Submitted to the **Annals of Statistics**

Estimating Deformations of Stationary Processes

Maureen Clerc
Ecole Nationale des Ponts et Chaussées, France

Stéphane Mallat
Ecole Polytechnique, France, and Courant Institute

June 21, 2000

Abstract

This paper studies classes of nonstationary processes, such as warped processes and frequency-modulated processes, that result from the deformation of stationary processes. Estimating deformations can often provide important information about an underlying physical process. A computational harmonic analysis viewpoint shows that the deformed autocorrelation satisfies a transport equation at small scales, with a velocity proportional to a deformation gradient. We derive an estimator of the deformation from a single realization of the deformed process, with a proof of consistency under appropriate assumptions.

Introduction

When a nonstationary process F results from the deformation of a stationary process R, estimating the deformation can provide important information about an underlying physical process of interest. From one realization of $F = D\,R$, we wish to recover the

[0] *AMS 1991 Subject Classifications: Primary 62M10 , 60G12 ; Secondary 60G35.*

Key words and phrases: nonstationary processes, inverse problem, autocorrelation, wavelets, warplets, warping; frequency modulation, scalogram, spectrogram.

deformation operator D, which is assumed to belong to a specified group $\mathcal{D}$. For example, a Doppler effect produces a warping deformation in time $F(x) = R(\alpha(x))$, where $\alpha'(x)$ depends upon velocity. The deformation of a stationary texture by perspective in an image also produces a warping, where $x \in \mathbb{R}^2$ is a spatial variable; recovering the Jacobian matrix of $\alpha(x)$ nearly characterizes the corresponding three-dimensional surface which is being viewed [8]. The frequency modulation of a stationary process $F(x) = R(x)\exp(i\alpha(x))$ corresponds to another class of deformations often encountered in signal processing; in transmissions by frequency modulation, the message is carried by $\alpha'(x)$.

Estimating the deformation $D \in \mathcal{D}$ from $F = D\,R$ is an inverse problem. As we suppose no prior knowledge about the stationary process R, the deformation D can only be recovered up to the subgroup $\mathcal{G}$ of $\mathcal{D}$ which leaves the set of stationary processes globally invariant. Rather than the deformation itself, we therefore seek to estimate the equivalence class of D in $\mathcal{D}/\mathcal{G}$. We consider cases where $\mathcal{G}$ is a finite-dimensional Lie group, and under appropriate assumptions, this equivalence class can be represented by a vector field on $\mathcal{G}$, which corresponds to a deformation gradient. A local analysis of the deformation is performed by decomposing the autocorrelation of F over an appropriate family of localized functions, called atoms in the harmonic analysis literature. The deformation gradient is shown to appear as a velocity vector in a transport equation satisfied by a localized autocorrelation. This general result is applied to one-dimensional warping and frequency modulation, where the atoms are wavelets, and multidimensional warping, where the atoms are called warplets.

Computing the deformation gradient requires estimating the autocorrelation of F projected over a family of localized atoms, from a single realization. Under certain conditions on the autocorrelation of the stationary process R, one can obtain consistent estimators for one-dimensional warping and frequency modulation. Numerical examples illustrate these results.

The paper is organized in three main sections: after discussing the well-posedness of the inverse problem in Section 1, we establish in Section 2 a transport equation for the localized autocorrelation of a deformed process; Section 3 introduces estimators and proves their consistency.

1 Inverse Problem

We want to estimate a deformation operator D which belongs to a known group $\mathcal{D}$, from a single realization of $F = D\,R$. The process R is wide-sense stationary, and is not known a priori. Since we are limited to a single realization, we concentrate on second-order moments. For this reason, stationarity will always be understood in the wide-sense, meaning that

$$\mathsf{E}\{R(x)\} = \mathsf{E}\{R(0)\} \ ,$$
$$\text{and } \mathsf{E}\{R(x)\, R^*(y)\} = c_R(x-y) \quad \text{with} \quad c_R(0) < +\infty \quad ,$$

where z^* denotes the complex conjugate of $z \in \mathbb{C}$. Although it is not normalized by the variance, the term $\mathsf{E}\{R(x)\, R^*(y)\}$ is called *autocorrelation* of R in the rest of this paper.

1.1 Class of Solutions

Nothing is known about the process R except for its stationarity, therefore the set of solutions to the inverse problem is the set of all operators $\tilde{D} \in \mathcal{D}$ such that $\tilde{D}^{-1}F$ is stationary. In general, this set is larger than $\{D\}$. Let $\mathcal{G}$ be the set of all operators $G \in \mathcal{D}$ such that if X is a wide-sense stationary process, then $G\,X$ is also wide-sense stationary. One can verify that $\mathcal{G}$ is a subgroup of $\mathcal{D}$, which we call *stationarity invariant group*. Clearly, if D is a solution of the inverse problem, any operator $\tilde{D} = D\,G$ with $G \in \mathcal{G}$ is also a solution. The set of solutions of the inverse problem therefore contains the equivalence class of D in the quotient group $\mathcal{D}/\mathcal{G}$. The equivalence class of D is equal to the set of solutions to the inverse problem if any deformation $\tilde{D} \in \mathcal{D}$ such that $\tilde{D}R$ is wide-sense stationary necessarily belongs to $\mathcal{G}$. This condition is not met by all stationary processes R, but we give sufficient conditions on the autocorrelation of R to guarantee uniqueness. In this paper we concentrate on four categories of deformation groups.

Example 1 The positive multiplicative group is a particularly simple example where

$$\mathcal{D} = \{D \ : \ D\,f(x) = \alpha(x)\,f(x) \ \text{ with } \ \alpha(x) > 0\} \ .$$

The stationarity invariant group corresponds to multiplicative functions $\alpha(x)$ which are constant:

$$\mathcal{G} = \{G_a \ : \ G_a\,f(x) = a\,f(x) \ \text{ with } \ a > 0\} \ .$$

Two operators D_1 and D_2 such that $D_1 f(x) = \alpha_1(x)\,f(x)$ and $D_2 f(x) = \alpha_2(x)\,f(x)$ belong to the same equivalence class in $\mathcal{D}/\mathcal{G}$ if $\alpha_1(x)/\alpha_2(x)$ is constant.
The equivalence class of D is computed from $F(x) = \alpha(x)R(x)$ by calculating

$$\mathsf{E}\{|F(x)|^2\} = \alpha^2(x)\,\mathsf{E}\{|R(x)|^2\} = \alpha^2(x)\,\mathsf{E}\{|R(0)|^2\} \ , \tag{1}$$

which specifies $\alpha(x) > 0$ up to a multiplicative constant.

Finding the equivalence class of D is in general much more complicated. In the following, we require the function $\alpha(x)$, which produces the deformation, to have a specified regularity. This regularity will play an important role in the estimation procedure.

Example 2 The frequency modulation group modifies signal frequency:

$$\mathcal{D} = \left\{ D \ : \ D\, f(x) = \mathrm{e}^{i\,\alpha(x)}\, f(x) \ \text{ where } \alpha(x) \text{ is real and } \mathbf{C}^4 \right\} \ . \tag{2}$$

In transmissions with frequency modulation, $\alpha'(x)$ is proportional to the signal to be transmitted, and the stationary process R is the carrier. The stationarity invariant group is

$$\mathcal{G} = \left\{ G_{(\phi,\xi)} \ : \ G_{(\phi,\xi)}\, f(x) = \mathrm{e}^{i(\phi+\xi x)} f(x) \ \text{ with } \ (\phi,\xi) \in \mathbb{R}^2 \right\} \ .$$

Two operators D_1 and D_2 such that $D_1 f(x) = \mathrm{e}^{i\alpha_1(x)}\, f(x)$ and $D_2 f(x) = \mathrm{e}^{i\alpha_2(x)}\, f(x)$ are in the same equivalence class in $\mathcal{D}/\mathcal{G}$ if and only if $\alpha_1(x) = \phi + \xi x + \alpha_2(x)$ and hence

$$\alpha_1''(x) = \alpha_2''(x) \ . \tag{3}$$

The following proposition gives a sufficient condition on the autocorrelation kernel $c_R(x)$ to identify $\alpha''(x)$ from the autocorrelation of $F = D\,R$. The proof is in Appendix A.1.

Proposition 1.1. *Let $F = D\,R$, where D belongs to the frequency modulation group $\mathcal{D}$ in (2). If there exists an $\varepsilon > 0$ such that*

$$\forall\, x \in]-\varepsilon, \varepsilon[\,, \quad c_R(x) > 0$$

then the equivalence class of D in $\mathcal{D}/\mathcal{G}$ is uniquely characterized by the autocorrelation of $F = D\,R$.

Example 3 The one-dimensional warping group is defined by

$$\mathcal{D} = \left\{ D \ : \ D\, f(x) = f(\alpha(x)) \ \text{ where } \alpha(x) \text{ is } \mathbf{C}^3 \text{ and } \alpha'(x) > 0 \right\} \ . \tag{4}$$

Such time warpings appear in many physical phenomena, such as the Doppler effect. We easily verify that the stationarity invariant group is the affine group:

$$\mathcal{G} = \left\{ G_{(u,s)} \ : \ G_{(u,s)}\, f(x) = f(u + sx) \ \text{ with } \ (u,s) \in \mathbb{R} \times \mathbb{R}_{+*} \right\} \ .$$

Two warping operators D_1 and D_2 are in the same equivalence class in $\mathcal{D}/\mathcal{G}$ if and only if there exists (u, s) such that $\alpha_1(x) = u + s\,\alpha_2(x)$, or equivalently

$$\frac{\alpha_1''(x)}{\alpha_1'(x)} = \frac{\alpha_2''(x)}{\alpha_2'(x)} \ . \tag{5}$$

The following proposition, whose proof is in Appendix A.2, gives a sufficient condition on R to characterize the equivalence class of D uniquely. Perrin and Senoussi [12] provide a similar result.

Proposition 1.2. *Let $F = D\,R$, with $D \in \mathcal{D}$, where $\mathcal{D}$ is the warping group (4). If there exists an $\varepsilon > 0$ such that c_R is $\mathbf{C}^1$ on $]0, \varepsilon[$ and*

$$\forall\, x \in]0, \varepsilon[\ ,\ c_R'(x) < 0\ , \tag{6}$$

then the equivalence class of D in $\mathcal{D}/\mathcal{G}$ is uniquely characterized by the autocorrelation of F.

Warping deformations are used in geostatistics [11, 13], to model nonstationary phenomena. Stationarizing the data $F(x)$ is suggested as an initial step before applying classical geostatistical methods such as kriging.

Example 4 The warping problem in two dimensions has an important application in image analysis, particularly in recovering a three-dimensional surface shape by analyzing texture deformations. More generally, we study a d-dimensional warping problem, specified by an invertible function $\alpha(x)$ from $\mathbb{R}^d$ to $\mathbb{R}^d$ with

$$\alpha(x_1, \dots, x_d) = \Big(\alpha_1(x_1, \dots, x_d), \dots, \alpha_d(x_1, \dots, x_d)\Big)\ .$$

The Jacobian matrix of α at position $x \in \mathbb{R}^d$ is written

$$J_\alpha(x) = \left(\frac{\partial \alpha_i(x)}{\partial x_j}\right)_{1 \le i,j \le d}\ . \tag{7}$$

If the Jacobian determinant $\det J_\alpha(x)$ does not vanish, $\alpha(x)$ is invertible and corresponds to a change of metric. We consider a group of regular warpings

$$\mathcal{D} = \big\{D \ :\ D\,f(x) = f(\alpha(x)) \ \text{ where } \alpha(x) \text{ is in } \mathbf{C}^3(\mathbb{R}^d) \text{ and } \det J_\alpha(x) > 0\big\}\ . \tag{8}$$

Let $GL^+(\mathbb{R}^d)$ be the group of linear operators in $\mathbb{R}^d$ with a strictly positive determinant. We easily verify that the stationarity invariant group is the affine group:

$$\mathcal{G} = \big\{G_{(u,S)} \ :\ G_{(u,S)}\,f(x) = f(u + Sx) \ \text{ with } (u, S) \in \mathbb{R}^d \times GL^+(\mathbb{R}^d)\big\}\ .$$

Two operators D and $\tilde{D}$ such that $Df(x) = f(\alpha(x))$ and $\tilde{D}f(x) = f(\tilde{\alpha}(x))$ are in the same equivalence class in $\mathcal{D}/\mathcal{G}$ if and only if

$$\exists\,(u, S) \in \mathbb{R}^d \times GL^+(\mathbb{R}^d) \quad , \quad \alpha(x) = u + S\,\tilde{\alpha}(x)\ . \tag{9}$$

The partial derivative of the Jacobian matrix in a fixed direction x_k is again a matrix:

$$\frac{\partial J_\alpha(x)}{\partial x_k} = \left(\frac{\partial^2 \alpha_i(u)}{\partial x_k\,\partial x_j}\right)_{1 \le i,j \le d} \ .$$

One can check that condition (9) is equivalent to the following matrix equalities, which generalize (5):

$$\forall\, k \in \{1, ..., d\} \quad , \quad J_\alpha^{-1}(x)\,\frac{\partial J_\alpha(x)}{\partial x_k} = J_{\tilde{\alpha}}^{-1}(x)\,\frac{\partial J_{\tilde{\alpha}}(x)}{\partial x_k}\ . \tag{10}$$

There are cases for which the inverse warping problem cannot be solved. For example, consider a stationary process $R(x) = R_1(x_1)$ which only depends on the first variable, and a warping which leaves x_1 invariant: $\alpha(x_1, ..., x_d) = (x_1, \alpha_1(x_2, ..., x_d))$. In this case

$$F(x) = R(x_1, \alpha_1(x_2, ..., x_d)) = R_1(x_1) = R(x)\ . \tag{11}$$

Hence we can not recover α. The following proposition, whose proof is in Appendix A.3, gives a sufficient condition on $c_R(x)$ to guarantee that the inverse warping problem has a unique solution in $\mathcal{D}/\mathcal{G}$.

Proposition 1.3. *Let $F = D\,R$, with $D \in \mathcal{D}$, where $\mathcal{D}$ is the multidimensional warping group (8). If there exists $h > 0$ and a function $\eta(x)$ such that*

$$c_R(0) - c_R(x) = |x|^h\,\eta(x)\ , \tag{12}$$

where $\eta(x)$ is $\mathbf{C}^2$ in a neighborhood of 0, then the equivalence class of D in $\mathcal{D}/\mathcal{G}$ is uniquely characterized by the autocorrelation of F.

The inverse warping problem can be applied to the reconstruction of three-dimensional surfaces from deformations of textures in images [8]. One can model the image of a three-dimensional surface, on which a texture is mapped, as

$$F(x) = R(\alpha(x))\ ,$$

where R is a stationary process, and $\alpha(x)$ is the two-dimensional warping due to the imaging process, which projects the surface onto the image plane [5].

We showed in (10) that solving the inverse warping problem is equivalent to computing normalized partial derivatives of the Jacobian matrix J_α:

$$J_\alpha^{-1}(x)\frac{\partial J_\alpha(x)}{\partial x_1} \text{ and } J_\alpha^{-1}(x)\frac{\partial J_\alpha(x)}{\partial x_2} \ . \tag{13}$$

Gårding [8], Malik and Rosenholtz [9] have proved that these matrices specify the local orientation and curvature of the three-dimensional surface in the scene. Knowing these surface parameters, it is then possible to recover the three-dimensional coordinates of the surface, up to a constant scaling factor. We will see in Section 2.4 that the Jacobian matrices (13) appear as velocity vectors in a transport equation satisfied by the autocorrelation of F.

1.2 Stationarity Invariant Group

The stationarity invariant group $\mathcal{G}$ specifies the class of solutions of the inverse problem $F = D\,R$, and Section 2 will show that it is also an important tool to identify the equivalence class of D in $\mathcal{D}/\mathcal{G}$. This section examines the properties of operators that belong to such a group. Recall that an operator G is said to be *stationarity invariant* if, for any wide-sense stationary process R, the process $F = G\,R$ is also wide-sense stationary.

The following theorem characterizes this class of operators. We denote by $x \cdot y$ the inner product of two vectors x and y of $\mathbb{R}^d$.

Theorem 1.1. *An operator G is stationarity invariant if and only if there exists $\hat{\rho}(\omega)$ from $\mathbb{R}^d$ to $\mathbb{C}$ and $\theta(\omega)$ from $\mathbb{R}^d$ to $\mathbb{R}^d$ such that*

$$G\,\mathrm{e}^{i\omega\cdot x} = \hat{\rho}(\omega)\,\mathrm{e}^{i\theta(\omega)\cdot x} , \tag{14}$$

with $\operatorname{ess\,sup}_{\omega\in\mathbb{R}^d} |\hat{\rho}(\omega)| < \infty$.

The proof is in Appendix A.4. This theorem proves that a stationarity invariant operator transposes the frequency of a sinusoid and modifies its amplitude. The examples detailed in the previous section correspond to particular classes of such operators, where $\theta(\omega)$ is affine in ω. Supposing that $\theta(\omega) = \overline{S}\omega + \xi$ with $\xi \in \mathbb{R}^d$ and where S is an invertible linear operator in $\mathbb{R}^d$, whose adjoint is denoted $\overline{S}$, the operator G in (14) can then be written

$$Gf(x) = \mathrm{e}^{i\xi\cdot x}\, f \star \rho(Sx) \ , \tag{15}$$

where $\rho(x)$ is the function whose Fourier transform is $\hat{\rho}(\omega)$.

Let us define a translation operator T_v for $v \in \mathbb{R}^d$ by

$$T_v f(x) = f(x - v) \ .$$

The following proposition proves that linear operators of the form (15) are characterized by a weak form of commutativity with T_v.

Proposition 1.4. *A linear operator G which is bounded in $\mathbb{L}^2(\mathbb{R}^d)$ satisfies*

$$\exists \xi \in \mathbb{R}^d, \ \exists S \in GL^+(\mathbb{R}^d), \ \forall v \in \mathbb{R}^d, \ G\, T_{Sv} = \mathrm{e}^{i\xi \cdot v} T_v \, G \tag{16}$$

if and only if G is stationarity invariant and can be written

$$\forall f \in \mathbb{L}^2(\mathbb{R}^d) \quad , \quad Gf(x) = \mathrm{e}^{i\xi\cdot x}\, f \star \rho(Sx) \ , \tag{17}$$

with $\operatorname{ess\,sup}_{\omega \in \mathbb{R}^d} |\hat{\rho}(\omega)| < \infty$.

The proof is in Appendix A.5. If $\rho(x) = \mathrm{e}^{i\phi}\, \delta(x - v)$ then the operator G defined in (17) represents frequency modulation and warping. In the rest of the paper, we concentrate on deformations where the stationarity invariant operators satisfy (16), which can be interpreted as a transport property.

2 Conservation and Transport

The stationarity of a random process R is a conservation property of its autocorrelation through translation. Because of the deformation, the process $F(x) = D\, R(x)$ is no longer stationary and its autocorrelation does not satisfy the same conservation property. Yet, we show that the stationarity of R implies a conservation of the autocorrelation of F, along characteristic curves in an appropriate parameter space. These characteristic curves, which identify the equivalence class of D in $\mathcal{D}/\mathcal{G}$, are computed by approximating D^{-1} by a "tangential" operator $G_{\beta(v)} \in \mathcal{G}$. If the operators of $\mathcal{G}$ satisfy the transport property (16), then the conservation equation can be rewritten as a transport equation whose velocity term depends upon $\vec{\nabla}_v \beta(v)$, called *deformation gradient*. This deformation gradient characterizes the equivalence class of D in $\mathcal{D}/\mathcal{G}$. Section 2.1 gives the general transport equation, and Sections 2.2, 2.3 and 2.4 apply this result to one-dimensional warpings, frequency modulations, and multidimensional warpings.

