cos(\theta - (\frac{\pi}{2} + \frac{\pi}{2k})))^{2l} \right] \right. \\ &\quad \left. + c_k \left[\sum_{l=0}^{\lfloor k/2 \rfloor} \binom{k}{2l} (-1)^l (\cos(\theta))^{k-2l} (\cos(\theta - \frac{\pi}{2}))^{2l} \right] \right). \end{aligned}$$

Without inflecting to the reader the pain of developing and simplifying these sums of sums of product of two powered cosines⁵⁹, we just observe that both the real and imaginary parts have been decomposed as a sum of product of cosines, all in the form $\cos(\theta - \alpha)$ for some $\alpha \in \mathbb{R}$ (the angles being $0, \frac{\pi}{2}$, or in the form $\frac{\pi}{2k}$ or $\frac{\pi}{2} + \frac{\pi}{2k}$ for $k \in \{1, \dots, m\}$). The coefficients appearing in these sums are all reals, composed of $2, b_k, c_k \in \mathbb{R}$, and $(-1)^l$ and $\binom{k}{2l}$.

We simply need to check that there is always even number of cosine terms for the real part and an odd number for the imaginary part. But it is indeed the case

⁵⁹ $(\cos(\theta))^{k-2l} \times (\cos(\theta - \frac{\pi}{2}))^{2l}$, and $(\cos(\theta - \frac{\pi}{2k}))^{k-2l} \times (\cos(\theta - (\frac{\pi}{2} + \frac{\pi}{2k})))^{2l}$ for the real part, and $(\cos(\theta - \frac{\pi}{2k}))^{k-2l} \times (\cos(\theta - (\frac{\pi}{2} + \frac{\pi}{2k})))^{2l}$ and $(\cos(\theta))^{k-2l} \times (\cos(\theta - \frac{\pi}{2}))^{2l}$ for the imaginary part.

as there is always $(k - 2l) + (2l) = k$ cosines in the product, and $k \in 2\mathbb{N}$ for the real part, and $k \in \mathbb{N} + 1$ for the imaginary part.

Therefore the real part of the polar dependency has been written as a sum of products of an even number ($m_i \leq m$) of terms of the form $j \cos(\theta - \alpha_l)$ (the j disappear as $j^k \in \{\pm 1\}$ if $k \in 2\mathbb{N}$), and the imaginary part of the polar dependency has been written as a sum of products of an odd number ($m_i \leq m$) of terms of the form $j \cos(\theta - \alpha_l)$ (the j stays as $j^k \in \{\pm j\}$ if $k \in 2\mathbb{N} + 1$).

We are done for the angular part, and the radial part was only a common factor of $r^{\gamma+m}$. Every term of the sum obtained for the angular part consist of a product of $k' \leq m$ cosines, so we can factorize the radial part $r^{\gamma+m}$ and split it, for each term, between a r^{m_i} and $r^{\gamma+m-m_i}$. The r^{m_i} term gets factorized as r for each $j \cos(\theta - \alpha_l)$, so we recognize the Fourier multiplier of D_{α_l} , and the remaining $r^{\gamma+m-m_i}$ is the Fourier multiplier of a fractional Laplacian with $\gamma_i = \gamma + m - m_i \geq 0$. It proves that $\gamma_i + m_i = \gamma + m$ as we announced.

Finally, we have indeed been able to decompose $\hat{g}(r, \theta)$, the Fourier multiplier of a 2D steerable $\gamma + m$ scale-invariant convolution, with $K_G = m$, as a sum of terms that are proportional (with real coefficients) to the Fourier multipliers of operators of the form $(-\Delta)^{\gamma_i/2} D_{\alpha_{i,1}} \dots D_{\alpha_{i,m_i}}$, with $\gamma_i + m_i = \gamma + m$.

Remark 4.47. The formula for the steerability order comes from two points:

- A block $C(-\Delta)^{\gamma_i/2} D_{\alpha_{i,1}} \dots D_{\alpha_{i,m_i}}$ as in (26) has $n_G = 1 + m_i$.
- And for this special case of composition of derivatives, the steerability order of a sum is the maximum of the steerability orders of each term (lemma 4.41).

Remark 4.48. And we also have a bound on the number of products needed to be included: there will be at most $(\lfloor \frac{m}{2} \rfloor)(\lfloor \frac{m}{2} \rfloor + 1)$ terms in the decomposition (26). It also comes from the number of terms in the right-hand side of the Viète formulas (lemma B.1). \square

We continue below with an important result which does not decompose G as a *sum* of simple terms, but as a *composition* of Laplacians, directional derivatives, and an invertible 0-scale-invariant factor G_0 , satisfying $\hat{g}_0(\omega) \neq 0$, for all $\omega \in \mathbb{R}^d$.

4.6 Main characterization for 2D steerable γ -SI convolution operators

In this subsection, we conclude the study of 2D steerable operators by giving a nicer characterization, based on the form $\hat{g}(r, \theta) = r^\gamma \sum_{-K_G \leq k \leq K_G} a_k e^{jk\theta}$. Instead of having a sum of $2K_G$ very simple terms, we will modify it to obtain a product of K_G simple factors, each consisting of only one or a sum of two terms.

4.6.1 First product decomposition

The exact form is left for the next section, here we will start by showing the first characterization as a product of a Laplacian, some directional derivatives, and a “nice part”, being invertible and 0-scale-invariant. This result states that steerable convolutions in 2D can be factorized into a composition of a fractional Laplacian, of directional derivatives, and of a certain operator G_0 which has a non-canceling Fourier multiplier \hat{g}_0 . The converse is also true, under some strong hypotheses on \hat{g}_0 .

It is one of the main contributions of the work presented in this report, as it is both an innovative result and a strong characterization. To the best of our knowledge, this theorem is the very first of its kind.

Theorem 4.49 (Invertible part of a steerable operator). *An operator G on $\mathbf{L}^2(\mathbb{R}^2, \mathbb{R})$ is a 2D steerable convolution, γ -scale-invariant with $\gamma \geq K_G$ iff G can be written*

$$G = C(-\Delta)^{(\gamma-m)/2} \circ D_{\alpha_1} \circ \dots \circ D_{\alpha_m} \circ G_0, \quad (27)$$

with $C \in \mathbb{R}$, $m \leq K_G$, $\alpha_1, \dots, \alpha_m \in \mathbb{R}$, and G_0 an invertible 0-scale-invariant steerable convolution, with $\widehat{g}_0(r, \theta) \neq 0$ for all r, θ . Furthermore, $\widehat{g}_0(r, \theta)$ is just a (Hermitian symmetric) trigonometric polynomial in $e^{j\theta}$, of degree $K_G - m$, that has no root of modulus 1, i.e., $\widehat{g}_0(r, \theta) = \prod_{z_i \in Z_2, |z_i| \neq 1} (e^{j\theta} - z_i)$, for a finite $Z_2 \subset \mathbb{C}$ of even size.

Remark 4.50. We tried to clean up and shorten its proof, but this result involves some unavoidable technical computations and “tricks”. Proving the symmetry $z_i \in Z_G \Leftrightarrow -1/\bar{z}_i \in Z_G$ is quick for roots of modulus 1, but really not obvious for the other roots.

Note that in all the following computations, θ is always a free variable, all the expressions involving a θ have to be understood with a leading “ $\forall \theta \in [0, 2\pi), \dots$ ”, and similarly for $r \in \mathbb{R}$.

Long and Hard Proof. • **Direct sense:**

Thanks to theorem 4.37, G is **LC TI** γ -**SI St** for $\gamma \geq K_G$ iff its Fourier multiplier can be written $\widehat{g}(r, \theta) = r^\gamma \sum_{-K_G \leq k \leq K_G} a_k e^{jk\theta}$. The radial dependency is very simple, so we only have to work on the polar part.

Let $P_G(e^{j\theta}) \stackrel{\text{def}}{=} \sum_{-K_G \leq k \leq K_G} a_k e^{jk\theta}$, which is a trigonometric polynomial of degree K_G ($a_{K_G} \neq 0$). We could consider the (complex) roots of modulus 1 of the polynomial $P_G(z) = \sum_{-K_G \leq k \leq K_G} a_k z^k$, and simply say that the other roots do not matter as long as $|z| = 1$ (i.e., $z = e^{j\theta}$), but this would not be good enough.

As $P_G(z)$ is not really a polynomial, we instead work with a second object,

$$\begin{aligned} \widetilde{P}_G(e^{j\theta}) &\stackrel{\text{def}}{=} e^{jK_G\theta} P_G(e^{j\theta}) = (e^{j\theta})^{K_G} P_G(e^{j\theta}) \\ &= e^{jK_G\theta} \sum_{-K_G \leq k \leq K_G} a_k e^{jk\theta} = \sum_{0 \leq k \leq 2K_G} a_{k-K_G} \underbrace{(e^{j\theta})^k}_{=z}, \end{aligned}$$

which is a (“true”) complex polynomial now: $\widetilde{P}_G(z) = \sum_{0 \leq k \leq 2K_G} a_{k-K_G} z^k$.

Note that P_G and \widetilde{P}_G are linked to \widehat{g} , by $\widehat{g}(r, \theta) = r^\gamma P_G(e^{j\theta}) = r^\gamma e^{-jK_G\theta} \widetilde{P}_G(e^{j\theta})$.

Because G is real, \widehat{g} is Hermitian symmetric, and so we have a symmetry on its coefficients ($a_{-k} = (-1)^k \overline{a_k}$) as well as on the trigonometric polynomial P_G . Indeed we have $\widehat{g}(-\omega) = \widehat{g}(\omega)$, which gives $\widehat{g}(r, \theta + \pi) = \widehat{g}(r, \theta)$ in polar coordinates, and so $P_G(e^{j(\theta+\pi)}) = P_G(-e^{j\theta}) = \overline{P_G(e^{j\theta})}$ on P_G .

This symmetry on P_G gives a similar symmetry on \widetilde{P}_G , with an extra factor:

$$\begin{aligned} \widetilde{P}_G(e^{j(\theta+\pi)}) &= \widetilde{P}_G(-e^{j\theta}) \\ \text{(By definition of } \widetilde{P}_G) &= (-1)^{K_G} e^{jK_G\theta} P_G(-e^{j\theta}) \\ \text{(By symmetry on } P_G) &= (-1)^{K_G} e^{jK_G\theta} \overline{P_G(e^{j\theta})} \end{aligned}$$

$$\begin{aligned}
& \text{(By definition of } \widetilde{P}_G) = (-1)^{K_G} e^{jK_G\theta} \overline{e^{-jK_G\theta} \widetilde{P}_G(e^{j\theta})} \\
& \text{So } \widetilde{P}_G(e^{j(\theta+\pi)}) = (-1)^{K_G} e^{j2K_G\theta} \overline{\widetilde{P}_G(e^{j\theta})}. \tag{28}
\end{aligned}$$

Let us consider the roots of \widetilde{P}_G . It is of degree exactly $2K_G$ as $a_{K_G} \neq 0$ (by definition), and so thanks to the Fundamental Theorem of Algebra, it has *exactly* $2K_G$ complex roots, and let Z_G be the set of its (complex) roots ($\#Z_G = 2K_G$). We can compute its value at 0: $\widetilde{P}_G(0) = a_{-K_G} = (-1)^{K_G} \overline{a_{K_G}} \neq 0$, so 0 is not a root of \widetilde{P}_G ($0 \notin Z_G$).

By definition of Z_G , we have $\widetilde{P}_G(e^{j\theta}) = a_{K_G} \prod_{z_i \in Z_G} (e^{j\theta} - z_i)$, as a_{K_G} is the highest-degree factor in \widetilde{P}_G .

So $\widetilde{P}_G(e^{j(\theta+\pi)}) = \widetilde{P}_G(-e^{j\theta}) = a_{K_G} \prod_{z_i \in Z_G} (-e^{j\theta} - z_i)$, and (28) gives

$$\begin{aligned}
\widetilde{P}_G(e^{j(\theta+\pi)}) &= (-1)^{K_G} e^{j2K_G\theta} \overline{\widetilde{P}_G(e^{j\theta})} \\
&= (-1)^{K_G} e^{j2K_G\theta} (\overline{a_{K_G}} \prod_{z_i \in Z_G} (e^{-j\theta} - \overline{z_i})) \\
(\#Z_G = 2K_G \text{ so}) &= (-1)^{K_G} \overline{a_{K_G}} \prod_{z_i \in Z_G} e^{j\theta} (e^{-j\theta} - \overline{z_i}) \\
&= (-1)^{K_G} \overline{a_{K_G}} \prod_{z_i \in Z_G} (1 - \overline{z_i} e^{j\theta})
\end{aligned}$$

And $0 \notin Z_G$ so $z_i \neq 0$, hence we can divide by $\overline{z_i}$:

$$\begin{aligned}
\widetilde{P}_G(e^{j(\theta+\pi)}) &= (-1)^{K_G} \overline{a_{K_G}} \prod_{z_i \in Z_G} \overline{z_i} (1/\overline{z_i} - e^{j\theta}) \\
a_{K_G} \prod_{z_i \in Z_G} (-e^{j\theta} - z_i) &= \left((-1)^{K_G} \overline{a_{K_G}} \prod_{z_i \in Z_G} \overline{z_i} \right) \prod_{z_i \in Z_G} (-1) (e^{j\theta} - 1/\overline{z_i}) \\
a_{K_G} (-1)^{\overbrace{\#Z_G}^{\in 2\mathbb{N}}} \prod_{z_i \in Z_G} (e^{j\theta} + z_i) &= \left((-1)^{K_G} \overline{a_{K_G}} \prod_{z_i \in Z_G} \overline{z_i} \right) (-1)^{\overbrace{\#Z_G}^{\in 2\mathbb{N}}} \prod_{z_i \in Z_G} (e^{j\theta} - 1/\overline{z_i}) \\
a_{K_G} \prod_{z_i \in Z_G} (e^{j\theta} + z_i) &= \left((-1)^{K_G} \overline{a_{K_G}} \prod_{z_i \in Z_G} \overline{z_i} \right) \prod_{z_i \in Z_G} (e^{j\theta} - 1/\overline{z_i}). \tag{29}
\end{aligned}$$

But we also know that $a_{-K_G} = a_{K_G} \prod_{z_i \in Z_G} -z_i = a_{K_G} \prod_{z_i \in Z_G} z_i$, from the fact that $a_{-K_G} = \widetilde{P}_G(0)$. So $(-1)^{K_G} \overline{a_{K_G}} \prod_{z_i \in Z_G} \overline{z_i} = (-1)^{K_G} \overline{a_{K_G}} \prod_{z_i \in Z_G} z_i = (-1)^{K_G} \overline{a_{-K_G}}$ and so it is a_{K_G} (by the symmetry on coefficients): therefore the constant term in the previous computation (29) vanishes. And so we have $\prod_{z_i \in Z_G} (e^{j\theta} + z_i) = \prod_{z_i \in Z_G} (e^{j\theta} - 1/\overline{z_i})$, which directly proves that $-z_i \in Z_G \Leftrightarrow 1/\overline{z_i} \in Z_G$, i.e.,

$$z_i \in Z_G \Leftrightarrow -1/\overline{z_i} \in Z_G. \tag{30}$$

The figure 2 illustrates this symmetry and shows some roots for a toy example of \widetilde{P}_G : z_1, z'_1, z_2, z'_2 are roots of modulus 1 (i.e., $\in \mathbb{U}$), and z_3, z'_3, z_4, z'_4 are roots of modulus $\neq 0, 1$ (i.e., $\notin \mathbb{U}$); and they are symmetric, $z'_i = -1/\overline{z_i}$.

The factorization $P_G(e^{j\theta}) = e^{-jK_G\theta} \widetilde{P}_G(e^{j\theta}) = a_{K_G} e^{-jK_G\theta} \prod_{z_i \in Z_G} (e^{j\theta} - z_i)$ can then be simplified thanks to this relation (30).

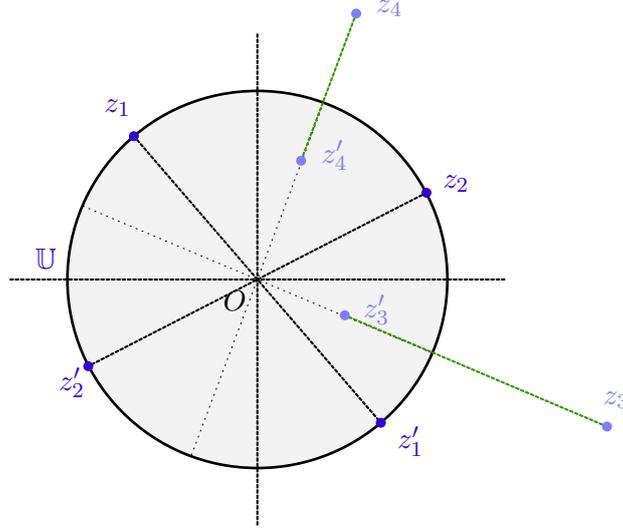


Figure 2: Some roots of the polynomial \widetilde{P}_G , $z_1, z'_1, z_2, z'_2 \in \mathbb{U}$, $z_3, z'_3, z_4, z'_4 \notin \mathbb{U}$.

- For the **roots of modulus 1**, if $z_i = e^{j\beta_i}$ is a root, then the relation (30) gives that $-1/\overline{z_i} = -z_i = e^{j(\beta+\pi)} = -e^{j\beta_i}$ is also a root.

So if we define $Z_1 \stackrel{\text{def}}{=} Z_G \cap \mathbb{U} = \{z_i \in Z_G, |z_i| = 1\}$ the set of roots of modulus 1, we just proved that it has an even size ($\#Z_1 \in 2\mathbb{N}$), and that $e^{j\beta_i} \in Z_1 \Leftrightarrow -e^{j\beta_i} \in Z_1$. Figure 2 shows $Z_1 = \{z_1, z'_1, z_2, z'_2\}$ and $Z'_1 = \{z_1, z_2\}$.

Let $Z'_1 = Z_1 \cap [0, \pi)$, and so $\prod_{z_i \in Z_1} (e^{j\theta} - z_i) = \prod_{\beta_i \in Z'_1} (e^{j\theta} - e^{j\beta_i})(e^{j\theta} + e^{j\beta_i})$.

For each term:

$$\begin{aligned} (e^{j\theta} - e^{j\beta_i})(e^{j\theta} + e^{j\beta_i}) &= e^{j2\theta} - e^{j2\beta_i} \\ &= e^{j(\theta+\beta_i)}(e^{j(\theta-\beta_i)} - e^{j-(\theta-\beta_i)}) \\ &= e^{j\theta} e^{j\beta_i} 2j \sin(\theta - \beta_i) \\ &= e^{j\theta} e^{j\beta_i} 2j \cos(\theta - (\beta_i - \frac{\pi}{2})). \end{aligned}$$

Remember that $P_G(e^{j\theta}) = e^{-jK_G\theta} \widetilde{P}_G(e^{j\theta})$, so we can factorize one $e^{-j\theta}$ for each term in order to simply keep $\prod_{\beta_i \in Z'_1} e^{j\beta_i} 2j \cos(\theta - (\beta_i - \frac{\pi}{2}))$ (and there is

$K_G - \#Z'_1$ terms $e^{-j\theta}$ left). Up-to a constant complex term $e^{j \sum_{\beta_i \in Z'_1} \beta_i}$ (that will be dealt with after), this product is the angular dependency of the Fourier multiplier of a composition of $m \stackrel{\text{def}}{=} \#Z'_1$ directional derivatives of angles $\alpha_i \stackrel{\text{def}}{=} \beta_i - \frac{\pi}{2} \in \mathbb{R}$ (without their radial dependencies r).

