

Université de Montréal

STATISTIQUE DE MAXIMA ET MODÈLES  
GRAPHIQUES MULTI-ÉCHELLES :  
APPLICATION À LA TURBULENCE

par

Philippe St-Jean

Département de physique  
Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures  
en vue de l'obtention du grade de  
Philosophiæ Doctor (Ph.D.)  
en Physique

avril 2003



QC  
3  
U57  
2003  
V.019

**Direction des bibliothèques**

**AVIS**

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

**NOTICE**

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal

Faculté des études supérieures

Cette thèse intitulée

**STATISTIQUE DE MAXIMA ET MODÈLES  
GRAPHIQUES MULTI-ÉCHELLES :  
APPLICATION À LA TURBULENCE**

présentée par

**Philippe St-Jean**

a été évaluée par un jury composé des personnes suivantes :

*Alain Vincent*

---

(président-rapporteur)

*Bernard Goulard*

---

(directeur de recherche)

*Jean-Marc Lina*

---

(co-directeur)

[REDACTED]

(examineur externe)

[REDACTED]

(représentant du doyen)

Thèse acceptée le:

---

# SOMMAIRE

---

Cette thèse est consacrée au développement de méthodes numériques et statistiques d'analyse de signaux, plus particulièrement dans le domaine de la turbulence pleinement développée.

Dans le premier chapitre nous présentons une méthode d'analyse de signaux basée sur la statistique d'excédence de champs stochastiques continus, à travers le calcul de l'espérance du nombre de maxima locaux d'un champ sur un domaine fini ; cette méthode a été introduite, développée et déployée par les travaux de Robert Adler et Keith Worsley. Nous introduisons une généralisation des travaux de Siegmund et Worsley sur les statistiques d'excédence de champ multi-échelle, permettant l'application de la méthode à des bruits corrélés ; nous proposons aussi une optimisation numérique permettant une réduction significative du temps de calcul nécessaire à l'implémentation d'une telle méthode.

Le deuxième chapitre couvre un problème récent en phénoménologie de la turbulence pleinement développée (i.e. à de grandes valeurs du nombre de Reynolds), celui du phénomène d'intermittence. Nous nous intéressons spécialement à l'approche multi-fractale et aux modèles "en cascades" qui en découlent. Nous soulignons une difficulté des modèles en cascades originaux, à savoir leur incapacité à respecter une symétrie fondamentale du système étudié, l'invariance translationnelle (ou stationnarité). Nous développons un modèle en cascade stationnaire qui conserve les propriétés d'auto-similarité et de multi-fractalité des processus en cascades, tout en respectant la symétrie translationnelle.

Le troisième chapitre, qui contient le résultat le plus important de la thèse, fait ressortir la relation entre des éléments présentés dans les deux premiers chapitres. Nous y présentons une interprétation de la phénoménologie proposée par She et

Leveque, à savoir leur “loi d’échelle universelle” (universal scaling law). Cette loi d’échelle, qui permet d’expliquer avec une excellente précision les données expérimentales, doit toutefois être postulée dans les travaux originaux de She et Leveque. Notre interprétation basée sur la statistique de maxima permet de déduire, plutôt que de postuler, cette loi d’échelle. Elle permet également de prédire d’autres solutions admissibles.

**MOTS CLEF : Statistique d’ordre, géométrie de champs stochastiques, modèles graphiques, phénoménologie de la turbulence, turbulence pleinement développée, modèles en casades, analyse multi-fractale.**

## SUMMARY

---

This thesis presents numerical and statistical methods devoted to signal analysis, with specific interest in the field of fully developed turbulence.

The first chapter presents a method of signal analysis based on exceedence statistics of continuous stochastic fields using expected value of the number of local maxima of the field over a finite region. This method was introduced mostly by Robert Adler, and later developed by Keith Worlsey. We introduce a generalization of Siegmund and Worlsey's work on exceedence statistics of multi-scale fields that allows to apply the method to signal detection in arbitrarily correlated noise. We also propose a numerical optimization allowing a significant reduction of computing time of the numerical implementation.

The second chapter covers a relatively recent problem regarding fully developed turbulence (i.e. at large values of the Reynolds number), namely the intermittency phenomenon. We focus on the multi-fractal approach and on the related cascade models. We bring up a difficulty regarding cascade models, namely the impossibility to generate models that respect translational invariance (stationarity). We develop a new stationary cascade model which respects self-similarity properties of the original models, while adding the stationarity property.

The third chapter contains the most important result of the thesis while joining elements developed in the first two chapters. We present a statistical interpretation of the phenomenological model originally proposed by She and Leveque and the *universal scaling law* presented therein. This scaling law is found to be in excellent agreement with experimental data; however it had to be postulated in

the original work of She et Leveque. Our interpretation is based on maxima statistics and allows to infer rather than postulate this law. It also allows to predict other admissible solutions.

**KEY WORDS :** Order statistics, geometry of random fields, graphical models, turbulence phenomenology, fully-developed turbulence, cascade models, multi-fractal analysis.

# Table des matières

---

<b>Sommaire</b> .....	iii
<b>Summary</b> .....	v
<b>Table des figures</b> .....	xi
<b>Remerciements</b> .....	xvii
<b>Chapitre 1. Introduction</b> .....	1
<b>Chapitre 2. Maxima de la Transformée en Ondelettes Continue de données corrélées</b> .....	8
2.1. Revue de littérature : Statistique d'excédence de champs stochastiques, approche topologique.....	9
2.1.1. Topologie, caractéristique d'Euler.....	11
2.1.2. Statistique d'excédence et caractéristique d'Euler : Applications	16
2.1.3. Espérance de la caractéristique d'Euler : développement formel.	18
Caractéristique d'Hadwiger, contribution volumique $\chi_V$ .....	21
Caractéristique d'Hadwiger, contribution frontière $\chi_E$ .....	23
2.1.4. Calcul de l'espérance de la caractéristique d'Hadwiger.....	24
2.1.5. Caractéristique d'Hadwiger d'un champ multi-résolution.....	27
2.2. Présentation de l'article.....	29
2.3. Introduction.....	31
2.4. Exceedence statistics of multi-dimensional fields.....	33
2.4.1. Smooth homogeneous fields.....	33

2.4.2.	Homogeneous fields with arbitrary spatial correlation.....	37
2.4.3.	Fractional brownian motion.....	44
2.5.	Dyadic wavelet representation.....	46
2.5.1.	Numerical evaluation of the Euler characteristic for a field defined on a lattice.....	46
2.5.2.	How does the scale-space field look like?.....	47
2.6.	Results.....	48
2.7.	Discussion.....	50
2.8.	Conclusion.....	51
<b>Chapitre 3. Modèles graphiques multi-échelle stationnaires.....</b>		<b>57</b>
3.1.	Revue de littérature : Phénomène d'intermittence en turbulence, approche multi-fractale et modèle en cascades.....	58
3.1.1.	Équation de Navier-Stokes et symétries.....	58
3.1.2.	Le modèle de Kolmogorov (K41).....	61
3.1.3.	Intermittence.....	64
3.1.3.1.	Dissipation d'énergie multi-fractale.....	68
3.1.4.	Modèles en cascades.....	70
3.2.	Formalisme multi-fractal.....	72
3.2.1.	Formalisme multi-fractal pour les fonctions et signaux.....	74
3.2.1.1.	W-cascades.....	80
3.2.1.2.	Approche thermodynamique.....	83
3.2.1.3.	Singularités oscillantes.....	85
3.2.2.	Cascades et stationnarité.....	89
3.3.	Présentation de l'article.....	91
3.4.	Introduction.....	92

3.5. Nearest-neighbor models .....	95
3.6. Stationary model ( $\alpha$ -model).....	97
3.7. Stationarity constraints .....	99
3.7.1. Symmetry of $\alpha$ .....	102
3.7.2. Interpretation of the “frequency shift” condition .....	102
3.8. Stationarity of higher order moments.....	104
3.9. Form of the covariance function .....	106
3.10. Estimators for $\hat{\eta}$ and $\hat{C}_W$ .....	107
3.11. Simulations and analysis.....	109
3.12. Discussion and conclusions.....	111
3.13. Appendix.....	113
<b>Chapitre 4. Une interprétation du modèle de She-Leveque basée sur la statistique d'ordre .....</b>	<b>116</b>
4.1. Revue de littérature : Le modèle de She-Leveque .....	117
4.1.1. La loi d'échelle universelle de She-Leveque .....	120
4.1.2. Distributions asymptotiques de maxima .....	121
4.2. Présentation de l'article.....	125
4.3. Introduction .....	127
4.4. The She-Leveque model.....	129
4.5. Statistical behavior of the hierarchical dynamics.....	131
4.5.1. Obtaining the She-Leveque universal scaling law from the Fisher- Tippett distribution .....	136
4.6. General study of the three asymptotic forms .....	137