2.1 Transport in Groups

We suppose that all operators G_β in the stationarity invariant group $\mathcal{G}$ satisfy the transport property (16) and can thus be written

$$G_\beta f(x) = \mathrm{e}^{i(\xi\cdot x+\phi)}\, f \star \rho_\gamma(Sx - v)\ .$$

The translation parameter v is isolated because of its particular role, and since the phase has no influence on the autocorrelation, ϕ is also set apart. We assume that ρ_γ belongs to a a finite-dimensional Lie group (convolution group), so $\mathcal{G}$ is also a finite-dimensional Lie group. We write

$$G_\beta = \mathrm{e}^{i\phi}\, \tilde{G}_{\tilde{\beta}}\, T_v$$

with

$$\tilde{G}_{\tilde{\beta}} f(x) = \mathrm{e}^{i\xi\cdot x}\, f \star \rho_\gamma(Sx) \quad \text{and } \tilde{\beta} = (\xi, S, \gamma)\ .$$

The group product and inverse are denoted

$$\tilde{G}_{\tilde{\beta}_1}\, \tilde{G}_{\tilde{\beta}_2} = \tilde{G}_{\tilde{\beta}_1 * \tilde{\beta}_2} \ \text{ and }\ \tilde{G}^{-1}_{\tilde{\beta}} = \tilde{G}_{\tilde{\beta}^{-1}}\ .$$

To identify the tangential deformation $G_{\beta(v)} \in \mathcal{G}$ which approximates D^{-1} for functions supported in a neighborhood of $v \in \mathbb{R}^d$, we use a family of test functions constructed from a single function $\psi(x)$ whose support is in $[-1,1]^d$. For $\sigma > 0$, $\psi_\sigma(x) = \psi(x/\sigma)$ has a support in $[-\sigma,\sigma]^d$. Let $\overline{\tilde{G}}_{\tilde{\beta}}$ be the adjoint of $\tilde{G}_{\tilde{\beta}}$. We define an *atomic decomposition* of a process $Y(x)$ by computing inner products in $\mathbb{L}^2(\mathbb{R}^d)$ with deformed and translated test functions, which are called *atoms*:

$$A^\sigma_Y(u,\tilde{\beta}) = \mathsf{E}\{|\langle Y, T_u\, \overline{\tilde{G}}_{\tilde{\beta}}\, \psi_\sigma\rangle|^2\}\ .$$

This atomic decomposition only depends on Y through its autocorrelation.

Let us now explain how to identify the tangential deformation $G_{\beta(v)}$ from a conservation property of atomic decompositions. If R is a stationary process, then

$$A^\sigma_R(u,\tilde{\beta}) = \mathsf{E}\{|\langle R, T_u\, \overline{\tilde{G}}_{\tilde{\beta}}\, \psi_\sigma\rangle|^2\}$$

does not depend upon u, hence $\vec{\nabla}_u A^\sigma_R(u,\tilde{\beta}) = 0$. This is not the case for the atomic decomposition of the deformed process $F = D\,R$:

$$A^\sigma_F(u,\tilde{\beta}) = \mathsf{E}\{|\langle F, T_u\, \overline{\tilde{G}}_{\tilde{\beta}}\, \psi_\sigma\rangle|^2\} = \mathsf{E}\{|\langle R, \overline{D}\, T_u\, \overline{\tilde{G}}_{\tilde{\beta}}\, \psi_\sigma\rangle|^2\}\ .$$

Yet, this atomic decomposition satisfies a conservation property along characteristic lines that depend upon D. The following proposition proves that if there exists $\overline{G}_{\beta(v)}$ which approximates $\overline{D^{-1}}$ for functions having a support in a neighborhood of v, then there exists $\tilde{\beta}(u)$ such that for all u and $\tilde{\beta}$

$$\vec{\nabla}_u A_F^\sigma(u, \tilde{\beta} * \tilde{\beta}(u)) \approx 0 \quad \text{for } \sigma \text{ small.}$$

Before stating the proposition, let us set some notation: if $f(x)$ and $g(x)$ are two functions with $x \in \mathbb{R}^d$, then $\vec{\nabla}_x g$ is a vector with d components, and the inner product $\langle f, \vec{\nabla}_x g\rangle$ is also a vector whose d components are the inner products $\left\langle f, \frac{\partial g}{\partial x_k}\right\rangle$. We denote $Re\langle f, \vec{\nabla}_x g\rangle$ the real part of this vector. We write $c(\sigma) = O(\sigma)$ if there exists a constant C such that for σ small, $|c(\sigma)| \le C\,\sigma$, without specifying the sign.

Proposition 2.1. *Let $\beta(v)$ and ψ be such that for each $v \in \mathbb{R}^d$ and each $\tilde{\beta}$, the function $\psi_{v,\tilde{\beta},\sigma} = \overline{G}_{\beta(v)} T_v \overline{\tilde{G}}_{\tilde{\beta}} \psi_\sigma$ satisfies*

$$|Re\langle K_F \psi_{v,\tilde{\beta},\sigma}\,,\, \overline{D^{-1}}(\vec{\nabla}_v + \vec{\nabla}_x)\overline{D}\,\psi_{v,\tilde{\beta},\sigma}\rangle| = O(\sigma)\,|Re\langle K_F \psi_{v,\tilde{\beta},\sigma}\,,\, \vec{\nabla}_x \psi_{v,\tilde{\beta},\sigma}\rangle|\;. \tag{18}$$

If there exists a differentiable invertible map $u(v)$ such that

$$\overline{G}_{\beta(v)}\, T_v = \mathrm{e}^{i\phi(u(v))}\, T_{u(v)}\, \overline{\tilde{G}}_{\tilde{\beta}(u(v))}\;, \tag{19}$$

then for each $(u, \tilde{\beta})$, at $t = u$,

$$\left|\vec{\nabla}_u A_F^\sigma(u, \tilde{\beta} * \tilde{\beta}(t)) + \vec{\nabla}_t A_F^\sigma(u, \tilde{\beta} * \tilde{\beta}(t))\right| = O(\sigma)\left|\vec{\nabla}_u A_F^\sigma(u, \tilde{\beta} * \tilde{\beta}(t))\right|\;. \tag{20}$$

The norms in (20) are Euclidean norms of d-dimensional vectors. The proof is in Appendix B.1. One can verify that if $\overline{G}_{\beta(v)} = \overline{D^{-1}}$, then the left-hand side of (18) vanishes. Condition (18) imposes a form of tangency between $\overline{G}_{\beta(v)}$ and $\overline{D^{-1}}$; however, it does not only depend on operators $\overline{G}_{\beta(v)}$ and $\overline{D^{-1}}$, but also on ψ and on the autocorrelation of R.

The partial differential equation (20) which results from the above proposition can be written as a transport equation in the $(u;\tilde{\beta})$ domain, by expanding the gradient with respect to t:

$$\vec{\nabla}_t A_F^\sigma(u, \tilde{\beta} * \tilde{\beta}(t)) = \vec{\nabla}_t(\tilde{\beta} * \tilde{\beta}(t)) \cdot \vec{\nabla}_{\tilde{\beta}} A_F^\sigma(u, \tilde{\beta} * \tilde{\beta}(t))\;,$$

where $\vec{\nabla}_{\tilde{\beta}} A_F^\sigma(u, \tilde{\beta})$ is a vector of partial derivatives with respect to each component of parameter $\tilde{\beta}$. Replacing the free variable $\tilde{\beta}$ by $\tilde{\beta} * \tilde{\beta}^{-1}(u)$ in (20) gives, at $t = u$,

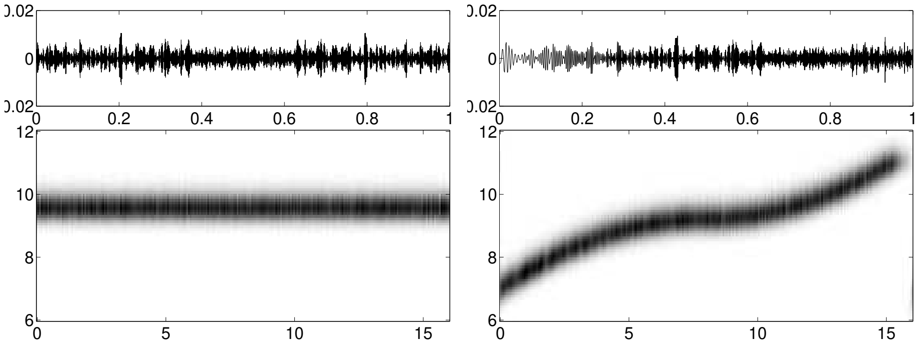
$$\left|\vec{\nabla}_u A_F^\sigma(u, \tilde{\beta}) + \vec{\nabla}_t(\tilde{\beta} * \tilde{\beta}^{-1}(u) * \tilde{\beta}(t)) \cdot \vec{\nabla}_{\tilde{\beta}} A_F^\sigma(u, \tilde{\beta})\right| = O(\sigma)\left|\vec{\nabla}_u A_F^\sigma(u, \tilde{\beta})\right|\;. \tag{21}$$

When σ is sufficiently small, the right-hand side is neglected, yielding a transport equation whose velocity term depends upon $\vec{\nabla}_u \tilde{\beta}(u)$. This is illustrated in the next three sections, which apply this proposition to recover warping deformations and frequency modulations. Section 3 will show how, from a single realization of F, we can estimate the partial derivatives of $A_F^\sigma(u, \tilde{\beta})$ and compute the deformation gradient.

2.2 Scale Transport.

If D is a one-dimensional warping deformation $Df(x) = f(\alpha(x))$ with $x \in \mathbb{R}$, then $\overline{D^{-1}} f(x) = \alpha'(x)\, f(\alpha(x))$. The stationarity invariant subgroup is the affine group, whose elements are $G_\beta f(x) = f(u + sx)$ with $\beta = (u, s)$. The adjoint of G_β is

$$\overline{G}_\beta f(x) = s^{-1}\, f\left((x-u)/s\right) = T_u\, \overline{\tilde{G}}_{\tilde{\beta}} f(x) \;\text{ with }\; \overline{\tilde{G}}_{\tilde{\beta}} f(x) = s^{-1}\, f\left(x/s\right)\;.$$



(a) (b)

Figure 1: (a) Scalogram $A_R(u, s)$ of a stationary process R. The horizontal and vertical axes respectively represent u and $-\log s$. Darkness of a point is proportional to the value of $A_R(u, s)$. (b) Scalogram $A_F(u, s)$ of a warped process F.

Let ψ be a function whose integral vanishes: $\int \psi(x) dx = 0$. It is called a *wavelet*. Using the above expression of the adjoint operator $\overline{\tilde{G}}_{\tilde{\beta}}$, the atomic decomposition $A_Y^\sigma(u, \tilde{\beta}) = \mathsf{E}\{|\langle Y,\, T_u\, \overline{\tilde{G}}_{\tilde{\beta}} \psi_\sigma\rangle|^2\}$ can be written, for $\tilde{\beta} = s$,

$$A_Y^\sigma(u, s) = \mathsf{E}\left\{\left|\left\langle Y(x)\,,\, s^{-1}\psi\left((s\,\sigma)^{-1}(x-u)\right)\right\rangle\right|^2\right\}.$$

We reduce the number of parameters by dividing $A_Y^\sigma(u, s)$ by σ^2, and replacing the product $s\sigma$ by a single scale parameter s. The resulting atomic decomposition

$$A_Y(u,s) = \mathsf{E}\left\{ \left| \langle Y(x)\,,\, s^{-1}\,\psi\left(s^{-1}(x-u)\right)\rangle \right|^2 \right\} \ . \tag{22}$$

is called a *scalogram*, and can be interpreted as the expected value of a squared wavelet transform. Figure 1(a) shows the scalogram $A_R(u,s)$ of a stationary process R. As expected, its value does not depend upon u. Figure 1(b) gives $A_F(u,s)$ for a warped process $F(x) = DR(x) = R(\alpha(x))$. The warping causes the values of the scalogram of R to migrate in the $(u;\log s)$ plane.

Let us now give the expression of $\beta(u)$ corresponding to the tangential approximation of Proposition 2.1. For regular functions f supported in a neighborhood of v, a tangential approximation of $\overline{D^{-1}}$ is calculated with a first order Taylor expansion of $\alpha(x)$ in a neighborhood of $u(v) = \alpha^{-1}(v)$:

$$\overline{D^{-1}}f(x) \approx \alpha'(u)\, f\left(v + \alpha'(u)(x-u)\right) = \overline{G}_{\beta(v)} f(x) \ . \tag{23}$$

Operators $\overline{D^{-1}}$ and $\overline{G}_{\beta(v)}$ both translate the support of f from a neighborhood of v to a neighborhood of $u(v)$.

In order to derive a transport equation from Proposition 2.1, we must make some assumptions on the autocorrelation of R, that will also guarantee uniqueness of the inverse warping problem. Proposition 1.2 shows that it is necessary to specify the behavior of the autocorrelation kernel $c_R(x)$ in a neighborhood of 0. The following theorem supposes that $c_R(x)$ is nearly h-homogeneous in a neighborhood of 0. Partial derivatives are denoted $\frac{\partial f}{\partial a} = \partial_a f$.

Theorem 2.1 (Scale Transport). *Let R be a stationary process such that there exists $h > 0$ with*

$$c_R(0) - c_R(x) = |x|^h\, \eta(x) \tag{24}$$

where η is $\mathbf{C}^1$ in a neighborhood of 0, and $\eta(0) > 0$. Let $\psi(x)$ be a $\mathbf{C}^1$ function supported in $[-1,1]$, such that

$$\int \psi(x)\,dx = 0 \quad \textit{and} \quad \iint |x-y|^h\, \psi^*(x)\,\psi(y)\,dx\,dy \neq 0 \ . \tag{25}$$

If

$$F(x) = R(\alpha(x)) \ ,$$

where $\alpha(x)$ is $\mathbf{C}^3$ and $\alpha'(x) > 0$, then for each $u \in \mathbb{R}$ such that $\alpha''(u) \neq 0$, when s tends to zero

$$\Big(1 + O(s)\Big)\, \partial_u A_F(u,s) - (\log\alpha')'(u)\, \partial_{\log s} A_F(u,s) = 0 \ . \tag{26}$$

The proof is in Appendix B.2. The conditions imposed on c_R and ψ in this theorem guarantee that $\partial_{\log s} A_F(u,s)$ does not vanish. The deformation gradient $(\log\alpha')'(u)$ which specifies the equivalence class of D in $\mathcal{D}/\mathcal{G}$ can thus be computed from (26) by letting s go to zero. It is therefore not surprising that (24) imposes a stronger condition on c_R than the uniqueness condition (6) of Proposition 1.2. The estimation of $(\log\alpha')'(u)$ from a single realization of F will be studied in Section 3.1.

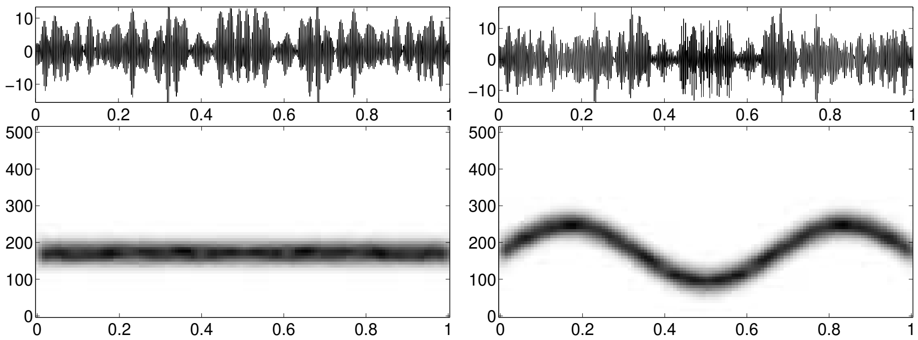
2.3 Frequency Transport

If the deformation operator D is a frequency modulation, $Df(x) = \mathrm{e}^{i\alpha(x)}\, f(x)$, the stationarity invariant subgroup $\mathcal{G}$ is composed of operators G_β such that

$$G_\beta f(x) = \mathrm{e}^{i(\phi+\xi x)}\, f(x)\ .$$

In this case $\tilde{G}_{\tilde{\beta}} f(x) = \mathrm{e}^{i\xi x}\, f(x)$ so $\tilde{\beta} = \xi$. Let us choose an even, positive window function $\psi(x) \geq 0$, with a support equal to $[-1,1]$. The atomic decomposition of process Y is the well-known *spectrogram*:

$$A_Y^\sigma(u,\xi) = \mathsf{E}\{|\langle Y(x)\,,\,\psi_\sigma(x-u)\mathrm{e}^{-i\xi(x-u)}\rangle|^2\} = \mathsf{E}\{|\langle Y(x)\,,\,\psi_\sigma(x-u)\mathrm{e}^{-i\xi x}\rangle|^2\}\ .$$



(a) (b)

Figure 2: (a) Spectrogram $A_R^\sigma(u,\xi)$ of a stationary process R. The horizontal and vertical axes respectively represent position u, and frequency ξ. The darkness of a point is proportional to the value of $A_R^\sigma(u,\xi)$. (b) Spectrogram $A_F^\sigma(u,\xi)$ of a frequency modulated process F.

Figure 2(a) shows a spectrogram $A_R^\sigma(u,\xi)$, whose values do not depend upon u because R is stationary. Figure 2(b) depicts $A_F^\sigma(u,\xi)$ for $F(x) = DR(x) = \mathrm{e}^{i\alpha(x)}\, R(x)$, with $\alpha(x) = \lambda_1 \cos(\lambda_2\, x)$ where λ_1 and λ_2 are two constants. The frequency modulation translates the spectrogram of R non-uniformly along the frequency axis.

Let us now give the expression of $\beta(v)$ corresponding to the tangential approximation of Proposition 2.1, when D is a frequency modulation. If f is supported in a neighborhood of v, a first order Taylor expansion of $\alpha(x)$ gives

$$\overline{D^{-1}} f(x) = \mathrm{e}^{i\alpha(x)}\, f(x) \approx \mathrm{e}^{i(\alpha(v)+\alpha'(v)(x-v))}\, f(x)$$

and one can define a tangential approximation of $\overline{D^{-1}}$ for functions supported in a neighborhood of v by

$$\overline{G}_{\beta(v)} f(x) = \mathrm{e}^{i(\alpha(v)+\alpha'(v)(x-v))}\, f(x)\ . \tag{27}$$

The following theorem uses this tangential approximation to derive from Proposition 2.1 a transport equation, satisfied by the spectrogram $A_F^\sigma(u,\xi)$ in the $(u;\xi)$ plane, when the window scale σ decreases to 0. The frequency ξ is chosen large enough so that the period of $\mathrm{e}^{i\xi x}$ is smaller than the support size σ of ψ_σ. We set $\xi = \xi_0/\sigma$ and select ξ_0 so that $\hat\psi(\omega)$ and its first $\lceil h \rceil + 2$ derivatives vanish at $\omega = \xi_0$ ($\lceil h \rceil$ denoting the smallest integer larger or equal to h).

Theorem 2.2 (Frequency transport). *Let R be a stationary process such that there exists $h > 0$ with*

$$c_R(0) - c_R(x) = |x|^h\, \eta(x) \tag{28}$$

where η is continuous in a neighborhood of 0, and $\eta(0) > 0$. Let ψ be an even, positive, $\mathbf{C}^1$ function supported in $[-1,1]$ and ξ_0 be such that $\hat\psi(\omega)$ and its first $\lceil h\rceil+2$ derivatives vanish at $\omega = \xi_0$ but

$$\iint |x-y|^h (x-y)\, \sin[\xi_0(x-y)]\, \psi(x)\, \psi(y)\, dx\, dy \neq 0\ .$$

If

$$F(x) = \mathrm{e}^{i\alpha(x)}\, R(x) \quad \textit{where } \alpha(x) \textit{ is } \mathbf{C}^{\lceil h\rceil+4},$$

then for each $u \in \mathbb{R}$ such that $\alpha''(u) \neq 0$ and for $\xi = \xi_0/\sigma$, when $\sigma \to 0$

$$\Big(1 + O(\sigma^2)\Big)\, \partial_u A_F^\sigma(u,\xi) - \alpha''(u)\, \partial_\xi A_F^\sigma(u,\xi) = 0\ . \tag{29}$$

The proof is in Appendix B.3. To satisfy the theorem hypothesis, one may choose $\psi(x)$ to be a box spline obtained by convolving the indicator function $1_{[-1/2m,1/2m]}$ with itself m times:

$$\hat\psi(\omega) = \left(\frac{\sin(\omega/(2m))}{\omega/(2m)}\right)^m\ ,$$

and $\xi_0 = 2\,m\,\pi$ with $m \geq \lceil h \rceil + 3$.
The deformation gradient $\alpha''(u)$ can be characterized from equation (29) by letting σ go to zero, and we proved in (3) that $\alpha''(u)$ specifies the equivalence class of D in $\mathcal{D}/\mathcal{G}$. Section 3.2 imposes some further conditions on c_R and α to obtain a consistent estimation of $\alpha''(u)$ from this partial differential equation.