- Now for the **roots not of modulus 1**, if $z_i = r_i e^{j\beta_i}$ is a root, then the relation (30) gives that $-1/\overline{z_i} = -(1/r_i)e^{j\beta_i}$ is also a root (and $r_i \neq 0, 1$ so $1/r_i \neq r_i$).

So if we define $Z_2 \stackrel{\text{def}}{=} Z_G \setminus Z_1 = \{z_i \in Z_G, |z_i| \neq 1\}$, it also has an even size ($\#Z_2 \in 2\mathbb{N}$), and that $r_i e^{j\beta_i} \in Z_2 \Leftrightarrow -(1/r_i)e^{j\beta_i} \in Z_2$.

Let $Z'_2 = \{|z_i| < 1, z_i \in Z_2\}$, and so $\prod_{z_i \in Z_2} (e^{j\theta} - z_i) = \prod_{\beta_i \in Z'_2} (e^{j\theta} - r_i e^{j\beta_i})(e^{j\theta} + (1/r_i)e^{j\beta_i})$. Figure 2 shows $Z_2 = \{z_3, z'_3, z_4, z'_4\}$ and $Z'_2 = \{z_3, z_4\}$.

For each term:

$$\begin{aligned}
(e^{j\theta} - r_i e^{j\beta_i})(e^{j\theta} + (1/r_i)e^{j\beta_i}) &= e^{j2\theta} + \underbrace{(1/r_i - r_i)}_{\neq 0, > 0 \text{ in fact}} e^{j(\theta+\beta_i)} - e^{j2\beta_i} \\
&= e^{j(\theta+\beta_i)}(e^{j(\theta-\beta_i)} + (1/r_i - r_i) - e^{j(-(\theta-\beta_i))}) \\
&= e^{j\theta} e^{j\beta_i} (2j \sin(\theta - \beta_i) + (1/r_i - r_i)) \\
&= e^{j\theta} e^{j\beta_i} \underbrace{(2j \cos(\theta - (\beta_i - \frac{\pi}{2})) + (1/r_i - r_i))}_{\notin j\mathbb{R}, \in \mathbb{C}, \text{ and } \neq 0, \forall \theta}.
\end{aligned}$$

Remember that there is still $(K_G - \#Z'_1)$ term $e^{-j\theta}$ left, and $K_G - \#Z'_1 = \#Z'_2$ (because $2K_G = \#Z_G = \#Z_1 + \#Z_2 = 2\#Z'_1 + 2\#Z'_2$), so we can factorize one $e^{-j\theta}$ for each term in order to simply keep $\prod_{\beta_i \in Z'_2} e^{j\beta_i} (2j \cos(\theta - (\beta_i - \frac{\pi}{2})) + (1/r_i - r_i))$. Up-to a constant complex term $e^{j \sum_{Z'_2} \beta_i}$ (that will be dealt with after), this product corresponds to a never-canceling Fourier multiplier of a steerable convolution. It will be studied more in detail in the next corollary 5.1.

For now, we can conclude by observing that this complex-valued Fourier multiplier is indeed never-canceling and is a trigonometric polynomial. \square

Remark 4.51. So we are almost done for the direct sense, we simply need to explain why the constant complex terms are not an issue. At first sight, they could mess up with the Hermitian symmetry of \hat{g} (if the leading factor is not a pure-real number). Remember that $a_{-K_G} = (-1)^{K_G} \overline{a_{K_G}}$, but also that $a_{-K_G} = \widehat{P}_G(0) = a_N \prod_{z_i \in Z_G} (-z_i) = a_N \prod_{z_i \in Z_G} z_i$ (as $\#Z_G \in 2\mathbb{N}$). Therefore, if $\psi \stackrel{\text{def}}{=} \arg a_{K_G}$, then $e^{j\psi}$ satisfies⁶⁰ $(-1)^{K_G} e^{-j\psi} = e^{j\psi} \prod_{\beta_i \in \arg Z_G} e^{j\beta_i}$. Indeed the modulus of the roots disappear in this product, by the symmetry we proved on Z_G , as $(r_i e^{j\beta_i})(-\frac{1}{r_i} e^{j\beta_i}) = -e^{j2\beta_i}$. So $K_G \pi - \psi = \psi + \sum_{\beta_i \in \arg Z_G} \beta_i \pmod{2\pi}$, but by symmetry, half of the $\beta_i \in \arg Z_G$ can be written $\beta_{i'} + \pi$ (for some i'), and $\#Z_G/2 = \#Z'_G = K_G$, so $K_G \pi = 2\psi + \sum_{\beta_i \in \arg Z'_G} (\beta_i + \beta_i + \pi) = 2\psi + 2 \sum_{\beta_i \in \arg Z'_G} \beta_i + K_G \pi \pmod{2\pi}$, i.e., $\psi + \sum_{\beta_i \in \arg Z'_G} \beta_i = 0 \pmod{\pi}$, which simply says that the constant complex term in the final factorization is purely real: $e^{j(\psi + \sum_{\beta_i \in \arg Z'_G} \beta_i)} = \pm 1 \in \mathbb{R}$. And so, we rechecked that the factorization we obtained is indeed Hermitian symmetric.

The direct sense of the proof was quite long, because we had to introduce different objects (P_G, \widehat{P}_G, Z_G etc), and prove some non-obvious results (mainly the symmetry $z_i \in Z_G \Leftrightarrow -1/\overline{z_i} \in Z_G$).

The converse sense is much simpler, as usual for such results.

Proof. • Converse sense:

Let G be written like the theorem 4.49 asks. We simply need to check that:

- It is obviously a convolution as a product of convolutions (corollary 2.25).

⁶⁰ We use the *unusual* notation $\arg Z_G$ for $\{\beta_i, z_i = r_i e^{j\beta_i} \in Z_G\} = \{\arg z_i, z_i \in Z_G\}$.

- G_0 is 0-SI, each derivative D_{α_i} is 1-SI and the $(\gamma - m)$ -fractional Laplacian is $(\gamma - m)$ -SI, so their composition, G , is $0 + (1 + \dots + 1) + (\gamma - m) = \gamma$ scale-invariant (proposition 2.39).
- $\widehat{g}_0(\omega)$ has a very specific form, which gives directly that $\widehat{g}(r, \theta)$ can be written as $r^\gamma \sum_{-K_G \leq k \leq K_G} a_k e^{jk\theta}$, and so thanks to the converse of theorem 4.31, G is steerable.
- We know that the fractional Laplacian and the derivatives are real, and $\widehat{g}_0(\omega)$ is Hermitian symmetric, so G is also real (differentiating indeed maps a real-valued function to a real-valued function). \square

4.6.2 Rotating a 2D steerable γ -SI convolution operators

This first product decomposition has the advantage of giving a very simple expression for a rotated version of G . A rotation R_{θ_0} on G gets applied separately on each derivatives D_{α_i} and on G_0 .

Proposition 4.52. For G written as in theorem 4.49, and a rotation $R = R_{\theta_0} \in \text{SO}(2)$,

$$r_R G r_{R^{-1}} = C(-\Delta)^{(\gamma-m)/2} \circ D_{\alpha_1 - \theta_0} \circ \dots \circ D_{\alpha_m - \theta_0} \circ (r_R G_0 r_{R^{-1}}).$$

Easy Proof. Quite obvious, as $C(-\Delta)^{(\gamma-m)/2}$ is **RI**, $r_{R_{\theta_0}} D_{\alpha_1} r_{R_{\theta_0}^{-1}} = D_{\alpha_1 - \theta_0}$, and the fact that rotating G is equivalent to rotating each block, remark 4.6. \square

This proposition 4.52 has to be put in perspective with a similar result we can prove for the sum decomposition, for which the rotation R_{θ_0} gets applied on the coefficients a_k .

Proposition 4.53. For G written as in theorem 4.37, and a rotation $R = R_{\theta_0} \in \text{SO}(2)$, and let $G' = r_R G r_{R^{-1}}$. G and its rotated version has the same frequency $K_{G'} = K_G$ and homogeneity degree γ , and the coefficients of G' are $a'_k = a_k e^{jk\theta_0}$.

In other words, the Fourier multiplier g' of G' satisfies

$$\forall \omega = (r, \theta) \in \mathbb{R}^2, \quad r_R \widehat{g}(r, \theta) = \widehat{g}'(r, \theta) = \widehat{g}(r, \theta + \theta_0) = r^\gamma \sum_{-K_G \leq k \leq K_G} a'_k e^{jk\theta}.$$

Easy Proof. Obvious,

$$\begin{aligned} r_R \widehat{g}(r, \theta) &= \widehat{g}(r, \theta + \theta_0) = r^\gamma \sum_{-K_G \leq k \leq K_G} a_k e^{jk(\theta + \theta_0)} \\ &= r^\gamma \sum_{-K_G \leq k \leq K_G} \underbrace{a_k e^{jk\theta_0}}_{\stackrel{\text{def}}{=} a'_k} e^{jk\theta}. \end{aligned} \quad \square$$

Conclusion

To the best of our knowledge, the theorem we just proved is the first decomposition of a steerable convolution operator as a **composition** of well-knowns operators (fractional Laplacian and directional derivatives) and an invertible 0-scale-invariant part (of non-canceling Fourier multiplier).

This theorem is quite strong, but still not explicit enough if we want to use a steerable convolution G in practice. Instead of this “nice” part G_0 , given by its Fourier multiplier $\hat{g}_0(r, \theta)$ parametrized as a non-canceling trigonometric polynomial in $e^{j\theta}$ (*i.e.*, a sum), we would like to have a decomposition in term of a composition of simple elementary blocks (*i.e.*, a product). The next section will explore this direction.

5 Implementing and inverting 2D steerable convolutions

In this short section, the last one, we still focus on steerable γ -**SI** convolution operators in 2D, and theorem 4.49 is our starting point. First, finishing the computations on G_0 will give an other characterization, which consists in a product of elementary blocks $G_{\lambda,\alpha}$ (corollary 5.1).

We then explain how to implement these operators for 2D images, by using a direct and an inverse Fast Fourier Transform (FFT), and the discrete-domain expression of the Fourier multipliers.

We also shortly present how this decomposition can be used to invert the 2D steerable convolutions. The same method can be used to implement their inverses, providing that we know how to inverse the block. In fact, this step turns out to be quite hard to obtain symbolically, and we expose why it is a difficult task that we have been able to finish in time. However, implementing an inverse elementary block was not too hard, and some examples illustrate its effect on a 2D Poisson sparse process are included.

5.1 Product decompositions vs sum decomposition

We gave and prove above a very strong characterization for 2D steerable convolutions, but we did not give much details about the invertible 0-scale-invariant G_0 and its Fourier multiplier \hat{g}_0 . The theorem 4.49 stated that \hat{g}_0 is a trigonometric polynomial, of degree $K_G - m$, that does not cancel. We can work on the expression of \hat{g}_0 in order to make G_0 appear as a composition of simpler terms.

It seemed weird to have, in one hand, two very simple elementary blocks ($(-\Delta)^{1/2}$ or $(-\Delta)^{\gamma/2}$, and D_{α_i}), and, on the other hand, an un-factorized G_0 . We will essentially prove that G_0 can be factorized as a composition of 1-**SI** steerable convolutions that are simple elementary blocks of the common form $\lambda_i D_{\alpha_i} + (1 - \lambda_i)(-\Delta)^{1/2}$, for $\lambda_i \in (0, 1)$, up to a real constant coefficient.

In other words, after factorizing a fractional Laplacian of the good order, G_0 is just a product of convex combinations of two well-known differential 1-scale-invariant convolutions, the directional derivative D_{α_i} and the half-Laplacian $(-\Delta)^{1/2}$.

It will again strengthen the fact that the two elementary blocks (D_{α_i} and $(-\Delta)^{1/2}$) are intrinsically important, and it will also have a practical consequence for the inversion of our convolutions, as it is much easier to compute the inverse and the Green's function of a product than of a sum.

5.1.1 Factorizing G_0

Remember that $\hat{g}_0(r, \theta) = \prod_{z_i=r_i e^{j\beta_i} \in Z'_2} (2j \cos(\theta - (\beta_i - \pi/2)) + (1/r_i - r_i))$ (cf. proof of the direct sense of theorem 4.49). We chose $r_i < 1$, so $1/r_i - r_i > 0$. Let $C_i \stackrel{\text{def}}{=} 2 + (1/r_i - r_i) \in \mathbb{R}$, > 2 and $\lambda_i = 2/C_i \in (0, 1)$. Then, with $\alpha_i \stackrel{\text{def}}{=} \beta_i - \pi/2$, we can write this term as a convex combination between the polar part of D_{α_i} and of $(-\Delta)^{1/2}$:

$$2j \cos(\theta - (\beta_i - \pi/2)) + (1/r_i - r_i) = C_i (\lambda_i \cos(\theta - \alpha_i) + (1 - \lambda_i)). \quad (31)$$

The *real* constant C_i is not an issue, they will be factorized into one real constant C .

Therefore, we can transform the \widehat{g}_0 into a composition of terms with a never-canceling Fourier multipliers, of the form $\lambda_j \cos(\theta - \alpha) + (1 - \lambda)$ for $\lambda \neq 0, 1$, as announced:

$$\widehat{G}_0(r, \theta) = C' \prod_{i=1}^m (\lambda_j \cos(\theta - \alpha) + (1 - \lambda)). \quad (32)$$

And so, up-to a term r^m , it will give the operator G_0 :

$$G_0 = C' \bigcirc_{i=1}^m (\lambda_i D_{\alpha_i} + (1 - \lambda_i)(-\Delta)^{1/2}). \quad (33)$$

We can add a radial 1-scale-invariant term r for each block, because the theorem below will assume $\gamma \geq K_G$ and $m \leq K_G$.

The form of the elementary block $\lambda_i D_{\alpha_i} + (1 - \lambda_i)(-\Delta)^{1/2}$, introduced here only for $\lambda \neq 0, 1$, is also interesting as it clearly generalizes the half-Laplacian (for $\lambda = 0$), and the directional derivatives (for $\lambda = 1$). The previous characterization 4.49 already introduced a Laplacian and some directional derivatives in the decomposition, and they can all also be written as $\lambda_i D_{\alpha_i} + (1 - \lambda_i)(-\Delta)^{1/2}$, for $\lambda_i \in \{0, 1\}$.

Therefore, we can decompose G into a product of well-known blocks and a “nice” part, and each term can be written as one elementary block (with $\lambda_i \in \{0, 1\}$) or a product of elementary blocks (with $\lambda_i \in (0, 1)$).

5.1.2 Elementary blocks decomposition for 2D LC TI γ -SI St operators

This can be formalized as the second characterization for 2D steerable γ -SI convolutions.

Corollary 5.1 (Elementary blocks decomposition for 2D LC TI γ -SI St operators).

An operator G on $\mathbf{L}^2(\mathbb{R}^2, \mathbb{R})$ is a 2D steerable convolution, γ -scale-invariant with $\gamma \geq K_G$, i.e., LC TI γ -SI St iff

$$G = C(-\Delta)^{(\gamma-K_G)/2} \circ \bigcirc_{i=1}^{K_G} (\lambda_i D_{\alpha_i} + (1 - \lambda_i)(-\Delta)^{1/2}). \quad (34)$$

with $C \in \mathbb{R}$, $\lambda_1, \dots, \lambda_{K_G} \in [0, 1]$ and $\alpha_1, \dots, \alpha_{K_G} \in \mathbb{R}$.

In this form, G has $n_G = 1 + \#\{i, \lambda_i \neq 0\} = 1 + \sum_{i=1}^{K_G} \mathbb{1}_{\lambda_i \neq 0}$.

Proof. It is simply a rewriting of theorem 4.49, thanks to the computations we did above⁶¹, combining the explicit form of G_0 (33) and the characterization 4.49. We factorized all the Laplacians in the leading term $(-\Delta)^{(\gamma-K_G)/2}$ (of scale-invariance order $\gamma - K_G$ that could be non-integer), and that is why $\lambda_i > 0$, and that is why there is exactly $N = K_G$ blocks G_{λ_i, α_i} . \square

⁶¹ For this last theorem, we preferred to include the proof beforehand, to clearly show how we came to this result. We could also have started directly with that characterization but we preferred our approach as it appeared more didactic and easier to follow.

Remark 5.2 (What if the homogeneity degree is *too small*?). This corollary assumes $\gamma \geq K_G$, but it can also be applied if it is not the case, at least for the implementation part, because we can write $\widehat{g}(r, \theta) = r^{-K_G} (r^{K_G} \widehat{g}(r, \theta)) = r^{-K_G} \widehat{g}'(r, \theta)$ (for $r \neq 0$), with G' of homogeneity degree $\gamma + K_G \geq K_G$. So we can apply (34) to G' , and obtain a similar decomposition for G , except that the leading Laplacian might have a negative order, *i.e.*, it can be an inverted fractional Laplacian.

Therefore, as soon as we know how to implement and apply in practice such inverted fractional Laplacian, we can be convinced that this characterization can be used for *any* steerable homogeneous convolutions, no matter their order γ (it can even be negative!).

5.1.3 A universality result for the elementary block $G_{\lambda, \alpha}$

This short result is interesting as it highlights the universality of our “elementary block” G_{λ, α_i} among the 2D steerable convolution 1-scale-invariant operators that have $K_G \leq 1$.

Proposition 5.3. *G is a non-zero 2D steerable convolution 1-scale-invariant operator, with $K_G \leq 1$ (or, equivalently, $n_G \leq 2$), iff there is $C \in \mathbb{R}, \lambda \in [0, 1], \alpha \in \mathbb{R}$ such that $G = CG_{\lambda, \alpha} = C \left(\lambda D_\alpha + (1 - \lambda)(-\Delta)^{1/2} \right)$. And G is **RI** iff $n_G = 1$ iff $K_G = 0$ iff $\lambda = 0$.*

Proof. The converse sense is obvious from what is already proved about the block $G_{\lambda, \alpha}$, so let us focus on the direct sense.

Thanks to the main characterization 4.37, we can write $\widehat{g}(r, \theta)$ as $r \sum_{-1 \leq k \leq 1} a_k e^{jk\theta}$, as $K_G \leq 1$. The Hermitian symmetry gives $a_0 \in \mathbb{R}$, and $a_{-1} = -\overline{a_1}$, and so if we write $a_1 = r_1 e^{j\phi_1}$, we can compute easily on $\widehat{g}(r, \theta)$ to obtain the desired form:

$$\begin{aligned} \widehat{g}(r, \theta) &= r(a_{-1}e^{-j\theta} + a_0 + a_1e^{j\theta}) = r(a_0 + a_1e^{j\theta} - \overline{a_1}e^{-j\theta}) \\ &= r(a_0 + r_1e^{j(\theta+\phi_1)} - r_1e^{-j(\theta+\phi_1)}) = r(a_0 + 2r_1j \sin(\theta + \phi_1)) \\ &= r \left(a_0 + 2r_1j \cos \left(\theta + \phi_1 - \frac{\pi}{2} \right) \right) \end{aligned}$$

$$\text{If } C \neq 0, \quad = \underbrace{(a_0 + 2r_1)}_{\stackrel{\text{def}}{=} C} \left(\underbrace{\frac{a_0}{a_0 + 2r_1}}_{=1-\lambda} r + \underbrace{\frac{2r_1}{a_0 + 2r_1}}_{\stackrel{\text{def}}{=} \lambda} jr \cos \left(\theta - \underbrace{\left(\frac{\pi}{2} - \phi_1 \right)}_{\stackrel{\text{def}}{=} \alpha} \right) \right)$$

$$\text{And so } \widehat{g}(r, \theta) = C \left(\lambda \widehat{D}_\alpha + (1 - \lambda)(-\Delta)^{1/2} \right) (r, \theta).$$

If $C \neq 0$ this computation gives the desired form. If $C = a_0 + 2r_1 = 0$, it means that $a_0 = -2r_1 \leq 0$, so with $\alpha = \left(\frac{\pi}{2} - \phi_1 \right) + \pi$, and $C = -2r_1$ and $\lambda = \frac{1}{2}$, we also have the desired form. \square

Example 5.4 (A simple example). We give one example, which at first sight look like a counter-example. Let consider $G = \frac{1}{2} \left(-D_x + (-\Delta)^{1/2} \right)$. It is indeed a 2D steerable convolution, and is clearly 1-scale-invariant, with $K_G = 2$ and $n_G = 1$, as $\widehat{g}(r, \theta) = \frac{1}{2} (r - jr \cos(\theta)) = r \left(-\frac{1}{4}e^{-j\theta} + \frac{1}{2} - \frac{1}{4}e^{j\theta} \right)$ in the form of (\heartsuit).