4.7. Discussion and conclusion.....	140
<b>Conclusion.....</b>	<b>143</b>
<b>Bibliographie .....</b>	<b>148</b>

## Table des figures

---

2.1	Exemple de réalisation d'un champ en 2 dimensions. L'ensemble d'excédence par rapport au plan (en gris foncé) comporte une zone annulaire (un lac).....	11
2.2	Caractéristique d'Euler de quelques polyèdres. a) Cube : $8 - 12 + 6 - 1 = 1$ b) 2 cubes joints par une face : $12 - 20 + 11 - 2 = 1$ c) "Tore" : $32 - 64 + 40 - 8 = 0$ .....	12
2.3	Topologie "en boulettes". Cette figure est obtenue en générant un bruit blanc en 3D filtré par un noyau gaussien, puis en extrayant une isosurface.....	15
2.4	Topologie "éponge".....	15
2.5	Topologie "bulles".....	16
2.6	Le même champ présenté à la figure 2.1, vu du dessus. En passant de l'hyperplan $u_1$ à l'hyperplan $u_2$ , l'intersection avec l'ensemble d'excursion garde la même caractéristique d'Hadwiger (un seul intervalle d'excédence, à gauche). En passant de $u_2$ à $u_3$ on observe un changement d'un à deux intervalles d'excédence. Le point marqué d'un (+) contribue à la somme (6).....	20
2.7	Le même champ présenté à la figure 2.1, vu d'un autre angle. Le point marqué d'un (+) compte dans la contribution frontière (11).....	24
2.8	A realization of a fBm noise for $H = 0.3$ .....	53
2.9	The same realization of a fBm noise for $H = 0.3$ , with a signal of the form $xe^{-x^2}$ added at $x = 320$ .....	53

2.10	Three realizations of fBm processes with $H = 0.2$ (top), $H = 0.5$ (classical Brownian motion, middle) and $H = 0.8$ (bottom).....	54
2.11	Euler characteristic of a fBm with $h = 0.7$ . The solid curve was computed from equation (37), while the circle points were obtained from an average of 50 fBm continuous wavelet transform simulations (with scale-normalized variance) from which the Euler characteristic was estimated. The cross have been computed on the discrete lattice described in Section 2.1. The simulations are in very good agreement with the theoretical expected value. For large values of $b$ , the characteristic goes down towards zero as expected, since no regions of $X(t, s)$ for which the field should take very high values. Similarly, for large negative values of $b$ , the whole field should be above that value over all its domain, which has an Euler characteristic of one.....	55
2.12	Scale-adapted tiling with examples of individual sites representing one point, two lines and a face.....	56
2.13	Evaluation of the Euler characteristic for the noise only and for noise+signal. The theoretical value is plotted as well, with the 95% confidence threshold at 2.74, i.e. $\chi(2.74) = 0.05$ . The noise+signal support goes beyond that value so that the signal is detected by the algorithm.....	56
3.1	Flot laminaire à bas Reynolds, $R = 1.54$ . Le flot moyen va de gauche à droite. Photo S. Taneda.....	59
3.2	Flot à des valeurs du nombre de Reynolds progressivement plus élevées ( $R = 9.6, 13.1, 26$ de haut en bas). Photo S. Taneda.....	60
3.3	Flot turbulent.....	61
3.4	Fonctions multi-fractales construites récursivement de manière a) déterministe ou b) aléatoire. (D'après [15].).....	75

3.5	Calcul de $\tau(q)$ et de $D(h)$ associés aux fonctions multi-fractales de la figure 3.4. a) $\log_2(a^{-\tau(1)}Z(q, a))/(q - 1)$ en fonction de $\log_2(a)$ . b) $\tau(q)$ en fonction de $q$ . c) $h(q, a)$ en fonction de $\log_2(a)$ . d) $D(h)$ en fonction de $h$ . Dans b) et d) : ( $\bullet$ ) fonction déterministe; ( $\Delta$ ) fonction aléatoire; (-) courbe théorique. (D'après [15],fig.35.).....	85
3.6	a) Exemple d'une singularité oscillante : $f(x) =  x ^\gamma \sin(2\pi/ x ^\beta)$ pour $\gamma = 4/3$ et $\beta = 1$ . b) Maxima du module de la T.O. de $f(x)$ . Les points ( $\bullet$ ) marquent la position des maxima globaux de la T.O. sur chaque ligne de maxima. c) $\log_2  T_\psi(b_n, a_n) $ vs $\log_2 a_n$ . La pente de ce graphe redonne $\alpha(x_0 = 0)$ . d) $\log_2 b_n$ vs $\log_2 a_n$ . Cette pente donne un estimé de $\phi(x_0 = 0)$ . (D'après [49].) .....	87
3.7	The single parent model. The solid line indicates ancestors of coefficient $a$ .....	95
3.8	The one-two parent model. Solid lines indicates all ancestors of coefficient $a$ .....	97
3.9	The $\alpha$ model.....	99
3.10	$\hat{C}[l]$ for $\hat{\eta}$ defined in (177).....	108
3.11	Correlations between nodes as a function of absolute position of the node ( $k$ ) and inter-nodal distance ( $\Delta k$ ) for the single-parent non-stationary model ( $j=8, \sigma_W = 1$ ). Blocky non-stationary artifacts are obvious. ....	110
3.12	Correlations between nodes as a function of absolute position of the node ( $k$ ) and inter-nodal distance ( $\Delta k$ ) for the two-parent non-stationary model ( $j=8, \sigma_W = 1$ ). Artifacts are subtler than the ones on figure (3.11), yet still visible at any inter-nodal distance. ....	111
3.13	Correlations for the stationary $\alpha$ -model as a function of absolute position of the node ( $k$ ) and inter-nodal distance ( $\Delta k$ ). ( $j=8, \sigma_W = 1$ )	112

- 4.1 Histograms of the maximum value of  $N$  gaussian variables for  $N = 2, 5, 50, 200, 1000$ . For  $N = 2$  the pdf is almost gaussian, as expected, while it tends to the asymmetrical  $\Lambda_3$  form for higher values of  $N$ . Each curve involves 100,000 realizations. . . . . 136

*À mes parents et à Josée*

*On dit : "Oui, ce calcul est bon" si on s'en est convaincu. Mais ce n'est pas quelque chose qu'on a inféré à partir de l'état de certitude dans lequel on est. On ne conclut pas à l'état des faits à partir de la certitude qu'on en a.*

*La certitude est comme un ton de voix selon lequel on constate un état de faits, mais on ne conclut pas de ce ton de voix que cet état est fondé.*

*Ludwig Wittgenstein, De la certitude*

# REMERCIEMENTS

---

Je veux remercier en premier lieu mes directeurs de recherche Bernard Goulard et Jean-Marc Lina, pour leur soutien inconditionnel tout au long des travaux qui ont mené à la rédaction de cette thèse, pour leur enthousiasme, leur curiosité, et surtout leur inexorable patience. Ils ont tous deux à cœur de voir leurs étudiants développer une expertise utile pour la suite des choses, je leur suis reconnaissant de m'y avoir sensibilisé. Il est impossible d'éviter un mauvais jeu de mots en remarquant que j'aurai peut-être été leur étudiant le plus turbulent. Le juste équilibre entre la latitude dont ils ont fait preuve et leur souci de voir mes travaux converger m'aura permis de mener à terme cette thèse.

Je remercie également mes collaborateurs au sein du groupe PhysNum, Gaël Sitzia, François Levac, Yan Basile-Bellavance, Marc Bergevin, Jean Daunizeau, Miguel Tremblay, Paul Turcotte, J.F. Muzy, Raphaël Béan, Cécile Amblard, Fahima Nekka et Frédéric Lesage. Je remercie plus spécialement Diego Clonda, Ervig Lapalme, Sébastien Tremblay et Réza Kasrai avec qui j'ai eu la chance de vivre une expérience de démarrage d'entreprise en parallèle avec mes travaux de thèse. J'aurai appris avec eux à relever le défi de poursuivre des travaux de recherche dans un cadre où il est essentiel que les résultats soient à la fois utiles et robustes.

Je tiens à remercier le professeur Habib Benali de m'avoir accueilli si chaleureusement dans son groupe de recherche lors de mes deux visites à Paris, et pour les discussions qui en ont résulté.

Je remercie le professeur Noel Cressie de m'avoir accordé du temps lors de sa visite à Montréal et de m'avoir permis de profiter de son expertise.

Je remercie tous les membres de l'équipe administrative du CRM, plus particulièrement Muriel Pasqualetti qui semble connaître la solution à tout problème.

Enfin, je remercie le Conseil de Recherches en Sciences Naturelles et en Génie du Canada (CRSNG), le Fonds pour la Formation des Chercheurs et l'Aide à la Recherche (Fonds FCAR) et le Réseau de Calcul et de Modélisation Mathématique (RCM<sup>2</sup>) pour leur support financier.

# Chapitre 1

---

## INTRODUCTION

L'évolution rapide et le déploiement efficace des technologies informatiques au courant des 30 dernières années a conduit à l'essor des méthodes numériques et statistiques en physique et en analyse de signaux. La contribution notoire de ces méthodes à la recherche en physique est manifeste dans l'étude de la turbulence, plus spécifiquement en rapport avec la turbulence pleinement développée. La modélisation des flots turbulents, vue par certains comme le dernier grand problème de la physique classique, est si complexe et subtile que plusieurs difficultés, même au niveau numérique, bloquent encore aujourd'hui l'évolution de cette modélisation.

En 1941, Kolmogorov [1] présente une théorie de la turbulence qui, encore à ce jour, constitue la contribution la plus importante à ce domaine, de l'avis unanime de la communauté scientifique. C'est cette théorie qui, par la suite, constituera la base de tous les développements de l'approche phénoménologique de la turbulence.

Peu de temps après, anticipant avec flair la suite des choses, von Neumann annonçait la nécessité d'aborder le problème de la turbulence numériquement, en proposant que "high-speed computing program... should be undertaken as soon as feasible".