2.4 Multidimensional scale transport

For a multidimensional warping where $Df(x) = f(\alpha(x))$ with $x \in \mathbb{R}^d$, the adjoint of D^{-1} is $\overline{D^{-1}}f(x) = \det J_\alpha(x)\, f(\alpha(x))$. The matrix $J_\alpha(x)$ is the Jacobian matrix (7) of α at position x. The stationarity invariant group $\mathcal{G}$ is the affine group, composed of operators G_β with $\beta = (u, S) \in \mathbb{R}^d \times GL^+(\mathbb{R}^d)$, such that

$$G_\beta f(x) = f(u + Sx) \ .$$

The adjoint of G_β is

$$\overline{G}_\beta f(x) = \det S^{-1}\, f\Big(S^{-1}(x-u)\Big) = T_u\, \overline{\tilde{G}}_{\tilde{\beta}} f(x)$$

where $\overline{\tilde{G}}_{\tilde{\beta}} f(x) = \det S^{-1}\, f(S^{-1}x)$ and

$$\tilde{\beta} = S = \Big(s_{l,m}\Big)_{1 \leq l,m \leq d} \ .$$

For a regular function f, a Taylor expansion of $\alpha(x)$ in a neighborhood of $u(v) = \alpha^{-1}(v)$ gives

$$\overline{D^{-1}}f(x) \approx \det J_\alpha(u)\, f\Big(\alpha(u) + J_\alpha(u)\,(x-u)\Big) = \overline{G}_{\beta(v)} f(x) \ . \tag{30}$$

The operators $\overline{D^{-1}}$ and $\overline{G}_{\beta(v)}$ both translate the support of f from a neighborhood of v to a neighborhood of $u(v) = \alpha^{-1}(v)$.

Let ψ be a function such that $\int_{\mathbb{R}^d} \psi(x)\,dx = 0$. A multidimensional extension of the scalogram is given by

$$\begin{aligned} A_Y^\sigma(u,S) &= \mathsf{E}\{|\langle Y(x)\,,\, \det S^{-1}\,\psi_\sigma\left(S^{-1}(x-u)\right)\rangle|^2\} \\ &= \mathsf{E}\{|\langle Y(x)\,,\, \det S^{-1}\,\psi\left(\sigma^{-1}S^{-1}(x-u)\right)\rangle|^2\} \ . \end{aligned}$$

As in the one-dimensional case, we divide $A_Y^\sigma(u,s)$ by σ^{2d} and replace the product σS by a matrix which we still denote S. The resulting atomic decomposition is

$$A_Y(u,S) = \mathsf{E}\{|\langle Y(x)\,,\, \det S^{-1}\,\psi\left(S^{-1}(x-u)\right)\rangle|^2\} \ . \tag{31}$$

It is similar to the scalogram (22) but since the scale parameter s is replaced by a warping matrix S, we call it a *warpogram*.

For a one-dimensional warping, the velocity term of transport equation (26) is $(\log \alpha')'(u) = \alpha''(u)/\alpha'(u)$. In two dimensions it becomes a set of matrices, indexed by the direction k of spatial differentiation:

$$\text{for } 1 \leq k \leq d, \;\; J_\alpha^{-1}(u)\,\frac{\partial J_\alpha(u)}{\partial u_k} = \left(\gamma_{l,m}^k(u)\right)_{1\leq l,m\leq d} . \tag{32}$$

This set of matrices has been shown in (10) to specify the equivalence class of D in $\mathcal{D}/\mathcal{G}$. It is denoted in a vectorial form:

$$\vec{\gamma}_{l,m}(u) = (\gamma_{l,m}^k(u))_{1\leq k\leq d} .$$

The partial derivative $\partial_{\log s} A_F(u,s) = s\,\partial_s A_F(u,s)$ which appears in the one-dimensional transport equation (26) now becomes a matrix product, between a partial derivatives matrix and the transpose S^t of S:

$$\left(\frac{\partial A_F(u,S)}{\partial s_{i,j}}\right)_{1\leq i,j\leq d} S^t = \left(a_{l,m}(u,S)\right)_{1\leq l,m\leq d} . \tag{33}$$

The following theorem isolates the scale parameter $\sigma = (\det S)^{1/d}$ by writing $S = \sigma\,\tilde{S}$ with $\det \tilde{S} = 1$, and gives a d-dimensional transport equation when σ goes to zero.

Theorem 2.3. *Suppose that $F(x) = R(\alpha(x))$, where $\alpha(x)$ is $\mathbf{C}^3$ and $\det J_\alpha(x) > 0$. Suppose that the autocorrelation kernel c_R of R satisfies*

$$c_R(0) - c_R(x) = |x|^h\,\eta(x) , \tag{34}$$

with $\eta(0) > 0$ and $\eta \in \mathbf{C}^2$ in a neighborhood of 0. For each $u \in \mathbb{R}^d$ and for each $\tilde{S}$ with $\det \tilde{S} = 1$, if there exists $C(u,\tilde{S}) > 0$ such that, for $S = \sigma\,\tilde{S}$ and σ small enough,

$$\left| Re \iint \vec{\nabla} c_R(S(x-y))\nabla J_\alpha(u) J_\alpha^{-1}(u) S(x-y)\,\psi^*(x)\,\psi(y)\,dx\,dy \right| \geq C(u,\tilde{S})\,\sigma^h , \tag{35}$$

then when σ goes to zero

$$\left| \vec{\nabla}_u A_F(u,S) - \sum_{l,m=1}^{d} \vec{\gamma}_{l,m}(u)\,a_{l,m}(u,S) \right| = O(\sigma)\,\left|\vec{\nabla}_u A_F(u,S)\right| . \tag{36}$$

The proof of this Theorem is in Appendix B.4.

If $c_R(0) - c_R(x) = \eta\,|x|^h$ for small $|x|$, with $\eta > 0$, and if $\alpha(x)$ is a separable warping function of the form

$$\alpha(x) = (\alpha_1(x_1),\, \dots,\, \alpha_d(x_d))$$

then one can verify that condition (35) is equivalent to

$$\sum_{i=1}^{d} \frac{\alpha_i''(u)}{\alpha_i'(u)} Re \iint |\tilde{S}(x-y)|^{h-2} (\sum_{j=1}^{d} \tilde{S}_{ij}(x_j - y_j))^2 \psi^*(x)\psi(y)\, dx\, dy \neq 0 \ .$$

For σ sufficiently small, neglecting the error term on the right-hand side of (36) yields d scalar equations:

$$\text{for } 1 \leq k \leq d, \quad \partial_{u_k} A_F(u,S) - \sum_{l,m=1}^{d} \gamma_{l,m}^k(u)\, a_{l,m}(u,S) = 0 \ .$$

For any (u, S), the values $\partial_{u_k} A_F(u,S)$ and $a_{l,m}(u,S)$ depend upon the autocorrelation of F, and have to be estimated. For each direction k, there are d^2 unknown coefficients $\gamma_{l,m}^k(u)$ equal to the d^2 matrix components of $J_\alpha^{-1}(u)\, \partial_{u_k} J_\alpha(u)$. To compute them we need to invert a linear system:

$$\begin{pmatrix} a_{1,1}(u,S_1) & a_{1,2}(u,S_1) & \dots & a_{d,d}(u,S_1) \\ \vdots & \vdots & \vdots & \vdots \\ a_{1,1}(u,S_{d^2}) & a_{1,2}(u,S_{d^2}) & \dots & a_{d,d}(u,S_{d^2}) \end{pmatrix} \begin{pmatrix} \gamma_{1,1}^k(u) \\ \gamma_{1,2}^k(u) \\ \vdots \\ \gamma_{d,d}^k(u) \end{pmatrix} = \begin{pmatrix} \partial_{u_k} A_F(u,S_1) \\ \vdots \\ \partial_{u_k} A_F(u,S_{d^2}) \end{pmatrix} \ . \quad (37)$$

Changing the direction index k only modifies the right-hand side of (37). Note that in order for the system to be invertible, the left-hand side matrix in (37) must have full rank. The matrices S_k must therefore be appropriately chosen, and the inverse warping problem must have a unique solution. This is not always the case, as shown by the example in (11).

3 Estimation of Deformations

The deformation gradient appears as a velocity vector in the transport (21). To recover it from a single realization of F, the derivatives $\vec{\nabla}_u A_F^\sigma(u,\tilde{\beta})$ and $\vec{\nabla}_{\tilde{\beta}} A_F^\sigma(u,\tilde{\beta})$ of the atomic decomposition of F have to be estimated. With a single realization, a sample mean estimator has a variance of the same order of magnitude as the term it estimates. This variance can be reduced with a spatial smoothing, while the bias, which is proportional to the width of the smoothing kernel, is controlled. The next three sections study the consistency of such smoothed estimator for one-dimensional warpings, frequency modulations and multidimensional warpings.

3.1 Warping in one dimension

The scalogram of F is defined as

$$A_F(u,s) = \mathsf{E}\{|\langle F\,,\,\psi_{u,s}\rangle|^2\}\ ,$$

with $\psi_{u,s}(x) = s^{-1}\psi((x-u)/s)$. If $F(x) = R(\alpha(x))$ then Theorem 2.1 proves that

$$\Big(1+O(s)\Big)\,\partial_u A_F(u,s) - (\log\alpha')'(u)\partial_{\log s}A_F(u,s) = 0\ . \tag{38}$$

To reduce the variance of empirical estimators, equation (38) is convolved with a smoothing kernel, which is chosen equal to

$$g(x) = \begin{cases} \Delta^{-1}(1-|x/\Delta|) & \text{if } |x|\leq\Delta \\ 0 & \text{if } |x|>\Delta \end{cases}\ . \tag{39}$$

Let a be a generic variable denoting either u or $\log s$. We define

$$\overline{\partial_a A_F}(u,s) = \int g(u-v)\,\partial_a A_F(v,s)\,dv\ . \tag{40}$$

The following proposition, whose proof is in Appendix C.1, shows that the bias introduced by convolving equation (38) with g is proportional to Δ.

Proposition 3.1. *Under the hypotheses of Theorem 2.1, for each $u\in\mathbb{R}$, when Δ tends to zero and $s<\Delta$,*

$$\Big(1+O(s)\Big)\,\overline{\partial_u A_F}(u,s) - \Big((\log\alpha')'(u)+O(\Delta)\Big)\,\overline{\partial_{\log s}A_F}(u,s) = 0\ . \tag{41}$$

An integration by parts shows that, for $a=u$,

$$\overline{\partial_u A_F}(u,s) = \Delta^{-2}\int_{u-\Delta/2}^{u+\Delta/2}\big[A_F(v+\Delta/2,s) - A_F(v-\Delta/2,s)\big]\,dv\ .$$

Given a discretized realization of F measured at a resolution N, wavelet coefficients $\langle F,\psi_{u,s}\rangle$ and $\langle F,\partial_a\psi_{u,s}\rangle$ can only be computed at scales $s\geq N^{-1}$ and at positions $u=k/N$ with $k\in\mathbb{Z}$. We therefore introduce the following empirical estimator for $\overline{\partial_u A_F}(u,s)$ at scale s:

$$\widehat{\partial_u A_F}(u,s) = \Delta^{-2}N^{-1}\sum_{|k/N-u|\leq\Delta/2}\big[|\langle F,\psi_{k/N+\Delta/2,s}\rangle|^2 - |\langle F,\psi_{k/N-\Delta/2,s}\rangle|^2\big]\ . \tag{42}$$

Noticing that

$$\partial_{\log s}A_F(u,s) = 2\,Re[\mathsf{E}\{\langle F,\psi_{u,s}\rangle\,\langle F,\partial_{\log s}\psi_{u,s}\rangle^*\}]$$

with

$$\partial_{\log s}\psi_{u,s}(x) = -\psi_{u,s}(x) - s^{-2}\,(x-u)\,\psi'(s^{-1}(x-u))\ , \tag{43}$$

we choose an empirical estimator of $\overline{\partial_{\log s}A_F}(u,s)$ at scale s given by

$$\widehat{\partial_{\log s}A_F}(u,s) = 2N^{-1}\sum_{|k/N-u|\le\Delta} g(u-k/N)\,Re\left[\langle F,\psi_{k/N,s}\rangle\,\langle F,\partial_{\log s}\psi_{k/N,s}\rangle^*\right]\ . \tag{44}$$

In view of equation (41), we suggest the following estimator for $(\log\alpha')'(u)$:

$$\widehat{(\log\alpha')}'(u) = \frac{\widehat{\partial_u A_F}(u,N^{-1})}{\widehat{\partial_{\log s}A_F}(u,N^{-1})}\ .$$

To guarantee that $\widehat{\partial_a A_F}(u,s)$ and $\overline{\partial_a A_F}(u,s)$ are close when $s=N^{-1}$ and N increases, we must ensure that the wavelet coefficients $\langle F,\psi_{k/N,s}\rangle$ and $\langle F,\partial_{\log s}\psi_{k/N,s}\rangle$ are sequences of random variables that have fast spatial decorrelation. This will depend upon the behavior of the autocorrelation kernel $c_R(x)$ of R in a neighborhood of 0, and on the number of vanishing moments of ψ. A wavelet $\psi(x)$ has p vanishing moments if

$$\int x^k\,\psi(x)\,dx = 0 \quad \text{for } 0\le k<p\ .$$

The following theorem proves the weak consistency of the above estimator $\widehat{(\log\alpha')}'(u)$ of $(\log\alpha')'(u)$.

Theorem 3.1 (Consistency, warping). *Let $F(x)=R(\alpha(x))$, where R is a stationary Gaussian process such that there exists $h>0$ with*

$$c_R(0)-c_R(x) = |x|^h\,\eta(x) \quad \textit{and} \quad \eta(0)>0\ . \tag{45}$$

Let ψ be a $\mathbf{C}^2$ wavelet supported in $[-1,1]$ with p vanishing moments, such that

$$2p-h>1/2 \ \textit{and}\ \iint |x-y|^h\,\psi^*(x)\,\psi(y)\,dx\,dy \neq 0\ .$$

If $\eta(x)$ is $\mathbf{C}^{2p}$ in a neighborhood of 0, and if $\alpha(x)\in\mathbf{C}^3\cap\mathbf{C}^{2p}$, then for each $u\in\mathbb{R}$ such that $\alpha''(u)\neq 0$, for $\Delta=N^{-1/5}$,

$$\mathrm{Prob}\left\{\left|\frac{\widehat{\partial_u A_F}(u,N^{-1})}{\widehat{\partial_{\log s}A_F}(u,N^{-1})} - (\log\alpha')'(u)\right| \le 2\,(\log N)\,N^{-1/5}\right\} \xrightarrow[N\to\infty]{} 1\ . \tag{46}$$

This theorem, whose proof is in Appendix C.2, relates the size Δ of the smoothing kernel to the resolution N. Although we supposed R to be stationary, since all estimations are based on wavelet coefficients, one can easily verify that the same results

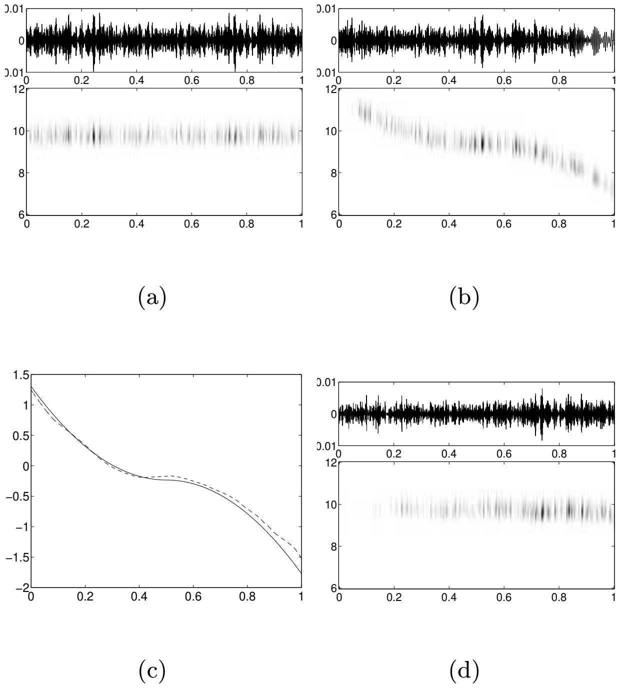


Figure 3: (a) Stationary signal R and its empirical scalogram $|\langle R, \psi_{u,s} \rangle|^2$. (b) Warped signal $F(x) = R(\alpha(x))$ and its empirical scalogram. (c) $\log \alpha'(x) = \lambda_1 + \lambda_2 \mathrm{sign}(1/2 - x)\,|x - 1/2|^2$, where λ_1 and λ_2 are two constants (full line) and its estimation from F (dashed line). (d) Stationarized signal and its empirical scalogram.

apply if R is not stationary but has stationary increments. This is the case of fractional Brownian motion [1, 6], for which $\eta(x) = 1$.

Figure 3 displays a numerical experiment conducted on a single realization of a warped process. The signal F in Figure 3(b) is obtained by warping a stationary signal R, depicted in Figure 3(a). Figure 3(c) shows in dotted lines the estimate $\widehat{\log \alpha'}$ of of $\log \alpha'$ obtained by integrating the estimate $(\widehat{\log \alpha'})'(u)$, and choosing the additive integration constant so that $\int_0^1 \exp \widehat{\log \alpha'} = \int_0^1 \alpha'$. An estimate $\widehat{\alpha}$ for the warping function can be obtained up to an additive constant by integrating $\exp \widehat{\log \alpha'}$. It is then possible to stationarize the deformed signal F by computing $F \circ (\widehat{\alpha})^{-1}$. Figure 3(d) displays such a stationarized signal.

3.2 Frequency Modulation

For a frequency modulated process, $F(x) = R(x) \mathrm{e}^{i\alpha(x)}$, Theorem 2.2 shows that the deformation gradient $\alpha''(u)$ can be computed from a spectrogram

$$A_F^\sigma(u, \xi) = \mathsf{E}\{|\langle F(x)\,,\, \psi_\sigma(x-u) \mathrm{e}^{i\xi(x-u)} \rangle|^2\}$$

with the equation

$$\Big(1 + O(\sigma)\Big)\, \partial_u A_F^\sigma(u, \xi) - \alpha''(u)\, \partial_\xi A_F^\sigma(u, \xi) = 0 \tag{47}$$

evaluated at a frequency $\xi = \xi_0/\sigma$. To estimate $\alpha''(u)$ from a single realization of F measured at a resolution N, the estimation is performed as in the previous section, with a spatial smoothing of equation (47).