But because of the difference of signs in front of the derivative and the half-Laplacian, we could think it would not fit as a *convex combination* ($\lambda \in [0, 1]$) of both. The trick here is that $-D_x = -D_{\alpha=0} = D_{\alpha=\pi}$, and so $G = G_{\lambda,\alpha} = \lambda D_\pi + (1 - \lambda)(-\Delta)^{1/2}$ with $\lambda = \frac{1}{2}$ and $C = 1$, in the form of proposition 5.3.

5.1.4 A natural interpretation for this last characterization?

This proposition 5.3 also shows that the form of the decomposition of corollary 5.1 is very natural: a N -scale-invariant steerable convolution G (with $K_G = N$) gets decomposed into a composition of N 1-scale-invariant steerable convolutions, which are exactly the terms written as $\lambda D_\alpha + (1 - \lambda)(-\Delta)^{1/2}$ (up to real multiplicative constant), and has $K_{G_{\lambda,\alpha}} \leq 1$.

We also found interesting that proposition 5.3 generalizes the universality of Laplacian among the 1-scale-invariant rotation-invariant steerable convolutions, *i.e.*, with $n_G = 1$ and $K_G = 0$ (and even if $n_G = 0$).

If the homogeneity order is not an integer or if it too small, corollary 5.1 can still be applied, and the resulting decomposition is still elegant and simple, but we no longer have this additional interpretation.

5.2 Implementation of 2D steerable γ -SI convolutions for images

As we already explain it, applying an operator known with its Fourier multiplier is easy on images, because the computations of \mathcal{F} and \mathcal{F}^{-1} does not need to be done symbolically, but can be done numerically. The Fast Fourier Transform (FFT) is the function `fft2` and its inverse is `ifft2`, and they are used like shown in figure 3.

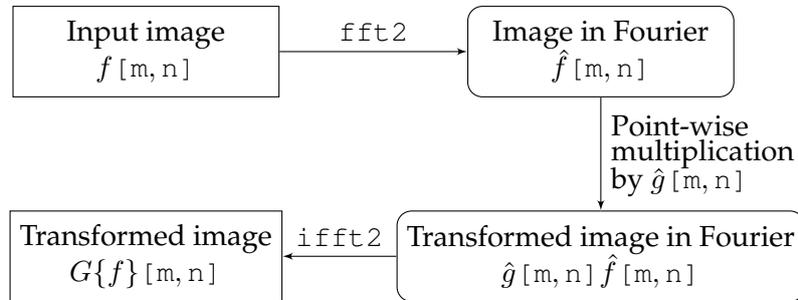


Figure 3: Applying a discrete operator G , given by its (continuous-domain) Fourier multiplier $\hat{g}[m, n]$, on a 2D image $f[m, n]$.

We implemented the elementary blocks $G_{\lambda,\alpha}$ in a GNU Octave script, also valid for MATLAB, using 2D matrices and the Fast Fourier Transform (FFT) and Inverse FFT (cf. `fft2` and `ifft2` functions in MATLAB). Our implementation is not yet open-source, as this part is still an on-going work in collaboration some BIGgers. It will probably be released soon.

The form we obtained with the characterization corollary 5.1 is very elegant because it appears as a composition (*i.e.*, a convolution for impulse responses or a product for Fourier multipliers) of N elementary blocks that now all have the same form: $G_{\lambda,\alpha} = \lambda D_\alpha + (1 - \lambda)(-\Delta)^{1/2}$, for $\lambda \in [0, 1]$ and $\alpha \in \mathbb{R}$.

As we announced it, having this is very interesting from a concrete point of view. We simply need to implement two things: a fractional Laplacian and a block $G_{\lambda,\alpha}$, and we just have to compose them in order to obtain any operators of the considered class. Additionally, the parameters λ_i, α_i have a physical interpretation, α being the direction of the derivative D_α and λ the trade-of between the directionality of D_α and the isotropy of the half-Laplacian.

5.3 Inverse steerable convolutions applied to Sparse Stochastic Processes

In this subsection, we present the results of some experiments we made on 2D images, generated as sparse stochastic processes, on two kinds of white noise, a Gaussian and compound Gaussian-Poisson noise.

A sparse process s is specified with three elements: a linear operator, usually written L in the literature and not G (usually it is a linear continuous differential operator), an initial “white” noise w , and the equation $Ls = w$. For most cases, s is defined as a solution of a linear stochastic differential equation, driven by the white Lévy noise w . To compute or plot s , one has to find a valid left-inverse for L , which can be well defined even if L is not invertible at first sight, cf. [FAU14, UT14]. This left-inverse might not be unique, and it is usually not easy to find one, especially when a stable or \mathbf{L}^p -stable inverse is required, cf. [UT14, Ch.5].

The case of the fractional Laplacians has been known from the 80s, cf. for instance [UT14, Sec.5.5.2], and the case of derivatives and directional derivatives was dealt with in [FAU14]. We tried to generalize these previous work for our elementary block $G_{\lambda,\alpha}$ but it turned out to be harder. We do not include more reminders on these stochastic processes, our reference being [UT14]⁶², mainly the chapters 5 and 7.

5.3.1 Details about the setting of our experiments

We were focusing on the effect of a simple operator, the **inverse** of an elementary block $G_{\lambda,\alpha}$. For $\lambda \in (0, 1)$, we already explained that its Fourier multiplier is non-canceling (except in $\mathbf{0}$), so its inverse can be applied by a simple point-wise division in the Fourier domain. We chose to handle the singularity by simply removing it: if $G_inv[i, j]$ was infinite (from a division by 0), we forced it to be 1. It only concerns one pixel (in Fourier), so hopefully it does not change the output image too much.

For our operators:

- We only consider 1-scale-invariant G , *i.e.*, $\gamma = 1$, so their inverse are all (-1) -SI.
- We tried many values λ , and the figures below use a value in $\lambda = 0, 0.25, 0.5, 0.75, 1$. It is a trade-of between the directionality of D_α and the isotropy of $(\Delta)^{1/2}$, as seen in practice, cf. the figures below. The higher the λ , the more directional the process will be:

$\lambda = 0$	can be interpreted as	purely isotropic
$\lambda = 0.25$	can be interpreted as	“almost isotropic”
$\lambda = 0.5$	can be interpreted as	“partly isotropic, partly directional”
$\lambda = 0.75$	can be interpreted as	“almost directional”
$\lambda = 1$	can be interpreted as	purely directional

⁶² If needed, cf. [UT14, Sec.7.5.3, p183] which is available on-line, *e.g.*, figure 7.9, or these slides (click).

- For the angle α , we tried values in $\{0, \pi/6, \pi/4, \pi/2, 3\pi/4, \pi\}$. Experiments with a turning α were done, on the same noise to really see the effect of α , in figure ??.
- We also experimented with the application of two inverse blocks, of same or different angle, α_1, α_2 , and the figure 18 and 19 below is for $\alpha_1 = 3\pi/4, \alpha_2 = 5\pi/4$.

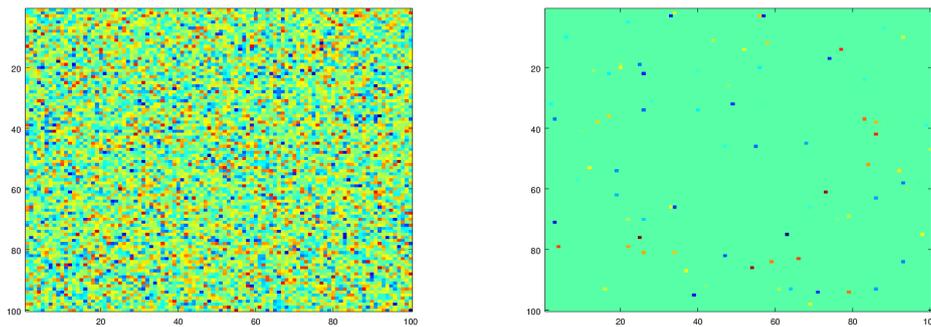
The setup of our sparse processes was the following:

- A 2D *i.i.d.* Gaussian “white” process, of scaling parameter $\sqrt{100}$, *i.e.*, every pixel of the image are all independent and identically distributed (*i.i.d.*), following the same (continuous) Gaussian law of parameter $\sigma = \sqrt{100}$.
- A 2D compound Gaussian Poisson⁶³ process of parameter $\nu = 0.01, \text{ or } 0.1, 1, 10$. In a nutshell, it consists in uniformly located spikes or deactivated pixels (*i.i.d.*). Each spikes is a sum of a certain number of Gaussians, all of parameter $\sigma = \sqrt{100}$, and the number of Gaussian (either 0 or more) follows a discrete Poisson law of parameter ν , which basically controls⁶⁴ the (average) number of spikes.
- Both were used to generate images of width and height of $N = 100$ pixels⁶⁵.
- Note that both Gaussian and Poisson are fully isotropic *in their formulation*, but a *realization* can be and *is usually not isotropic*.

For these sparse stochastic processes, we used an implementation of Gaussian and Gaussian compound Poisson noises in GNU Octave⁶⁶, that is still internal to the BIG team at EPFL, and not yet open-sourced.

5.3.2 List of experiments and their figures

Gaussian and compound Poisson noise



(a) Gaussian “white” noise

(b) Compound Poisson noise with $\nu = 0.1$

Figure 4: Two examples of realizations of Sparse Stochastic Processes

⁶³ For more details on compound Poisson and compound Gaussian-Poisson processes, refer for instance to this Wikipedia article, or some explanations can be found in [UT14, Ch.7].

⁶⁴ When Gaussian compound Poisson are used to models spikes for neurons, ν is usually referred as a “firing rate”. Cf. this course by J.-P. Nadal (MVA, 2016) or this article by S. Shinomoto.

⁶⁵ Our code is not yet optimized so it was slow for bigger images ($\simeq 3$ minutes by figure for $N = 200$).

⁶⁶ Which is also valid for MATLAB.

– First, let see two realizations of noise in figure 4, a white Gaussian on the left⁶⁷, and a compound Poisson on the right (with medium “firing rate” $\nu = 0.1$).

Illustrations with one inverse block

– One block $G_{\lambda,0}^{-1}$, with $\alpha = 0$, and an increasing $\lambda = 0, 0.25, 0.5, 0.75, 1$, applied on both noise, cf. figures 5–9. We start to observe some nice-looking patterns.

And for the higher values of λ , figures 8 and 9 shows very clearly a directional behavior. We observe an effect on this line ($\alpha = 0$, *i.e.*, east, along \rightarrow), which looks like a discrete integration: the positive spikes (red-orange) or the negative spikes (blue-cyan) get propagated, like if the spikes were (discretely) summed along the direction. The summation appears on the *same direction* as it is the *inverse* of a discrete derivative, generalizing the result known in 1D: the inverse of the discrete derivative D is a cumulative sum from the left to right (a discrete integration).

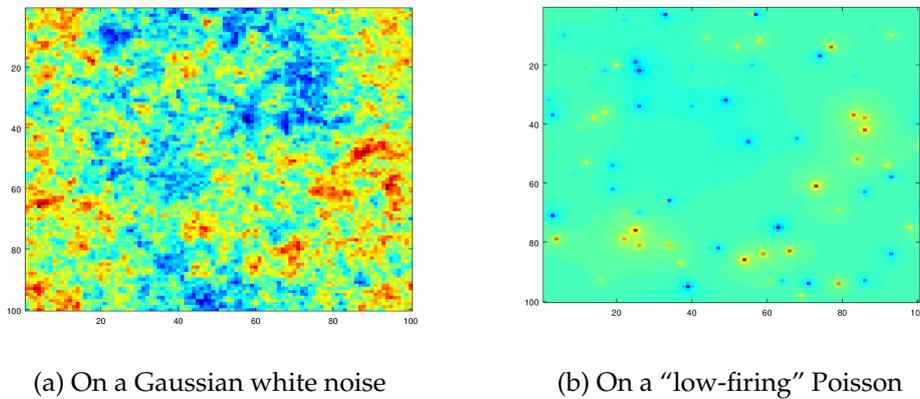


Figure 5: $G_{0,0}^{-1}$: purely isotropic, like a local integration

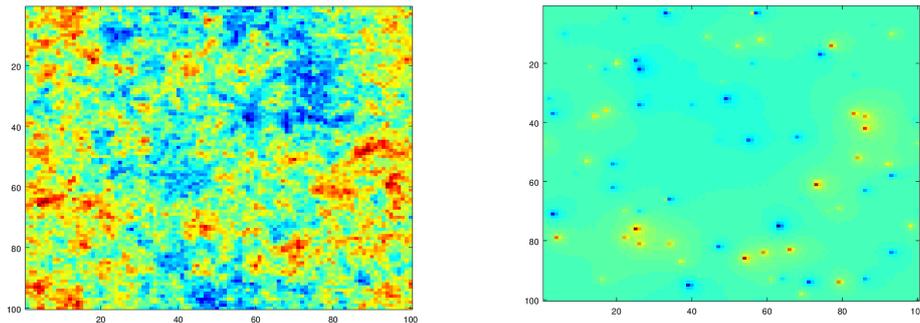
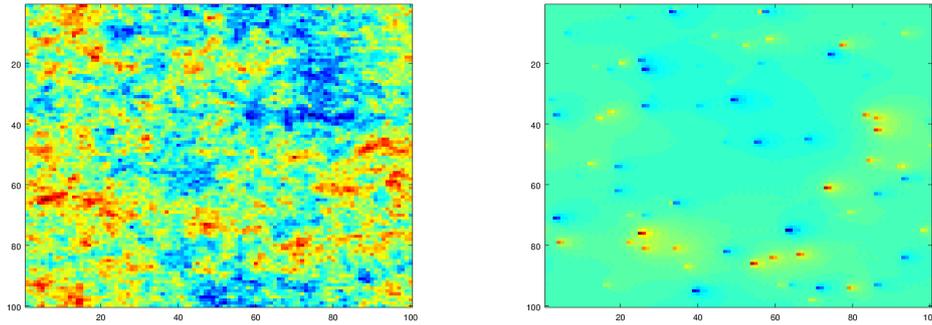
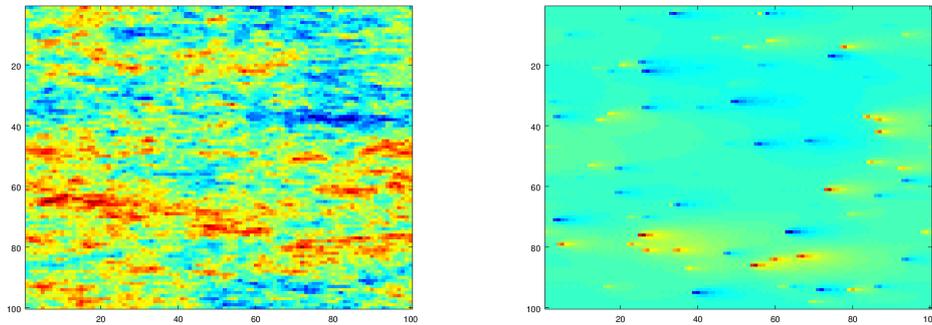
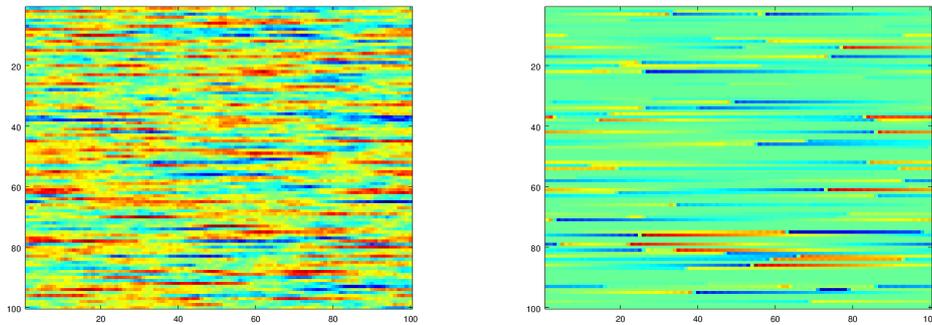


Figure 6: $G_{0.25,0}^{-1}$: not yet directional

– We also applied $G_{\lambda,5\pi/4}^{-1}$ to the same Gaussian white noise, for three different values of $\lambda = 0.25, 0.5, 0.75$, to really highlight the trade-of isotropic/directional, in figure 10. The directionality started to appear strongly for $\lambda = 0.75$, when we observed the same phenomenon, the emergence of a directional behavior, indicating again a direction of

⁶⁷ In all the next figures, the Gaussian will be on the left and the Poisson on the right.

Figure 7: $G_{0.50,0}^{-1}$: not much directionalFigure 8: $G_{0.75,0}^{-1}$: more directional, along \rightarrow Figure 9: $G_{1,0}^{-1}$: purely directional, along \rightarrow

integration (*i.e.*, a discrete cumulative sum), except it is along the line of angle $\alpha = \pi/6$ (north-west) this time.

This is exactly what we expected an inverse directional derivative (of angle $\pi/6$) would do, as in 1D the derivative is the inverse of the integration. It is very satisfying to check that it works similarly along a line in 2D. Remember that we implemented the inverse operator $G_{\lambda, 5\pi/6}^{-1}$ in the Fourier domain, with a point-wise multiplier by its inverse Fourier multiplier, and **not** by a space-domain numerical integration, but it is quite interesting to observe that they have a similar behavior!