Il expliqua cette proposition plus en détail dans son rapport à l'**Office of Naval Research** en 1949, qui constitue l'une des premières revues exhaustives des travaux jusqu'alors effectués en turbulence [2] :

*Under these conditions there might be some hope to “break the deadlock” by extensive, but well-planned, computational efforts. It must be admitted that the problems in question are too vast to be solved by a direct computational attack, that is, by an outright calculation of a representative family of special cases. There are, however, strong indications that one could name certain strategic points in this complex, where relevant information must be obtained by direct calculations. If this is properly done, and the operation is then repeated on the basis of broader information then becoming available, etc., there is a reasonable chance of effecting real penetrations in this complex of problems and gradually developing a useful, intuitive relationship to it. This should, in the end, make an attack with analytical methods, that is truly more mathematical, possible.*

Le vœu de Von Neumann ne se réalisa que vingt ans plus tard, par les travaux précurseurs d’Orszag et de Patterson en 1972 [3].

Bien que cette thèse ne porte pas directement sur la simulation numérique de flots turbulents, elle participe du même effort de “débroussaillage phénoménologique” qui nous conduit vers une explication de plus en plus saisissable de la dynamique turbulente.

Les outils numériques et statistiques qui sont présentés dans cette thèse n’ont pas été développés avec en tête l’objectif unique de régler tel problème spécifique à la turbulence. Le but premier était plutôt d’augmenter le répertoire habituel de méthodes numériques et statistiques pour une variété de domaines faisant appel à l’analyse numérique et statistique de signaux, d’images, de volumes, etc., et plus spécifiquement à des problèmes faisant appel à la statistique d’excédence ou de maxima, avec des caractéristiques multi-échelles : détection d’anomalie, modélisation de processus auto-similaires, de processus fractals ou multi-fractals. Cependant, il est apparu en cours de route que c’est en turbulence pleinement développée où les outils en question menaient naturellement à la contribution la plus intéressante.

Avant de présenter le contenu des chapitres, il est utile d'énoncer les lignes directrices de l'ensemble des travaux couverts dans cette thèse. C'est la notion d'*analyse multi-échelle* qui joue ici ce rôle fédérateur. Une quantité physique observable variant dans l'espace ou dans le temps peut généralement être décrite par une fonction (en 1, 2, 3, 4, 6 dimensions) définie sur un domaine (spatial, temporel, de phase), ou par un objet similaire à une fonction (mesure, distribution). Ces fonctions présentent dans un grand nombre de cas des propriétés d'invariance d'échelle, i.e. les équations qui régissent leur comportement demeurent vraies sous des transformations du type  $x \rightarrow \lambda x$ , où  $\lambda$  est un scalaire positif. Dans d'autres cas, ce sont des phénomènes mal expliqués par un modèle (des anomalies) qui viendront modifier la fonction, et cette modification pourra s'effectuer à différentes échelles.

Des outils mathématiques et statistiques furent développées pour permettre la description et l'analyse de propriétés multi-échelle, et ce plus particulièrement au courant des 25 dernières années. On rassemble librement plusieurs de ces outils sous l'appellation *analyse en ondelettes* [4, 5, 6]. Des définitions formelles de certains de ces outils sont énoncées tout au long de cette thèse, lorsque le contexte l'exige ; pour l'instant il suffit d'en présenter l'esprit général.

L'analyse en ondelettes est analogue à bien des égards à l'analyse de Fourier ; comme cette dernière, elle s'interprète comme la décomposition d'une fonction sur une base de fonctions qui engendre un espace fonctionnel particulier ( $L^1$ ,  $L^2$ ...). Par exemple, dans l'espace de Fourier réel cette base sera formée de sinusoïdales. Dans une certaine mesure, l'analyse de Fourier peut également être considérée comme une analyse multi-échelle puisque ses différents modes (ses différentes fréquences) caractérisent les fluctuations de la fonction analysée à différentes échelles. Cette analyse, cependant, est complètement délocalisée puisque tous les modes de Fourier s'étendent sur l'ensemble du domaine de la fonction étudiée. L'analyse en ondelettes se distingue de l'analyse de Fourier sur ce point, offrant une analyse

multi-échelle *locale* de la fonction ; c'est dans une large mesure ce qui en justifie l'intérêt qu'on lui porte.

Il est utile pour la suite du propos de présenter dès maintenant une définition partielle de la Transformée en Ondelettes (T.O.) continue  $T_\psi[f](b, a)$  en  $b$  d'une fonction  $f(\cdot) : \mathbb{R} \rightarrow \mathbb{R}$  selon l'ondelette  $\psi(\cdot)$  à l'échelle  $a$  :

$$T_\psi[f](b, a) = \frac{1}{a} \int \psi\left(\frac{x-b}{a}\right) f(x) dx. \quad (1)$$

Dans ce contexte, la T.O. peut-être interprétée comme une simple convolution locale d'une fonction  $f(\cdot)$  avec un noyau  $\psi(\cdot)$  dilaté par un facteur d'échelle  $a > 0$  et ayant subi une translation  $b$ . On choisira pour  $\psi(\cdot)$  une fonction d'intégrale nulle ; nous reviendrons sur ce point un peu plus loin.

Cette représentation de la fonction  $f(\cdot)$  est de toute évidence redondante ; il suffit d'observer que la représentation est décrite par deux paramètres,  $a$  et  $b$ , impliquant le passage à un espace en deux dimensions alors que la fonction originale provient d'un espace à une seule dimension. Cette redondance peut être levée de deux façons.

La manière la plus simple consiste à effectuer un échantillonnage discret de la T.O. continue en des points fixes tels que la représentation devienne minimale, i.e. en utilisant seulement un sous-ensemble discret de "positions"  $a$  et  $b$  tel que les ondelettes associées constituent une base orthonormale de l'espace des fonctions. C'est ainsi que l'on définira plus loin la Transformée en Ondelettes dyadique discrète.

On peut également *sélectionner* un sous-ensemble de coordonnées  $a$  et  $b$  qui dépendra explicitement de la fonction étudiée, i.e. dont les valeurs varieront d'une fonction à une autre. Par exemple, on s'attend à pouvoir représenter la fonction en se concentrant sur des coordonnées  $a$  et  $b$  telles que la T.O. continue prennent des valeurs significatives en ces points. Autrement dit, puisque la représentation en T.O. continue est très redondante, on s'attend à ce que l'énergie de la fonction se concentre sur certaines régions de l'espace bi-dimensionnel engendré par les

coordonnées  $a$  et  $b$ . À cet égard il existe un résultat très intéressant concernant les Maxima (locaux) du Module de la Transformée en Ondelette continue (MMTO) ; sous certaines conditions [7], les maxima locaux du module suffisent à reconstruire la fonction originale. Même lorsque ces conditions ne sont pas respectées, on peut obtenir numériquement une très bonne approximation à la fonction originale [8, 9, 10, 11, 12]. Cette représentation est bien différente de la T.O. dyadique discrète, car la position des coordonnées  $a$  et  $b$  dépend explicitement de la fonction.

La comparaison de ces deux approches est éclairante. L'avantage majeur de la T.O. dyadique discrète est son implémentation numérique, bien plus simple et bien plus stable. Il existe des algorithmes très performants effectuant à la fois l'analyse et la synthèse de fonctions, et ce en 1, 2 et même 3 dimensions. La méthode des maxima du module, quant à elle, est très lourde numériquement. La T.O. dyadique discrète garantit aussi une reconstruction exacte de la fonction originale à partir de ces coefficients en ondelettes grâce à un algorithme efficace, ce qui n'est pas le cas pour la MMTO dont l'algorithme de reconstruction n'est exact que si l'on dispose d'un temps infini.

La MMTO, cependant, offre une représentation plus spécifique de la fonction. Bien que la T.O. dyadique discrète ne soit pas redondante, il arrive qu'elle "diffuse" une structure locale de la fonction sur plusieurs de ces coefficients. La MMTO constitue également une représentation invariante sous translation ; les maxima ne s'en trouvent que simplement décalés, alors qu'une translation aura sur la T.O. dyadique discrète l'effet d'une redistribution complète de l'énergie sur tous ses coefficients. Bien que plutôt localisée, cette redistribution crée certaines difficultés quant à la localisation précise de l'information.

Au-delà de ces considérations, la MMTO permet également l'étude formelle de la régularité (ou fractalité) de la fonction, mieux que ne le fait la T.O. dyadique discrète. Le théorème de Jaffard, présenté dans le chapitre 3, explique clairement que la caractérisation de la régularité d'une fonction (en terme de ses singularités) s'obtient sans ambiguïté par l'étude des maxima locaux du module de la T.O.

La structure des maxima porte donc une information tangible sur les propriétés fondamentales de fonctions. Ainsi cette “concentration” de l’information tangible concernant une fonction dans ses maxima du module de la T.O. justifie dans une large mesure l’importance qui y est accordée tout au long de cette thèse ; la façon dont cette information peut être extraite et analysée y est présentée en détail.

Toutes ces questions seront discutées dans les chapitres qui suivent. Pour l’instant, il suffit de garder en tête cette dichotomie représentationnelle. Le passage de l’une à l’autre, dépendamment des besoins propres à un contexte donné, constitue le fil conducteur de cette thèse.

La thèse présente trois résultats (sous forme de publication par article) qui constituent le corps des trois chapitres. Chaque article est précédé d’une revue de littérature élaborée de manière à présenter tous les éléments nécessaires à la compréhension de l’article et qui ne font pas typiquement partie de la connaissance scientifique générale. Les éléments couverts dans cette partie ainsi que la structure de la présentation sont choisis de sorte que les résultats novateurs se retrouvant dans les articles soient compris comme suite logique à la présence de questions ouvertes dans leur champ respectif. Cette façon de faire est complétée par une brève présentation en français du contenu de chaque article ; cette présentation suit la revue de littérature et précède l’article lui-même.