Let $g(x)$ be the smoothing kernel defined in (39). For a generic variable a denoting either u or ξ, we define

$$\overline{\partial_a A_F^\sigma}(u, \xi) = \int g(u - v)\, \partial_a A_F^\sigma(v, \xi)\, dv \ . \tag{48}$$

Similarly to Proposition 3.1, we prove in Appendix C.3.1 that

$$\Big(1 + O(\sigma^2)\Big)\, \overline{\partial_u A_F^\sigma}(u, \xi_0/\sigma) - \Big(\alpha''(u) + O(\Delta)\Big)\, \overline{\partial_\xi A_F^\sigma}(u, \xi_0/\sigma) = 0 \ . \tag{49}$$

To compute an estimator of the smoothed partial derivatives of the spectrogram, we relate the spectrogram coefficients to a particular wavelet transform. Observe that

$$\psi_\sigma(x - u)\, \exp\left(i\xi_0\, \frac{x-u}{\sigma}\right) = \psi^1\left(\frac{x-u}{\sigma}\right) \tag{50}$$

where

$$\psi^1(x) = \psi(x)\, \mathrm{e}^{i\xi_0 x} \ . \tag{51}$$

If $\hat{\psi}(\omega)$ has a zero of order $\lceil h \rceil + 3$ at $\omega = \xi_0$, since ψ is real, $\hat{\psi}(\omega)$ is even, and hence

$$\int x^k \, \psi^1(x) dx = (-i)^k \, \frac{d^k \hat{\psi}^1}{d\omega^k}(-\xi_0) = 0 \quad \text{for } k \leq \lceil h \rceil + 2 \; .$$

This means that ψ^1 is a wavelet with $\lceil h \rceil$+3 vanishing moments [10]. We write $\psi^1_{u,\sigma}(x) = \sigma^{-1} \, \psi^1(\sigma^{-1}(x-u))$. The scalogram associated to this wavelet is defined by $A_F(u,\sigma) = \mathsf{E}\{|\langle F, \psi^1_{u,\sigma} \rangle|^2\}$. It results from (50) that

$$A^{\sigma}_F(u, \xi_0/\sigma) = \mathsf{E}\left\{ \left| \langle F(x), \psi^1\left(\sigma^{-1}(x-u)\right) \rangle \right|^2 \right\} = \sigma^2 \, A_F(u,\sigma) \; ,$$

and hence

$$\overline{\partial_u A^{\sigma}_F}(u, \xi_0/\sigma) = \sigma^2 \, \overline{\partial_u A_F}(u,\sigma) \; .$$

Let $\widehat{\partial_u A_F}(u,\sigma)$ be the empirical estimator defined in (42): we choose to estimate $\overline{\partial_u A^{\sigma}_F}(u, \xi_0/\sigma)$ with

$$\widehat{\partial_u A^{\sigma}_F}(u, \xi_0/\sigma) = \sigma^2 \, \widehat{\partial_u A_F}(u,\sigma) \; .$$

To compute an empirical estimator of the other partial derivative, $\overline{\partial_\xi A^{\sigma}_F}(u, \xi_0/\sigma)$, observe that

$$\partial_\xi A^{\sigma}_F(u,\xi) = 2 \, Re[\mathsf{E}\{\langle F(x) \, , \, \psi_\sigma(x-u) \mathrm{e}^{i\xi(x-u)} \rangle \langle F(x) \, , \, \partial_\xi \left[\psi_\sigma(x-u)\mathrm{e}^{i\xi(x-u)}\right] \rangle^* \}] \; .$$

Introducing a new wavelet

$$\psi^2(x) = x \, \psi^1(x) = x \, \psi(x) \, \mathrm{e}^{i\xi_0 x} \; , \tag{52}$$

and $\psi^2_{u,\sigma}(x) = \sigma^{-1} \, \psi^2(\sigma^{-1}(x-u))$, this partial derivative can be rewritten, for $\xi = \xi_0/\sigma$:

$$\partial_\xi A^{\sigma}_F(u, \xi_0/\sigma) = 2 \, \sigma^3 \, Im[\mathsf{E}\{\langle F \, , \, \psi^1_{u,\sigma} \rangle \, \langle F \, , \, \psi^2_{u,\sigma} \rangle^* \}] \; .$$

Similarly to (44), for $\sigma = N^{-1}$ we suggest the empirical estimator

$$\widehat{\partial_\xi A^{\sigma}_F}(u, \xi_0/\sigma) = 2 \, \sigma^3 \, N^{-1} \sum_{|k/N-u| \leq \Delta} g(u-k/N) \, Im\left[\langle F, \psi^1_{k/N,\sigma} \rangle \, \langle F, \psi^2_{k/N,\sigma} \rangle^* \right] \; . \tag{53}$$

The following theorem proves that for $\sigma = N^{-1}$,

$$\widehat{\alpha''}(u) = \frac{\widehat{\partial_u A^{\sigma}_F}(u, \xi_0/\sigma)}{\widehat{\partial_\xi A^{\sigma}_F}(u, \xi_0/\sigma)}$$

is a weakly consistent estimator of $\alpha''(u)$ as $N \to \infty$.

Theorem 3.2 (Consistency, frequency modulation). *Let $F(x) = R(x)\, e^{i\alpha(x)}$, where R is a Gaussian process such that there exists $h > 0$ with*

$$c_R(0) - c_R(x) = |x|^h\, \eta(x) \quad \textit{and} \quad \eta(0) > 0\ . \tag{54}$$

Suppose that $\psi^1(x) = \psi(x) e^{i\xi_0 x}$ is a compactly supported wavelet with $p \geq \lceil h \rceil + 3$ vanishing moments, such that

$$\iint |x-y|^h (x-y)\, \sin[\xi_0(x-y)]\, \psi(x)\, \psi(y)\, dx\, dy \neq 0\ .$$

If $\eta \in \mathbf{C}^{2p}$ in a neighborhood of 0 and if $\alpha \in \mathbf{C}^{2p}$, then for each $u \in \mathbb{R}$, for $\Delta = N^{-1/5}$,

$$\mathrm{Prob}\left\{ \left| \frac{\widehat{\partial_u A_F^\sigma}(u, N\xi_0)}{\widehat{\partial_\xi A_F^\sigma}(u, N\xi_0)} - \alpha''(u) \right| \leq 2\,(\log N)\, N^{-1/5} \right\} \xrightarrow[N\to\infty]{} 1\ . \tag{55}$$

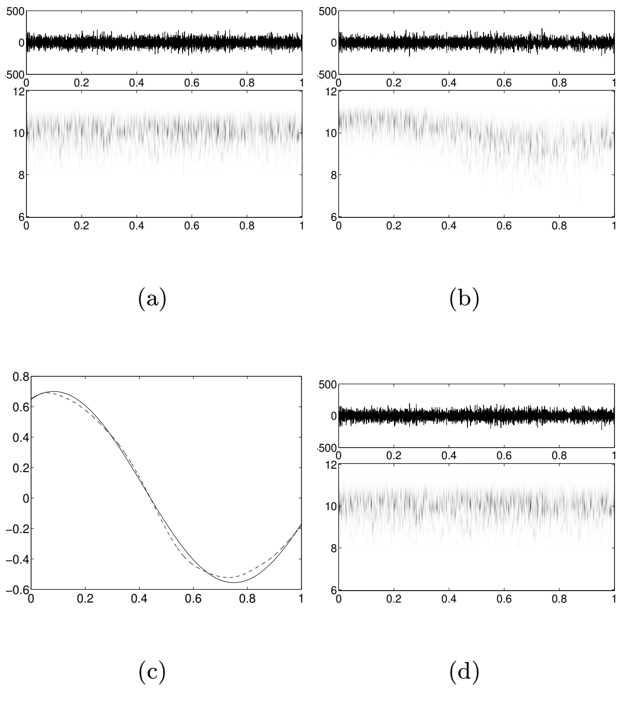


Figure 4: (a) Stationary signal R and its empirical scalogram $|\langle R, \psi^1_{u,s}\rangle|^2$. (b) Frequency modulated signal $F(x) = R(x)\exp(i\alpha(x))$ and its empirical scalogram. (c) Frequency modulation $\alpha'(x)$ (full line), and its estimation from F (dashed line). (d) Stationarized signal and its empirical scalogram.

The proof is in Appendix C.3.

The numerical example in Figure 4 shows the estimation of a frequency modulation. We explained that the empirical estimator $\widehat{\alpha''}(u)$ is in fact computed from wavelet coefficients associated to the two wavelets ψ^1 and ψ^2 defined in (51) and (52). Figure 4(a) shows a realization of a stationary signal $R(x)$ and the corresponding empirical scalogram $\widehat{A_R}(u,s) = |\langle R, \psi^1_{u,s}\rangle|^2$. The frequency modulated signal $F(x) = R(x)\ \exp(i\alpha(x))$ and its empirical scalogram are in Figure 4(b). The derivative α' of the frequency modulation is plotted in Figure 4(c) (full line). An estimate $\widehat{\alpha'}$ of α' is obtained by integrating $\widehat{\alpha''}$, and choosing the additive integration constant so that $\int_0^1 \widehat{\alpha'} = \int_0^1 \alpha'$. Figure 4(c) plots $\widehat{\alpha'}$ (dashed line), superposed on the theoretical function α' (full line). Lastly, Figure 4(d) represents the stationarized process $F(x)\exp(-i\widehat{\alpha}(x))$ and its empirical scalogram.

3.3 Warping in higher dimension

For a multidimensional warping, at each position u, the deformation gradient corresponds to a set of d matrices $\vec{\gamma}_{l,m}(u) = (\gamma^k_{l,m}(u))_{1\leq k\leq d}$ defined in (32). Theorem 2.3 shows that these coefficients appear in the velocity term of the transport equation (36) satisfied by the warpogram of F:

$$A_F(u,S) = \mathsf{E}\{|\langle F, \psi_{u,S}\rangle|^2\}\ ,$$

with

$$\psi_{u,S}(x) = (\det S^{-1})\, \psi(S^{-1}(x-u))\ .$$

At a sufficiently small scale σ, the error on the right-hand side of the transport equation (36) can be neglected. The vector transport equation can then be written as a linear system

$$\begin{pmatrix} a_{1,1}(u,S_1) & a_{1,2}(u,S_1) & \dots & a_{d,d}(u,S_1) \\ \vdots & \vdots & \vdots & \vdots \\ a_{1,1}(u,S_{d^2}) & a_{1,2}(u,S_{d^2}) & \dots & a_{d,d}(u,S_{d^2}) \end{pmatrix} \begin{pmatrix} \gamma^k_{1,1}(u) \\ \vdots \\ \gamma^k_{d,d}(u) \end{pmatrix} = \begin{pmatrix} \partial_{u_k} A_F(u,S_1) \\ \vdots \\ \partial_{u_k} A_F(u,S_{d^2}) \end{pmatrix} \tag{56}$$

where

$$\Big(a_{l,m}(u,S)\Big)_{1\leq l,m\leq d} = \big(\partial_{s_{i,j}} A_F(u,S)\big)_{1\leq i,j\leq d}\, S^t\ .$$

If the process F is measured at a resolution N, we can only compute the warpogram with functions $\psi_{u,S}$ whose support in any direction is larger than N^{-1}. We therefore require that $S = \sigma\,\tilde{S}$ where $\sigma \geq KN^{-1}$, and all the eigenvalues of $\tilde{S}$ are greater than K^{-1}. The location parameter is also restricted to a uniform grid $u = N^{-1}k$ with $k \in \mathbb{Z}^d$. To estimate the deformation gradient from a single realization of F, as in the one-dimensional case, the system of equations (56) is convolved with a d-dimensional kernel of radius Δ. The smoothed matrix coefficients are

$$\left(\overline{a_{l,m}}(u,S)\right)_{1\leq l,m\leq d} = \left(\overline{\partial_{s_{i,j}} A_F}(u,S)\right)_{1\leq i,j\leq d} S^t$$

where, for any variable a, we have defined

$$\overline{\partial_a A_F}(u,S) = \int g(u-v)\,\partial_a A_F(v,S)\,dv\ .$$

Since $\partial_a A_F(u,S) = 2\,Re\mathsf{E}\{\langle F,\psi_{u,S}\rangle\langle F,\partial_a\psi_{u,S}\rangle^*\}$, an empirical estimator of $\overline{\partial_a A_F}(u,S)$ is

$$\widehat{\partial_a A_F}(u,S) = 2N^{-d} \sum_{|N^{-1}k-u|\leq\Delta} g(u-N^{-1}k)\,Re\{\langle F,\psi_{N^{-1}k,S}\rangle\langle F,\partial_a\psi_{N^{-1}k,S}\rangle^*\}\ ,$$

and we define

$$\left(\widehat{a_{l,m}}(u,S)\right)_{1\leq l,m\leq d} = \left(\widehat{\partial_{s_{i,j}} A_F}(u,S)\right)_{1\leq i,j\leq d} S^t\ .$$

If the warping $\alpha(x)$ is sufficiently regular, then the averaged values $\overline{a_{l,m}}(u,S)$ and $\overline{\partial_{u_k} A_F}(u,S)$ are related by a system of equations identical to (56), but with an additional bias equal to $O(\Delta)$.

The k^{th} component of the deformation gradient, $(\gamma^k_{l,m}(u))$, can be thus estimated by

$$\begin{pmatrix} \widehat{\gamma^k_{1,1}}(u) \\ \vdots \\ \widehat{\gamma^k_{d,d}}(u) \end{pmatrix} = \begin{pmatrix} \widehat{a_{1,1}}(u,S_1) & \dots & \widehat{a_{d,d}}(u,S_1) \\ \vdots & \vdots & \vdots \\ \widehat{a_{1,1}}(u,S_{d^2}) & \dots & \widehat{a_{d,d}}(u,S_{d^2}) \end{pmatrix}^{-1} \begin{pmatrix} \widehat{\partial_{u_k} A_F}(u,S_1) \\ \vdots \\ \widehat{\partial_{u_k} A_F}(u,S_{d^2}) \end{pmatrix}\ . \tag{57}$$

Extending the consistency Theorem 3.1 to this d-dimensional case is possible, but requires technical hypotheses that are not yet well understood.

In Section 1.1, we mentioned that the warping of textures in images specify the three-dimensional coordinates of the corresponding surfaces appearing in the scene. The estimator defined in (57) is studied numerically in [4, 5] to compute surfaces from texture gradients in images.

The Matlab routines which reproduce the numerical illustrations of this section can be downloaded from
`http://www.cmap.polytechnique.edu/~mallat/Deform.html`.

A Proofs of Section 1

A.1 Proof of Proposition 1.1

Let R be a stationary process, and suppose that there exists an $\varepsilon > 0$ such that $c_R(x) > 0$ for $|x| < \varepsilon$. Let $\tilde{R}$ be another stationary process. We want to show that if the autocorrelations of $R(x)\exp[i\alpha(x)]$ and of $\tilde{R}(x)\exp[i\tilde{\alpha}(x)]$ are equal, i.e. if

$$c_R(x-y)\exp(i[\alpha(x)-\alpha(y)]) = c_{\tilde{R}}(x-y)\exp(i[\tilde{\alpha}(x)-\tilde{\alpha}(y)]), \tag{58}$$

then $\alpha''(x) = \tilde{\alpha}''(x)$. The functions α and $\tilde{\alpha}$ are assumed $\mathbf{C}^4$, therefore $\theta = \alpha - \tilde{\alpha}$ is also $\mathbf{C}^4$. Let us fix $x \in \mathbb{R}$; our goal is to prove that $\theta''(x) = 0$. We choose $y \in \mathbb{R}$ such that $|x-y| < \varepsilon$. After dividing both sides of (58) by $c_R(x-y) > 0$, it appears that $\mathrm{e}^{i[\theta(x)-\theta(y)]}$ is a function of $x-y$. Therefore $\theta(x)-\theta(y)$ is also a function of $x-y$, and in particular for all a,

$$\theta(x) - \theta(y) = \theta(x+a) - \theta(y+a) \ .$$

Differentiating this expression with respect to x shows that $\theta'(x) = \theta'(x+a)$, thus $\theta''(x) = 0$.

A.2 Proof of Proposition 1.2

Let R be a stationary process and let $\varepsilon > 0$ such that $c_R(x)$ is $\mathbf{C}^1$ for $0 < |x| < \varepsilon$, with $c'_R(x) < 0$. Let $\tilde{R}$ denote another stationary process, and let us suppose that the autocorrelations of $R(\alpha(x))$ and of $\tilde{R}(\tilde{\alpha}(x))$ are equal. The functions α and $\tilde{\alpha}$ are assumed $\mathbf{C}^3$, therefore $\theta = \alpha \circ \tilde{\alpha}^{-1}$ is also $\mathbf{C}^3$. Proving the proposition amounts to proving that θ is linear, or equivalently, that θ'' vanishes everywhere. By definition of θ,

$$c_{\tilde{R}}(x-y) = c_R(\theta(x)-\theta(y)) \ . \tag{59}$$

Let us fix $x \in \mathbb{R}$, and choose $y \neq x$, but sufficiently close to x so that $|\theta(x)-\theta(y)| < \varepsilon$. Differentiating (59) with respect to x and y shows that

$$c'_R(\theta(x)-\theta(y))\theta'(y) = c'_R(\theta(x)-\theta(y))\theta'(x) \ .$$

Since $c'_R(\theta(x)-\theta(y)) < 0$, we obtain $\theta'(x) = \theta'(y)$, therefore $\theta''(x) = 0$.