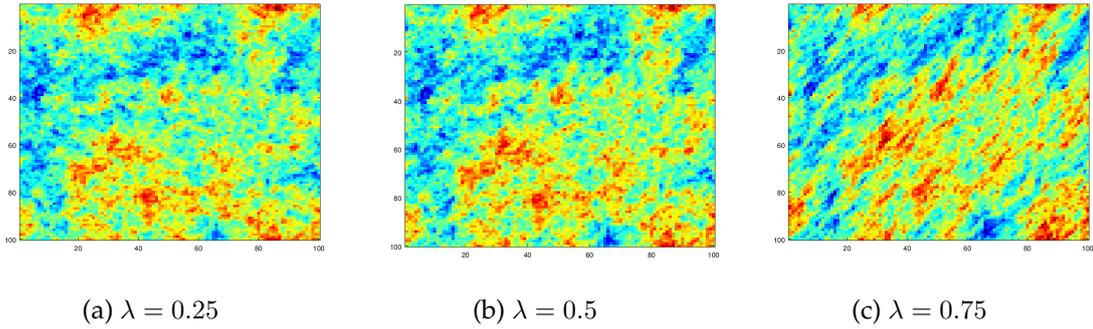


Figure 10: $G_{\lambda, \pi/6}^{-1}$ for $\lambda = 0.25, 0.5, 0.75$, on the same Gaussian white noise

– We also tried to observe the effect of the angle α , by applying $G_{0.75, \alpha}^{-1}$ and $G_{0.95, \alpha}^{-1}$, for three different values of $\alpha = 0, \pi/4, \pi/2$, applied to the same Gaussian, and with $\lambda = 0.75$ and then $\lambda = 0.95$ to see a strong but not full directionality. The figures 11 and 12 clearly shows the effect of the rotation when α increases (in the trigonometric sense, *i.e.*, anti-clockwise, ↶).

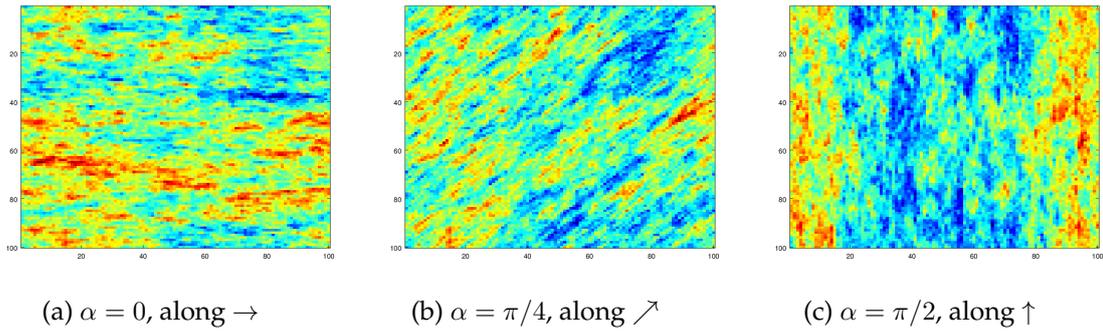


Figure 11: $G_{0.75, \alpha}^{-1}$ for $\alpha = 0, \pi/4, \pi/2$, on the same Gaussian white noise

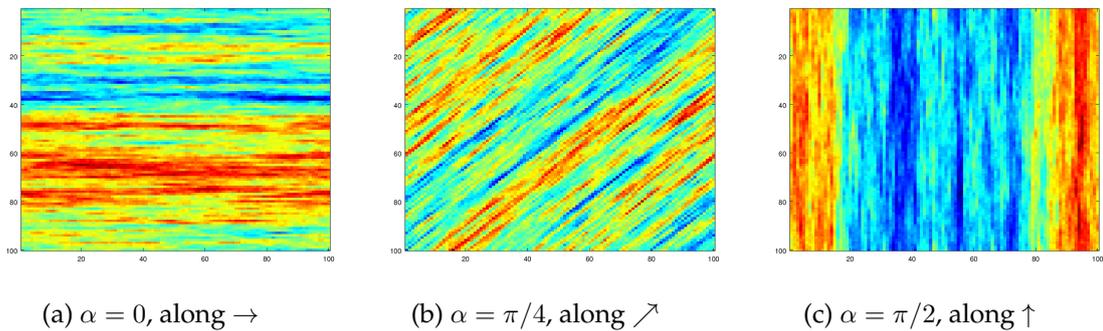


Figure 12: Very directional, $G_{0.95, \alpha}^{-1}$ for $\alpha = 0, \pi/4, \pi/2$, on the same noise

Mondrian and partially-directional Mondrian process

– Applying to both of them the same operator, $D_x^{-1}D_y^{-1}$, composed of two *purely-directional* blocks, $G_{1,0}^{-1}G_{1,\pi/2}^{-1}$, gives the “Mondrian process”⁶⁸, in figure 13. The left and right figures are similar, but slightly different: the (low-firing) Poisson has less vertices than the Gaussian.

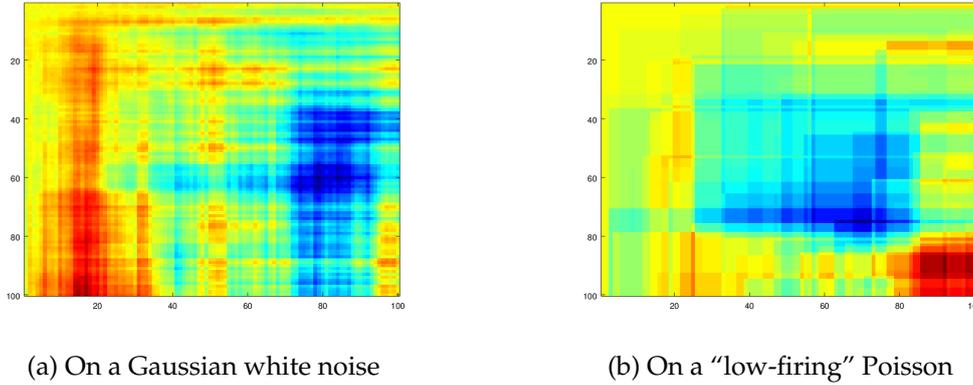


Figure 13: Purely directional, two orthogonal integrations, this is the “Mondrian” process

– And with $\lambda = 0.5$ and then 0.75 , applying the two *partly-directional* blocks, $G_{\lambda,0}^{-1}G_{\lambda,\pi/2}^{-1}$, gives a partly-directional “Mondrian” process, not too different from the purely-directional one, cf. figure 14.

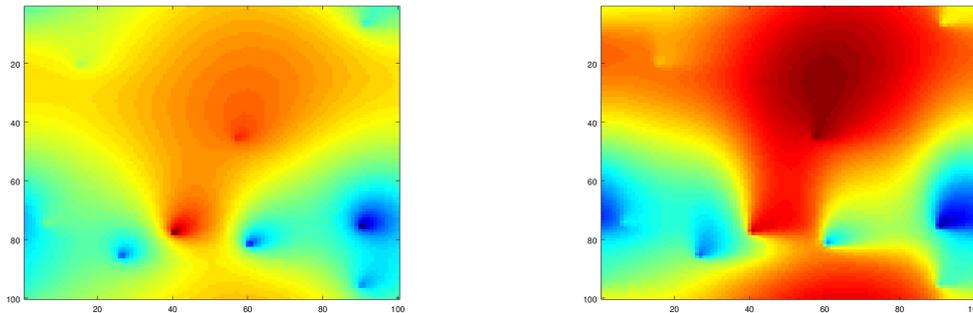


Figure 14: Partly directional “Mondrian” process, on a “low-firing” Poisson, for $\lambda = 0.5, 0.75$

This example is a very interesting one, as figure 14 shows how easy it is to generate new families of process, based on well-known examples, by adding a “fractional-isotropy” (*i.e.*, $\lambda < 1$) to purely-directional processes (*e.g.*, the Mondrian), or by adding a “fractional-directionality” (*i.e.*, $\lambda > 0$) to purely-isotropic processes (*e.g.*, white noise), thanks to our framework of 2D steerable convolutions.

We have great hope that this idea can lead to new and interesting stochastic sparse processes, and maybe to *fractional-directional fractals*⁶⁹

⁶⁸ It is a famous example in the BIG team, *e.g.*, [UT14, Fig.7.9], or the examples on the book website sparseprocesses.org. The “Mondrian process” consists in two discrete integration, along the x and y axis, of some sparsely localized spikes, and that is why it looks like rectangle aligned with the two axes.

⁶⁹ If the Lévy noise w is self-similar, then the sparse process s solution of $Ls = w$ can be seen as a *self-similar process*, *i.e.*, a fractal.

More illustrations with two inverse blocks

– We also tried to chain two elementary blocks, of the same angle, by applying twice the operator $G_{\lambda, -\pi/6}^{-1}$ to the same Gaussian noise (fig. 4a), for three different values of $\lambda = 0.3, 0.5, 0.8$, to continue to highlight the trade-of isotropic/directional, in figure 15.

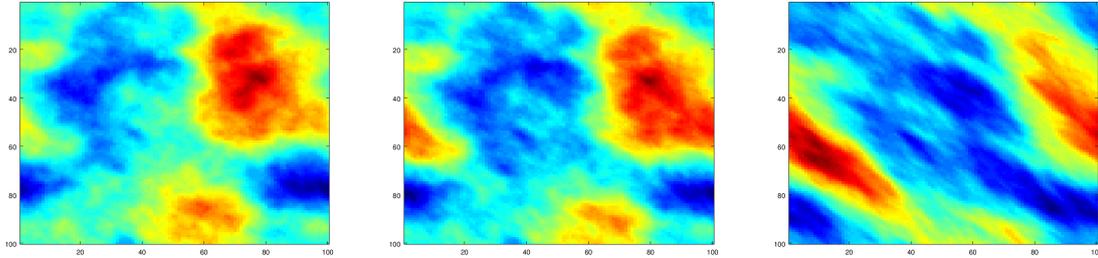


Figure 15: $G_{\lambda, -\pi/6}^{-1} G_{\lambda, -\pi/6}^{-1}$ for $\lambda = 0.3, 0.5, 0.8$ on the same Gaussian noise

– We applied twice the operator $G_{\lambda, -\pi/4}^{-1}$ to the same Gaussian noise, for higher values of $\lambda = 0.5, 0.8, 0.9$, in figure 16. As before, The isotropy of the Laplacian is almost entirely lost for $\lambda = 0.9$, showing a highly pronounced directionality. $\alpha = -\pi/4$ gives an image aligned along the line of angle $3\pi/4 = -\pi/4 \pmod{\pi}$, because inverting the operator $G_{\lambda, \alpha}$ implies inverting the underlying direction (*i.e.*, $\alpha \mapsto \alpha + \pi$):

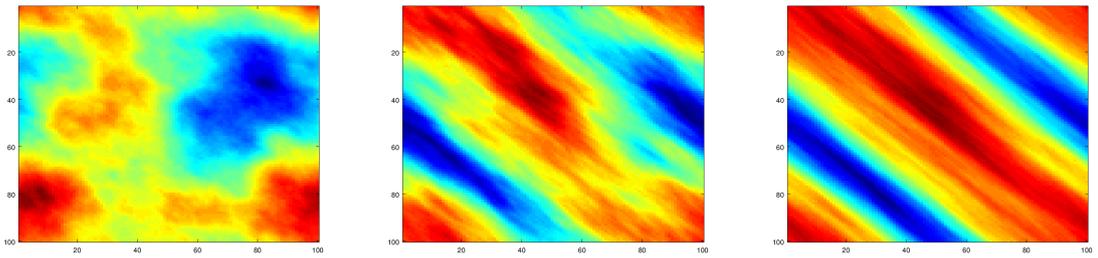
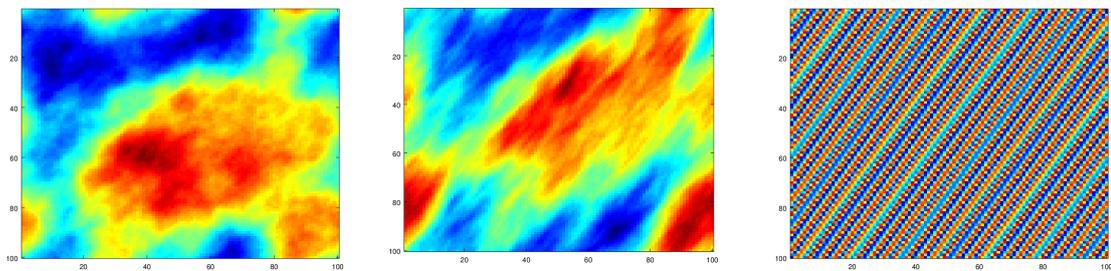


Figure 16: $G_{\lambda, -\pi/4}^{-1} \circ G_{\lambda, -\pi/4}^{-1}$ for $\lambda = 0.5, 0.8, 0.9$ on the same Gaussian noise

– Figure 17 shows the application of two blocks of same angle, $G_{\lambda, \pi/3}^{-1} G_{\lambda, \pi/3}^{-1}$ with $\alpha_1 = \alpha_2 = \pi/3$ and an increasing $\lambda = 0.5, 0.75, 1$, on the same Gaussian.



(a) “Not very directional”

(b) “Partly directional”

(c) “Very directional”

Figure 17: $G_{\lambda, \pi/3}^{-1} G_{\lambda, \pi/3}^{-1}$ for $\lambda = 0.5, 0.75, 1$ on the same Gaussian noise

– We also tried to chain two elementary blocks, of different angles, $G_{\lambda,3\pi/4}^{-1} \circ G_{\lambda,5\pi/4}^{-1}$, again for $\lambda = 0.3, 0.5, 0.8$, applied to a low-firing compound Poisson noise ($\nu = 0.01$, *i.e.*, very few spikes), in figure 18. Having a small number of spikes help to see the directionality increases when λ increases. The difference between the three sub-figures is obvious: the few discrete spikes get integrated along the two directions $3\pi/4$ and $5\pi/4$, and the resulting effect look like an integrated cone of opening angle $\pi/2 = 5\pi/4 - 3\pi/4$ in the left direction (west).

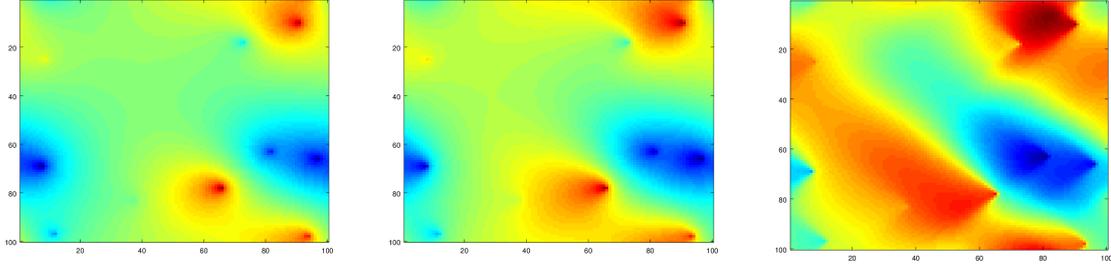


Figure 18: $G_{\lambda,3\pi/4}^{-1} \circ G_{\lambda,5\pi/4}^{-1}$ for $\lambda = 0.3, 0.5, 0.8$ on a low-firing Poisson noise ($\nu = 0.01$)

Remark 5.5. The patterns in figure 18 are quite similar to the result of some fluid-dynamic simulation on 2D diffusion⁷⁰..., and it not surprising as we simply integrated twice some spikes: this discrete integration is a cumulative sum, which behaves numerically like if fluid particles were diffusing around an obstacle.

– Our last illustration will be the application of the same operator $G_{\lambda,3\pi/4}^{-1} \circ G_{\lambda,5\pi/4}^{-1}$ for $\lambda = 0.3, 0.5, 0.8$, to the same *high-firing* compound Poisson noise ($\nu = 10$, *i.e.*, lots of spikes), in figure 19. This time, there is so many spikes that the realization of the Poisson noise was (almost) isotropic, and the inverse operator of different “directionality trade-off” λ (almost) all have the same visual effect.

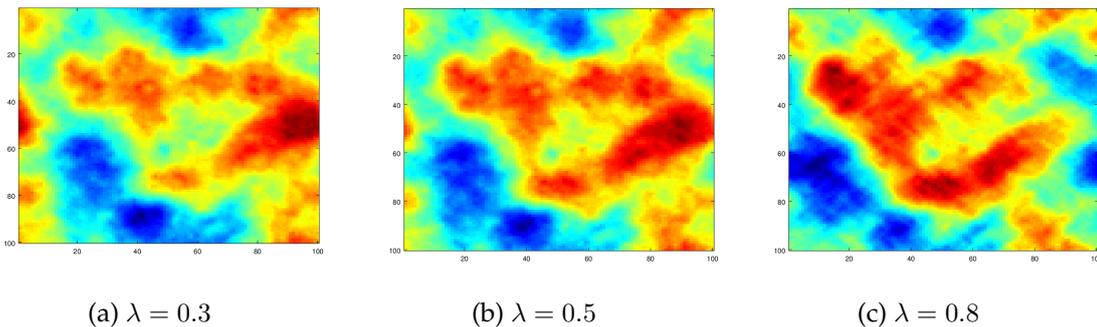


Figure 19: $G_{\lambda,3\pi/4}^{-1} \circ G_{\lambda,5\pi/4}^{-1}$ for $\lambda = 0.3, 0.5, 0.8$ on a high-firing Poisson noise ($\nu = 10$)

⁷⁰ For example, cf. this video youtu.be/vOFcHqImXJ8, based on the Navier-Stokes equations.

We cannot reasonably include more experiments or explanations, but we tried to include the most interesting ones. Each illustration was making a point that was explained shortly, and they all backed up the different intuitions we had about our elementary block $G_{\lambda,\alpha}$ and its inverse.

In particular, we really saw that $\lambda \in [0, 1]$ is a simple trade-of between isotropy and directionality, and that α is the direction of the (discrete) derivation (from $G_{\lambda,\alpha}$) and integration (from $G_{\lambda,\alpha}^{-1}$).

Remark 5.6. More figures are available on-line, see here (goo.gl/EJ7LFa).

Remark 5.7. We are thankful to Virginie Uhlmann (BIG team, EPFL) for her great help about the implementation of these 2D Sparse Stochastic Processes, and the γ -SI steerable convolutions and their inverses.

Conclusion

This last part presented a refinement on the main characterization of γ -SI 2D steerable convolution operators. It has several advantage, as its simplicity is appealing from a parametric point of view, and its modularity is useful for the implementation. Defining rich steerable operators is as simple as picking some directions α_i , and some values for λ_i , indicating the trade-of between the directionality of the derivative D_{α_i} and the isotropy of the half-Laplacian, and composing the blocks G_{λ_i,α_i} .

We explained how to implement both the fractional Laplacian and the block $G_{\lambda,\alpha}$, and their inverse numerically, in order to “integrate” some 2D isotropic compound Poisson noise, to illustrate the action of our operators. We have been able to observe the role of both parameters λ and α , and to start to understand the effect of an inverse block $G_{\lambda,\alpha}$.

Our experiments and implementation are still on-going research, and we will keep working on them in the future, especially for the article version of this report.

6 Conclusion

We defined in section 2 our functional operators, on $\mathbf{L}^2(\mathbb{R}^d, \mathbb{R})$ and on $\mathbf{S}(\mathbb{R}^d, \mathbb{R})$, and we explained what kind of structural as well as geometric properties they can have: linearity, continuity, translation-, scaling-, rotation-invariance or unity. We studied extensively the links between all these properties in section 3, and we presented many characterizations of different properties. Some were obvious and already known, but some were new and quite interesting. Combining properties usually gives elegant characterizations, on k_G for integral operators (*i.e.*, **LC**), and on g and \hat{g} for convolution operators (*i.e.*, **LC TI**). The universality of the fractional Laplacians among **LC TI RI γ -SI** (theorem 3.40), for instance, was not too hard to obtain, but it highlights the importance of this family of isotropic differential operators. To the best of our knowledge, this is the first attempt to summarize all these characterizations.