Dans le premier chapitre nous présentons une méthode d’analyse de signaux et de détection d’anomalie basée sur la statistique d’excédence de champs stochastiques continus, par le calcul de l’espérance du nombre d’optima locaux d’un champ multi-échelle sur un domaine fini. Nous reprenons dans la section 2.1 les éléments fondamentaux des travaux d’Adler [13] concernant la géométrie des champs stochastiques, en présentant d’abord quelques notions simples de topologie et le lien qui les unit au problème qui nous intéresse.

Le deuxième chapitre couvre un problème récent en phénoménologie de la turbulence pleinement développée (i.e. à de grandes valeurs du nombre de Reynolds), celui du phénomène d’intermittence. Nous nous intéressons spécialement

à l'approche multi-fractale (développée largement par Frisch et Parisi [14]) et aux modèles "en cascades" qui en découlent, en l'occurrence par l'entremise des travaux de Muzy et Arnéodo et de leur "méthode des maxima de la transformée en ondelettes" (MMTO) [15]. Nous présentons alors dans la section 3.1 un survol historique des travaux en turbulence depuis le modèle de Kolmogorov en 1941, et plus particulièrement nous nous attardons au problème d'intermittence. Dans la section 3.2 nous décrivons les notions fondamentales de l'analyse multi-fractale en portant un intérêt plus spécifique à la Méthode des Maxima de la Transformée en Ondelettes (MMTO) proposée par Muzy et Arnéodo. Nous présentons aussi la définition du modèle en cascades multiplicatif dans cette section.

Le troisième chapitre, qui contient le résultat le plus important de la thèse, associe des éléments présentés dans les deux premiers chapitres pour introduire un modèle statistique expliquant le comportement d'échelle des moments de la distribution des incréments de vitesse et de la dissipation d'énergie dans la dynamique turbulente. Nous y présentons une réinterprétation de la phénoménologie proposée par She et Lévêque à travers leur "loi d'échelle universelle" (universal scaling law) [16]. La section 4.1 reprend les grandes lignes du modèle de She-Lévêque. Nous présentons dans la section 4.1.2 certains éléments de statistique de maxima sur lesquels les résultats du chapitre 3 reposent.

## Chapitre 2

---

### MAXIMA DE LA TRANSFORMÉE EN ONDELETTES CONTINUE DE DONNÉES CORRÉLÉES

## 2.1. REVUE DE LITTÉRATURE : STATISTIQUE D'EXCÉDENCE DE CHAMPS STOCHASTIQUES, APPROCHE TOPOLOGIQUE

Nous présentons dans cette section une introduction générale à la théorie multi-dimensionnelle des probabilités d'excédence pour des champs stochastiques. Nous nous attarderons plus spécifiquement à l'application de cette théorie aux champs obtenus par transformée en ondelettes continue d'un champ stochastique.

Le problème qui nous intéresse ici s'exprime et se résout trivialement pour un champ stochastique (un signal, une fonction) en une dimension et c'est sa généralisation en plusieurs dimensions qui renferme sa complexité et son intérêt. Nous commençons tout de même par énoncer le problème en une dimension et nous verrons ensuite comment doit se faire cette généralisation.

Soit un champ stochastique uni-dimensionnel  $X(t)$ ,  $0 \leq t \leq T$  suffisamment régulier, possédant un certain nombre de dérivées bornées (nous spécifierons des conditions précises plus loin), dont nous connaissons la fonction de distribution en tout point, ainsi que la fonction de corrélation spatiale. On s'intéresse à la statistique décrivant le nombre de fois, en moyenne, où ce champ excédera un seuil  $b$  donné. On cherche alors à obtenir l'espérance du nombre d'intervalles distincts sur lesquels la fonction (le champ) prend des valeurs supérieures ou égales à  $b$ . Cette statistique sera intéressante, par exemple, pour la détection d'un signal pollué par un bruit. Dans ce cas, on considère le bruit comme le champ étudié. S'il existe un intervalle (ou plusieurs) sur lequel une réalisation du processus dépasse un seuil pour lequel l'espérance du nombre d'intervalle d'excédence (pour le bruit seul) est largement inférieure à 1, on pourra en déduire la présence dans cet intervalle d'un signal non-expliqué par une simple fluctuation normale du bruit. Nous reviendrons plus tard à des exemples précis.

On cherche à exprimer cette espérance en terme de paramètres connus du champ. Il serait souhaitable de pouvoir l'écrire comme l'espérance d'un processus ponctuel si l'on connaît déjà la statistique du champ "point par point". En fait, on

se convainc facilement que le nombre d'intervalles sera égal au nombre de points prenant exactement la valeur  $b$  et pour lesquels la dérivée est positive, i.e. le nombre de fois que l'on dépasse le seuil en se promenant le long de la dimension  $t$ , auquel on doit ajouter la probabilité que le champ soit plus grand que  $b$  au début de l'intervalle, i.e.  $X(t=0) \geq b$ . Le problème uni-dimensionnel est résolu.

Comment généraliser cette technique pour un champ sur un sous-ensemble  $C$  de  $\mathbb{R}^N$ ? Voyons d'abord ce qui se passe en deux dimensions, pour fixer les idées. On peut alors voir une réalisation du champ comme une fonction de  $\mathbb{R}^2$  dans  $\mathbb{R}$ , i.e. comme un ensemble de montagnes et de vallées. Pour un seuil suffisamment élevé tel qu'un seul sommet (le maximum global de cette réalisation) soit supérieur à ce seuil, le champ dépassera le seuil sur une petite région fermée de  $C$ . Par analogie avec le cas uni-dimensionnel, nous voudrions que cette petite région "compte" comme un intervalle en une dimension. En diminuant la valeur du seuil, d'autres maxima locaux du champ vont se manifester, et encore une fois on voudra les compter comme des intervalles supplémentaires pour lesquels on a  $X(t) \geq b$ .

Que se passe-t-il lorsque, pour une valeur de  $b$  donnée, on trouve un "lac", autrement dit lorsque le champ prend une valeur supérieure à  $b$  sur une région annulaire? On peut considérer que le champ a alors dépassé le seuil une seule fois, ou encore deux fois puisque le long d'un axe traversant l'anneau le champ excédera le seuil deux fois (figure 2.1). Avant de répondre à cette question, examinons quelques caractéristiques souhaitables de la valeur que nous tentons de définir. Nous aimerions d'abord qu'elle puisse être définie comme l'espérance d'un processus ponctuel, afin qu'elle soit calculable à partir de la statistique du champ. Nous voudrions aussi que cette représentation soit invariante sous une rotation ou une translation du domaine  $C$  du champ, et idéalement sous toute déformation continue de celui-ci. Nous sommes donc à la recherche d'une caractéristique topologique du champ stochastique. La section qui suit présente un survol de quelques notions simples de topologie qui permettront d'élaborer par la suite l'outil recherché.

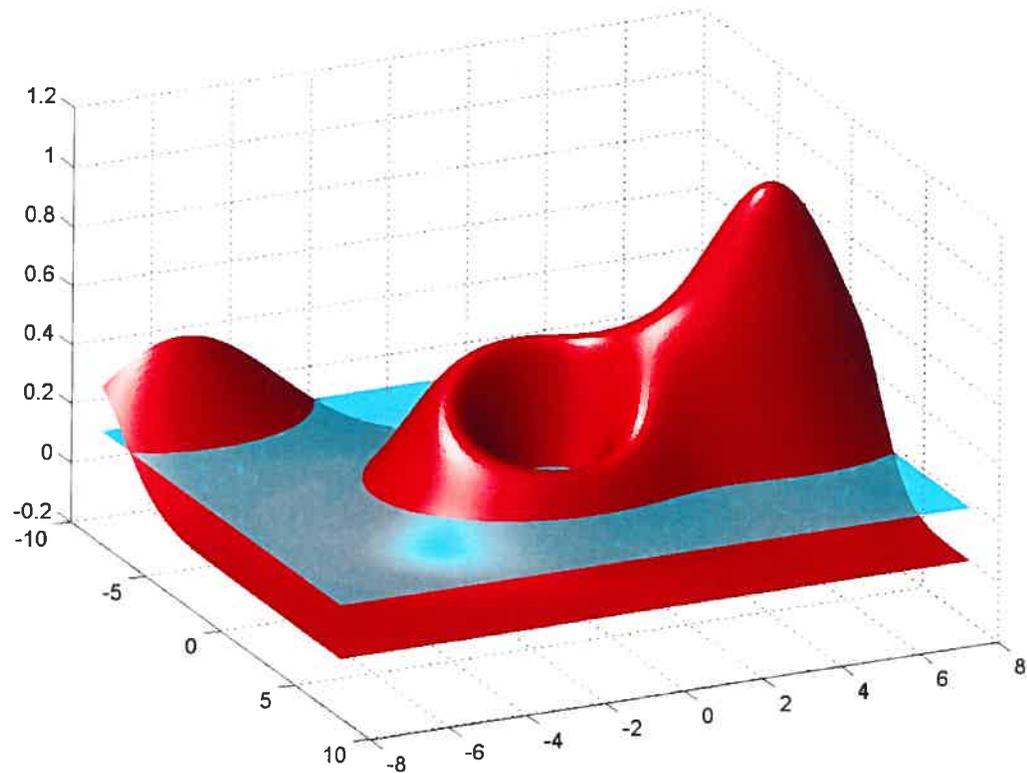


FIG. 2.1. Exemple de réalisation d'un champ en 2 dimensions. L'ensemble d'excédence par rapport au plan comporte une zone annulaire (un lac).