A.3 Proof of Proposition 1.3

Let R be a stationary process such that c_R satisfies (12). Let $\tilde{R}$ denote another stationary process, and suppose that the autocorrelations of $R(\alpha(x))$ and of $\tilde{R}(\tilde{\alpha}(x))$ are equal. Let $\theta = \alpha \circ \tilde{\alpha}^{-1}$: by definition of θ, $R(\theta(x)) = \tilde{R}(x)$, thus

$$c_R(\theta(x) - \theta(y)) = c_{\tilde{R}}(x - y) \ .$$

Differentiating this expression with respect to x and y, for $x \neq y$, shows that

$$\vec{\nabla} c_R(\theta(x) - \theta(y)) J_\theta(y) = \vec{\nabla} c_R(\theta(x) - \theta(y)) J_\theta(x) \ . \tag{60}$$

Let us fix $x \in \mathbb{R}^d$, and prove that $\vec{\nabla} J_\theta(x) = 0$. Let $\varepsilon > 0$ such that η is $\mathbf{C}^2$ on $\{z,\ |z| < \varepsilon\}$. Let us choose $y \in \mathbb{R}^d$ such that $|\theta(x) - \theta(y)| < \varepsilon$ and let $z = \theta(x) - \theta(y)$:

$$\vec{\nabla} c_R(z) = -h\,|z|^{h-2}\,(\eta(z)\,z + |z|^2 \vec{\nabla}\eta(z)) \ .$$

Replacing this expression in (60) and dividing both sides by $-h\,|z|^{h-2}\,\eta(z)$ proves that

$$(z + h^{-1}|z|^2 \vec{\nabla} \log \eta(z))\, J_\theta(y) = (z + h^{-1}|z|^2 \vec{\nabla} \log \eta(z))\, J_\theta(\theta^{-1}(z + \theta(y))) \ ,$$

so

$$(z + h^{-1}|z|^2 \vec{\nabla} \log \eta(z))\, J_\theta(y) J_\theta^{-1}(\theta^{-1}(z + \theta(y))) = z + h^{-1}|z|^2 \vec{\nabla} \log \eta(z) \ .$$

Introducing a function $\tilde{\theta}$ such that

$$\tilde{\theta}(z) = J_\theta(y)\,\theta^{-1}(z + \theta(y)) \ , \tag{61}$$

this can be rewritten

$$(z + h^{-1}|z|^2 \vec{\nabla} \log \eta(z))\, J_{\tilde{\theta}}(z) = (z + h^{-1}|z|^2 \vec{\nabla} \log \eta(z)) \ .$$

Noticing that $z\, J_{\tilde{\theta}}(\lambda z) = \frac{d}{d\lambda} \tilde{\theta}(\lambda\, z)$, we have, for $\lambda \geq 0$,

$$\frac{d}{d\lambda} \tilde{\theta}(\lambda\, z) = z + h^{-1}|z|^2\, \lambda\, \vec{\nabla} \log \eta(z) (J_{\tilde{\theta}}(\lambda z) - Id) \ ,$$

which, when integrated between $\lambda = 0$ and $\lambda = 1$, gives

$$\tilde{\theta}(z) - \tilde{\theta}(0) = z + h^{-1}|z|^2 \int_0^1 \lambda \vec{\nabla} \log \eta(\lambda z) (J_{\tilde{\theta}}(\lambda z) - Id)\, d\lambda \ .$$

After replacing $\tilde{\theta}$ with (61), and noticing that $\tilde{\theta}(0) = J_\theta(y)\, y$, we obtain

$$\theta^{-1}(z + \theta(y)) = J_\theta^{-1}(y)\, z + y + J_\theta^{-1}(y) h^{-1}|z|^2 \int_0^1 \lambda \vec{\nabla} \log \eta(\lambda z) (J_{\tilde{\theta}}(\lambda z) - Id)\, d\lambda \ . \tag{62}$$

Since c_R is even, $\vec{\nabla}\eta(0) = 0$ and so $\vec{\nabla}\log\eta(0) = 0$. Let us denote $\vec{\nabla}\log\eta(\lambda z) = |\lambda z|\,\vec{a}(\lambda z)$. Recalling that η is twice continuously differentiable in a neighborhood of λz for $0 \le \lambda \le 1$, the vector $\vec{a}(\lambda z)$ is differentiable, with a bounded gradient. Differentiating (62) with respect to z shows that

$$J_\theta^{-1}(\theta^{-1}(z+\theta(y))) = J_\theta^{-1}(y)(Id + |z|^2\,A(z)) \ ,$$

and one can check that $A(z)$ is uniformly bounded for $|z| < \varepsilon$. Replacing z by $\theta(x)-\theta(y)$ gives

$$J_\theta^{-1}(x) = J_\theta^{-1}(y)\left(Id + |\theta(x)-\theta(y)|^2\,A(\theta(x)-\theta(y))\right) \ ,$$

therefore for any unit-length vector $x_k \in \mathbb{R}^d$,

$$\begin{aligned}
\frac{\partial}{\partial x_k} J_\theta^{-1}(x) &= \lim_{\lambda\to 0} \frac{J_\theta^{-1}(x+\lambda x_k) - J_\theta^{-1}(x)}{\lambda} \\
&= -\lim_{\lambda\to 0} \frac{J_\theta^{-1}(x+\lambda x_k)|\theta(x+\lambda x_k)-\theta(x)|^2}{\lambda}\,A(\theta(x+\lambda x_k)-\theta(x)) \\
&= 0 \ .
\end{aligned}$$

This proves that $\vec{\nabla} J_\theta^{-1}(x) = 0$, and therefore $\vec{\nabla} J_\theta(x) = 0$. As a consequence, for each direction x_k, $\frac{\partial}{\partial x_k} J_\theta(\tilde{\alpha}(x)) = 0$. Since $J_\theta(x) = J_\alpha(\tilde{\alpha}^{-1}(x)) J_{\tilde{\alpha}}^{-1}(\tilde{\alpha}^{-1}(x))$, we obtain

$$\frac{\partial}{\partial x_k}(J_\alpha(x)\,J_{\tilde{\alpha}}^{-1}(x)) = 0 \ ,$$

and expanding the above differential expression then proves (10).

A.4 Proof of Theorem 1.1

Let us consider a specific family of zero-mean wide-sense stationary processes defined by

$$R_\omega(x) = X\,\mathrm{e}^{i\omega\cdot x}$$

where X is a zero-mean random variable with variance σ^2. Then

$$\mathsf{E}\{R_\omega(x)\,R_\omega^*(y)\} = \sigma^2\,\exp(i\omega\cdot(x-y)) = c_{R_\omega}(x-y) \ .$$

Let G be a stationarity invariant operator. If $F_\omega(x) = GR_\omega(x)$ then

$$\mathsf{E}\{F_\omega(x)\,F_\omega^*(y)\} = \sigma^2\,f_\omega(x)\,f_\omega^*(y) = c_{F_\omega}(x-y) \ .$$

with $f_\omega(x) = G\mathrm{e}^{i\omega\cdot x}$. This implies that for any (x, y) the product $f_\omega(x)\, f_\omega(y)^*$ is a function of $x - y$. One can thus conclude that there exists $\hat\rho(\omega) \in \mathbb{C}$ and $\theta(\omega) \in \mathbb{R}^d$ such that

$$f_\omega(x) = G\mathrm{e}^{i\omega\cdot x} = \hat\rho(\omega)\, \mathrm{e}^{i\theta(\omega)\cdot x} \ . \tag{63}$$

Let us now prove that such operators are stationarity invariant. A zero-mean process is wide-sense stationary if and only if it admits a spectral representation:

$$R(x) = \int_{\mathbb{R}^d} \mathrm{e}^{i\omega\cdot x}\, dZ(\omega) \ ,$$

where $Z(\omega)$ is an orthogonal process. Let $dP(\omega) = \mathsf{E}\{|dZ(\omega)|^2\}$, we have

$$c_R(0) = \int_{\mathbb{R}^d} dP(\omega) < +\infty \ .$$

We derive from (63) that

$$F(x) = \int_{\mathbb{R}^d} \mathrm{e}^{i\theta(\omega)\cdot x}\, \hat\rho(\omega)\, dZ(\omega) \ ,$$

which shows that F is wide-sense stationary.
For any wide-sense stationary process R, one can write

$$GR(x) = G\mathsf{E}\{R(0)\} + G\left(R(x) - \mathsf{E}\{R(0)\}\right) \ .$$

Since $R(x) - \mathsf{E}\{R(0)\}$ is zero-mean and wide-sense stationary, $G\left(R(x) - \mathsf{E}\{R(0)\}\right)$ is wide-sense stationary, therefore so is $GR(x)$.

For any positive integrable measure $dP(\omega)$ we must have

$$c_F(0) = \int_{\mathbb{R}^d} |\hat\rho(\omega)|^2\, dP(\omega) < +\infty \ ,$$

and a necessary and sufficient condition is that $\operatorname{ess\,sup}_{\omega\in\mathbb{R}^d} |\hat\rho(\omega)| < \infty$.

A.5 Proof of Proposition 1.4

We denote by K_Y the autocorrelation operator of a process Y, defined by

$$K_Y f(x) = \int \mathsf{E}\{Y(x)Y^*(y)\}\, f(y)\, dy \ .$$

Let G be a bounded linear operator and $F = G\,R$. The autocorrelation operators of F and R satisfy

$$K_F = G\, K_R\, \bar G \ .$$

Since R is stationary, K_R commutes with the translation operator T_v for any $v \in \mathbb{R}^d$. We derive from (16) that K_F also commutes with T_v and hence that F is wide-sense stationary. The operator G is therefore stationarity invariant and Theorem 1.1 proves that

$$G\,\mathrm{e}^{i\omega\cdot x} = \hat{\rho}(\omega)\,\mathrm{e}^{i\theta(\omega)\cdot x}\,. \tag{64}$$

Inserting this expression in the equality $GT_{Sv}f(x) = \mathrm{e}^{i\xi\cdot v}T_v\,Gf(x)$ for $f(x) = \mathrm{e}^{i\omega\cdot x}$ implies

$$\hat{\rho}(\omega)\,\mathrm{e}^{i\theta(\omega)\cdot x}\,\mathrm{e}^{-iSv\cdot\omega} = \hat{\rho}(\omega)\,\mathrm{e}^{i\theta(\omega)\cdot(x-v)}\,\mathrm{e}^{-i\xi\cdot v}\quad,$$

from which we derive that $\theta(\omega) = \overline{S}\omega + \xi$ for all ω where $\hat{\rho}(\omega) \neq 0$. For ω such that $\hat{\rho}(\omega) = 0$, (64) clearly holds with $\theta(\omega) = \overline{S}\omega + \xi$. So G can indeed be written as in (17).

Conversely, if G satisfies (17) then a direct calculation shows that (16) holds.

B Proofs of Section 2

B.1 Proof of Proposition 2.1 (Transport)

The autocorrelation operator of $F = D\,R$ satisfies $K_F = D\,K_R\,\overline{D}$, therefore

$$\langle K_F\,\psi_{v,\tilde{\beta},\sigma}\,,\,\psi_{v,\tilde{\beta},\sigma}\rangle = \langle K_R\overline{D}\psi_{v,\tilde{\beta},\sigma}\,,\,\overline{D}\psi_{v,\tilde{\beta},\sigma}\rangle\,.$$

Let us compute

$$\begin{aligned}\vec{\nabla}_v\langle K_F\psi_{v,\tilde{\beta},\sigma}\,,\,\psi_{v,\tilde{\beta},\sigma}\rangle &= 2\,Re\,\langle K_R\;\overline{D}\psi_{v,\tilde{\beta},\sigma},\vec{\nabla}_v\overline{D}\psi_{v,\tilde{\beta},\sigma}\rangle\\ &= 2\,Re\,\langle K_R\;\overline{D}\psi_{v,\tilde{\beta},\sigma},\left(\vec{\nabla}_v+\vec{\nabla}_x\right)\overline{D}\psi_{v,\tilde{\beta},\sigma}\rangle - 2\,Re\,\langle K_R\;\overline{D}\psi_{v,\tilde{\beta},\sigma},\vec{\nabla}_x\overline{D}\psi_{v,\tilde{\beta},\sigma}\rangle\,.\end{aligned}$$

Since R is stationary, for any g we have $\langle K_R g,\vec{\nabla}_x g\rangle = 0$, so

$$\vec{\nabla}_v\langle K_F\,\psi_{v,\tilde{\beta},\sigma}\,,\,\psi_{v,\tilde{\beta},\sigma}\rangle = 2\,Re\,\langle K_F\;\psi_{v,\tilde{\beta},\sigma}\,,\,\overline{D^{-1}}\left(\vec{\nabla}_v+\vec{\nabla}_x\right)\overline{D}\psi_{v,\tilde{\beta},\sigma}\rangle\,.$$

Hypothesis (18) thus implies that

$$\left|\vec{\nabla}_v\langle K_F\,\psi_{v,\tilde{\beta},\sigma}\,,\,\psi_{v,\tilde{\beta},\sigma}\rangle\right| = O(\sigma)\,\left|Re\langle K_F\,\psi_{v,\tilde{\beta},\sigma}\,,\,\vec{\nabla}_x\,\psi_{v,\tilde{\beta},\sigma}\rangle\right|\,. \tag{65}$$

Since $\psi_{v,\tilde{\beta},\sigma} = \overline{G}_{\beta(v)}T_v\overline{\tilde{G}}_{\tilde{\beta}}\psi_\sigma$, transport property (19) shows that

$$\psi_{v,\tilde{\beta},\sigma} = \mathrm{e}^{i\phi(u(v))}\,T_{u(v)}\,\overline{\tilde{G}}_{\tilde{\beta}(u)}\overline{\tilde{G}}_{\tilde{\beta}}\psi_\sigma = \mathrm{e}^{i\phi(u(v))}\,T_{u(v)}\,\overline{\tilde{G}}_{\tilde{\beta}*\tilde{\beta}(u)}\psi_\sigma\,.$$

The phase $e^{i\phi(u(v))}$ disappears in all calculations of $\langle K_F \psi_{v,\tilde\beta,\sigma}, \psi_{v,\tilde\beta,\sigma}\rangle$. By definition, $A_F^\sigma(u,\tilde\beta) = \langle K_F T_u \overline{\tilde G}_{\tilde\beta} \psi_\sigma, T_u \overline{\tilde G}_{\tilde\beta}\psi_\sigma\rangle$, and since $\vec\nabla_v f = \vec\nabla_u f \, J_v^{-1}(u)$,

$$\vec\nabla_v \langle K_F\, \psi_{v,\tilde\beta,\sigma}, \psi_{v,\tilde\beta,\sigma}\rangle = \vec\nabla_v A_F^\sigma(u(v), \tilde\beta * \tilde\beta(u(v))) = \vec\nabla_u A_F^\sigma(u(v), \tilde\beta * \tilde\beta(u(v)))\, J_v^{-1}(u) \ .$$

This implies that

$$\left|\vec\nabla_u A_F^\sigma(u(v), \tilde\beta * \tilde\beta(u(v)))\right| \leq \|J_v(u)\| \, \left|\vec\nabla_v \langle K_F\, \psi_{v,\tilde\beta,\sigma} \, , \, \psi_{v,\tilde\beta,\sigma}\rangle\right| \ ,$$

where $\|J_v(u)\|$ is the operator sup norm of $J_v(u)$. Inserting this in (65) shows that for u fixed

$$\left|\vec\nabla_u A_F^\sigma(u, \tilde\beta * \tilde\beta(u))\right| = O(\sigma) \, \left| Re\langle K_F\, \psi_{v,\tilde\beta,\sigma} \, , \, \vec\nabla_x \, \psi_{v,\tilde\beta,\sigma}\rangle\right| \ . \tag{66}$$

Since $\vec\nabla_u T_u = -\vec\nabla_x T_u$, using the symmetry of K_F we get

$$2Re\langle K_F\, \psi_{v,\tilde\beta,\sigma} \, , \, \vec\nabla_x \psi_{v,\tilde\beta,\sigma}\rangle = -\vec\nabla_u A_F^\sigma(u, \tilde\beta * \tilde\beta(t)) \quad \text{at } t = u. \tag{67}$$

Inserting (67) in (66) finally proves that

$$\left|\vec\nabla_u A_F^\sigma(u, \tilde\beta * \tilde\beta(u))\right| = O(\sigma) \, \left|\vec\nabla_u A_F^\sigma(u, \tilde\beta * \tilde\beta(t))\right| \quad \text{at } t = u,$$

which implies (20).

B.2 Proof of Theorem 2.1

This theorem is proved as a consequence of Proposition 2.1. Operator $G_{\beta(v)}$ is given by (23)

$$\overline{G}_{\beta(v)} f(x) = \alpha'(u)\, f(v + \alpha'(u)(x-u))$$

with $u = \alpha^{-1}(v)$. Transport property (19) holds because $u(v) = \alpha^{-1}(v)$ is differentiable and invertible and

$$\overline{G}_{\beta(v)} T_v = T_{u(v)} \, \overline{\tilde G}_{\tilde\beta(u(v))} \ ,$$

where $\overline{\tilde G}_{\tilde\beta(u)} f(x) = \alpha'(u)\, f(\alpha'(u)x)$ with $\tilde\beta(u) = 1/\alpha'(u)$.

Let us now verify hypothesis (18) concerning

$$\psi_{v,\tilde\beta,\sigma} = \overline{G}_{\beta(v)} T_v \overline{\tilde G}_{\tilde\beta} \psi_\sigma$$

with $\overline{\tilde G}_{\tilde\beta} f(x) = 1/s f(x/s)$. The scalogram renormalization (22) is equivalent to dividing $\psi_\sigma(x)$ by σ, which yields $\psi_\sigma(x) = 1/\sigma\, \psi(x/\sigma)$, and replacing σs by s which gives

$$\psi_{v,\tilde\beta,\sigma}(x) = \theta_{v,s}(x) = \frac{\alpha'(u)}{s}\, \psi\left(\frac{\alpha'(u)}{s}(x-u)\right) \ .$$

Let us compute

$$\phi_{v,s} = \overline{D^{-1}}\,(\partial_v + \partial_x)\,\overline{D}\,\theta_{v,s}\ .$$

Since $\overline{D}f(x) = (\alpha'(\alpha^{-1}(x)))^{-1}\,f(\alpha^{-1}(x))$ and $\overline{D^{-1}}f(x) = \alpha'(x)\,f(\alpha(x))$, a direct calculation gives

$$\overline{D^{-1}}\partial_x\overline{D}\,\theta_{v,s}(x) = -\frac{\alpha'(u)\,\alpha''(u)}{s\,|\alpha'(x)|^2}\psi\left(\frac{\alpha'(u)}{s}(x-u)\right) + \frac{|\alpha'(u)|^2}{s^2\,\alpha'(x)}\psi'\left(\frac{\alpha'(u)}{s}(x-u)\right)$$

and

$$\overline{D^{-1}}\partial_v\overline{D}\,\theta_{v,s}(x) = \frac{\alpha''(u)}{s\alpha'(u)}\psi\left(\frac{\alpha'(u)}{s}(x-u)\right) + \frac{1}{s^2}\Big((x-u)\alpha''(u) - \alpha'(u)\Big)\psi'\left(\frac{\alpha'(u)}{s}(x-u)\right)\ ,$$

therefore

$$\phi_{v,s}(x) = \frac{\alpha''(u)}{s\alpha'(u)|\alpha'(x)|^2}(|\alpha'(u)|^2 - |\alpha'(x)|^2)\psi\left(\frac{\alpha'(u)}{s}(x-u)\right) + \frac{1}{s^2}\left(\frac{\alpha'(u)}{\alpha'(x)}(\alpha'(u) - \alpha'(x) - (u-x)\alpha''(u))\right)\psi'\left(\frac{\alpha'(u)}{s}(x-u)\right)\ .$$

Since ψ is supported in $[-1,1]$, $\phi_{v,s}$ is supported in $[u-s/\alpha'(u), u+s/\alpha'(u)]$. Since $\alpha \in \mathbf{C}^3$, a Taylor series expansion of $|\alpha'(x)|^2$ and of $\alpha'(x)$ around position u prove that, for small s,

$$\phi_{v,s}(x) = \left(\frac{2\alpha''(u)^2}{\alpha'(u)^3} + O(s)\right)\psi\left(\frac{\alpha'(u)}{s}(x-u)\right) + \left(\frac{\alpha'''(u)}{2\alpha''(u)^2} + O(s)\right)\psi'\left(\frac{\alpha'(u)}{s}(x-u)\right)$$

The autocorrelation kernel of $F(x)$ is $c_F(x,y) = c_R(\alpha(x)-\alpha(y))$, hence

$$\langle K_F\theta_{v,s}\,,\,\phi_{v,s}\rangle = \iint c_R(\alpha(x)-\alpha(y))\,\theta^*_{v,s}(x)\,\phi_{v,s}(y)\,dx\,dy\ .$$

Since $\int \theta_{v,s}(x)\,dx = \int \psi(x)\,dx = 0$,

$$\langle K_F\theta_{v,s}\,,\,\phi_{v,s}\rangle = -\iint \Big(c_R(0) - c_R(\alpha(x)-\alpha(y))\Big)\,\theta^*_{v,s}(x)\,\phi_{v,s}(y)\,dx\,dy\ .$$

The supports of $\theta_{v,s}$ and $\phi_{v,s}$ are in $[u-s/\alpha'(u), u+s/\alpha'(u)]$ and for z in a neighborhood of 0, the continuity of η implies that $c_R(0) - c_R(z) = \eta(0)\,|z|^h + o(|z|^h)$. Since α' is continuous at u, a Taylor series expansion of α around u combined with a change of variables $x' = (x-u)\alpha'(u)/s$ and $y' = (y-u)\alpha'(u)/s$ yield, for s sufficiently small,

$$\langle K_F\theta_{v,s}\,,\,\phi_{v,s}\rangle = -\iint (\eta(0)|s(x-y)|^h + o(|s(x-y)|^h))\psi^*(x)\,\times$$
$$\times\,\frac{s}{\alpha'(u)}\left[\left(\frac{2\alpha''(u)^2}{\alpha'(u)^3}+O(s)\right)\psi(y)+\left(\frac{\alpha'''(u)}{2\alpha''(u)^2}+O(s)\right)\psi'(y)\right]dx\,dy\ , \quad (68)$$

and therefore

$$|Re\langle K_F\theta_{v,s}\,\phi_{v,s}\rangle| = O(s^{h+1}). \quad (69)$$

Let us now compute

$$\langle K_F\theta_{v,s}\,,\,\partial_x\theta_{v,s}\rangle = \iint c_R(\alpha(x)-\alpha(y))\,\theta^*_{v,s}(x)\,\frac{d}{dy}\theta_{v,s}(y)\,dx\,dy\ .$$