We then focused in section 4 on steerable convolution operators G , mainly in 2D, for which our main results consist on some characterizations, to first write them as a sum of modulated and iterated real Riesz transforms. This form was already known, and too rich to be used in practice, so we restrict to steerable homogeneous (*i.e.*, **γ -SI**) convolutions. Adding the scale-invariance gave a nicer form, as a composition of a fractional Laplacian $(-\Delta)^{\gamma/2}$, some directional derivatives D_{α_i} and an invertible 0-scale-invariant factor G_0 (of never-canceling Fourier multiplier, given as a certain trigonometric polynomial).

We also obtained a simpler form, writing G as a composition of elementary blocks G_{λ_i, α_i} , all being convex combination of the $(-\Delta)^{\gamma/2}$ and D_{α_i} . This last form is very appealing for implementation, and defining new steerable operators with a simple parametric interface (only need to pick the directions $\alpha_i \in [0, 2\pi]$, and the trade-of parameters $\lambda_i \in [0, 1]$). It is also interesting from our theoretical point-of-view, as we showed that G_{λ_i, α_i} are the only steerable **1-SI** operators of order less than 2. It yields a deep interpretation: a steerable G of order $n_G \leq K_G + 1$ is first written as a *sum* of modulated iterated Riesz up to a max frequency of K_G , and then as a *composition* of exactly K_G elementary blocks, all of order 1 or 2.

We did not have the time to do a lot of experiments, but we concluded by explaining the ideas behind an implementation of these steerable operators, and by presenting the results of some experiments on 2D stochastic processes in section 5.3. It helped to illustrate the effects of our elementary blocks, and of more complicated operators.

Our approach has the advantage of providing a nice theoretical framework for studying operators, especially steerable operators and 2D steerable **γ -SI** convolutions. We strongly believe that our contribution is plural, as this report summed up many known theoretical results and presented some new theorems, but also explained and illustrated some examples in 2D on sparse stochastic processes (on Gaussian or compound Poisson noise). Our main future direction of research is to conclude these experiments, and hopefully we will try to publish our main theoretical results on 2D steerable convolutions, illustrated by these experiments.

We only evoked about the other possible future works, of applying our 2D steerable **γ -SI** convolution operators to sampling (with new splines) and denoising and data recovery algorithms (with new Green's functions). We did not have the time to explore these directions, but most surely someone will continue on this topic in the BIG team at EPFL.

6.1 Other questions, possible future works

We also thought of some additional questions⁷¹ on our theory of steerable operators with geometric properties. We have not worked a lot on these questions, but however we wanted to mention them.

- Can we study more than “just” translation-invariance, maybe *invariance by affine transforms* ? It seems harder.
- For steerability, we mentioned quickly that $r_R \circ G \circ r_{R^{-1}}$ was the form of a group action, the conjugation action. Could we study *other group actions* ? Possibly, but we did not find any interesting one so far.
- We defined steerability in a general setting but only studied it in 2D. What happens for *steerability in 3D* ? It is known to be less easy, but is the difficulty coming from non-commutativity (of 3D rotations), from non-compactness ?
- For a non-steerable operator, can we “approximate” it with a steerable operator of a certain order? Is it possible to define a *projection on the set of steerable operators of bounded order* ?, like it is possible to project a continuous function on the set of polynomial of bounded degrees (e.g., with Lagrange interpolation). It seems a reasonable result, but we ran out of time to investigate it.
- Can we define *steerability for non-linear operator*? The same define will work, as the linearity of G is not needed for V_G to be a (finite dimension) vector space. However, we do not have Schwartz’s kernel for non-linear operator, so the theory will probably be harder to develop.
- Any *discrete theory of steerability*?

⁷¹ Thanks to Gabriel Peyré for his suggestions.

A Appendix – On unitary operators (U)

This additional section is devoted to the study of unitary operators, from their definition (on Euclidean or Hilbert spaces), to characterizations when some geometric properties are added. We also present and study an interesting family of unitary operators, generalizing the well-known Hilbert transform. Some additional results can be proved for unitary steerable and γ -SI steerable operators in 2D.

A.1 Unitary operators (U), definitions and examples

For the special case when $E = F = \mathbf{L}^2(\mathbb{R}^d, \mathbb{R})$, or more generally if both E and F have an inner product (*i.e.*, are Euclidean spaces), we can ask our operators $G : E \rightarrow F$ to be unitary, in the sense of the following definition⁷²:

Definition A.1 (Unitary operators). For $G \in \mathcal{L}(E, F)$, G is unitary (**U**) when

$$\forall f_1, f_2 \in E, \quad \langle G\{f_1\}, G\{f_2\} \rangle = \langle f_1, f_2 \rangle. \quad (35)$$

Proposition A.2. A unitary operator is also called an isometry.

Easy Proof. In an Euclidean space, preserving inner products is equivalent to preserving norms, by the polarization identity ($\langle \mathbf{x}, \mathbf{y} \rangle = (\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2)/4$). Therefore, an operator is unitary *iff* it preserves norms, which is the definition of an isometry. \square

We will denote **U** the class of unitary operators (not studied by itself), and by **LC TI U** the class of linear, continuous and unitary operators (on E).

Proposition A.3. This class **L U** is *not* a linear sub-space of $\mathcal{L}(E, F)$, but is stable by composition if $F = E$.

Easy Proof. A simpler counter example is $G_1 = \text{Id}$ and $G_2 = -\text{Id}$, which are both linear and unitary, but their sum is the null operator, which clearly does not conserve inner products. Stability by composition is easy. For two **L U** operators G_1, G_2 ,

$$\begin{aligned} \forall f_1, f_2 \in E, \quad & \langle G\{f_1\}, G\{f_2\} \rangle \\ &= \langle (G_1 \circ G_2)\{f_1\}, (G_1 \circ G_2)\{f_2\} \rangle \\ &= \langle G_1\{G_2\{f_1\}\}, G_1\{G_2\{f_2\}\} \rangle \\ &= \langle G_2\{f_1\}, G_2\{f_2\} \rangle = \langle f_1, f_2 \rangle. \end{aligned} \quad \square$$

A.2 Characterizations for unitary convolution operators

On $E = F = \mathbf{L}^2(\mathbb{R}^d, \mathbb{R})$, thanks to their Fourier multiplier, **LC TI** operators have a very simple characterization for being unitary.

Theorem A.4. $G \text{ LC TI}$ is unitary (**U**) *iff* $|\hat{g}(\omega)| = 1$ for (almost) all $\omega \in \mathbb{R}^d$.

⁷² Unitary operators only have a meaning on normed spaces, or on Euclidean spaces.

Proof. By definition, G is unitary iff it preserves the inner products. And on L^2 , we know that the Fourier transform is an isometry⁷³, so \mathcal{F} preserves the inner products. Let f_1, f_2 be two test functions (in $L^2(\mathbb{R}^d)$).

$$\begin{aligned}\langle G\{f_1\}, G\{f_2\} \rangle &= \langle \mathcal{F}\{G\{f_1\}\}, \mathcal{F}\{G\{f_2\}\} \rangle \\ &= \langle \widehat{g}f_1, \widehat{g}f_2 \rangle \\ &= \langle |\widehat{g}|^2 f_1, f_2 \rangle.\end{aligned}$$

And so, G is unitary iff $\langle G\{f_1\}, G\{f_2\} \rangle = \langle f_1, f_2 \rangle$ for all f_1, f_2 iff $\langle |\widehat{g}|^2 f_1, f_2 \rangle = \langle f_1, f_2 \rangle$ for all f_1, f_2 iff $|\widehat{g}(\omega)|^2 - 1 = 0$ for (almost) all $\omega \in \mathbb{R}^d$ iff $|\widehat{g}(\omega)| = 1$ for (almost) all $\omega \in \mathbb{R}^d$.

Note that for usual examples of operators, \widehat{g} is a true function, and this result will be $\forall \omega \in \mathbb{R}^d, |\widehat{g}(\omega)| = 1$. \square

Proposition A.5. *Let G be LC TI and unitary, if it is γ -SI then $\gamma = 0$.*

Easy Proof. Thanks to the previous theorem, and the characterization for LC TI γ -SI operators (theorem 3.35), it is obvious, as $\widehat{g}(\omega) = \|\omega\|^\gamma \widehat{g}(\omega/\|\omega\|)$ for $\omega \neq 0$, so $|\widehat{g}(\omega)| = \|\omega\|^\gamma |\widehat{g}(\omega/\|\omega\|)|$ and $|\widehat{g}(\omega)| = 1 = |\widehat{g}(\omega/\|\omega\|)|$, which implies that $\|\omega\|^\gamma = 1$ for all $\omega \neq 0$, so $\gamma = 0$. \square

Remark A.6. Note that we do **not** have that converse: a unitary convolution can be **not** scale-invariant! We could imagine for instance a Fourier multiplier like $\widehat{g}(r, \theta) = e^{jr \cos(\theta)}$, which is Hermitian symmetric ($\widehat{g}(r, \theta + \pi) = e^{jr \cos(\theta + \pi)} = e^{-jr \cos(\theta)} = \overline{e^{jr \cos(\theta)}} = \overline{\widehat{g}(r, \theta)}$). It clearly has a modulus equal of 1, but it is not γ -scale-invariant for any $\gamma \in \mathbb{R}_+$: $e^{jr \cos(\theta)} \neq r^\gamma e^{j \cos(\theta)}$. Indeed, for $\theta = 0$, it would give $e^{jr} = r^\gamma e^j$ which can only be true for $r \in 1 + 2\pi\mathbb{Z}$ but the relation has to be true for any $r \in \mathbb{R}_+$.

We tried for quite some time to obtain a characterization for unitary LC TI 0-SI operators, but we failed to obtain anything more explicit or more powerful than this (obvious) proposition:

Proposition A.7. *G is LC TI 0-SI and unitary iff \widehat{g} is scale-invariant and of modulus 1.*

Proof. Thanks to the previous proposition, and the characterization for LC TI 0-SI operators (theorem 3.35) (iff \widehat{g} is scale-invariant), it is obvious.

In such cases, \widehat{g} has its image in \mathbb{S}^1 (complex unit circle), and can be restricted to \mathbb{S}^{d-1} (unit sphere of \mathbb{R}^d) thanks to its scale-invariance. We tried to characterize such $\widehat{g}|_{\mathbb{S}^{d-1}} : \mathbb{S}^{d-1} \rightarrow \mathbb{S}^1$, but we failed. The more promising direction was the use of some *path lifting theorem*⁷⁴ [Bil04], but in this general setting there is no hope of imposing an hypothesis of continuity on \widehat{g} , making it useless in general. \square

⁷³ With the good choice of constant for \mathcal{F} , as explained in the introduction 1.3.1.

⁷⁴ For a continuous (or C^k) function $f : \Omega \rightarrow \mathbb{C}$, if f has module 1, there exists a continuous (or C^k) “path lifting” function Θ_f such that $f(\mathbf{x}) = \exp(j\Theta_f(\mathbf{x}))$.

And quite directly, we also obtain a very narrow result, showing that the identity is the only γ -SI convolution to be both unitary and rotation-invariant.

Proposition A.8. *If G is LC TI γ -SI and unitary, it is RI iff it is the identity.*

Proof. The converse is obvious, and the direct sense is quick thanks to the previous theorem, and the characterizations of rotation-invariant and 0-scale-invariant operators, theorems 3.35 and 3.39. Combining the three results gives directly that $\forall \omega \in \mathbb{R}^d, \hat{g}(\omega) = 1$, i.e., $G = \text{Id}$. \square

In other words, the only operator to be LC TI RI γ -SI and unitary is the identity.

We tried to obtain more general results on unitary operators, but nothing else interesting came out of this research in the general setting. The next part introduces a parametric family of unitary operators in any dimension (the fractional-directional Hilbert transforms), and we prove that it is universal in dimension $d = 1$, and that no parametric family can be universal in higher dimension. However, the closure of this family will turn out to be a dense family in dimension $d = 2$.

A.3 An extended example: the fractional-directional Hilbert transforms fdHT

We start by reminding what is the classic Hilbert transform, and some of its earlier generalizations, and then we introduce the family of fractional-directional Hilbert transform, in any dimension d . We focus on the specific case of dimension $d = 2$, where the family will not be universal but its closure for composition will be dense in the class of unitary LC TI 0-SI operators on $L(\mathbb{R}^2, \mathbb{R})$.

We will explain why, but this is the best result that can be achieved when we look for a parametrization of such operators. This negative result should be opposed to the universality of the fractional Hilbert transform in dimension $d = 1$, but it should not be seen as a failure: the class of possible Fourier multiplier for unitary LC TI 0-SI in higher dimensions is simply too rich to be parametrized.

A.3.1 Classic Hilbert transform and earlier generalizations in dimension 1 and 2

In dimension $d = 1$, the classic *Hilbert transform* H is defined on $L^2(\mathbb{R}, \mathbb{R})$ by its simple Fourier multiplier⁷⁵:

$$H\{f\}(x) \xleftrightarrow{\mathcal{F}} -j \text{sign}(w) \hat{f}(w). \quad (36)$$

It is well-known, for instance in optics, and it has been used since the seventies. We will not detail how it is used and implemented, and why it is interesting, but the interested reader could consult for instance [LMZ96].

Let us expose two earlier generalizations in dimension 1 and 2.

⁷⁵ One can check quickly that this 1D Fourier multiplier $\hat{H}(w) = -j \text{sign}(w)$ is indeed Hermitian symmetric.

In dimension $d = 1$

A well-known generalization is the *fractional Hilbert transform* H_τ , as presented for instance in [CU10, Part 2.1] or [LMZ96], and it is defined on $\mathbf{L}^2(\mathbb{R}, \mathbb{R})$ by:

$$H_\tau \stackrel{\text{def}}{=} \cos(\pi\tau)\text{Id} - \sin(\pi\tau)H.$$

It contains the special cases $\text{Id} = H_0$ and $H = H_{-1/2}$, it satisfies a nice composition property: $H_{\tau_1} \circ H_{\tau_2} = H_{\tau_1 + \tau_2}$ and its Fourier multiplier is given by $w \mapsto \exp(j\pi\tau \text{sign}(w))$.

In dimension $d = 2$

A generalization was introduced and studied in [CU10, Part 4, p15], with a parameter $\theta \in [0, 2\pi)$, and by writing $\vec{u}_\alpha \stackrel{\text{def}}{=} [\cos(\alpha) \ \sin(\alpha)]^T$ the unitary vector of angle α , the 2-dimension *directional Hilbert transform* H_α is defined on $\mathbf{L}^2(\mathbb{R}^2, \mathbb{R})$ by:

$$H_\alpha\{f\}(\mathbf{x}) \xleftrightarrow{\mathcal{F}} -j \text{sign}(\vec{u}_\alpha^T \boldsymbol{\omega}) \hat{f}(\boldsymbol{\omega}).$$

It has also nice properties, but we leave the reader refers to [CU10] for more details.

A.3.2 Definition of the fdHT

The fractional-directional Hilbert transform is a parametric operator, with two parameters, that generalize the Hilbert transform H (and its previously introduced generalization) to operate on $\mathbf{L}^2(\mathbb{R}^d, \mathbb{R})$ and not simply on $\mathbf{L}^2(\mathbb{R}, \mathbb{R})$ or $\mathbf{L}^2(\mathbb{R}^2, \mathbb{R})$. The form of its Fourier multiplier is very simple, but the resulting family is surprisingly interesting.

Definition A.9 (fdHT). For $\tau \in \mathbb{R}$ a scalar and $\vec{u} \in \mathbb{S}^{d-1}$ a unitary vector (in \mathbb{R}^d).

We denote by $H_{\tau, \vec{u}}$ the **LC TI** operator on $E = \mathbf{L}^2(\mathbb{R}^d)$, defined by its Fourier multiplier:

$$\forall \boldsymbol{\omega} \in \mathbb{R}^d, \widehat{H_{\tau, \vec{u}} f}(\boldsymbol{\omega}) \stackrel{\text{def}}{=} \exp(j\pi\tau \underbrace{\text{sign}(\vec{u}^T \boldsymbol{\omega})}_{\in \{-1, 1\}}). \quad (37)$$

It is called the fractional directional Hilbert transform⁷⁶ of parameters τ, \vec{u} (**fdHT**).

Easy Proof. This operator is obviously **LC TI** as we define it by its (measurable) Fourier multiplier $\widehat{H_{\tau, \vec{u}} f}(\boldsymbol{\omega})$. We need to check that it maps real-valued functions to real-valued functions, and thanks to theorem 3.17, it is enough to check that its Fourier multiplier is Hermitian symmetric:

$$\begin{aligned} \widehat{H_{\tau, \vec{u}} f}(-\boldsymbol{\omega}) &= \exp(j\pi\tau \text{sign}(\vec{u}^T(-\boldsymbol{\omega}))) = \exp(-j\pi\tau \text{sign}(\vec{u}^T \boldsymbol{\omega})) \\ &= \overline{\exp(j\pi\tau \text{sign}(\vec{u}^T \boldsymbol{\omega}))} = \overline{\widehat{H_{\tau, \vec{u}} f}(\boldsymbol{\omega})}. \quad \square \end{aligned}$$

Warning A.10. The $\text{sign}(\dots)$ term has to be **inside** the exponential, not outside, in order to have the Hermitian symmetry. Indeed, $\overline{e^{j\theta}} = e^{-j\theta} \neq -e^{j\theta} = e^{j(\theta+\pi)}$.

⁷⁶ "Fractional" because of the factor $\tau \in \mathbb{R}$, and "directional" because of the direction vector $\vec{u} \in \mathbb{S}^{d-1}$.

A.3.3 Elementary properties of the fdHT

Some properties of the fdHT operators are easy to obtain.

Proposition A.11. Any fdHT operator $H_{\tau, \vec{u}}$ is also 0-SI, and unitary (**U**).

Easy Proof. • ω only appears in $\text{sign}(\vec{u}^T \omega)$ in the expression of its Fourier multiplier, and $\text{sign}(\vec{u}^T(a\omega)) = \text{sign}(a(\vec{u}^T \omega)) = \text{sign}(\vec{u}^T \omega)$ for any $a > 0$, so its Fourier multiplier is scale-invariant, hence the operator $H_{\tau, \vec{u}}$ is 0-SI.

- By definition, $|\widehat{H_{\tau, \vec{u}}}(\omega)| = 1$. By the characterization A.4 of unitary LC TI operators, this gives directly that $H_{\tau, \vec{u}}$ is unitary. \square

Lemma A.12. Our fractional-directional Hilbert transform $H_{\tau, \vec{u}}$ can also be seen as:

$$H_{\tau, \vec{u}} = \cos(\pi\tau)\text{Id} - \sin(\pi\tau)H_{\vec{u}}. \quad (38)$$

With $H_{\vec{u}}$ the (purely) directional Hilbert transform ($\widehat{H_{\vec{u}}}(\omega) = -j \text{sign}(\vec{u}^T \omega)$), cf. [CU10, Part 4, p15].