### 2.1.1. Topologie, caractéristique d'Euler

Nous reprenons dans cette section les grandes lignes de la présentation intuitive de Worsley [17]. Il est utile de présenter la définition de la caractéristique d'Euler en passant par l'observation originelle d'Euler à propos de la relation entre le nombre de sommets, d'arêtes et de faces d'un polyèdre quelconque. Cette relation est donnée par

$$S - A + F = 2, \quad (2)$$

où  $S$  représente le nombre de sommets,  $A$  le nombre d'arêtes et  $F$  le nombre de faces. On vérifie facilement pour un cube qu'on a bien  $8 - 12 + 6 = 2$  (figure 2.2a).

Si l'on considère plutôt un solide constitué de  $P$  polyèdres partageant au moins une face (comme par exemple deux cubes ayant une face commune), on peut généraliser la relation précédente ainsi :

$$S - A + F - P = 1, \quad (3)$$

où les sommets, arêtes et faces communes ne sont évidemment comptés qu'une seule fois. Dans le cas de deux cubes, on a bien  $12 - 20 + 11 - 2 = 1$  (figure 2.2b). La relation précédente n'est cependant pas toujours vérifiée. En effet, un solide

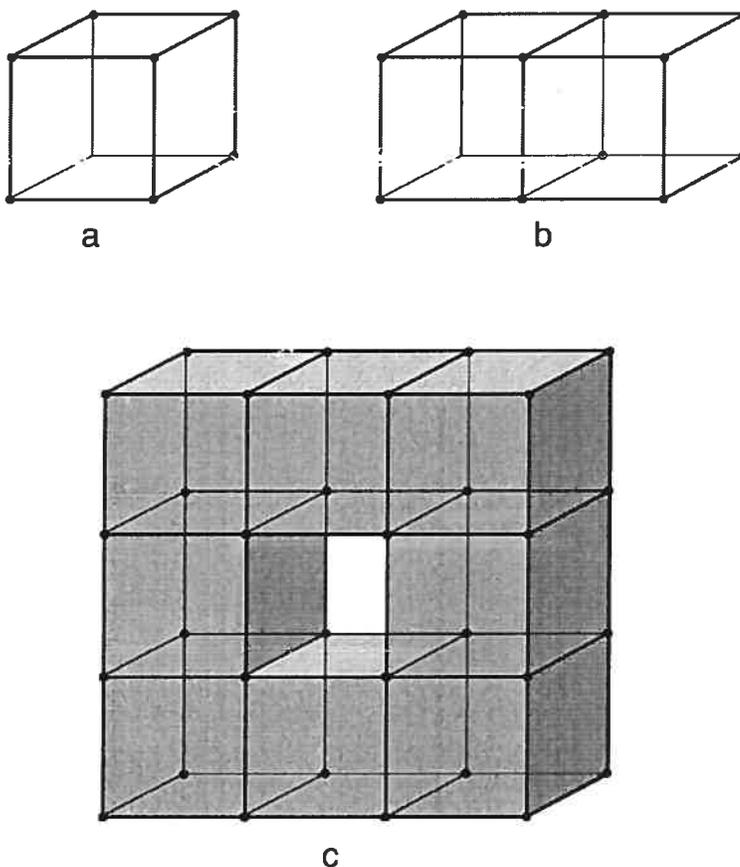


FIG. 2.2. Caractéristique d'Euler de quelques polyèdres.

a) Cube :  $8 - 12 + 6 - 1 = 1$  b) 2 cubes joints par une face :

$12 - 20 + 11 - 2 = 1$  c) "Tore" :  $32 - 64 + 40 - 8 = 0$ .

prenant la forme d'un tore (donc avec un "trou" ou une "poignée", figure 2.2c)

respectera plutôt la relation

$$S - A + F - P = 0. \quad (4)$$

De même, on aura  $-1$  pour un solide à deux trous, et ainsi de suite. Si on s'intéresse à un ensemble de polyèdres non-connexes, on peut sommer les relations précédentes. On en conclut que la quantité  $S - A + F - P$  compte le nombre de solides distincts, moins le nombre total de trous pour tous les solides considérés. Lorsque les solides considérés sont tous sans trous, on comptera simplement le nombre de solides distincts. Voilà qui nous rapprochent de notre objectif. Notons également qu'un solide "creux", i.e. analogue à une balle de tennis, respectera  $S - A + F - P = 2$ . On interprète ce résultat par le fait que les sommets, arêtes et faces internes ne "s'annulent" pas entre elles, et ainsi que la surface intérieure joue le même rôle qu'un solide sans trous, augmentant le compte à 2.

Nous n'avons considéré pour l'instant que des solides formés de polyèdres. Il est utile d'envisager des volumes "continus" (n'importe quel sous-ensemble dans  $\mathbb{R}^3$ , en fait) dont la surface, comme la sphère, n'est pas nécessairement formé de polygones. On peut alors toujours se représenter le volume en l'approximant par l'union d'un grand nombre de petits polyèdres (e.g. des cubes), et la relation demeurera juste.

Cette quantité qui nous intéresse, dont la valeur est donnée par  $S - A + F - P$ , se nomme *caractéristique d'Euler*. Elle possède toutes les propriétés d'invariance souhaitée, i.e. elle est conservée même sous une transformation continue (sans *déchirure*) de tout volume. Nous avons donc décrit une caractéristique topologique pour tout ensemble dans  $\mathbb{R}^3$ .

On peut maintenant faire le lien avec le problème original. Pour une réalisation d'un champ stochastique, on considèrera la caractéristique d'Euler de l'ensemble des points qui excèdent un seuil donné. Notons que cette caractéristique se généralise aussi à des dimensions supérieures, comme nous le verrons plus loin. Pour un espace de dimension 2, c'est encore plus simple. La caractéristique d'Euler

compte le nombre de régions distinctes et non-connexes, moins le nombre de *lacs* se trouvant dans ces régions. Nous avons donc également répondu à la question ci-haut concernant une région annulaire : sa caractéristique d'Euler vaut zéro. En dimension 1, tel que décrit au tout début de ce chapitre, chaque intervalle augmente la caractéristique d'Euler de 1.

C'est par cette caractéristique que nous généralisons la notion de *comptage* du nombre de fois qu'un champ stochastique excède un seuil. Évidemment le fait qu'une région annulaire (en 2 dimensions) ou qu'un tore possède une caractéristique d'Euler nulle peut paraître embêtant à première vue. En fait, pour la détection de signaux bruités, on s'intéressera en général à des seuils suffisamment élevés pour que la probabilité qu'une seule région ne dépasse le seuil (en présence uniquement du bruit) soit déjà faible. À ces seuils on n'observe pas, en général, de région annulaire ou de tore mais plutôt une petite région connexe de caractéristique d'Euler égale à 1.

Au-delà de cette subtilité, la caractéristique d'Euler est intéressante en soi pour décrire la nature topologique du champ stochastique. Pour un ensemble de dimension 3, on peut définir de manière informelle trois types de caractérisation topologique. Pour un ensemble formé de plusieurs composantes distinctes non-connexes, on parlera d'une topologie "en boulettes" ("meatball" en anglais), qui possèdera une caractéristique d'Euler positive, égale au nombre de composantes. Lorsque l'ensemble est formé d'un réseau complexe d'interconnexion, on parlera plutôt d'une topologie "éponge" et l'ensemble des "trous" ou "poignées" rendront la caractéristique d'Euler négative. Finalement, pour un ensemble presque plein avec des régions vides à l'intérieur, on aura une topologie "bulles" avec une caractéristique d'Euler à nouveau positive. On observe souvent ces trois classes (dans l'ordre) en diminuant progressivement le seuil. Pour une réalisation donnée d'un champ stochastique, on pourra toujours calculer la caractéristique d'Euler de l'ensemble des points qui excèdent un seuil donné. Notre objectif cependant sera de calculer l'espérance de cette caractéristique pour un champ connaissant

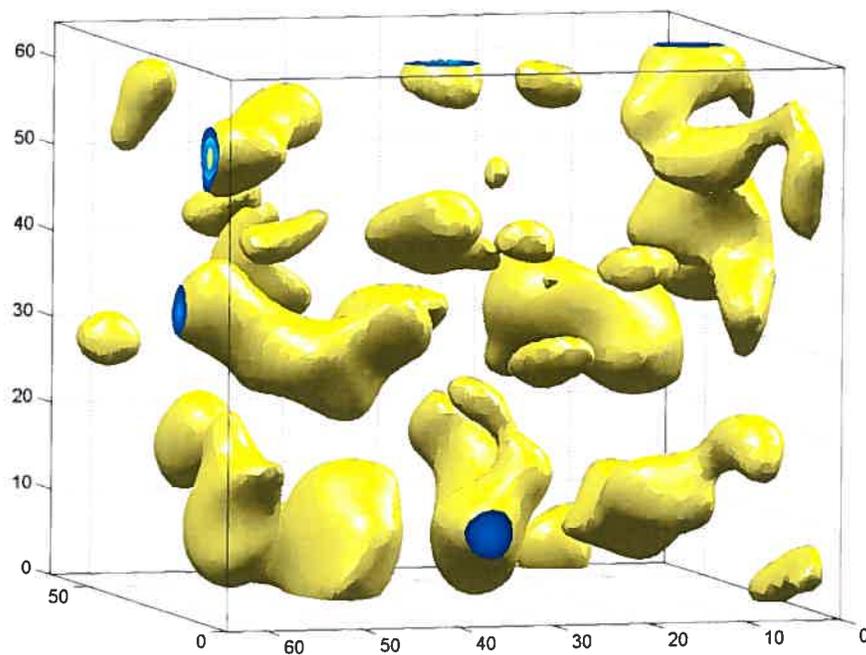


FIG. 2.3. Topologie “en boulettes”. Cette figure est obtenue en générant un bruit blanc en 3D filtré par un noyau gaussien, puis en extrayant une isosurface.