With an integration by parts,

$$\langle K_F\theta_{v,s}\,,\,\partial_x\theta_{v,s}\rangle = -\iint \alpha'(y)\,c'_R(\alpha(x)-\alpha(y))\,\theta^*_{v,s}(x)\,\theta_{v,s}(y)\,dx\,dy\ ,$$

and since $c'_R(z)$ is antisymmetric

$$Re\,\langle K_F\theta_{v,s}\,,\,\partial_x\theta_{v,s}\rangle = \tfrac{1}{2}\iint\Big(\alpha'(x)-\alpha'(y)\Big)\,c'_R(\alpha(x)-\alpha(y))\,\theta^*_{v,s}(x)\,\theta_{v,s}(y)\,dx\,dy\ .$$

A change of variable $x' = (x-u)\alpha'(u)/s$ and $y' = (y-u)\alpha'(u)/s$ gives

$$Re\,\langle K_F\theta_{v,s}\,,\,\partial_x\theta_{v,s}\rangle = \tfrac{1}{2}\iint\Big(\alpha'(u+sx/\alpha'(u))-\alpha'(u+sy/\alpha'(u))\Big)$$
$$c'_R\Big(\alpha(u+sx/\alpha'(u))-\alpha(u+sy/\alpha'(u))\Big)\,\psi^*(x)\,\psi(y)\,dx\,dy\ .$$

Because of assumption (24), since η is $\mathbf{C}^1$ in a neighborhood of 0, $c'_R(z) = h\,\eta(0)\,\mathrm{sign(z)}\,|z|^{h-1} + o(|z|^{h-1})$. With a Taylor expansion for α, we get, for s small enough,

$$Re\,\langle K_F\theta_{v,s}\,,\,\partial_x\theta_{v,s}\rangle = \tfrac{1}{2}\,\frac{\alpha''(u)}{\alpha'(u)}\iint (h\,\eta(0)\,s^h\,|x-y|^h + o(s^h|x-y|^h))\,\psi^*(x)\,\psi(y)\,dx\,dy\ . \quad (70)$$

Since $\iint |x-y|^h\psi^*(x)\,\psi(y)\,dx\,dy \neq 0$ and $\alpha''(u)\neq 0$, there exists $a(u)>0$ such that

$$|Re\,\langle K_F\theta_{v,s}\,,\,\partial_x\theta_{v,s}\rangle| \geq a(u)\,s^h + o(s^h)\ ,$$

and (69) implies that

$$|Re\,\langle K_F\theta_{v,s}\,\phi_{v,s}\rangle| = O(s)\,|Re\,\langle K_F\theta_{v,s}\,,\,\partial_x\theta_{v,s}\rangle|\ .$$

Since all the conditions of Proposition 2.1 have been verified, we can apply the resulting transport equation (21) with $\tilde\beta = s$, $\tilde\beta(u) = 1/\alpha'(u)$ and $\tilde\beta * \tilde\beta(t) = s/\alpha'(t)$:

$$\left|\partial_u A_F(u,s) - s\,\alpha'(u)\,\frac{\alpha''(t)}{(\alpha'(t))^2}\,\partial_s A_F(u,s)\right| = O(s)\,|\partial_u A_F(u,s)| \quad \text{at } t=u,$$

which proves (26).

B.3 Proof of Theorem 2.2.

This theorem is proved as a consequence of Proposition 2.1. The operator $G_{\beta(v)}$ is defined in (27) by:

$$\overline{G}_{\beta(v)} f(x) = \mathrm{e}^{i(\alpha(v)+\alpha'(v)(x-v))} f(x) \ . \tag{71}$$

Transport property (19) holds because, for $u(v) = v$, we have

$$\overline{G}_{\beta(v)} T_v = \mathrm{e}^{i\alpha(u(v))}\, T_{u(v)}\, \overline{\tilde{G}}_{\tilde{\beta}(u(v))}$$

with $\overline{\tilde{G}}_{\tilde{\beta}(u)} f(x) = \mathrm{e}^{i\alpha'(u)x}\, f(x)$ and $\tilde{\beta}(u) = -\alpha'(u)$.

Let now verify hypothesis (18):

$$\left| Re \left\langle K_F \psi_{v,\xi,\sigma}, \overline{D^{-1}}\,(\partial_v + \partial_x)\,\overline{D}\psi_{v,\xi,\sigma} \right\rangle \right| = O(\sigma) \left| Re \left\langle K_F \psi_{v,\xi,\sigma}, \frac{\partial}{\partial x}\psi_{v,\xi,\sigma} \right\rangle \right| \ , \tag{72}$$

for

$$\psi_{v,\xi,\sigma}(x) = \overline{G}_{\beta(v)} T_v \overline{\tilde{G}}_{\xi} \psi_\sigma(x) = \exp[i(\alpha(v) + \alpha'(v)(x-v))]\ \exp[i\xi(x-v)]\, \psi\left(\frac{x-v}{\sigma}\right) \ . \tag{73}$$

A direct calculation shows that

$$\begin{aligned}\left\langle K_F \psi_{v,\xi,\sigma}, \overline{D^{-1}}\,(\partial_v + \partial_x)\,\overline{D}\psi_{v,\xi,\sigma} \right\rangle = \\ \iint c_R(x-y) \exp[i(\alpha(x) - \alpha(y) - \alpha'(v)(x-y))] \exp[i\xi(y-x)] \times \\ \times\, i(-\alpha''(v)(y-v) + \alpha'(y) - \alpha'(v))\psi\left(\frac{x-v}{\sigma}\right)\psi\left(\frac{y-v}{\sigma}\right) dx\, dy \ ,\end{aligned}$$

and with a change of variables $x' = (x-v)/\sigma$ and $y' = (y-v)/\sigma$,

$$\begin{aligned}\left\langle K_F \psi_{v,\xi,\sigma}, \overline{D^{-1}}\,(\partial_v + \partial_x)\,\overline{D}\psi_{v,\xi,\sigma} \right\rangle = \\ \iint c_R(\sigma(x'-y')) \exp[i(\alpha(v+\sigma x') - \alpha(v+\sigma y') - \sigma\alpha'(v)(x'-y'))] \times \\ \times\, i(\alpha'(v+\sigma y') - \alpha'(v) - \alpha''(v)\sigma y')\mathrm{e}^{i\xi_0(y'-x')}\, \psi(x')\psi(y')\, dx'\, dy' \ .\end{aligned} \tag{74}$$

Because $\int \mathrm{e}^{-i\xi_0 x}\psi(x)\, dx = 0$,

$$\iint c_R(0) i(\alpha'(v+\sigma y) - \alpha'(v) - \alpha''(v)\sigma y)\mathrm{e}^{i\xi_0(y-x)}\, \psi(x)\psi(y)\, dx\, dy = 0 \ . \tag{75}$$

After subtracting (75) from (74),

$$
\begin{aligned}
&\left\langle K_F \psi_{v,\xi,\sigma}, \overline{D^{-1}}\,(\partial_v + \partial_x)\,\overline{D}\psi_{v,\xi,\sigma}\right\rangle = \\
&\qquad \iint c_R(\sigma(x-y))\,(\exp[i(\alpha(v+\sigma x) - \alpha(v+\sigma y) - \sigma\alpha'(v)(x-y))] - 1) \times \\
&\qquad\qquad \times\, i(\alpha'(v+\sigma y) - \alpha'(v) - \alpha''(v)\sigma y)\tilde{\psi}^*(x)\tilde{\psi}(y)\,dx\,dy\, + \\
&\iint (c_R(\sigma(x-y)) - c_R(0))\,i(\alpha'(v+\sigma y) - \alpha'(v) - \alpha''(v)\sigma y)\,\mathrm{e}^{i\xi_0(y-x)}\,\psi(x)\psi(y)\,dx\,dy\ .
\end{aligned} \tag{76}
$$

Since $\alpha \in \mathbf{C}^{4+\lceil h \rceil}$, we can perform the following Taylor expansions, where a_k, b_k and c_k are real parameters which depend on the derivatives $\alpha^{(k)}(v)$, for $(x,y) \in [0,1]^2$:

$$
\begin{aligned}
&\exp[i(\alpha(v+\sigma x) - \alpha(v+\sigma y) - \sigma\alpha'(v)(x-y))] = \\
&\qquad 1 + i\sum_{k=2}^{2+\lceil h \rceil} a_k \sigma^k (x-y)^k + \sum_{k=4}^{2+\lceil h \rceil} b_{k-2}\sigma^k (x-y)^k + O(\sigma^{3+\lceil h \rceil})
\end{aligned} \tag{77}
$$

$$
\alpha'(v+\sigma y) - \alpha'(v) = \sum_{k=1}^{2+\lceil h \rceil} c_{k+1}\sigma^k y^k + O(\sigma^{3+\lceil h \rceil})\ . \tag{78}
$$

In particular, $a_2 = \alpha''(v)/2$ and $c_k = \alpha^{(k)}(v)/(k-1)!$.
Replacing these Taylor expansions in (76), we obtain

$$
\begin{aligned}
&\left\langle K_F \psi_{v,\xi,\sigma}, \overline{D^{-1}}\,(\partial_v + \partial_x)\,\overline{D}\psi_{v,\xi,\sigma}\right\rangle = \\
&\quad -\iint c_R(\sigma(x-y)) \left[i\sum_{k=2}^{2+\lceil h \rceil} a_k \sigma^k (x-y)^k + \sum_{k=4}^{2+\lceil h \rceil} b_{k-2}\sigma^k (x-y)^k \right] \times \\
&\qquad\qquad \times \sum_{k=2}^{2+\lceil h \rceil} c_{k+1}\sigma^k y^k\,\mathrm{e}^{i\xi_0(y-x)}\,\psi(x)\psi(y)\,dx\,dy \\
&\quad - i\iint \sigma^h |x-y|^h \eta(\sigma(x-y)) \left[\sum_{k=2}^{2+\lceil h \rceil} c_{k+1}\sigma^k y^k \right] \mathrm{e}^{i\xi_0(y-x)}\,\psi(x)\psi(y)\,dx\,dy \\
&\qquad\qquad\qquad\qquad\qquad\qquad + o(\sigma^{3+\lceil h \rceil})\ .
\end{aligned}
$$

In the first of these two integrals, one can replace $c_R(\sigma(x-y))$ by $c_R(0) - \sigma^h|x-y|^h\eta(\sigma(x-y))$. Since $\hat{\psi}$ and its first $\lceil h \rceil + 2$ derivatives vanish at ξ_0, we derive that $\mathrm{e}^{i\xi_0 t}\,\psi(t)$ is a function with $\lceil h \rceil + 3$ vanishing moments [10], so the first integral is of the order of $O(\sigma^{4+h})$.

For the second integral, because ψ is even, exchanging x and y shows that

$$\iint |x-y|^h \eta(\sigma(x-y)) y^2 \sin(\xi_0(y-x)) \psi(x)\, \psi(y)\, dx\, dy = 0\ .$$

Since $2 + \lceil h \rceil \geq 3$, the second integral has a real part which is equal to

$$-\sigma^{3+h} \eta(0) \frac{\alpha^{(4)}(v)}{6} \iint \sin(\xi_0(x-y))\, |x-y|^h y^3\, \psi(x)\, \psi(y)\, dx\, dy + o(\sigma^{3+h})\ .$$

As a consequence,

$$\begin{aligned} &Re \left\langle K_F \psi_{v,\xi,\sigma}, \overline{D^{-1}}\, (\partial_v + \partial_x)\, \overline{D} \psi_{v,\xi,\sigma} \right\rangle = \\ &\quad - \sigma^{3+h} \eta(0) \frac{\alpha^{(4)}(v)}{6} \iint \sin(\xi_0(x-y))\, |x-y|^h y^3\, \psi(x)\, \psi(y)\, dx\, dy + o(\sigma^{3+h})\ . \end{aligned} \tag{79}$$

Let us now estimate $|Re\, \langle K_F \psi_{v,\xi,\sigma}, \partial_x \psi_{v,\xi,\sigma} \rangle|$. After a change of variables,

$$\begin{aligned} &\langle K_F \psi_{v,\xi,\sigma}, \partial_x \psi_{v,\xi,\sigma} \rangle = \\ &\quad \iint c_R(\sigma(x-y)) \exp[i(\alpha(v+\sigma x) - \alpha(v+\sigma y) - \sigma \alpha'(v)(x-y))] \times \\ &\qquad \times \exp[i\xi_0(y-x)]\, i(\alpha'(v) + \xi_0/\sigma)\, \psi(x)\psi(y)\, dx\, dy \\ &\quad + \iint c_R(\sigma(x-y)) \exp[i(\alpha(v+\sigma x) - \alpha(v+\sigma y) - \sigma \alpha'(v)(x-y))] \times \\ &\qquad \times \exp[i\xi_0(y-x)] \psi(x) \frac{1}{\sigma} \psi'(y)\, dx\, dy\ . \end{aligned}$$

Using Taylor expansions (77) and (78), we obtain

$$\begin{aligned} &\langle K_F \psi_{v,\xi,\sigma}, \partial_x \psi_{v,\xi,\sigma} \rangle = \\ &\quad \iint c_R(\sigma(x-y)) \left[1 + i \sum_{k=2}^{2+\lceil h \rceil} a_k \sigma^k (x-y)^k + \sum_{k=4}^{2+\lceil h \rceil} b_{k-2} \sigma^k (x-y)^k \right] \times \\ &\qquad \times \exp[i\xi_0(y-x)]\, i(\alpha'(v) + \xi_0/\sigma)\, \psi(x)\psi(y)\, dx\, dy \\ &\quad + \iint c_R(\sigma(x-y)) \left[1 + i \sum_{k=2}^{2+\lceil h \rceil} a_k \sigma^k (x-y)^k + \sum_{k=4}^{2+\lceil h \rceil} b_{k-2} \sigma^k (x-y)^k \right] \times \\ &\qquad \times \exp[i\xi_0(y-x)] \psi(x)\, \frac{1}{\sigma} \psi'(y)\, dx\, dy + O(\sigma^{2+\lceil h \rceil})\ . \end{aligned} \tag{80}$$

Exchanging x and y shows that

$$\iint c_R(\sigma(x-y)) \sin[\xi_0(y-x)] \psi(x)\, \psi(y)\, dx\, dy = 0\ ;$$

and since ψ is even and ψ' is odd, changing x to $-x$ and y to $-y$ shows that

$$\iint c_R(\sigma(x-y)) \cos[\xi_0(y-x)] \psi(x) \, \psi'(y) \, dx \, dy = 0 \ .$$

Writing $c_R(\sigma(x-y)) = c_R(0) - \sigma^h |x-y|^h \eta(\sigma(x-y))$, and noticing that $e^{i\xi_0 t} \, \psi(t)$ is a function with $\lceil h \rceil + 3$ vanishing moments, the first integral in (80) has a real part equal to

$$\sigma^{1+h} \frac{\alpha''(v)}{2} \eta(0) \iint |x-y|^{2+h} \xi_0 \cos[\xi_0(y-x)] \, \psi(x) \, \psi(y) \, dx \, dy + o(\sigma^{1+h}) \ .$$

Because ψ is even, the second integral in (80) has a real part equal to

$$\sigma^{1+h} \frac{\alpha''(v)}{2} \eta(0) \iint |x-y|^{2+h} \sin[\xi_0(y-x)] \, \psi(x) \, \psi'(y) \, dx \, dy + o(\sigma^{1+h}) \ .$$

An integration by parts with respect to y shows that

$$\begin{aligned} \iint |x-y|^{2+h} \sin[\xi_0(y-x)] \, \psi(x) \, \psi'(y) \, dx \, dy = \\ - \xi_0 \iint |x-y|^{2+h} \cos[\xi_0(y-x)] \, \psi(x) \, \psi(y) \, dx \, dy + \\ + (2+h) \iint |x-y|^{1+h} \mathrm{sign}(x-y) \, \sin[\xi_0(y-x)] \, \psi(x) \, \psi(y) \, dx \, dy \ . \end{aligned}$$

Summing up the two contributions, we see that

$$\begin{aligned} & Re \, \langle K_F \psi_{v,\xi,\sigma}, \partial_x \psi_{v,\xi,\sigma} \rangle = \\ & -\sigma^{1+h} \, \eta(0) \, (1+h/2) \, \alpha''(v) \iint |x-y|^h (x-y) \, \sin[\xi_0(x-y)] \, \psi(x) \, \psi(y) \, dx \, dy + o(\sigma^{1+h}) \ . \end{aligned} \tag{81}$$

Because of the hypothesis that

$$\iint |x-y|^h (x-y) \, \sin[\xi_0(x-y)] \, \psi(x) \, \psi(y) \, dx \, dy \neq 0 \ ,$$

comparing (81) and (79) proves a result which is stronger than (18), because the right-hand side has an $O(\sigma^2)$ instead of $O(\sigma)$:

$$|Re \left\langle K_F \psi_{v,\xi,\sigma}, \overline{D^{-1}} \, (\partial_v + \partial_x) \, \overline{D} \psi_{v,\xi,\sigma} \right\rangle | = O(\sigma^2) \ |Re \, \langle K_F \psi_{v,\xi,\sigma}, \partial_x \psi_{v,\xi,\sigma} \rangle | \ .$$

With a slight modification of Proposition 2.1 to account for the $O(\sigma^2)$ term, we obtain a transport equation (21) with $\tilde{\beta} = \xi$, $\tilde{\beta}(u) = -\alpha'(u)$ and $\tilde{\beta}_1 * \tilde{\beta}_2 = \tilde{\beta}_1 + \tilde{\beta}_2$: for u such that $\alpha''(u) \neq 0$,

$$|\partial_u A_F^\sigma(u,\xi) - \alpha''(u) \, \partial_\xi A_F^\sigma(u,\xi)| = O(\sigma^2) \, |\partial_u A_F^\sigma(u,\xi)| \ ,$$

which proves (29).

B.4 Proof of Theorem 2.3

The proof of this theorem follows the same lines as the proof of Theorem 2.1. The hypotheses of Proposition 2.1 are verified in order to apply (21) in d dimensions.