Easy Proof. We simply observe that the real and imaginary parts of the Fourier multiplier of $H_{\tau, \vec{u}}$ are respectively even and odd:

$$\begin{aligned} \widehat{H_{\tau, \vec{u}}}(\omega) &= \exp(j\pi\tau \text{sign}(\vec{u}^T \omega)) \\ &= \underbrace{\cos(\pi\tau \text{sign}(\vec{u}^T \omega))}_{\text{is even}} + j \underbrace{\sin(\pi\tau \text{sign}(\vec{u}^T \omega))}_{\pm 1} \\ &= \cos(\pi\tau) + j \sin(\pi\tau) \text{sign}(\vec{u}^T \omega) \\ &= \cos(\pi\tau) - \sin(\pi\tau) (-j \text{sign}(\vec{u}^T \omega)). \\ &= \cos(\pi\tau)\widehat{\text{Id}}(\omega) - \sin(\pi\tau)\widehat{H_{\vec{u}}}(\omega). \end{aligned} \quad (39) \quad \square$$

Corollary A.13 (In dimension 2). Remember that H_α is the (purely) directional Hilbert transform in dimension 2, defined by $\widehat{H_\alpha}(r, \theta) = -j \text{sign}(\cos(\theta - \alpha))$, then if $\vec{u} = \vec{u}_\alpha$,

$$H_{\tau, \alpha} \stackrel{\text{def}}{=} H_{\tau, \vec{u}_\alpha} = \cos(\pi\tau)\text{Id} - \sin(\pi\tau)H_\alpha.$$

So we recognize in $\widehat{H_{\tau, \alpha}}(r, \theta)$ a term extremely similar to the radial dependency of one of our “elementary block” $G_{\lambda, \alpha'} = \lambda D_{\alpha'} + (1 - \lambda)(-\Delta)^{1/2}$ (up-to a real constant C).

Easy Proof. We can manipulate directly its Fourier multiplier:

$$\begin{aligned} |\omega| \widehat{H_{\tau, \alpha}}(\omega) &= r \widehat{H_{\tau, \alpha}}(r, \theta) \\ &= r (\cos(\pi\tau) + j \sin(\pi\tau) \text{sign}(\cos(\theta - \alpha))) \\ &= \cos(\pi\tau)r + \sin(\pi\tau)jr \text{sign}(\cos(\theta - \alpha)) \\ &= \underbrace{C}_{\in \mathbb{R}} \left((1 - \lambda) \underbrace{r}_{= (-\Delta)^{1/2}(r, \theta)} + \underbrace{\lambda}_{\in [0, 1]} \underbrace{(jr \text{sign}(\cos(\theta - \alpha')))}_{= \widehat{D_{\alpha'}}(r, \theta)} \right). \end{aligned}$$

With $c = \cos(\pi\tau) + \sin(\pi\tau)$, and

- If $c = 0$, $C = 2 \cos(\pi\tau)$, $\lambda = \frac{1}{2} \in [0, 1]$, $\alpha' = \alpha + \pi$.

- Else if $c \neq 0$, $C = c \operatorname{sign}(\frac{\sin(\pi\tau)}{c})$, $\lambda = \frac{\sin(\pi\tau)}{c} \operatorname{sign}(\frac{\sin(\pi\tau)}{c}) \in [0, 1]$, $\alpha' = \alpha + \pi$ if $\frac{\sin(\pi\tau)}{c} \leq 0$ else $\alpha' = \alpha$. \square

Remark A.14. Unfortunately, we have not been able to exploit this similarity. The previous corollary is not very useful, but we wanted at least to expose this similarity in 2D between the **fdHT** $H_{\tau, \alpha}$ and the elementary steerable convolutions $G_{\lambda, \alpha \dots}$

The **fdHT** family also has an interesting algebraic property, similar to the one of the group of rotation:

Proposition A.15. For a fixed \vec{u} , every family $\{H_{\tau, \vec{u}}, \tau \in \mathbb{R}\}$ is a semi-group:

- $H_{\tau_1, \vec{u}} \circ H_{\tau_2, \vec{u}} = H_{\tau_1 + \tau_2, \vec{u}}$ for the composition law⁷⁷,
- $\text{Id} = H_{0, \vec{u}}$ is neutral for the composition,
- And the inverses⁷⁸ of $H_{\tau, \vec{u}}$ is **any** operator of the set $\{H_{-\tau + 2k, \vec{u}}, k \in \mathbb{Z}\}$.

Easy Proof. The three points are easy to check, we let the interested reader do it. \square

There is no similar result of stability by compositions when two different directions ($\vec{u}_1 \neq \vec{u}_2$) can be used, unfortunately, but we also have this result.

Lemma A.16. $H_{\tau, \vec{u}}$ also has an inverse with the same τ but of inverse direction: $H_{\tau, -\vec{u}}$.

Easy Proof. It is obvious from

$$\widehat{H_{\tau, \vec{u}}}(\omega) = \widehat{H_{\tau, \vec{u}}}(-\omega) = \widehat{H_{\tau, -\vec{u}}}(\omega),$$

and the fact that $\widehat{H_{\tau, \vec{u}}}(\omega)$ has modulus 1, so $\widehat{H_{\tau, \vec{u}}}(\omega) \widehat{H_{\tau, \vec{u}}}(\omega) = 1$. \square

We will then consider the family of operators composed of any finite compositions of **fdHT**. Considering limits of such compositions will be interesting, it will give the closure by composition and inverse of the family of composed **fdHT**. The closure by inverse is obvious from lemma A.16 but it is a nice property. It turns out that this last family is dense in 2D, but not in higher dimension.

A.3.4 Implementation of the **fdHT**

We implemented the **fdHT** operators, in a Python v2/v3 script, using `numpy` arrays and the Fast Fourier Transform (FFT) and Inverse FFT implementation from `numpy` (cf. `fft`, `ifft`). Our implementation is open-source, available on-line, and comes with a documentation and examples.

A **fdHT** is represented by its parameter $\tau \in \mathbb{R}$ and direction $\vec{u} \in \mathbb{S}^{d-1}$, and we apply it to real-valued images using the usual process, presented above in figure 3.

⁷⁸ It highlights the presence of a semi-group morphism between $\{H_{\tau, \vec{u}}, \tau \in \mathbb{R}\}$ for \circ and \mathbb{R} for $+$, or a group morphism if τ is restricted to $\mathbb{R}/2\mathbb{Z}$ the 2-periodic torus, instead of the real line \mathbb{R} .

⁷⁸ This set has a structure of a semi-group only, because there is not a unique inverse.

The implementation was quick to obtain, but we spent quite some time experimenting and trying to understand the behavior of the **fdHT** operators. We applied them to square images (in 2D only), of different kind, firstly on purely isotropic images, then on simple patterns and finally on black-and-white image (*e.g.*, Lenna, of course). For a reason we failed to understand, the discrete function we implemented from our formulas was not a real operator, even if the continuous formulas ensures that it is (as $\widehat{H_{\tau, \vec{u}}}$ is Hermitian symmetric). That is why the images and animations we got are not simple black-and-white images, but complex-valued images. We used this other script to associate a clear color-map to complex values. The figures we obtained are on-line, see here (goo.gl/P5huxH).

Unfortunately, the results of our experiments are not so satisfactory, hence we prefer to not spend much time detailing them. A curious reader can refer to the code and the figures on-line.

A.3.5 Non-universality but density of the fdHT among unitary 0-SI convolutions

In dimension 1, there is a strong characterization of **LC TI 0-SI** and unitary operators as fractional Hilbert transforms. Indeed, [CU10, Th.3.1, p8] showed that the fractional Hilbert transform on $\mathbf{L}^2(\mathbb{R}, \mathbb{R})$ are exactly the unitary linear operators invariant to translations and scaling. In other words, on $\mathbf{L}^2(\mathbb{R}, \mathbb{R})$, a unitary **LC TI St** operator G can be written as $G = H_{\tau}$ for some $\tau \in \mathbb{R}$ (and τ is unique if is taken in $(-1, 1]$).

We would like to get the same kind of result in any dimension d , and that was the motivation behind our study of this d -dimensional generalization of the 1-d fractional Hilbert transform. Without explaining all the research process, we can simply say that we have been able to answer this question negatively.

Theorem A.17 (No universal parametric family in dimension $d \geq 2$).

*In any dimension $d \geq 2$, any family \mathcal{A} of operators parametrized by a finite number of parameters (reals or integers) cannot be universal in the class of unitary **LC TI 0-SI** operators on $\mathbf{L}^2(\mathbb{R}^d, \mathbb{R})$.*

Proof. As we already seen it, the class of possible Fourier multipliers is very rich, as the only condition always imposed on \widehat{g} is its modulus being 1, its measurability and 0-scale-invariance, *i.e.*, it can be restricted to $\widehat{g} : \mathbb{S}^{d-1} \rightarrow \mathbb{S}^1$. The set of measurable functions from \mathbb{S}^{d-1} to $\mathbb{S}^1 = \mathbb{U}$ is not finite dimensional, as for instance it contains the infinite-dimensional family $\omega \rightarrow e^{jk\omega_1}$ (for $k \in \mathbb{Z}$). \square

Instead of a universal family (a basis), we can try to exhibit a dense family. In dimension bigger than 2, we failed to find a family of unitary convolutions that was both simple to work with and dense, and we strongly believe that it is unachievable. But in dimension 2, the family introduced above, constituted of any finite number of compositions of **fdHT**, turns out to be dense in the unitary convolutions.

Theorem A.18 (Density of composed **fdHT** among 2D unitary **LC TI 0-SI** operators).

*Any 2D unitary **LC TI 0-SI** operator G can be approximated by a sequence G_n , each being a composition of at most n **fdHT** operators, in the sense of $g_n \rightarrow g$ in \mathbf{S}' or $\widehat{g}_n \rightarrow \widehat{g}$ as a simple limit of measurable functions (*i.e.*, point-wise limit).*

Long and Hard Proof. ⁷⁹

We work with Fourier multipliers instead of impulse response. The core idea of the proof is to be able to compose two **fdHT** Fourier multipliers to obtain a Hermitian-symmetric function that acts like the indicator function of a small angular sector on \mathbb{S}^1 (as small as needed).

More formally, let $\alpha \in \mathbb{R}$ and $\varepsilon > 0$. Let $A_{\alpha,\varepsilon}$ be the (possibly small) angular sector, surrounding α , of size ε : $A_{\alpha,\varepsilon} = (\alpha - \varepsilon/2, \alpha + \varepsilon/2) \bmod 2\pi$ (an open interval). And let $B_{\alpha,\varepsilon} = A_{\alpha,\varepsilon} \uplus (A_{\alpha,\varepsilon} + \pi)$ (the later being a Minkowsky sum). $B_{\alpha,\varepsilon}$ is stable by $\theta \mapsto \theta + \pi$. See figure 20 for an illustration of the two angular sectors.

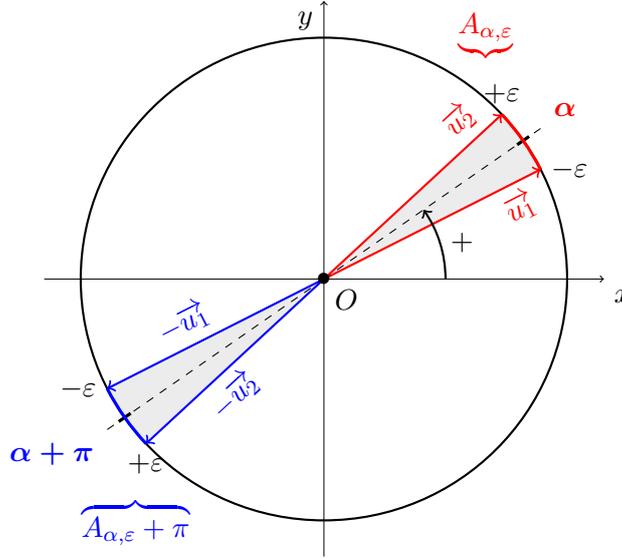


Figure 20: Two angular sectors $A_{\alpha,\varepsilon}$ and $A_{\alpha,\varepsilon} + \pi (= A_{\alpha+\pi,\varepsilon})$, in 2D.

Let $\beta \in \mathbb{R}$, and let $\widehat{h_{\alpha,\varepsilon}}(\omega) = \widehat{h_{\alpha,\varepsilon}}(r, \theta) = \exp(j\beta (\mathbb{1}_{A_{\alpha,\varepsilon}}(\theta) - \mathbb{1}_{A_{\alpha,\varepsilon}}(\theta + \pi)))$, it has a constant modulus of 1, and is Hermitian symmetric (*H-s*):

$$\begin{aligned} \widehat{h_{\alpha,\varepsilon}}(-\omega) &= \widehat{h_{\alpha,\varepsilon}}(r, \theta + \pi) = \exp(j\beta (\mathbb{1}_{A_{\alpha,\varepsilon}}(\theta + \pi) - \mathbb{1}_{A_{\alpha,\varepsilon}}(\theta + 2\pi))) \\ &= \overline{\exp(j\beta (\mathbb{1}_{A_{\alpha,\varepsilon}}(\theta) - \mathbb{1}_{A_{\alpha,\varepsilon}}(\theta + \pi)))} = \overline{\widehat{h_{\alpha,\varepsilon}}(r, \theta)} = \overline{\widehat{h_{\alpha,\varepsilon}}(\omega)}. \end{aligned}$$

And $\widehat{h_{\alpha,\varepsilon}}$ can be written as a product of two Fourier multipliers of **fdHT**. Let $u_1 = \vec{u}_{\alpha-\varepsilon/2}$ and $u_2 = \vec{u}_{\alpha+\varepsilon/2}$ be two unitary directional vectors, corresponding to the left- and right-limit of $B_{\alpha,\varepsilon}$. Then we can check that

$$\widehat{h_{\alpha,\varepsilon}}(\omega) = \exp\left(j\frac{\beta}{2} \text{sign}(u_1 \omega^T)\right) \exp\left(-j\frac{\beta}{2} \text{sign}(u_2 \omega^T)\right). \quad (40)$$

Proof. To prove (40), we simply need to check that it is true in the four different domains: between u_1 and u_2 , on the left of both, between $-u_2$ and $-u_1$, and on the right of both (in the trigonometric sense, cf. figure 20). \square

⁷⁹ **Beware** that this proof is long and technical, and make use of some notions coming from Lebesgue's theory of integration. For a reminder, see for instance these Wikipedia articles: on Lebesgue integration or on simple functions.

So we can write h as a composition of two **fdHT**, $h_{\alpha,\varepsilon} = H_{\beta/2,u_1} \circ H_{\beta/2,-u_2} = H_{\beta/2,\vec{u}_{\alpha-\varepsilon/2}} \circ H_{\beta/2,\vec{u}_{-(\alpha+\varepsilon/2)}}$. Now, let us explain how we can use $\widehat{h_{\alpha,\varepsilon}}$ as an elementary building block. We aim at approximating any \widehat{g} , that is known to be H -s, of modulus 1, and measurable.

We know⁸⁰ that the simple functions⁸¹ are dense among the measurable functions (for the point-wise limit), and it is still true for H -s functions on \mathbb{S}^1 to \mathbb{S}^1 (*i.e.*, functions of θ only). And H -s step functions are dense among the H -s simple functions. So we can just show that our building block can be compose to approximate a H -s unitary simple function (for the point-wise limit).

A H -s unitary simple function f on $[0, 2\pi]$ will be written $f(\theta) = \sum_{k=-K}^K e^{j\beta_k} \mathbb{1}_{A_k}(\theta)$, for $K \in \mathbb{N}$ and $\beta_k \in \mathbb{R}$, and a partition of measurable sets $\bigsqcup_{k=-K}^K A_k$ of $[0, 2\pi]$. The coefficients in the linear combination are of modulus 1, so they can be written as $e^{j\beta_k}$. The Hermitian symmetry imposes that $f(\theta) \neq 0 \Leftrightarrow f(\theta + \pi) \neq 0$, so, up-to a new indexing, we can have $A_0 = \emptyset$, and $\theta \in A_k \Leftrightarrow \theta + \pi \in A_{-k}$ for all $k > 0$, and so the symmetry gives $\beta_k = -\beta_{-k}$. Now, we can write (or at least approximate) A_k as a (limit of) finite union(s) of open intervals because the measurable sets can be written as (possibly infinite, but always countable⁸²) unions of Borelian sets, which can be taken as left- and right-open intervals (in dimension 1, on $[0, 2\pi]$).

Now the trick is to write $f(\theta)$ as

$$\begin{aligned} f(\theta) &= \sum_{k=-K}^K e^{j\beta_k} \mathbb{1}_{A_k}(\theta) = \sum_{k=1}^K e^{j\beta_k} \mathbb{1}_{A_k}(\theta) + e^{-j\beta_k} \mathbb{1}_{A_{-k}}(\theta) \\ &= \sum_{k=1}^K e^{j\beta_k} \mathbb{1}_{A_k}(\theta) + e^{-j\beta_k} \mathbb{1}_{A_k}(\theta - \pi) \\ &= \sum_{k=1}^K \exp(j\beta_k \mathbb{1}_{A_k}(\theta) - j\beta_k \mathbb{1}_{A_k}(\theta + \pi)) \\ &= \exp\left(j \sum_{k=1}^K \beta_k (\mathbb{1}_{A_k}(\theta) - \mathbb{1}_{A_k}(\theta + \pi))\right) \end{aligned}$$

This last step is wrong in the general case, but true here because we have a partition: only one term is non-zero at a time.

$$f(\theta) = \prod_{k=1}^K \exp(j\beta_k (\mathbb{1}_{A_k}(\theta) - \mathbb{1}_{A_k}(\theta + \pi)))$$

We recognize the product of K terms with the same form, $\exp(j\beta_k (\mathbb{1}_{A_k}(\theta) - \mathbb{1}_{A_k}(\theta + \pi)))$. So if $A_k = \bigsqcup_{i=1}^{N_k} I_{i,k}$ is partitioned with a finite number N_k of open disjoint intervals, $f(\theta)$ becomes a (big) product of $M = \sum_{k=1}^K N_k$ terms,

$$f(\theta) = \prod_{k=1}^K \exp(j\beta_k (\mathbb{1}_{A_k}(\theta) - \mathbb{1}_{A_k}(\theta + \pi)))$$

⁸⁰ Cf. this Wikipedia article if needed, or in French.

⁸¹ Definition: a simple function is a finite linear combination of indicator functions of measurable sets.

⁸² It is very important to have (in the worse scenario) a countable family, and not any infinite family, as we want at the end to chain all these approximations and get one sequence, indexed by \mathbb{N} , *i.e.*, a countable sequence. This result appeared to not be known with a unique name, one name we found was the Lindelöf lemma.

$$\begin{aligned}
&= \prod_{k=1}^K \exp \left(j\beta_k \left(\sum_{i=1}^{N_k} \mathbb{1}_{I_{i,k}}(\theta) - \sum_{i=1}^{N_k} \mathbb{1}_{I_{i,k}}(\theta + \pi) \right) \right) \\
&= \prod_{k=1}^K \left(\prod_{i=1}^{N_k} \exp(j\beta_k (\mathbb{1}_{I_{i,k}}(\theta) - \mathbb{1}_{I_{i,k}}(\theta + \pi))) \right) \\
&= \prod_{l=1}^M \exp(j\beta_l (\mathbb{1}_{I_l}(\theta) - \mathbb{1}_{I_l}(\theta + \pi))) = \prod_{l=1}^M h_{\alpha,\varepsilon}(r, \theta).
\end{aligned}$$

As a open interval I_l can always⁸³ be seen as a angular sector $A_{\alpha,\varepsilon}$, and in this case, (40) implies that $h_{\alpha,\varepsilon}$ is $\exp(j\beta_l (\mathbb{1}_{I_l}(\theta) - \mathbb{1}_{I_l}(\theta + \pi)))$.