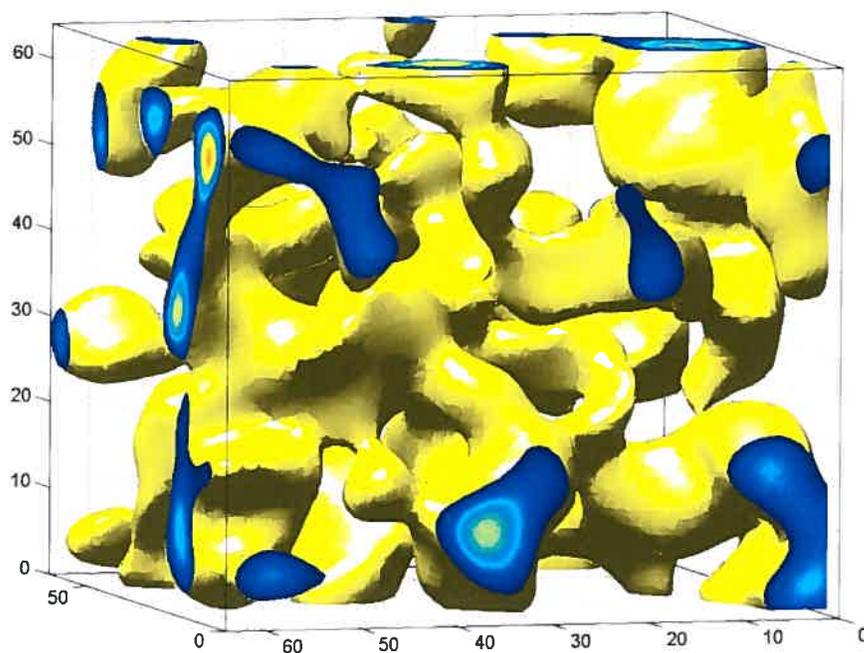


FIG. 2.4. Topologie “éponge”.

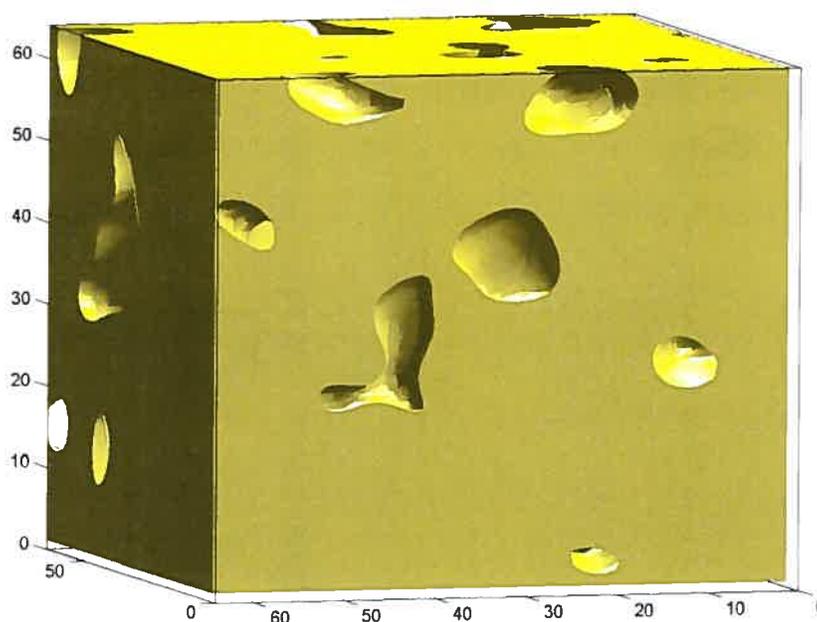


FIG. 2.5. Topologie "bulles".

sa statistique, i.e. ses paramètres (e.g. moyenne, variance, fonction de covariance spatiale).

Avant de passer au calcul de l'espérance de la caractéristique d'Euler, nous présentons brièvement quelques exemples d'applications en justifiant l'utilité.

### 2.1.2. Statistique d'excédence et caractéristique d'Euler : Applications

En couverture du Scientific American de juillet 1992 on trouve une image créée à partir des données du radio-télescope COBE, résultat des travaux de 28 chercheurs [18]. Cette "carte du ciel" présente les fluctuations du rayonnement fossile obtenues par moyennage de données cumulées pendant un an de fonctionnement du télescope. La question primordiale était de déterminer si on peut déduire de ces données bruitées la présence d'anomalies dans les radiations du rayonnement fossile. La méthode de la caractéristique d'Euler a été appliquée à ces données ; on en a conclu qu'il s'y trouvait bel et bien des anomalies, avec une probabilité

significative qu'il ne s'agissait pas simplement de fluctuations du bruit en présence [19, 20]. Depuis, de meilleures mesures ont été effectuées et ont confirmé l'existence de ces anomalies.

On fait face à un problème similaire en imagerie cérébrale fonctionnelle. Il existe maintenant plusieurs appareils permettant d'obtenir des cartes fonctionnelles du cerveau humain; parmi celles-ci, les deux plus utilisées à ce jour sont l'Imagerie par Résonance Magnétique fonctionnelle (IRMf) [21], et la Tomographie par Émission de Positrons (TEP) [22]. Pour ces deux modalités on rencontre le même problème, à savoir la détection d'une activation réelle du cerveau lorsque le sujet exécute une tâche pré-définie, à partir d'une carte fortement bruitée du cerveau. Cette carte est représentée par un ensemble de petits cubes (voxels) couvrant tout le volume du cerveau, auxquels on associe une valeur d'intensité qui, en l'absence de bruit, ne devrait être élevée que dans les régions activées.

Une approche naïve consisterait à utiliser un test statistique pour chaque élément de volume (voxel), et de considérer comme activés ceux pour lesquels on observe une intensité supérieure à un seuil de confiance de 95%, par exemple. Or un échantillonnage typique peut compter plus d'un million de voxels. En considérant une expérience sans activation (i.e. où les intensités ne sont redevables qu'aux fluctuations du bruit), on aura quand même 50000 voxels faussement interprétés comme activés (faux positifs). En fait, le test serait valable si les intensités de tous les voxels étaient fortement corrélés, de manière à ce que tous s'activent (ou ne s'activent pas) simultanément. En pratique, ça n'est jamais le cas.

À l'autre extrême, si on considérait que les valeurs d'intensité de tous les voxels sont indépendantes les unes des autres, on corrigerait le seuil de manière à s'assurer qu'aucun des voxels non-activés n'excèdent ce seuil, avec une probabilité de 95%. Cette méthode est connue sous le nom de *correction de Bonferroni*. Typiquement, cette correction est trop sévère, car il existe toujours un certain degré de corrélation entre des voxels voisins. La caractéristique d'Euler, elle, est une quantité globale dont le calcul fait intervenir la structure de corrélation spatiale

du bruit et la taille (et la géométrie) du domaine échantillonné. Elle permet donc d'établir de manière fiable le seuil d'activation recherché puisqu'elle évite cette sous-estimation (ou sur-estimation) du degré de corrélation entre les valeurs des différents voxels.

### 2.1.3. Espérance de la caractéristique d'Euler : développement formel

Bien que le calcul de l'espérance de la caractéristique d'Euler pour un champ stochastique donné ne soit pas trivial, il est grandement simplifié grâce à l'approche développée par Robert Adler [13]. Cette approche fait appel à la *géométrie intégrale* qui permet de décrire les caractéristiques topologiques des ensembles de niveau d'une fonction à partir de ses points critiques. Nous ne présenterons dans cette section uniquement les éléments nécessaires à notre discussion ; une étude exhaustive de la géométrie intégrale peut être trouvée dans [23, 24, 25, 26].

Pour aller plus avant, il est maintenant nécessaire de considérer le calcul de l'espérance de la caractéristique d'Euler de "l'ensemble d'excédence" d'un champ stochastique dans un cadre plus formel. Nous débutons par quelques définitions qui nous rapprochent de cet objectif.

Soit  $X(t)$ ,  $t = (t_1, \dots, t_N) \in \mathbb{R}^N$  un champ stochastique, que l'on suppose stationnaire pour l'instant, et soit  $C$  un sous-ensemble compact de  $\mathbb{R}^N$ . On considère pour toute définition de la stationnarité que la statistique du champ ne dépende pas de la position  $t \in C$ , une définition suffisante pour les fins de cette discussion. On définit alors *l'ensemble d'excursion*  $A_b$  de  $X(t)$  comme l'ensemble des points de  $C$  où  $X(t)$  excède un seuil  $b$ , i.e.

$$A_b = \{t \in C; X(t) \geq b\}. \quad (5)$$

$A_b$  est bien sûr un sous-ensemble de  $C$ , et donc de  $\mathbb{R}^N$ . L'ensemble d'excursion n'est autre que "l'ensemble d'excédence" présenté ci-haut.