The transport property (19) clearly holds. Let us verify hypothesis (18) concerning

$$\psi_{v,\tilde{\beta},\sigma} = \overline{G}_{\beta(v)} T_v \overline{\tilde{G}}_{\tilde{\beta}} \psi_\sigma$$

with $\overline{\tilde{G}}_{\tilde{\beta}} f(x) = \det S^{-1} f(S^{-1}x)$. The warpogram renormalization (31) is equivalent to dividing $\psi_\sigma(x)$ by σ^d and replacing σS by S. Recalling the definition (30) of $\overline{G}_{\beta(v)}$, we introduce

$$\psi_{v,\tilde{\beta},\sigma}(x) = \theta_{v,S}(x) = \det(S^{-1} J_\alpha(u)) \psi(S^{-1} J_\alpha(u)(x-u)) \ . \tag{82}$$

Let us define the vector of functions

$$\vec{\phi}_{v,S} = \overline{D^{-1}} (\vec{\nabla}_v + \vec{\nabla}_x) \overline{D} \, \theta_{v,S} \ .$$

We now prove that for any fixed u and $\tilde{S}$ such that $\det \tilde{S} = 1$, if $S = \sigma \, \tilde{S}$ then

$$\left| Re\langle K_F \theta_{v,S} \, , \vec{\phi}_{v,S} \, \rangle \right| = O(\sigma) \, \left| Re\langle K_F \theta_{v,S} \, , \, \vec{\nabla}_x \theta_{v,S} \rangle \right| \ . \tag{83}$$

Let us first compute an upper bound for $\left| Re\langle K_F \theta_{v,S} \, , \vec{\phi}_{v,S} \, \rangle \right|$. Since

$$\overline{D^{-1}} f(x) = \det(J_\alpha(x)) f(\alpha(x))$$

and

$$\overline{D} f(x) = \det(J_\alpha^{-1}(\alpha^{-1}(x))) f(\alpha^{-1}(x)) \ ,$$

we have

$$\overline{D^{-1}} \vec{\nabla}_x \overline{D} \theta_{v,S}(x) = \sigma^{-d} \left[-\frac{\det J_\alpha(u)}{\det J_\alpha(x)} \vec{\nabla} \det J_\alpha(x) J_\alpha^{-1}(x) \psi(S^{-1} J_\alpha(u)(x-u)) \right.$$
$$\left. + \det J_\alpha(u) \vec{\nabla} \psi(S^{-1} J_\alpha(u)(x-u)) S^{-1} J_\alpha(u) J_\alpha^{-1}(x) \right]$$

and

$$\overline{D^{-1}} \vec{\nabla}_v \overline{D} \theta_{v,S}(x) = \sigma^{-d} \left[\vec{\nabla} \det J_\alpha(u) \psi(S^{-1} J_\alpha(u)(x-u)) J_\alpha^{-1}(u) \right.$$
$$\left. + \det J_\alpha(u) \vec{\nabla} \psi(S^{-1} J_\alpha(u)(x-u)) S^{-1} (\vec{\nabla} J_\alpha(u)(x-u) - J_\alpha(u)) J_\alpha^{-1}(u) \right] \ .$$

After summing these two expressions, a Taylor expansion of $\det J_\alpha$, J_α^{-1} and of $\vec{\nabla} \det J_\alpha$ in the vicinity of position u shows that for $S = \sigma \tilde{S}$ and σ small, there exists $C(u, \tilde{S})$ such that

$$|\vec{\phi}_{v,S}| \leq C(u, \tilde{S}) \, \sigma^{1-d} \ . \tag{84}$$

By definition of K_F,

$$\langle K_F \theta_{v,S}, \vec{\phi}_{v,S} \rangle = \iint c_R(\alpha(x) - \alpha(y)) \theta^*_{v,S}(x) \vec{\phi}_{v,S}(y)\, dx\, dy \ .$$

The wavelet ψ has one vanishing moment, so $\int \theta_{v,S}(x) dx = 0$, and therefore

$$\langle K_F \theta_{v,S}, \vec{\phi}_{v,S} \rangle = \iint \left[c_R(\alpha(x) - \alpha(y)) - c_R(0) \right] \theta^*_{v,S}(x) \vec{\phi}_{v,S}(y)\, dx\, dy \ ,$$

which implies that

$$|\langle K_F \theta_{v,S}, \vec{\phi}_{v,S} \rangle| \leq \iint |c_R(\alpha(x) - \alpha(y)) - c_R(0)|\, |\theta_{v,S}(x)|\, |\vec{\phi}_{v,S}(y)|\, dx\, dy \ .$$

Inserting (84) and (82), and using condition (34) on c_R, after a change of variable and a Taylor expansion of α around u, we obtain

$$|\langle K_F \theta_{v,S}, \vec{\phi}_{v,S} \rangle| = O(\sigma^{h+1}) \ .$$

To prove (83), we now show that there exists $K(u, \tilde{S}) > 0$ such that

$$\left| Re \langle K_F \theta_{v,S} \ , \ \vec{\nabla}_x \theta_{v,S} \rangle \right| \geq K(u, \tilde{S})\, \sigma^h \ . \tag{85}$$

With an integration by parts, and using the fact that $\vec{\nabla} c_R(x)$ is antisymmetric, we get

$$\begin{aligned} \langle K_F \theta_{v,S}, \vec{\nabla}_x \theta_{v,S} \rangle &= \iint \vec{\nabla} c_R(\alpha(x) - \alpha(y)) J_\alpha(y) \theta^*_{v,S}(x) \theta_{v,S}(y)\, dx\, dy \\ &= -\tfrac{1}{2} \iint \vec{\nabla} c_R(\alpha(x) - \alpha(y)) (J_\alpha(x) - J_\alpha(y)) \theta^*_{v,S}(x) \theta_{v,S}(y)\, dx\, dy \ . \end{aligned}$$

Therefore

$$\begin{aligned} &\langle K_F \theta_{v,S}, \vec{\nabla}_x \theta_{v,S} \rangle + \tfrac{1}{2} \iint \vec{\nabla} c_R(S(x-y)) \vec{\nabla} J_\alpha(u) J_\alpha^{-1}(u) S(x-y) \psi^*(x) \psi(y)\, dx\, dy \\ &= -\tfrac{1}{2} \iint (\nabla c_R(\alpha(u + J_\alpha^{-1}(u) Sx) - \alpha(u + J_\alpha^{-1}(u) Sy)) - \nabla c_R(S(x-y))) \times \\ &\qquad \times (J_\alpha(u + J_\alpha^{-1}(u) Sx) - J_\alpha(u + J_\alpha^{-1}(u) Sy))\, \psi^*(x)\, \psi(y)\, dx\, dy \\ &-\tfrac{1}{2} \iint \nabla c_R(S(x-y)) (J_\alpha(u + J_\alpha^{-1}(u) Sx) - J_\alpha(u + J_\alpha^{-1}(u) Sy) - \vec{\nabla} J_\alpha(u) J_\alpha^{-1}(u) S(x-y)) \times \\ &\qquad \times\ \psi^*(x)\, \psi(y)\, dx\, dy \end{aligned}$$

Because $\vec{\nabla} c_R$ is $\mathbf{C}^1$ in a neighborhood of 0 excluding 0, for small σ, second order Taylor series expansions for α and for J_α around position u prove that

$$\langle K_F \theta_{v,S}, \vec{\nabla}_x \theta_{v,S} \rangle + \tfrac{1}{2} \iint \vec{\nabla} c_R(S(x-y)) \vec{\nabla} J_\alpha(u) J_\alpha^{-1}(u) S(x-y) \psi^*(x) \psi(y)\, dx\, dy = o(\sigma^h)$$

Hypothesis (35) on c_R guarantees that (85) holds, and therefore (83) is satisfied. Now that conditions (18) and (19) of Proposition 2.1 have been verified, the resulting transport equation (21) can be applied, with $\tilde{\beta} = S$, $\tilde{\beta}(u) = J_\alpha^{-1}(u)$ and $S_1 * S_2 = S_2\, S_1$. This yields:

$$\left|\vec{\nabla}_u A_F(u,S) + \left[J_\alpha(u)^{-1}\,\vec{\nabla}_u J_\alpha(u)\, S\right] \cdot \vec{\nabla}_S A_F(u,S)\right| = O(\sigma)\left|\vec{\nabla}_u A_F(u,S)\right| \ .$$

The final result (36) is derived from this equation by noting that

$$\left[J_\alpha^{-1}(u)\,\vec{\nabla}_u J_\alpha(u)\, S\right] \cdot \vec{\nabla}_S A_F(u,S) = \left[J_\alpha^{-1}(u)\vec{\nabla}_u J_\alpha(u)\right] \cdot \left[\vec{\nabla}_S A_F(u,S)\, S^t\right] \ .$$

C Proofs of Section 3

C.1 Proof of Proposition 3.1

With a slight modification of the proof of Theorem 2.1, one can prove a stronger result than (26), which is stated in the following lemma:

Lemma C.1. *Under the hypotheses of Theorem 2.1,*

$$\partial_u A_F(u,s) - (\log\alpha')'(u)\partial_{\log s} A_F(u,s) = s(C(u)+o(1))\,\partial_u A_F(u,s) \ , \tag{86}$$

where C is continuous.

Proof. In one dimension, the proof of Proposition 2.1 can be adapted to show that, if (18) is replaced by

$$Re\langle K_F\psi_{v,\tilde{\beta},\sigma}\,,\,\overline{D^{-1}}\,(\partial_v+\partial_x)\,\overline{D}\,\psi_{v,\tilde{\beta},\sigma}\rangle = c(u,\sigma)\,Re\langle K_F\psi_{v,\tilde{\beta},\sigma}\,,\,\partial_x\psi_{v,\tilde{\beta},\sigma}\rangle \ , \tag{87}$$

and if (19) holds, then the resulting transport equation (21) is replaced by

$$\partial_u A_F^\sigma(u,\tilde{\beta}) + \partial_t(\tilde{\beta} * \tilde{\beta}^{-1}(u) * \tilde{\beta}(t))\partial_{\tilde{\beta}} A_F^\sigma(u,\tilde{\beta}) = c(u,\sigma)\partial_u A_F^\sigma(u,\tilde{\beta}) \ .$$

Now, in the proof of Theorem 2.1, (68) proves that

$$Re\langle K_F\theta_{v,s},\phi_{v,s}\rangle = s^{h+1}(B(u)+o(1)) \ , \tag{88}$$

where B is continuous. On the other hand, (70) proves that

$$Re\langle K_F\theta_{v,s},\partial_x\theta_{v,s}\rangle = \tfrac{1}{2}\frac{\alpha''(u)}{\alpha'(u)}s^h(1+o(1))h\eta(0)\iint |x-y|^h\psi^*(x)\psi(y)\,dx\,dy \ , \tag{89}$$

where $\alpha''(u)/\alpha'(u)$ is continuous in u.
Comparing (88) and (89) shows that (87) holds with

$$c(u,\sigma) = s(C(u)+o(1))$$

and C continuous. This proves that (86) is indeed satisfied. □

Convolving both sides of (86) with g, we obtain

$$\overline{\partial_u A_F}(u,s) - \int g(u-v)\,(\log\alpha')'(v)\,\partial_{\log s} A_F(v,s)\,dv$$
$$= s\int (C(v)+o(1))g(u-v)\partial_u A_F(v,s)\,dv \ . \quad (90)$$

The hypotheses of Theorem 2.1 imply that $\partial_{\log s} A_F(u,s)$ does not vanish. By continuity, $\partial_{\log s} A_F(v,s)$ therefore keeps a constant sign for v in $[u-\Delta, u+\Delta]$. Moreover,

$$\left|\int g(u-v)((\log\alpha')'(v) - (\log\alpha')'(u))\,\partial_{\log s} A_F(v,s)\,dv\right|$$
$$\leq \max_{|v-u|\leq\Delta} |(\log\alpha')'(v) - (\log\alpha')'(u)|\,|\overline{\partial_{\log s} A_F}(u,s)|$$
$$= O(\Delta)\,|\overline{\partial_{\log s} A_F}(u,s)| \ . \quad (91)$$

because $(\log\alpha')''$ is bounded over $[u-\Delta, u+\Delta]$. If u is such that $\alpha''(u)\neq 0$, and if Δ is small enough, $\partial_u A_F(v,s)$ also keeps a constant sign over $[u-\Delta, u+\Delta]$. Since C is continuous,

$$\left|\int g(u-v)(C(v)+o(1)-C(u))\,\partial_u A_F(v,s)\,dv\right|$$
$$\leq \max_{|v-u|\leq\Delta} |C(v)+o(1)-C(u)|\ |\overline{\partial_u A_F}(u,s)|$$
$$= o(1)\,|\overline{\partial_u A_F}(u,s)| \text{ when } \Delta\to 0. \quad (92)$$

Combining (90), (91), and (92) proves that

$$(1+O(s))\overline{\partial_u A_F}(u,s) - ((\log\alpha')'(u) + O(\Delta))\overline{\partial_{\log s} A_F}(u,s) = 0 \ .$$

C.2 Proof of Theorem 3.1

The following lemma, whose proof is in Appendix C.2.1, shows that estimators $\widehat{\partial_u A_F}(u,s)$ and $\widehat{\partial_{\log s} A_F}(u,s)$ are consistent.

Lemma C.2. *Let $F(x) = R(\alpha(x))$, where R is a stationary Gaussian process such that there exists $h>0$ with*

$$c_R(0) - c_R(x) = |x|^h\,\eta(x) \quad \textit{and} \quad \eta(0) > 0 \ . \quad (93)$$

Let ψ be a $\mathbf{C}^2$ wavelet supported in $[-1,1]$ and with p vanishing moments, such that

$$2p - h > 1/2, \ \textit{and} \ \iint |x-y|^h\,\psi^*(x)\,\psi(y)\,dx\,dy \neq 0 \ .$$

If $\eta(x)$ is $\mathbf{C}^{2p}$ in a neighborhood of 0, and if $\alpha(x) \in \mathbf{C}^{2p} \cap \mathbf{C}^3$, then for each u, for s small enough,

$$\mathrm{Prob}\left\{|\widehat{\partial_{\log s} A_F}(u,s) - \overline{\partial_{\log s} A_F}(u,s)| \geq C\,|\overline{\partial_{\log s} A_F}(u,s)|\right\} \leq \varepsilon_1 \ , \tag{94}$$

$$\mathrm{Prob}\left\{|\widehat{\partial_u A_F}(u,s) - \overline{\partial_u A_F}(u,s)| \geq C\,|\overline{\partial_u A_F}(u,s)|\right\} \leq \varepsilon_2 \ , \tag{95}$$

where $C = \frac{\log(N\Delta)}{\Delta\sqrt{N\Delta}}$, $\varepsilon_1 = \frac{C_1(u)\Delta^2}{(\log(N\Delta))^2}$ and $\varepsilon_2 = 6\,(N\Delta)^{-1/(2C_2(u))}$.

The parameters $C_1(u)$ and $C_2(u)$, which are defined in the proof of the lemma, are both positive.

The weak consistency of

$$\frac{\widehat{\partial_u A_F}(u,N^{-1})}{\widehat{\partial_{\log s} A_F}(u,N^{-1})}$$

as an estimator of $(\log\alpha')'(u)$ then results from the following lemma, whose proof is straightforward:

Lemma C.3. *If X_1 and X_2 are two random variables, and $C < 1$ a constant such that*

$$\mathrm{Prob}\{|X_1 - \mathsf{E}\{X_1\}| \leq C\,|\mathsf{E}\{X_1\}|\} \geq 1 - \varepsilon_1 \,,$$
$$\mathrm{Prob}\{|X_2 - \mathsf{E}\{X_2\}| \leq C\,|\mathsf{E}\{X_2\}|\} \geq 1 - \varepsilon_2 \ ,$$

then

$$\mathrm{Prob}\left(\left|\frac{X_2}{X_1} - \frac{\mathsf{E}\{X_2\}}{\mathsf{E}\{X_1\}}\right| \leq \frac{2\,C}{1-C}\right) \geq 1 - \varepsilon_1 - \varepsilon_2 \ .$$

In view of Lemma C.2, one can apply Lemma C.3 to $X_1 = \widehat{\partial_{\log s} A_F}(u,N^{-1})$ and $X_2 = \widehat{\partial_u A_F}(u,N^{-1})$ with $C = \frac{\log(N\Delta)}{\Delta\sqrt{N\Delta}}$ yielding

$$\mathrm{Prob}\left\{\left|\frac{\widehat{\partial_u A_F}(u,N^{-1})}{\widehat{\partial_{\log s} A_F}(u,N^{-1})} - \frac{\overline{\partial_u A_F}(u,N^{-1})}{\overline{\partial_{\log s} A_F}(u,N^{-1})}\right| \leq \frac{2\log(N\Delta)}{\Delta\sqrt{N\Delta} - \log(N\Delta)}\right\} \geq 1 - \varepsilon_1 - \varepsilon_2 \ .$$

Because of the averaged transport equation (41),

$$(\log\alpha')'(u) = O(\Delta) + \frac{\overline{\partial_u A_F}(u,N^{-1})}{\overline{\partial_{\log s} A_F}(u,N^{-1})}(1 + O(N^{-1})).$$

Since $\Delta > N^{-1}$ and $(\log\alpha')'(u)$ is bounded, we derive

$$(\log\alpha')'(u) = O(\Delta) + \frac{\overline{\partial_u A_F}(u,N^{-1})}{\overline{\partial_{\log s} A_F}(u,N^{-1})}$$

therefore

$$\mathrm{Prob}\left(\left|\frac{\widehat{\partial_u A_F}(u,N^{-1})}{\widehat{\partial_{\log s} A_F}(u,N^{-1})} - (\log\alpha')'(u)\right| \leq \frac{2\log(N\Delta)}{\Delta\sqrt{N\Delta} - \log(N\Delta)} + O(\Delta)\right) \geq 1 - \varepsilon_1 - \varepsilon_2 \ .$$

We pick Δ such that $\Delta^{-1}(N\Delta)^{-1/2} = \Delta$, i.e. $\Delta = N^{-1/5}$. When $N \to \infty$, ε_1 and ε_2, whose expressions are given in Lemma C.2, both tend to 0. Moreover, for N large enough,

$$\frac{2\log(N\Delta)}{\Delta\sqrt{N\Delta} - \log(N\Delta)} + O(\Delta) \leq 2(\log N)\, N^{-1/5} \ .$$

Therefore

$$\lim_{N\to\infty} \operatorname{Prob}\left(\left| \frac{\widehat{\partial_u A_F}(u, N^{-1})}{\widehat{\partial_{\log s} A_F}(u, N^{-1})} - (\log\alpha')'(u) \right| \leq 2(\log N)\, N^{-1/5} \right) = 1 \ .$$

C.2.1 Proof of Lemma C.2 We start by proving (94). Let $n = N\Delta$ denote the number of discrete samples covered by the support of g. We seek an upper bound for the variance of $\widehat{\partial_{\log s} A_F}(u,s)$, which is defined by

$$V_{\log s}(u) = \mathsf{E}\{|\widehat{\partial_{\log s} A_F}(u,s) - \overline{\partial_{\log s} A_F}(u,s)|^2\} \ .$$

Let us choose $u = 0$ without loss of generality. One can see that

$$\left| \frac{\partial^2}{\partial u\, \partial \log s} A_F(u,s) \right| = O(s^h)$$

and a Riemann series approximation shows that

$$\int g(v)\, \partial_{\log s} A_F(v,s)\, dv - N^{-1} \sum_{k=-n}^{n} g(k/N)\, \partial_{\log s} A_F(k/N, s) = O(s^h/N) \ .$$

Replacing $\widehat{\partial_{\log s} A_F}(0,s)$ by its expression in (44), and noticing that the real part is smaller than the modulus, we obtain

$$V_{\log s}(0) \leq \frac{4}{N^2} \mathsf{E}\left\{ \left| \sum_{|k|\leq n} g_k X_k Y_k - g_k \mathsf{E}\{X_k Y_k\} \right|^2 \right\} + O(s^{2h}/N^2)$$

where g_k, X_k and Y_k respectively denote $g(k/n)$, $\langle F, \psi_{k/N,s}\rangle$ and $\langle F, \partial_{\log s}\psi_{k/N,s}\rangle^*$. Expanding $\left|\sum_{|k|\leq n}\right|^2$ under the form $\left(\sum_{|k|\leq n}\right) \cdot \left(\sum_{|l|\leq n}\right)^*$,

$$\begin{aligned} V_{\log s}(0) &\leq \frac{4}{N^2} \mathsf{E}\{ \sum_{|k|\leq n} [g_k X_k Y_k - g_k \mathsf{E}\{X_k Y_k\}] \sum_{|l|\leq n} [g_l X_l Y_l - g_l \mathsf{E}\{X_l Y_l\}]^* \} + O(s^{2h}/N^2) \\ &\leq \frac{4}{N^2} \sum_{\substack{|k|\leq n \\ |l|\leq n}} [g_k g_l \mathsf{E}\{X_k Y_k X_l^* Y_l^*\} - g_k g_l \mathsf{E}\{X_k Y_k\} \mathsf{E}\{X_l^* Y_l^*\}] + O(s^{2h}/N^2) \ . \end{aligned}$$

Since R is Gaussian, so is F, as well as the random variables X_k and Y_k. A classical result on Gaussian random variables shows that

$$\mathsf{E}\{X_k Y_k X_l^* Y_l^*\} = \mathsf{E}\{X_k Y_k\}\, \mathsf{E}\{X_l^* Y_l^*\} + \mathsf{E}\{X_k X_l^*\}\, \mathsf{E}\{Y_k Y_l^*\} + \mathsf{E}\{X_k Y_l^*\}\, \mathsf{E}\{Y_k X_l^*\} \ .$$

Therefore

$$V_{\log s}(0) \leq \frac{4}{N^2}\sum_{\substack{|k|\leq n\\|l|\leq n}} g_k g_l [\mathsf{E}\{X_k X_l^*\}\,\mathsf{E}\{Y_k Y_l^*\} + \mathsf{E}\{X_k Y_l^*\}\,\mathsf{E}\{Y_k X_l^*\}] + O(s^{2h}/N^2)$$

$$\leq \frac{4}{n^2}\sum_{\substack{|k|\leq n\\|l|\leq n}} [|\mathsf{E}\{X_k X_l^*\}|\,|\mathsf{E}\{Y_k Y_l^*\}| + |\mathsf{E}\{X_k Y_l^*\}|\,|\mathsf{E}\{Y_k X_l^*\}|] + O(s^{2h}/N^2)\ . \quad (96)$$

Each of the terms appearing in the sum above can be bounded thanks to the following decorrelation lemma:

Lemma C.4. *Let $F(x) = R(\alpha(x))$, let $X_k = \langle F, \psi_{k/N,s}\rangle$ and $Y_k = \langle F, \partial_{\log s}\psi_{k/N,s}\rangle^*$. Under the hypotheses of Lemma C.2, for s small enough, there exist continuous functions M_1 and M_2 such that, for $|k-l| \leq 2$,*

$$|\mathsf{E}\{X_k X_l^*\}| \leq\ M_1(sk)s^h\ , \quad (97a)$$

$$|\mathsf{E}\{X_k Y_l^*\}| \leq\ M_1(sk)s^h\ , \quad (97b)$$

$$|\mathsf{E}\{Y_k Y_l^*\}| \leq\ M_1(sk)s^h\ , \quad (97c)$$

and for $|k-l| > 2$,

$$|\mathsf{E}\{X_k X_l^*\}| \leq\ M_2(sk)\frac{s^{2p}}{(s(|k-l|-2))^{2p-h}}\ , \quad (98a)$$

$$|\mathsf{E}\{X_k Y_l^*\}| \leq\ M_2(sk)\frac{s^{2p}}{(s(|k-l|-2))^{2p-h}}\ , \quad (98b)$$

$$|\mathsf{E}\{Y_k Y_l^*\}| \leq\ M_2(sk)\frac{s^{2p}}{(s(|k-l|-2))^{2p-h}}\ . \quad (98c)$$

The proof of the above lemma is in Appendix C.2.2.