Finally, if we put everything together, we obtain the complete story:

1. If G is a 2D real 0-**SI** unitary convolution, its Fourier multiplier \hat{g} is a unitary Hermitian symmetry measurable function, whose domain can be restricted to $\mathbb{S}^1 = \mathbb{U}$ (each hypothesis is important).
2. \hat{g} can be approximated⁸⁴ by unitary H -s simple functions on \mathbb{S}^1 (approximated in the sense of point-wise limit for functions, *i.e.*, a simple limit).
3. A simple function f is a *finite* sum of indicator functions of measurable sets A_k (on the circle \mathbb{S}^1 or the angles $[0, 2\pi]$). A unitary simple function will have complex coefficients of modulus 1: $f(\theta) = \sum_{k=-K}^K e^{j\beta_k} \mathbb{1}_{A_k}(\theta)$.
4. Such a measurable set A_k is decomposed as disjoint union of open intervals $I_{k,i}$ (possibly infinite, but always countable).
5. So the simple function f can be written as $\sum_{l=-M}^M e^{j\beta_l} \mathbb{1}_{I_l}(\theta)$.
6. A trick is to then write $f(\theta) = \prod_{l=-M}^M e^{j\beta_l \mathbb{1}_{I_l}(\theta)}$ instead⁸⁵.
7. The Hermitian symmetry gives a certain symmetry on I_l (and A_k), and so $f(\theta) = \prod_{l=1}^M e^{j\beta_l (\mathbb{1}_{I_l}(\theta) - \mathbb{1}_{I_l}(\theta + \pi))}$.
8. For intervals I_l , this term $e^{j\beta_l (\mathbb{1}_{I_l}(\theta) - \mathbb{1}_{I_l}(\theta + \pi))}$ is exactly what we called our elementary block $h_{\alpha,\varepsilon}$ of parameters α, ε depending⁸⁶ on I_l .
9. Composing two simple⁸⁷ **fdHT** operators gives this elementary block $h_{\alpha,\varepsilon}$.

Conclusion: so by going backward (from 9 to 1), we see that our **fdHT** operators can indeed be used to *approximate* G , in the sense of the point-wise limit of their Fourier multiplier. \square

Remark A.19. This proof was tedious, and one should keep in mind that we only proved the *density* of the composed **fdHT** operators, in particular we did not give an explicit representation of any 2D unitary **LC TI 0-SI** operator as a *finite* composition of fractional-directional Hilbert transforms (it is unachievable!).

Even worse, the density result is not fully explicit either. Each step we developed in the proof is explicit, but it would be impossible to sum them up in a simple and explicit scheme.

⁸³ Simply take α the middle of the interval, and ε its length.

⁸⁴ The approximation can be as precise as we want (there is convergence of a family $(g_n)_{n \in \mathbb{N}}$).

⁸⁵ It seems odd but this is true, as exactly one $\mathbb{1}_{I_l}(\theta)$ is non-zero for any θ , and for the others: adding a $e^{j\beta_l} \mathbb{1}_{I_l}(\theta) = 0$ or multiplying by a $e^{j\beta_l \mathbb{1}_{I_l}(\theta)} = 1$ does nothing.

⁸⁶ $\alpha \in [0, 2\pi]$ is the *middle* of the interval, and $\varepsilon > 0$ is its *length*.

⁸⁷ Of common parameter $\tau = \frac{1}{2}$ and directions $u_1 = \vec{u}_{\alpha-\varepsilon/2}$ and $u_2 = \vec{u}_{-(\alpha+\varepsilon/2)}$.

This being said, our theorem of density (theorem A.18) is still a very interesting result, as to the best of our knowledge it is the first of its kind for dimension bigger than 1. As theorem A.17 highlighted it, we have no hope of finding a parametrization of 2D unitary **LC TI 0-SI** operators with a finite number of parameters, so the density of a simple family of operators is the best one can hope to achieve. The **fdHT** have the advantage of being very easy to manipulate, parametrize and implement. Note that we do not know of any other dense family, so we cannot compare ours to another one. . .

And in higher dimension?

The **fdHT** was introduced in any dimension (definition A.9), but our theorem A.18 can only work in 2D. The key property we used was that the **fdHT** are constants on half-circles, and combining two of them gives a block similar to the indicator function of a (small) convex angular sector (*i.e.*, an interval). The same trick has no chance to work in higher dimensions, as **fdHT** are constants on half hyper-spheres, and combining two (or more) of them will never give a small surface element on \mathbb{S}^{d-1} .

Figure 21 is an illustration of this sphere in 3D, showing two symmetric angular sectors (denoted $A_{\phi,\psi,\varepsilon}$ (ϕ, ψ) having their usual meaning as spherical coordinates). They are obtained by combining two angular sectors, for instance on ϕ (in blue) and ψ (in orange), and we see that the nice behavior we used in 2D cannot be obtained in 3D: there is two “bands” (blue and orange, outside the red sectors) that cannot be neglected. It would be tedious to fully explain why our 2D trick does not generalize in higher dimension, but we hope this illustration would be enough for the reader to get a glimpse of this issue.

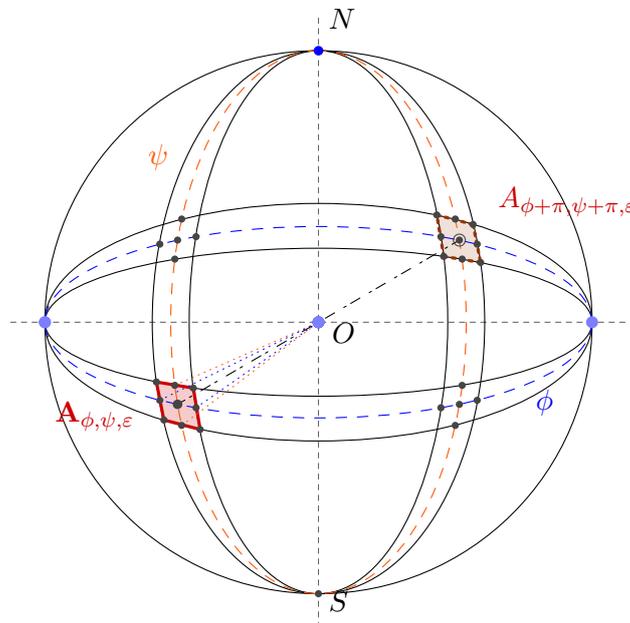


Figure 21: Two angular sectors $A_{\phi,\psi,\varepsilon}$ (front) and $A_{\phi+\pi,\psi+\pi,\varepsilon} + \pi$ (back), in 3D.

A.3.6 Inverting a fdHT

We will simply interpret the fact that the family of **fdHT** is closed under the inversion. For instance, in dimensions 1 and 2.

Lemma A.20. For a **fdHT** in dimension 1 (i.e., a purely-fractional Hilbert transform), $H_\tau = \cos(\pi\tau)\text{Id} - \sin(\pi\tau)H$, its Fourier multiplier is $\widehat{H}_\tau(w) = \cos(\pi\tau) + j \sin(\pi\tau) \text{sign}(w)$. It has an inverse, given by $H_\tau^{-1} = H_{-\tau}$.

Lemma A.21. For a **fdHT** in dimension 2, $H_{\tau,\alpha} = \cos(\pi\tau)\text{Id} - \sin(\pi\tau)H_\alpha$, its Fourier multiplier is $\widehat{H}_{\tau,\alpha}(w) = \cos(\pi\tau) + j \sin(\pi\tau) \text{sign}(\cos(\theta - \alpha))$. It has an inverse, given by $H_{\tau,\alpha}^{-1} = H_{\tau,\alpha+\pi}$.

Corollary A.22. Therefore, there is need of Green's function for our **fdHT**.

Easy Proof. Indeed, if we know how to apply (theoretically or practically, e.g., on an image) a **fdHT**, then we directly know how to apply an inverse **fdHT**. \square

A.4 Attempts of characterizing unitary 2D steerable convolutions

After a long and detailed study of 2D steerable convolutions in the core of this work (section 4.3), one could wonder what can be said if we mix the unity and the steerability properties. We do not have any rich characterization for steerability in dimension $d \in \mathbb{N}$ but only in dimension 1 and 2, so this last part focus only on steerable convolutions in dimension 2.

Surprisingly, a characterization of unitary **LC TI St** operators appeared very hard to obtain, while adding the scale-invariance reduces greatly the richness of the class and gives a very narrow result (as before). This will be the last characterization of operators with geometric properties.

Theorem A.23. An operator G **LC TI St** on $L^2(\mathbb{R}^2, \mathbb{R})$ is unitary iff $\widehat{g}(r, \theta)$ can be written as $\sum_{-K_G \leq k \leq K_G} \rho_k(r) a_k e^{jk\theta}$, for $\rho_k \in L^2(\mathbb{R}, \mathbb{C})$ and $\rho_{-k} a_{-k} = (-1)^k \overline{\rho_k a_k}$, and $|\widehat{g}(r, \theta)| = 1$ (almost) everywhere.

Easy Proof. We simply put together two characterizations obtained earlier, proposition 4.39 and theorem A.4. \square

Remark A.24 (Can we say more?). This theorem is quite limited, but the presence of **unknown** radial dependencies $\rho_k(r)$ makes it impossible to characterize the fact that $|\widehat{g}(r, \theta)| = 1$ on the different terms $\rho_k(r) a_k(r)$.

As we said before, adding the γ -**SI** property simplifies greatly the computations as the radial dependencies $\rho_k(r) = r^\gamma$ can be factorized. In the case of unitary steerable γ -**SI** operators, it turns out that they are the ones with only *one* term $a_k e^{jk\theta}$ in their additive decomposition (\heartsuit), and the Hermitian symmetry for real operators imposes a condition on the coefficients so the only non-zero term can be for $k = 0$.

Theorem A.25. *The identity is the only real unitary γ -SI 2D steerable convolution operator.*

Computational Proof. The converse is obviously true, as we mentioned it earlier: the identity operator is indeed a convolution, it is steerable as it is rotation-invariant, and it is unitary.

For the direct sense, we already know that γ -SI + U implies $\gamma = 0$, i.e., 0-SI. Then we know that γ -SI steerable convolutions have a specific form for their Fourier multiplier: $\hat{g}(r, \theta) = r^\gamma \sum_{-K_G \leq k \leq K_G} a_k e^{jk\theta}$, so $\gamma = 0$ implies that G is a (Hermitian symmetric) sum of iterated complex Riesz transforms. Now $|\hat{g}(r, \theta)|^2 = 1$ implies $\hat{g}(1, \theta)\hat{g}(1, \theta) = 1$, for all θ . But this product can be developed:

$$\begin{aligned}
|\hat{g}(r, \theta)|^2 &= \left(\sum_{-K_G \leq k \leq K_G} a_k e^{jk\theta} \right) \overline{\left(\sum_{-K_G \leq k \leq K_G} a_k e^{jk\theta} \right)} \\
&= \left(\sum_{-K_G \leq k \leq K_G} \bar{a}_k e^{-jk\theta} \right) \left(\sum_{-K_G \leq k \leq K_G} a_k e^{jk\theta} \right) \\
&= \left(\sum_{-K_G \leq k \leq K_G} (-1)^k a_{-k} e^{-jk\theta} \right) \left(\sum_{-K_G \leq k \leq K_G} a_k e^{jk\theta} \right) \\
&= \left(\sum_{-K_G \leq k' \leq K_G} (-1)^{k'} a_{k'} e^{jk'\theta} \right) \left(\sum_{-K_G \leq k \leq K_G} a_k e^{jk\theta} \right) \\
|\hat{g}(r, \theta)|^2 &= \sum_{-K_G \leq k' \leq K_G} \sum_{-K_G \leq k \leq K_G} (-1)^{k'} a_{k'} e^{jk'\theta} a_k e^{jk\theta} \\
(\text{With } l = k + k') &= \sum_{-2K_G \leq l \leq 2K_G} \left(\sum_{-K_G \leq k \leq K_G} (-1)^{l-k} a_{l-k} a_k \right) e^{jl\theta} \\
\text{So } |\hat{g}(r, \theta)|^2 = 1 &= \sum_{-2K_G \leq l \leq 2K_G} \left(\sum_{-K_G \leq k \leq K_G} (-1)^{l-k} a_{l-k} a_k \right) e^{jl\theta}, \forall \theta
\end{aligned}$$

But the functions $\theta \mapsto e^{jl\theta}$ are free for $-2K_G \leq l \leq 2K_G$, so this last equality gives that

$$\sum_{-K_G \leq k \leq K_G} (-1)^{l-k} a_{l-k} a_k = \begin{cases} 0 & \text{if } l \neq 0 \\ 1 & \text{if } l = 0. \end{cases}$$

These $2K_G + 1$ relations on the coefficients $(a_k)_{-K_G \leq k \leq K_G}$ can then be used to prove that $a_0 = 1$ and $a_k = 0$ if $k \neq 0$ (by using the formula for the determinant of a circulant matrix). And so $\hat{g}(r, \theta) = 1$, i.e., G is indeed the identity. \square

Conclusion

In this appendix we defined the notion of unitary operators, and we started by giving some naive examples. We have a characterization for unitary convolution operators (theorem A.4), but it turns out that the condition on $\hat{g}(\omega)$ ($|\hat{g}(\omega)| = 1$) is too rich to be characterized more specifically. In particular, we explained that it is impossible to obtain a finite-dimensional parametrization of the class of unitary convolution operators. However, when G satisfies other geometric properties (γ -**SI** or **RI**), the richness of \hat{g} is greatly reduced, giving for example a very narrow result: the only unitary **LC TI RI** γ -**SI** operator is the identity. Instead, we studied a specific family, the fractional-directional Hilbert transforms (**fdHT**), and despite not being a basis, its closure by composition gives a dense family among the 2D unitary convolution operators. It is the best we can achieve as in dimension $d \geq 2$, the class of valid Fourier multipliers is simply too rich to be parametrized with a finite number of parameters, so having a basis is impossible. This result is false in higher dimension, and it even seems impossible to obtain a dense parametric family for $d > 2$. We concluded by an easy, but limited, characterization of unity for steerable and γ -**SI** steerable operators in 2D.

B Appendix – Additional proofs

This short appendix contains 3 proofs that were not important enough to be included above, or part of some long proofs. Their order follows the outline of the report.

B.1 A composition of m directional derivatives has $n_G = m + 1$

We moved the longer part of the proof of lemma 4.41 here. Let $m \in \mathbb{N}$ and $\alpha_1, \dots, \alpha_m \in \mathbb{R}$, and $G = D_{\alpha_1} \circ \dots \circ D_{\alpha_m}$. Let us show that $n_G = m + 1$.

Long and Hard Proof. ⁸⁸ We proceed by induction, we will first prove that the following property⁸⁹ (P_m) is true for any $m \in \mathbb{N}$:

$$\forall R_{\theta_0} \in \text{SO}(2), \forall \theta \in \mathbb{R}, \quad r_{R_{\theta_0}} \widehat{g}(r, \theta) = \widehat{g}(r, \theta + \theta_0) = j^m r^m \sum_{i=0}^m f_i^{[m]}(\theta_0) \phi_i^{[m]}(\theta).$$

For $f_i^{[m]}(\theta_0) \stackrel{\text{def}}{=} \cos^{m-i}(\theta_0) \sin^i(\theta_0)$ and $\phi_i^{[m]}(\theta) \stackrel{\text{def}}{=} \sum_{J_1 \uplus J_2 = \{1, \dots, m\}} (\prod_{k \in J_1} \cos(\theta - \alpha_k)) (\prod_{k' \in J_2} -\sin(\theta - \alpha_{k'}))$ with $\#J_1 = i$ and $\#J_2 = m - i$, for any $0 \leq i \leq m$.

- (Optional) If $m = 0$, $G = \text{Id}$, and so $n_G = 1 = m + 1$, and $f_0^{[0]}(\theta_0) = 1$ and $\phi_0^{[0]}(\theta) = 1$, so (P_0) is true.
- If $m = 1$, $G = D_{\alpha_1}$, and so $n_G = 2$ (proposition 4.26). and $f_0^{[1]}(\theta_0) = \sin(\theta_0)$, $f_1^{[1]}(\theta_0) = \cos(\theta_0)$, and $\phi_0^{[1]}(\theta) = -\sin(\theta - \alpha_1)$, $\phi_1^{[1]}(\theta) = \cos(\theta - \alpha_1)$. Indeed, we already proved that

$$\begin{aligned} r_{R_{\theta_0}} \widehat{D_{\alpha_1}}(r, \theta) &= jr \cos((\theta - \alpha_1) + \theta_0) \\ &= jr \cos(\theta_0) \cos(\theta - \alpha_1) - jr \sin(\theta_0) \sin(\theta - \alpha_1) \quad (41) \\ &= j^1 r^1 \sum_{i=0}^1 f_i^{[1]}(\theta_0) \phi_i^{[1]}(\theta). \end{aligned}$$

So (P_1) is true.

- Now if $m \geq 1$, and if (P_m), let us show it for $m + 1$. We can write G as $G' \circ D_{\alpha_{m+1}}$, with $G' = D_{\alpha_1} \circ \dots \circ D_{\alpha_m}$. We can apply (P_m) to G' (the first m derivatives), and use the formula (41) for the last derivative $r_{R_{\theta_0}} \widehat{D_{\alpha_{m+1}}}(r, \theta)$.

$$\begin{aligned} r_{R_{\theta_0}} \widehat{g}(r, \theta) &= (r_{R_{\theta_0}} \widehat{g}'(r, \theta)) \left(r_{R_{\theta_0}} \widehat{D_{\alpha_{m+1}}}(r, \theta) \right) \\ &= \left(j^m r^m \sum_{i=0}^m f_i^{[m]}(\theta_0) \phi_i^{[m]}(\theta) \right) (jr \cos(\theta_0) \cos(\theta - \alpha_{m+1}) - jr \sin(\theta_0) \sin(\theta - \alpha_{m+1})) \\ &= j^{m+1} r^{m+1} \left(\sum_{i=0}^m f_i^{[m]}(\theta_0) \phi_i^{[m]}(\theta) (\cos(\theta_0) \cos(\theta - \alpha_{m+1}) - \sin(\theta_0) \sin(\theta - \alpha_{m+1})) \right). \end{aligned}$$

The leading $j^{m+1} r^{m+1}$ is good, and by identifying the dependencies in θ_0 (for $f_i^{[m+1]}$) and in θ (for $\phi_i^{[m+1]}$), we check easily that:

$$f_i^{[m]}(\theta_0) \cos(\theta_0) = (\cos^{m-i}(\theta_0) \sin^i(\theta_0)) \cos(\theta_0)$$

⁸⁸ We included colors, blue for θ and green for θ_0 , hoping it will improve the readability.

⁸⁹ This property is quite rich, as we have been able to fully explicit the basis of $V_{\widehat{g}}$, but we did not find it any use except here.

$$\begin{aligned}
&= \cos^{(m+1)-i}(\theta_0) \sin^i(\theta_0) = f_i^{[m+1]}(\theta_0), \\
f_i^{[m]}(\theta_0) \sin(\theta_0) &= (\cos^{m-i}(\theta_0) \sin^i(\theta_0)) \sin(\theta_0) \\
&= \cos^{(m+1)-(i+1)}(\theta_0) \sin^{i+1}(\theta_0) = f_{i+1}^{[m+1]}(\theta_0).
\end{aligned}$$

And for θ_0 , we recognized the long formula (without expanding it):

$$\underbrace{\phi_{i-1}^{[m]}(\theta_0)}_{\stackrel{\text{def}}{=} 0 \text{ if } i=0} \cos(\theta - \alpha_{m+1}) - \underbrace{\phi_i^{[m]}(\theta_0)}_{\stackrel{\text{def}}{=} 0 \text{ if } i=m+1} \sin(\theta - \alpha_{m+1}) = \phi_i^{[m+1]}(\theta_0).$$

And so we proved the formulas for $f_{i+1}^{[m+1]}$ and $\phi_i^{[m+1]}$, *i.e.*, (P_{m+1}) .