Soit  $B$  un sous-ensemble compact de  $\mathbb{R}^N$  tel que l'intersection de  $B$  avec tout hyperplan de dimension  $k$ ,  $k = 1, \dots, N$  soit connexe. On entend par hyperplan un ensemble formé par les combinaisons linéaires d'un sous-ensemble des vecteurs canoniques dans  $\mathbb{R}^N$ , par exemple le plan X-Y ou l'axe Z en trois dimensions. On dira alors de  $B$  qu'il est un sous-ensemble *basique* de  $\mathbb{R}^N$ . C'est le cas par exemple d'une sphère ou d'un cube. On interprète librement cette définition comme analogue à la convexité d'un ensemble. En fait un ensemble convexe est toujours également un basique, mais l'inverse n'est pas vrai. Toute union d'un nombre fini d'ensembles basiques constituera un *complexe basique*.

L'ensemble  $A_b$  associé à un champ  $X(t)$  sera presque sûrement un complexe basique, en admettant que le champ  $X(t)$  respecte certaines conditions de régularité ([13], chap. 3).

Notée  $\chi(A)$ , la *caractéristique d'Hadwiger* est la caractéristique d'Euler d'un complexe basique. Cette dernière est plus générale, car elle est aussi définie pour des ensembles qui ne sont pas des complexes basiques. Ces ensembles ne nous intéressent pas, puisque les conditions que nous imposerons aux champs étudiés feront en sorte que leur ensemble d'excursion sera presque sûrement un complexe basique ; nous utiliserons désormais la caractéristique d'Hadwiger, tout en gardant en tête qu'elle ne diffère pas ici de la caractéristique d'Euler.

Donnons maintenant la définition de la caractéristique d'Hadwiger. En une seule dimension,  $\chi(A)$  correspond au nombre d'intervalles disjoints de  $A$ . Pour  $N > 1$ , on a une définition récursive. Considérons d'abord l'hyperplan  $\mathcal{E}_u = \{t \in C; t_N = u\}$  qui définit un sous-espace de dimension  $N - 1$  pour lequel la  $N$ -ième coordonnée prend la valeur  $u$ . En se "promenant" le long de l'axe  $t_N$  (donc en variant  $u$ ), on percevra l'apparition ou la disparition d'un objet dès que la caractéristique d'Hadwiger de l'intersection de  $\mathcal{E}_u$  et du complexe basique  $A$  changera de valeur (voir figure 2.6).

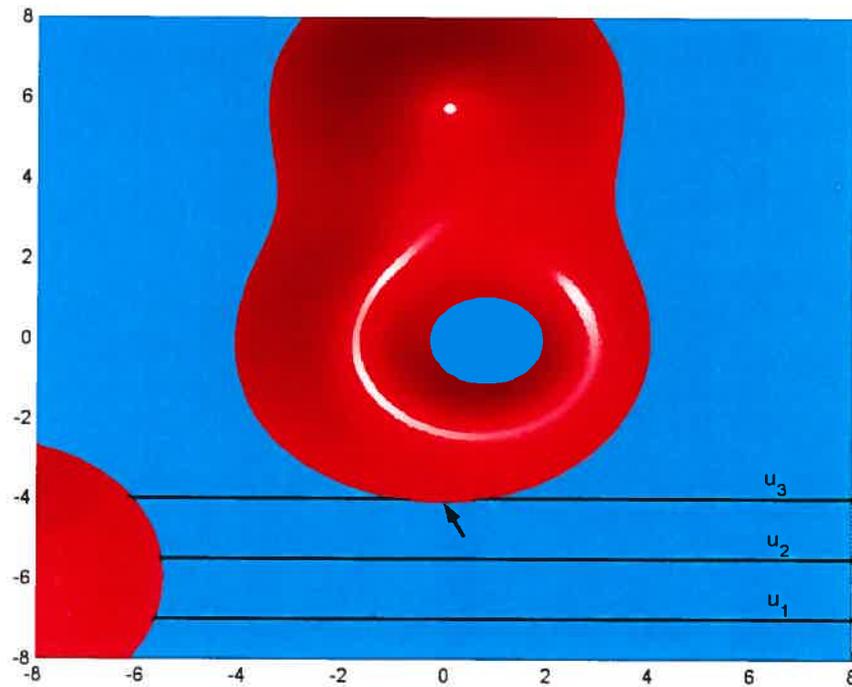


FIG. 2.6. Le même champ présenté à la figure 2.1, vu du dessus. En passant de l'hyperplan  $u_1$  à l'hyperplan  $u_2$ , l'intersection avec l'ensemble d'excursion garde la même caractéristique d'Hadwiger (un seul intervalle d'excédence, à gauche). En passant de  $u_2$  à  $u_3$  on observe un changement d'un à deux intervalles d'excédence. Le point marqué d'un (+) contribue à la somme (6).

Ainsi on définit

$$\chi(A) = \sum_u [\chi(A \cap \mathcal{E}_u) - \chi(A \cap \mathcal{E}_u^-)], \quad (6)$$

où  $\mathcal{E}_u = \{t \in C; t_N = u\}$  et

$$\chi(A \cap \mathcal{E}_u^-) = \lim_{v \uparrow u} \chi(A \cap \mathcal{E}_v). \quad (7)$$

La somme sur  $u$  est une somme discrète sur l'ensemble des points pour lesquels l'expression  $\chi(A \cap \mathcal{E}_u) - \chi(A \cap \mathcal{E}_u^-)$  est non-nulle.

Notons que la caractéristique d'Hadwiger satisfait la propriété d'additivité suivante :

$$\chi(A \cup B) = \chi(A) + \chi(B) - \chi(A \cap B). \quad (8)$$

Rappelons que l'objectif était d'obtenir une caractéristique à la fois invariante sous rotation, mais d'abord et avant tout qui possède une représentation ponctuelle. En observant plus en détail la définition de  $\chi(A)$ , on s'aperçoit que cette caractéristique peut être obtenue itérativement dans les dimensions, i.e. on voit qu'elle mesure en fait un changement de la caractéristique d'Hadwiger définie sur l'intersection de l'ensemble initial (de dimension  $N$ ) avec un hyperplan de dimension inférieure ( $N - 1$ ). Revenons à l'ensemble d'excédence  $A_b$  d'un champ  $X(t)$ . Il devient alors évident que cette caractéristique de dimension inférieure ne peut être modifiée que pour un point  $t$  où  $X(t)$  prend précisément la valeur  $b$ ; ce point doit aussi constituer un maximum local de  $X(t)$  dans ce sous-ensemble de dimension inférieure. De plus, puisque la limite dans la définition est prise pour des valeurs de  $v$  inférieure à  $u$ , on doit avoir que la dérivée du champ dans la direction  $t_N$  soit positive (nous illustrerons cela avec un exemple). Finalement, la caractéristique ne pourra (presque sûrement) être modifiée en un point que par une valeur de  $\pm 1$  et c'est le nombre de valeurs propres négatives de la matrice des dérivées secondes de  $X(t)$  qui en contrôlera le signe. Toute autre contribution à la caractéristique  $\chi$  ne pourra provenir que de points situés sur la frontière de l'ensemble  $C$ ; dans le cas uni-dimensionnel, c'est le rôle que jouait le point  $t = 0$ . On écrit donc  $\chi(A) = \chi_V(A) + \chi_E(A)$  où  $\chi_V(A)$  est la contribution volumique à la caractéristique d'Hadwiger et  $\chi_E(A)$  la contribution de la frontière ("edge"), dont nous donnons les expressions maintenant.

*Caractéristique d'Hadwiger, contribution volumique  $\chi_V$*

Nous introduisons la notation suivante : soit  $X_i = \partial X / \partial t_i$ , et  $X_{ij} = \partial^2 X / \partial t_i \partial t_j$ . Soit  $D_{N-1}$  la matrice des dérivées secondes de taille  $(N - 1) \times (N - 1)$ , ayant

comme éléments les  $X_{ij}$ ,  $i, j = 1 \dots N - 1$ . La contribution  $\chi_V$  des points situés à l'intérieur de l'ensemble  $C$  sera donnée par

$$\chi_V(A_b) = \sum_C (X = b)(X_1 = 0) \cdots (X_{N-1} = 0)(X_N > 0) \text{ sign}[\det(-D_{N-1})], \quad (9)$$

où la somme sur  $C$  est à nouveau une somme discrète sur l'ensemble des points pour lesquels l'argument de la somme est non-nul. Nous utilisons également la notation de Knuth [27] où une expression logique entre parenthèse prend la valeur 1 si elle est vraie et 0 si elle fausse.

Cette équation demeure valable quelque soit le nombre de dimensions. La notation  $\chi_V$  signifie "contribution volumique" à la caractéristique. Cette contribution peut aussi s'écrire de la manière suivante :

$$\chi_V(A_b) = \sum_{t \in C} (X \geq b)(\dot{X} = 0) \text{ sign}[\det(-\ddot{X})]. \quad (10)$$

où  $\dot{X}$  est le gradient de  $X$  et  $\ddot{X}$  la matrice de toutes les dérivées secondes de  $X$ .

Il est assez simple de comparer ces deux représentations (9,10). La première "compte" chaque objet à partir des points situés à l'extrémité de l'objet selon la  $N$ -ième dimension, des points pour lesquels il est facile de vérifier que l'on a bien  $X = b$  puisque l'on est alors à la frontière de l'ensemble d'excursion  $A_b$ ; on a aussi  $X_i = 0$ ,  $i = 1 \dots N - 1$  puisque la tangente en ces points est nécessairement perpendiculaire à la direction  $t_N$ , et donc que le gradient de  $X$  est dans cette direction; finalement on a aussi  $X_N > 0$  puisque l'on ne veut compter chaque objet qu'une seule fois (on peut oublier cette dernière condition, mais l'on devra diviser par deux la valeur obtenue). Le signe du déterminant de  $\ddot{X}$  nous indique si l'on "gagne" ou si l'on "perd" un objet en se promenant le long de la  $N$ -ième dimension. La figure 2.6 illustre bien cette représentation en deux dimensions où on a pour le point (+) une dérivé nulle dans la direction horizontale et positive dans la direction verticale. Le signe du déterminant de  $\ddot{X}$  en ce point est positif.