Replacing (97) and (98) in (96), we see that, since M_1 and M_2 are continuous and since $k/N = \Delta \to 0$ when $N \to \infty$,

$$V_{\log s}(0) \leq \frac{4}{n^2}\sum_{\substack{|k-l|\leq 2\\|k|,|l|\leq n}} 2(M_1(0)+o(1))^2 s^{2h}$$

$$+ \frac{4}{n^2}\sum_{\substack{|k-l|>2\\|k|,|l|\leq n}} \frac{2(M_2(0)+o(1))^2 s^{4p}}{(s(|k-l|-2))^{4p-2h}} + O(s^{2h}/N^2)\ . \quad (99)$$

Since $4p-2h > 1$,

$$\sum_{\substack{|k-l|>2\\|k|,|l|\leq n}} (|k-l|-2)^{2h-4p} = K_p n\ . \quad (100)$$

Replacing (100) in (99), we obtain

$$V_{\log s}(0) \leq 8C^2 \frac{s^{2h}}{n} (3\, M_1(0)^2 + K_p M_2(0)^2) + o(s^{2h}/n) \ .$$

In the proof of Theorem 2.1, (70) proves that there exists $a(u) > 0$ such that

$$|\partial_u A_F(u,s)| \geq a(u) s^h + o(s^h) \ .$$

For Δ small enough, $\partial_u A_F(v,s)$ does not change sign for $|v-u| \leq \Delta$ thus, after convolution with g,

$$|\overline{\partial_u A_F}(u,s)| \geq a(u) s^h + o(s^h) \ .$$

Because of transport equation (41), the same applies to $\overline{\partial_{\log s} A_F}(u,s)$, therefore there exists a constant $C_1(u)$ such that

$$V_{\log s}(u) \leq C_1(u) \left[\frac{|\overline{\partial_{\log s} A_F}(u,s)|}{\sqrt{n}} \right]^2 \ .$$

Applying Chebyshev's Lemma [3] then proves that, for all $\varepsilon > 0$,

$$\text{Prob} \left\{ |\widehat{\partial_{\log s} A_F}(u,s) - \overline{\partial_{\log s} A_F}(u,s)| \geq \frac{\sqrt{C_1(u)}\, |\overline{\partial_{\log s} A_F}(u,s)|}{\varepsilon \sqrt{n}} \right\} \leq \varepsilon^2 \ ,$$

and (94) follows by choosing $\varepsilon = \frac{\sqrt{C_1(u)}\Delta}{\log n}$ and $\varepsilon_1 = \frac{C_1(u)\Delta^2}{(\log n)^2}$.

Let us now prove (95). We denote $D_u = |\widehat{\partial_u A_F}(0,s) - \overline{\partial_u A_F}(0,s)|$. One can see that

$$|\partial_u A_F(u,s)| = O(s^h)$$

and a Riemann series approximation once again shows that

$$\Delta^{-2} \int_0^\Delta A_F(v,s)\, dv - \Delta^{-1} n^{-1} \sum_{k=0}^{n} A_F(k/N, s) = O(s^h/n) \ .$$

Therefore, using once more the notation $X_k = \langle F, \psi_{k/N,s} \rangle$,

$$D_u = \frac{1}{n\Delta} \Big| \sum_{k=0}^{n} (|X_k|^2 - \mathsf{E}\{|X_k|^2\}) - \sum_{k=-n}^{0} (|X_k|^2 - \mathsf{E}\{|X_k|^2\}) \Big| + O(s^h/n) \ .$$

Denoting

$$\widetilde{X}^- = \sum_{k=-n+1}^{0} |X_k|^2 \quad \text{and} \quad \widetilde{X}^+ = \sum_{k=1}^{n} |X_k|^2 \ ,$$

we have

$$D_u \leq \frac{1}{n\Delta} \left(|\widetilde{X}^+ - \mathsf{E}\{\widetilde{X}^+\}| + |\widetilde{X}^- - \mathsf{E}\{\widetilde{X}^-\}| \right) + O(s^h/n) \qquad (101)$$

We are now going to prove that there exists a strictly positive constant C_2 such that

$$\forall y\ ,\quad \text{Prob}\{D_u > y\,C_2 \frac{s^h}{\Delta\sqrt{n}}\} \leq 6\,\mathrm{e}^{-y/2} \tag{102}$$

and since $|\overline{\partial_u A_F}(u,s)| \geq a(u)s^h + o(s^h)$ with $a(u) > 0$, choosing $y = \log n / C_2$ will then imply (95).

Let us consider the random vector $X = (X_1, X_2, \ldots, X_n)$, let K_X denote the covariance operator of X, and $(e_j)_{j=1,\ldots,n}$ its Karhunen-Loève basis. If $(\alpha_j)_{j=1,\ldots,n}$ are the eigenvalues of K_X corresponding to the eigenvectors $(e_j)_{j=1,\ldots,n}$, then

$$X = \sum_{j=1}^{n} \sqrt{\alpha_j}\, Z_j\, e_j$$

where Z_j are independent random variables with variance 1. As a consequence,

$$\widetilde{X}^+ = \|X\|^2 = \sum_{j=1}^{n} \alpha_j Z_j^2\ .$$

The following lemma, which is proved in [7], relies on a theorem by Bakirov [2].

Lemma C.5. *If* $\widehat{X} = \sum_j \beta_j Z_j^2$ *where* Z_j *are independent Gaussian random variables with variance one, and* $\sum_j \beta_j^2 = 1$, *then*

$$\forall y\ ,\quad \text{Prob}\{|\widehat{X} - \mathsf{E}\{\widehat{X}\}| > y\} \leq 6\,\mathrm{e}^{-y/2}\ .$$

The random variable $\widehat{X}^+ = \left(\sum_j \alpha_j^2\right)^{-1/2} \widetilde{X}^+$ satisfies the requirements of Lemma C.5, therefore

$$\forall y,\ \text{Prob}\{|\widetilde{X}^+ - \mathsf{E}\{\widetilde{X}^+\}| > y\,(\sum_j \alpha_j^2)^{1/2}\} \leq 6\,\mathrm{e}^{-y/2}$$

but $\sum_j \alpha_j^2$ is equal to the Hilbert-Schmidt norm of K_X:

$$\sum_j \alpha_j^2 = \sum_{j,k} \mathsf{E}\{X_j\, X_k^*\}\ ,$$

which is bounded by $Bs^{2h}n$ because of (97a) and (98a). Hence

$$\forall y\ ,\quad \text{Prob}\{|\widetilde{X}^+ - \mathsf{E}\{\widetilde{X}^+\}| > y\,\sqrt{B}s^h\sqrt{n}\} \leq 6\,\mathrm{e}^{-y/2}\ .$$

The same applies to $\widetilde{X}^-$, and by combining the two and using (101) we obtain (102).

C.2.2 Proof of Lemma C.4

The three terms $\mathsf{E}\{X_k X_l^*\}$, $\mathsf{E}\{X_k Y_l^*\}$ and $\mathsf{E}\{Y_k Y_l^*\}$ can be written as

$$I = \iint c_R(\alpha(u+sx) - \alpha(v+sy))\psi(x)\,\tilde{\psi}(y)\,dx\,dy\ ,$$

where $(u,v) = (sk, sl)$, and ψ and $\tilde{\psi}$ are two wavelets with p vanishing moments. Clearly,

$$I = \iint \left[c_R(\alpha(u+sx) - \alpha(v+sy)) - c_R(0)\right]\psi(x)\tilde{\psi}(y)\,dx\,dy\ . \tag{103}$$

For $|u-v| \leq \Delta$, $|x| \leq 1$ and $|y| \leq 1$, we have

$$|\alpha(u+sx) - \alpha(v+sy)| \leq (\Delta + 2s) \sup_{|x-u|\leq \Delta+2s} [\alpha'(x)] \leq (\Delta+2s)\,C_u$$

because α is continuously differentiable. For Δ small enough, $|\alpha(u+sx) - \alpha(v+sy)|$ is therefore in a neighborhood of 0. Since η is assumed continuous in a neighborhood of 0,

$$|\eta(\alpha(u+sx) - \alpha(v+sy))| \leq B$$

for $|u-v| \leq \Delta$, $|x| \leq 1$ and $|y| \leq 1$.
Hence

$$\begin{aligned}|I| &\leq \iint |\alpha(u+sx) - \alpha(v+sy)|^h B\,|\psi(x)|\,|\tilde{\psi}(y)|\,dx\,dy \\ &\leq Cs^h + o(s^h)\ .\end{aligned}$$

This proves (97a), (97b) and (97c).

Let us now prove (98). Since ψ and $\tilde{\psi}$ in (103) are compactly supported and have p vanishing moments, there exist two compactly supported functions θ and $\tilde{\theta}$ such that $\psi(x) = \theta^{(p)}(x)$ and $\tilde{\psi}(y) = \tilde{\theta}^{(p)}(y)$. Integrating (103) by parts p times with respect to x and to y gives

$$I = \iint \frac{\partial^p}{\partial x^p}\frac{\partial^p}{\partial y^p}\left\{|\alpha(u+sx) - \alpha(v+sy)|^h \eta(\alpha(u+sx) - \alpha(v+sy))\right\}\theta(x)\,\tilde{\theta}(y)\,dx\,dy\ .$$

But for $|u-v| > 2s$, one can show that

$$\left|\frac{\partial^p}{\partial x^p}\frac{\partial^p}{\partial y^p}\left\{|\alpha(u+sx) - \alpha(v+sy)|^h \eta(\alpha(u+sx) - \alpha(v+sy))\right\}\right| \leq \frac{M(u)\ s^{2p}}{(|u-v| - 2s)^{2p-h}}\ ,$$

where $M(u)$ depends on h, on derivatives of α up to order $2p$ in a neighborhood of u, and on derivatives of η up to order $2p$ in a neighborhood of 0. Therefore there exists a continuous $M_2(u)$ such that

$$|I| \leq M_2(u)\frac{s^{2p}}{(s(|k-l|-2))^{2p-h}}\ ,$$

which proves (98a), (98b), and (98c).

C.3 Proofs of Section 3.2

C.3.1 Proof of (49)

Lemma C.6. *Under the hypotheses of Theorem 2.2,*

$$\partial_u A_F^\sigma(u,\xi_0/\sigma) - \alpha''(u)\,\partial_\xi A_F^\sigma(u,\xi_0/\sigma) = \sigma^2(C(u)+o(1))\partial_u A_F^\sigma(u,\xi_0/\sigma)\ , \tag{104}$$

where C is continuous.

Proof. The proof mimicks the proof of Lemma C.1. In the proof of Theorem 2.2, we showed in (79) that

$$Re\left\langle K_F\psi_{v,\xi,\sigma}, \overline{D^{-1}}\,(\partial_v+\partial_x)\,\overline{D}\psi_{v,\xi,\sigma}\right\rangle = \sigma^{3+h}(A(v)+o(1))\ , \tag{105}$$

and in (81) that

$$Re\,\langle K_F\psi_{v,\xi,\sigma}, \partial_x\psi_{v,\xi,\sigma}\rangle = \sigma^{1+h}(B(v)+o(1))\ , \tag{106}$$

with $B(v)$ continuous. Comparing (105) and (106) shows that

$$Re\left\langle K_F\psi_{v,\xi,\sigma}, \overline{D^{-1}}\,(\partial_v+\partial_x)\,\overline{D}\psi_{v,\xi,\sigma}\right\rangle = \sigma^2(C(v)+o(1))\,Re\,\langle K_F\psi_{v,\xi,\sigma}, \partial_x\psi_{v,\xi,\sigma}\rangle\ ,$$

with $C(v)$ continuous. This implies, by repeating the argument of Lemma C.1, that (104) is satisfied. □

Using Lemma C.6, the arguments of Proposition 3.1 can be repeated to prove (49).

C.3.2 Proof of Theorem 3.2 As in the proof of Theorem 3.1, one can combine the following lemma with Lemma C.3 to prove the weak consistency result (55).

Lemma C.7. *Let $F(x) = R(x)\,e^{i\alpha(x)}$, where R is a stationary Gaussian process such that there exists $h>0$ with*

$$c_R(0) - c_R(x) = |x|^h\,\eta(x) \quad and \quad \eta(0) > 0\ .$$

Let ψ be a $\mathbf{C}^2$ even, positive function supported in $[-1,1]$ such that $\psi^1(x) = \mathrm{e}^{i\xi_0 x}\psi(x)$ has $p \geq \lceil h\rceil + 3$ vanishing moments and such that

$$\iint |x-y|^h(x-y)\,\sin[\xi_0(x-y)]\,\psi(x)\,\psi(y)\,dx\,dy \neq 0\ .$$

If η is $\mathbf{C}^{2p}$ in a neighborhood of 0, and if $\alpha \in \mathbf{C}^{2p}$, then

$$Prob\left\{|\widehat{\partial_\xi A_F^\sigma}(u,N\xi_0) - \overline{\partial_\xi A_F^\sigma}(u,N\xi_0)| \geq C|\overline{\partial_\xi A_F^\sigma}(u,N\xi_0)|\right\} \leq \varepsilon_1\ ,$$

$$Prob\left\{|\widehat{\partial_u A_F^\sigma}(u,N\xi_0) - \overline{\partial_u A_F^\sigma}(u,N\xi_0)| \geq C|\overline{\partial_u A_F^\sigma}(u,N\xi_0)|\right\} \leq \varepsilon_2\ .$$

The proof of Lemma C.7 is almost identical to the proof of Lemma C.2; the only difference is that ψ^2 has $p-1$ vanishing moments instead of p, so that Lemma C.4 must be replaced with the following lemma, which is proved by using the same method.

Lemma C.8. *Let $F(x) = R(x)\,\mathrm{e}^{i\alpha(x)}$, let $X_k = \langle F, \psi^1_{k/N,\sigma}\rangle$ and $Y_k = \langle F, \psi^2_{k/N,\sigma}\rangle^*$. Under the hypotheses of Lemma C.7, for σ small enough, there exist two continuous functions M_1 and M_2 such that*
for $|k-l| \leq 2$,

$$\begin{aligned}
|\mathsf{E}\{X_k X_l^*\}| &\leq\ M_1(\sigma k)\sigma^h\ , \\
|\mathsf{E}\{X_k Y_l^*\}| &\leq\ M_1(\sigma k)\sigma^h\ , \\
|\mathsf{E}\{Y_k Y_l^*\}| &\leq\ M_1(\sigma k)\sigma^h\ ,
\end{aligned}$$

and for $|k-l| > 2$,

$$\begin{aligned}
|\mathsf{E}\{X_k X_l^*\}| &\leq\ M_2(\sigma k)\frac{\sigma^{2p}}{(\sigma(|k-l|-2))^{2p-h}}\ , \\
|\mathsf{E}\{X_k Y_l^*\}| &\leq\ M_2(\sigma k)\frac{\sigma^{2p-1}}{(\sigma(|k-l|-2))^{2p-1-h}}\ , \\
|\mathsf{E}\{Y_k Y_l^*\}| &\leq\ M_2(\sigma k)\frac{\sigma^{2p-2}}{(\sigma(|k-l|-2))^{2p-2-h}}\ .
\end{aligned}$$

Since $p \geq \lceil h \rceil + 3$, we have $2(2p-2-h) > 1$, therefore the variance term

$$\mathsf{E}\{|\widehat{\partial_\xi A_F^\sigma}(u,\xi) - \overline{\partial_\xi A_F^\sigma}(u,\xi)|^2\}$$

can be controlled as in the proof of Lemma C.2.

References

[1] P. Abry, P. Gonçalves, and P. Flandrin. Wavelets, spectrum analysis and $1/f$ processes. In *Wavelets and Statistics*, volume 103 of *Lecture Notes in Statistics*, pages 15–29. Springer-Verlag, Berlin, 1995. A. Antoniadis and G. Oppenheim eds.

[2] N. K. Bakirov. Extrema of the distributions of quadratic forms of Gaussian variables. *Theory Probab. Appl.*, 34, 1989.

[3] P. Billingsley. *Probability and Measure*. John Wiley & Sons, 1995.

[4] M. Clerc. *Estimation de processus localement dilatés et application au gradient de texture*. PhD thesis, Ecole Polytechnique, Paris, September 1999. (in French).

[5] M. Clerc and S. Mallat. Shape from texture through deformations. In *Proc. 7th Int. Conf. on Computer Vision*, Kerkyra, Greece, 1999.

[6] P. Flandrin. On the spectrum of fractional brownian motions. *IEEE Trans. Inform. Theory*, 35(1):197–199, 1989.

[7] H.-Y. Gao. *Wavelet Estimation of Spectral Densities in Time Series Analysis*. PhD thesis, U.C. Berkeley, 1993.

[8] J. Gårding. Shape from texture for smooth curved surfaces in perspective projection. *J. Math. Imaging Vision*, 2:327–350, 1992.

[9] J. Malik and R. Rosenholtz. Computing local surface orientation and shape from texture for curved surfaces. *Int. J. Comp. Vision*, 23(2):149–168, 1997.

[10] S. Mallat. *A wavelet tour of signal processing*. Academic Press, 1997.

[11] G.P. Patil and C.R. Rao, editors. *Handbook of Statistics*, volume 12, chapter 20, pages 661–689. Elsevier Science, 1994. Methods for Estimating Heterogeneous Spatial Covariance Functions with Environmental Applications.

[12] O. Perrin and R. Senoussi. Reducing non-stationary stochastic processes to stationarity by a time deformation. *Statist. Probab. Lett.*, 43(4):393–397, 1999.

[13] P.D. Sampson and P. Guttorp. Nonparametric estimation of nonstationary spatial covariance structure. *J. Am. Statist. Ass.*, 87:108–119, 1992.