Now that (P_m) has been proved (for any m), let us use it to prove the required result of steerability of G , *i.e.*, on the dimension of $V_{\hat{g}}$ (as a sub-space of functions of $\mathbb{R}^2 \rightarrow \mathbb{R}$). But (P_m) gives a nice formula for $r_{R_{\theta_0}} \hat{g}$, as its dependency in θ_0 are only in the functions $f_i^{[m]}$, $0 \leq i \leq m$.

Let us prove (Q_m) , that $(f_i^{[m]})_{0 \leq i \leq m}$ is a free family (as functions of θ_0), again by induction on m (of step size 2 this time). For one example, with $m = 4$, see figure 22.

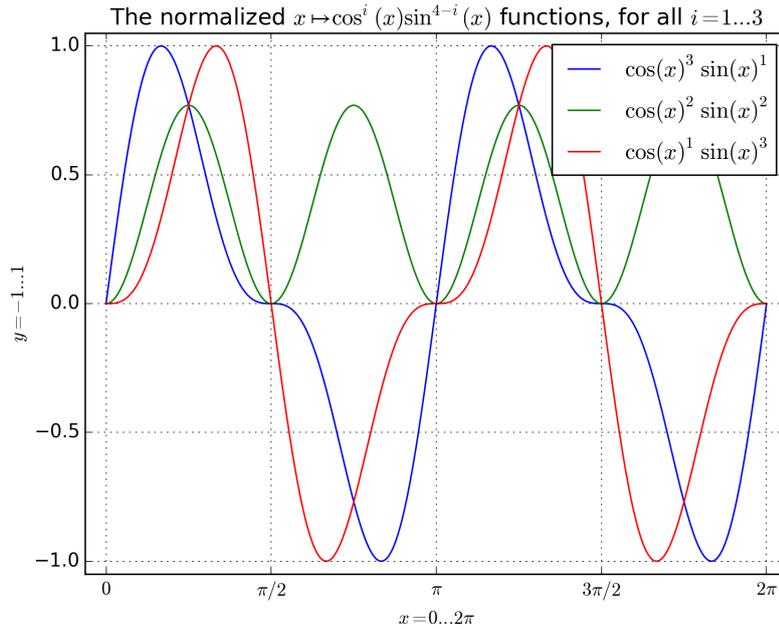


Figure 22: The functions $f_i^{[m]}(\theta_0)$ for $1 \leq i \leq m - 1$, for $m = 4$ (normalized to bound their values in $[-1, 1]$). Obtained with this Python script (click). We can observe the linear independence we just proved.

- For $m = 0, 1$ it is obvious, from the form of the functions (1 for $m = 0$, and \cos, \sin for $m = 2$). So we have (Q_0) and (Q_1) .
- Now, for $m \geq 2$, if we have (Q_{m-2}) , let us prove (Q_m) ,

Let $(\lambda_i)_{0 \leq i \leq m} \in \mathbb{R}^{m+1}$ so that $\sum_{i=0}^m \lambda_i f_i^{[m]}(\theta_0) = 0$ (for all θ_0). Taking the special values at $\theta_0 = 0$ and $\pi/2$ gives $\lambda_0 = 0$ (only one $\cos^m(0) = 1$ stays)

and $\lambda_m = 0$ (only one $\sin^m(\pi/2) = 1$ stays). Now, on $(0, \pi/2)$, we have

$$\begin{aligned} 0 &= \cos(\theta_0) \left(\sum_{i=1}^{m-1} \lambda_i \cos^{m-i-1}(\theta_0) \sin^{i-1}(\theta_0) \right) \sin(\theta_0) \\ 0 &= \underbrace{\cos(\theta_0)}_{\neq 0} \left(\sum_{i'=0}^{m-2} \lambda'_{i'} \cos^{m-2-i'}(\theta_0) \sin^{i'}(\theta_0) \right) \underbrace{\sin(\theta_0)}_{\neq 0} \\ \implies 0 &= \sum_{i'=0}^{m-2} \lambda'_{i'} f_i^{[m-2]}(\theta_0), \forall \theta_0 \in (0, \pi/2). \end{aligned}$$

So $\sum_{i'=0}^{m-2} \lambda'_{i'} f_i^{[m-2]}(\theta_0) = 0$ for all θ_0 , except for 4 points, $0, \pi/2, \pi, 3\pi/2$, but as we are dealing with continuous functions, discrete values do not matter. And so we find 0 written as a linear combination of the functions $f_i^{[m-2]}$, which are free by the induction hypothesis (Q_{m-2}) , so all the $\lambda'_{i'}$ are 0, and so $\lambda_i = 0, \forall 0 \leq i \leq m$. And so we proved (Q_m) .

Now, to conclude, the linear independence of this family, of size $m + 1$, is enough to show that $n_G = \dim_{\mathbb{R}} V_{\hat{g}} \geq m + 1$. But we already knew that $n_G \leq m + 1$, so we have $n_G = m + 1$ as wanted. \square

B.2 Viète formulas to develop $\cos(k\theta)$ and $\sin(k\theta)$

We include here a small lemma used in the proof of proposition 4.46, and its proof.

Lemma B.1 (Viète formula for $\cos(k\theta)$). *For any $\theta, \alpha \in \mathbb{R}, k \in \mathbb{N}^*$, we have*

$$\cos(k\theta) = \sum_{l=0}^{\lfloor k/2 \rfloor} \binom{k}{2l} (-1)^l (\cos(\theta))^{k-2l} \left(\cos\left(\theta - \frac{\pi}{2}\right)\right)^{2l}. \quad (42)$$

Proof. It is a classic computation from the binomial theorem applied to $e^{k\theta} = (e^\theta)^k$ and by taking the real part of the left-hand side, $\Re(e^{k\theta}) = \cos(k\theta)$, and the real part of the right-hand side, $\Re((e^\theta)^k) = \Re\left(\sum_{l=0}^k \binom{k}{l} \cos(\theta)^{k-l} (j \sin(\theta))^l\right)$. $\lfloor k/2 \rfloor$ is the lower integer part of $k/2$ (i.e., the biggest integer to be smaller or equal to $k/2$).

We have a similar formula for $\sin(k\theta)$, as a direct application of cos (42):

$$\begin{aligned} \sin(k\theta) &= \cos\left(k\left(\theta - \frac{\pi}{2k}\right)\right) \\ &= \sum_{l=0}^{\lfloor k/2 \rfloor} \binom{k}{2l} (-1)^l \left(\cos\left(\theta - \frac{\pi}{2k}\right)\right)^{k-2l} \left(\cos\left(\theta - \left(\frac{\pi}{2} + \frac{\pi}{2k}\right)\right)\right)^{2l}. \quad \square \end{aligned}$$

B.3 The Fourier coefficients of $\mathbf{c}(\theta)$ are mutually orthogonal projections

We include here the proof of an intermediate result used in the proof of theorem 4.28.

Proof. Let us check these two identities on the Fourier coefficients of the $q \times q$ matrix-valued function $\mathbf{c}(\theta)$, defined by $\mathbf{c}_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{c}(\theta) e^{-jk\theta} d\theta$:

$$\begin{cases} \mathbf{c}_k^2 = \mathbf{c}_k & \text{for every } k \in \mathbb{Z}, \\ \mathbf{c}_{k_1} \mathbf{c}_{k_2} = \mathbf{0} & \text{for every } k_1 \neq k_2 \in \mathbb{Z}. \end{cases}$$

- For $k \in \mathbb{Z}$,

$$\begin{aligned} \mathbf{c}_k^2 &= \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{c}(\theta) e^{-jk\theta} d\theta \right)^2 \\ &= \left(\frac{1}{2\pi} \right)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \mathbf{c}(\theta) e^{-jk\theta} \mathbf{c}(\theta') e^{-jk\theta'} d\theta d\theta' \\ \text{(Thanks to (20))} &= \left(\frac{1}{2\pi} \right)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \mathbf{c}(\theta + \theta') e^{-jk(\theta + \theta')} d\theta d\theta' \\ \text{(With } \theta'' = \theta + \theta') &= \left(\frac{1}{2\pi} \right)^2 \int_{-\pi}^{\pi} \mathbf{c}(\theta'') e^{-jk\theta''} \underbrace{\left(\int_{-\pi}^{\pi} 1 d\theta \right)}_{=2\pi} d\theta'' \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{c}(\theta'') e^{-jk\theta''} d\theta'' = \mathbf{c}_k. \end{aligned}$$

- For $k_1 \neq k_2 \in \mathbb{Z}$,

$$\begin{aligned} \mathbf{c}_{k_1} \mathbf{c}_{k_2} &= \left(\frac{1}{2\pi} \right)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \mathbf{c}(\theta) e^{-jk_1\theta} \mathbf{c}(\theta') e^{-jk_2\theta'} d\theta d\theta' \\ \text{(Thanks to (20))} &= \left(\frac{1}{2\pi} \right)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \mathbf{c}(\theta + \theta') e^{-jk_1\theta} e^{-jk_2\theta'} d\theta d\theta' \\ \text{(With } \theta'' = \theta + \theta') &= \left(\frac{1}{2\pi} \right)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \mathbf{c}(\theta'') e^{-jk_2\theta''} e^{-j(k_1 - k_2)\theta} d\theta d\theta'' \\ &= \left(\frac{1}{2\pi} \right)^2 \int_{-\pi}^{\pi} \mathbf{c}(\theta'') e^{-jk_2\theta''} \underbrace{\left(\int_{-\pi}^{\pi} e^{-j(k_1 - k_2)\theta} d\theta \right)}_{=0 \text{ as } k_1 - k_2 \neq 0} d\theta'' = \mathbf{0}. \end{aligned}$$

□

C Acknowledgments

I would like to first thank warmly Prof. Dr. Michael Unser, director of the BIG team at EPFL, for trusting me and accepting me in his team for this 5-month internship, and both him and Julien Fageot for their supervision, their advice and their constant help. Julien was very present and helpful, on every aspect of our research. Many thanks also to Nadia Macor for her kind and efficient support for all administrative aspects, and for her always cheerful attitude.

In France, thanks to Prof. Gabriel Peyré for accepting to be my internship advisor for the master MVA, reviewing my report and attending to the oral presentation; and at ENS de Cachan thanks to Delphine Laverne and Prof. Alain Trouvé for their administrative and financial support. Thanks also to Prof. Nicolas Vayatis for his advice while I was trying to choose my internship, and thanks to him and all the other professors I had in 2015–2016 during the master MVA.

At EPFL, thanks to my friend and desktop-mate⁹⁰ Thibault Groueix; and thanks also to all the members of the BIG team, especially to Harshit Gupta, Dr. Ferréol Soulez, Dr. Denis Fortun, Anaïs Badoual, Virginie Uhlmann, Silvia Colabrese, Laurène Donati and Dr. Daniel Sage for many fruitful discussions. Thanks to my roommates for their useful \LaTeX -related suggestions, and many good moments we shared together.

My internship was partly funded by the mathematical department of ENS de Cachan in France, and by the *Swiss European Mobility Program* and the BIG team at EPFL in Switzerland – many thanks to them as well. Working at EPFL was a wonderful experience, made possible thanks to the *Swiss European Mobility Program*, and the help of Aurore Patey from the international relations office at ENS de Cachan.

Quick personal feeling about the internship

TL;DR: I enjoyed living in Lausanne and in Switzerland, I liked working with the “BIG-gers”, and I loved working on this report during my last weeks at EPFL.

I enjoyed very much working on this project, especially because it was a unique chance of working on such theoretical considerations of functional analysis and signal processing. It was exciting to have the time to work on the report and make it as complete and as detailed as I wanted. I hope it was not too long or too hard to follow!

We ran out of time to fully implement everything and to experiment in depth with stochastic sparse processes, and real images, but this is an interesting direction of research. Virginie, Julien and I will most surely keep working on this aspect, and hopefully we will soon try to publish our results.

Working in the BIG team at EPFL was very nice and intellectually stimulating, and I enjoyed staying in Lausanne during the spring and summer 2016. EPFL and its campus are quiet and nice to live in, and every aspect of my stay was made easy thanks to their great infrastructures.

This work is dedicated to the memory of the late Nicolas Pajor and Bogdana Neuville. ☹

⁹⁰ For his daily help on basic maths, the regular experiments on Sublime Text, and command-line- or Ubuntu-related exploratory hacks that we used to enjoy together.

D List of references

- [Bil04] E. Biland (2004). *Quelques applications du théorème de relèvement*. URL http://erwanbiland.fr/recherche_pdf/relevement.pdf, notes in French.
- [Bou81] Nicolas Bourbaki (1981). *Espaces Vectoriels Topologiques : Chapitres 1 à 5*. Masson.
- [BU07] T. Blu and M. Unser (2007). *Self-similarity: Part II — Optimal estimation of fractal processes*. In *IEEE Transactions on Signal Processing* [UB07], pages 1364–1378. URL <http://bigwww.epfl.ch/publications/blu0701>.
- [Cha11] K.N. Chaudhury (February 2011). *Optimally Localized Wavelets and Smoothing Kernels*. EPFL thesis no. 4968 (2011), 239 p., Swiss Federal Institute of Technology Lausanne (EPFL). URL <https://infoscience.epfl.ch/record/162113>.
- [CU09] K.N. Chaudhury and M. Unser (2009). *The fractional Hilbert transform and Dual-Tree Gabor-like wavelet analysis*. In *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 3205–3208. IEEE. URL <http://bigwww.epfl.ch/publications/chaudhury0901>.
- [CU10] K.N. Chaudhury and M. Unser (2010). *On the Shiftability of Dual-Tree Complex Wavelet transforms*. *IEEE Transactions on Signal Processing*, 58(1):221–232. URL <http://bigwww.epfl.ch/publications/chaudhury0904>.
- [D’A13] John P. D’Angelo (2013). *Hermitian Analysis: From Fourier series to Cauchy-Riemann Geometry*, volume 10. Springer, 12 pages.
- [Des11] L. Desvillettes (2011). *Lecture notes on Functional Analysis*. URL http://desvillettes.perso.math.cnrs.fr/analyse_fonctionnelle59.pdf, in French.
- [FAU14] J. Fageot, A. Amini, and M. Unser (December 2014). *On the Continuity of Characteristic Functionals and Sparse Stochastic Modeling*. *The Journal of Fourier Analysis and Applications*, 20(6):1179–1211. URL <http://bigwww.epfl.ch/publications/fageot1401>.
- [FR13] S. Foucart and H. Rauhut (2013). *A Mathematical Introduction to Compressive Sensing*, volume 1. Springer.
- [Gro73] A. Grothendieck (1973). *Topological Vector Spaces*. Taylor & Francis.
- [Hör83] L. Hörmander (1983). *Distribution Theory and Fourier Analysis*. Volume I of *The Analysis of Linear Partial Differential Operators* [Hör85].
- [Hör85] L. Hörmander (1985). *Differential Operators with Constant Coefficients*. Volume II of *The Analysis of Linear Partial Differential Operators* [Hör83].
- [HOT98] Y. Hel-Or and P. C. Teo (1998). *Canonical decomposition of steerable functions*. *Journal of Mathematical Imaging and Vision*, 9(1):83–95.
- [KUW13] I. Khalidov, M. Unser, and J.P. Ward (2013). *Operator-Like Wavelet Bases of $L^2(\mathbb{R}^d)$* . *The Journal of Fourier Analysis and Applications*, 19:1294–1322. URL <http://bigwww.epfl.ch/publications/khalidov1301>.
- [LMZ96] A.W. Lohmann, D. Mendlovic, and Z. Zalevsky (Feb 1996). *Fractional Hilbert transform*. *Optics Letters*, 21(4):281–283.
- [LWU13] S. Lefkimmiatis, J.P. Ward, and M. Unser (2013). *Hessian Schatten-Norm Regularization for Linear Inverse Problems*. *IEEE Trans. Image Process.*, 22(5):1873–1888. URL <http://arxiv.org/abs/1209.3318>.

- [Pey11] G. Peyré (2011). *The Numerical Tours of Signal Processing – Advanced Computational Signal and Image Processing*, volume 13. IEEE Computing in Science and Engineering, 94–97 pages. URL <http://www.numerical-tours.com/>.
- [Püs16] S. Püspöki (June 2016). *Local Geometric Transformations in Image Analysis*. EPFL thesis no. 7014 (2011), 212 p., Swiss Federal Institute of Technology Lausanne (EPFL). URL <https://infoscience.epfl.ch/record/218634>.
- [Rud91] W. Rudin (1991). *Functional Analysis*. McGraw-Hill, Inc., New York.
- [SW71] E. M. Stein and G. L. Weiss (1971). *Introduction to Fourier analysis on Euclidean spaces*, volume 1. Princeton University Press.
- [UB07] M. Unser and T. Blu (2007). *Self-similarity: Part I — Splines and operators*. In *IEEE Transactions on Signal Processing* [BU07], pages 1352–1363. URL <http://bigwww.epfl.ch/publications/unser0701>.
- [UC13] M. Unser and N. Chenouard (2013). *A unifying parametric framework for 2D steerable wavelet transforms*. *SIAM Journal on Imaging Sciences*, 6(1):102–135. URL <http://bigwww.epfl.ch/publications/unser1301>.
- [UFW16] M. Unser, J. Fageot, and J.P. Ward (2016). *Splines are Universal Solutions of Linear Inverse Problems with Generalized-TV Regularization*. *arXiv preprint arXiv:1603.01427*. URL <http://arxiv.org/abs/1603.01427>.
- [Uns16] M. Unser (2016). *Foundations of Digital Signal Processing - Discrete Spaces Operators and Kernel Theorems*. Unpublished notes.
- [UT14] M. Unser and P.D. Tafti (2014). *An Introduction to Sparse Stochastic Processes*. Cambridge University Press. URL <http://www.sparseprocesses.org/>.
- [VSU15] C. Vonesch, F. Stauber, and M. Unser (2015). *Steerable PCA for Rotation-Invariant Image Recognition*. *SIAM Journal on Imaging Sciences*, 8(3):1857–1873. URL <http://bigwww.epfl.ch/publications/vonesch1501>.
- [War11] J.P. Ward (2011). *L^p error estimates for approximation by Sobolev splines and Wendland functions on \mathbb{R}^d* . *Advances in Computational Mathematics*, 38(4):873–889.
- [WCU13] J.P. Ward, K.N. Chaudhury, and M. Unser (2013). *Decay properties of Riesz transforms and steerable wavelets*. *SIAM Journal on Imaging Sciences*, 6(2):984–998. URL <http://bigwww.epfl.ch/publications/ward1301>.
- [WU14] J.P. Ward and M. Unser (2014). *Harmonic Singular Integrals and Steerable Wavelets in $L^2(\mathbb{R}^d)$* . *Applied Computational Harmonic Analysis*, 36(2):183–197. URL <http://bigwww.epfl.ch/publications/ward1401>.

Note: the bibliography is also available on-line, in HTML, plain text, PDF and BibTeX.

Not yet published

Note that this report has **not** been published on any conference, journal or preprint platform. It is simply the result of a 5-month research internship, required to conclude the Master program I followed in 2015–2016. **But an article version is on the way.**

License?

This report (and the additional resources – including slides, code, images etc) are publicly published under the terms of the open-source *CC-BY 4.0 License*.

Copyright 2016, © Lilian Besson.

The end. □