La deuxième représentation identifie un objet par ses maxima centraux, pour lesquels on aura bien entendu  $X \geq b$ , et  $\dot{X} = 0$ . Le signe du déterminant de

$\ddot{X}$  nous assure par exemple qu'un objet connexe sans "trous" qui possède deux maxima ne comptera pas pour deux objets, puisque entre ces deux maxima se trouvera nécessairement un point de selle (pour lequel  $\det[\ddot{X}] < 0$ ) qui viendra annuler un de ces deux maxima.

*Caractéristique d'Hadwiger, contribution frontière  $\chi_E$*

Pour les contributions provenant de la frontière de l'ensemble  $C$ , il existe aussi une représentation par points, un peu moins intuitive cependant :

$$\chi_E(A_b) = \sum_{t \in \partial C} (X \geq b)(\dot{X}_T = 0)(\dot{X}_\perp < 0) \text{sign}[\det(-\ddot{X}_T - \dot{X}_\perp c)]. \quad (11)$$

où  $\dot{X}_T$  correspond au vecteur des dérivées partielles dans un plan tangent à la frontière  $\partial C$  de  $C$ , et où  $\dot{X}_\perp$  représente la dérivée de  $X$  le long de la normale intérieure de  $C$ . Nous utilisons toujours la notation de Knuth introduite ci-haut.

L'expression (11) est similaire à la deuxième représentation du terme volumique, où l'on s'intéresse à des points qui ne sont pas nécessairement situés à la frontière même de  $A_b$ . Il est également possible de réécrire cette dernière équation de manière similaire à la première représentation du terme volumique ; nous renvoyons à [29] pour sa formulation explicite.

La raison pour laquelle le terme  $\chi_E$  n'est généralement pas nul et, par conséquent, que certains points situés sur la frontière  $\partial C$  de l'ensemble  $C$  participent à la caractéristique d'Hadwiger, s'illustre par un exemple simple. Supposons que l'on ait un objet (une région de  $A_b$ ) qui soit en contact avec la frontière  $\partial C$ . Il est possible qu'à l'intérieur de cet objet ne se trouve aucun maximum (pour lequel on aurait  $\dot{X} = 0$ ). Revenant au cas bi-dimensionnel et à l'analogie topographique, on peut imaginer que la frontière de l'ensemble  $C$ , décrivant une région montagneuse, passe justement à travers la base d'une montagne, et que son sommet se trouve à l'extérieur de  $C$ . Alors, le maximum d'un ensemble d'excursion incluant la base de cette montagne sera situé quelque part sur la frontière  $\partial C$ . En fait, ce point sera aussi le maximum des points appartenant à la fois à  $A_b$  et à  $\partial C$  dans

cette région. On voit facilement que pour ce point, on aura  $(X \geq b)$ ,  $(\dot{X}_T = 0)$  et  $(\dot{X}_\perp < 0)$ . Ce cas est illustré sur la figure 2.7.

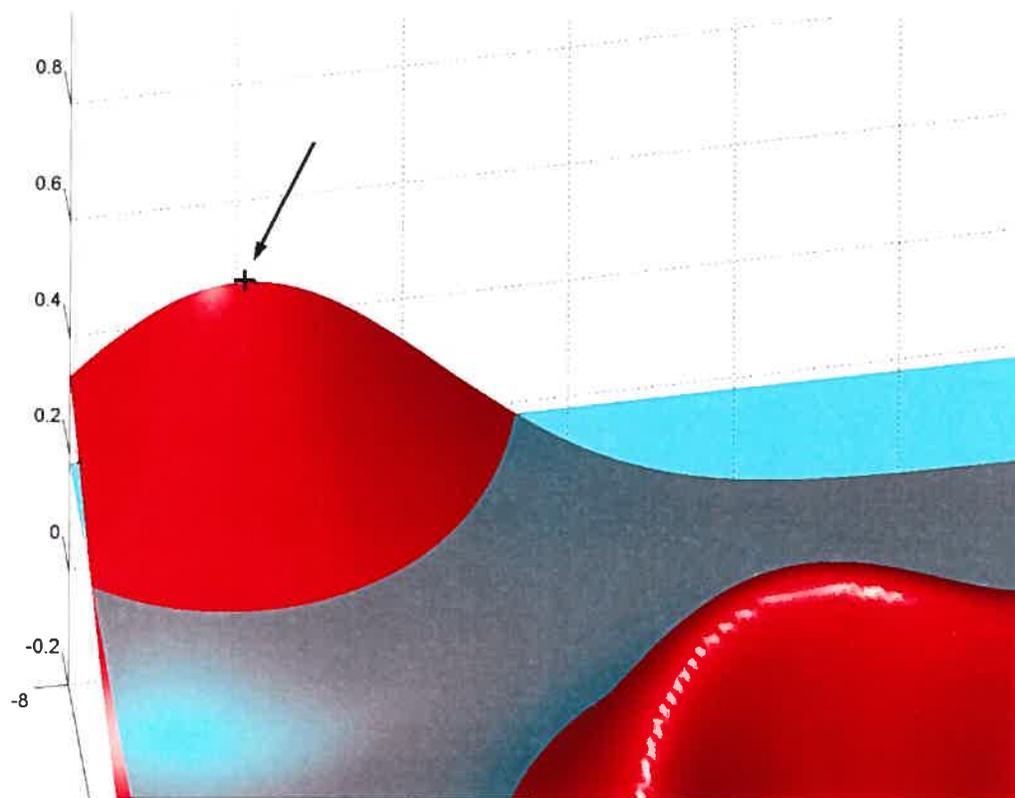


FIG. 2.7. Le même champ présenté à la figure 2.1, vu d'un autre angle. Le point marqué d'un (+) compte dans la contribution frontière (11).

#### 2.1.4. Calcul de l'espérance de la caractéristique d'Hadwiger

Ayant obtenu une représentation ponctuelle de la caractéristique d'Hadwiger pour un ensemble d'excursion  $A_b$  d'une réalisation d'un champ stochastique quelconque, nous cherchons maintenant à écrire l'espérance de cette caractéristique. Le résultat qui suit se trouve dans [19]; sa preuve s'inspire cependant largement du théorème 5.1.1 de Adler [13]. Parce qu'elle est très longue, nous ne reproduisons pas cette preuve ici, mais énonçons seulement le résultat suivi de son

interprétation intuitive.

$$E[\chi(A_b)] = E[\chi_V(A_b) + \chi_E(A_b)] \quad (12)$$

$$= \int_C E[(X \geq b) \det(-\ddot{X}) | \dot{X} = 0] \theta(0) dt \quad (13)$$

$$+ \int_{\partial C} E[(X \geq b)(\dot{X}_\perp < 0) \det(-\ddot{X}_T - \dot{X}_\perp c) | \dot{X}_T = 0] \theta_T(0) dt, \quad (14)$$

où  $\theta$  correspond à la fonction de densité de  $\dot{X}$ . L'espérance en (12) est prise sur la distribution jointe de  $X$ ,  $\dot{X}$  et  $\ddot{X}$ . L'espérance à l'intérieur de l'intégrale en (13) n'est prise que sur  $X$  et  $\ddot{X}$ , car nous avons conditionné sur  $\dot{X}$ . Le premier terme de cette équation, qui correspond à la contribution volumique, doit être donné par l'espérance du nombre de points satisfaisant les conditions suivantes :  $(\dot{X} = 0)$ ,  $(X \geq b)$ , et on doit lui assigner le signe du déterminant de  $\ddot{X}$ . On voit que le premier terme (la première intégrale) calcule l'espérance conditionnelle d'une quantité étant donnée la condition  $\dot{X} = 0$ ; c'est là que nous avons conditionné sur  $\dot{X}$ . Puisque l'on intègre sur la variable  $t$  plutôt que sur les valeurs possibles de  $\dot{X}$ , on a effectué un changement de variable qui doit faire apparaître un jacobien

$$J = \left| \det \left( \frac{\partial \dot{X}}{\partial t} \right) \right| = |\det(\ddot{X})|. \quad (15)$$

Puisque l'on doit multiplier ce jacobien par le signe de  $\det(\ddot{X})$ , on obtient bien la première intégrale. Le raisonnement est semblable pour le deuxième terme.

C'est précisément dans ce deuxième terme que cette formulation très générale cache une grande complexité calculatoire. Cependant, sous l'hypothèse d'isotropie et de stationnarité du champ, il devient possible de développer ce terme de manière très élégante. Pour cela, il nous faut d'abord écrire quelques définitions.

Soit  $\rho(X, b)$  l'intensité de la caractéristique d'Euler de l'ensemble d'excursion  $A_b$  par unité de volume, définie ainsi :

$$\rho(X, b) = E[(X \geq b) \det(-\ddot{X}) | \dot{X} = 0] \theta(0). \quad (16)$$

Le terme volumique dans l'expression (12) s'écrit maintenant

$$E[\chi_V(A_b)] = |C|\rho(X, b). \quad (17)$$

On a fait appel ici à l'hypothèse de stationnarité du champ. Cette formulation met en évidence la pertinence du terme "intensité" dans la définition de  $\rho(X, b)$ .

La restriction d'un champ à ces  $j$  premières variables est notée par

$$X_{|j}(t_1, \dots, t_j) = X(t_1, \dots, t_j, 0, \dots, 0). \quad (18)$$

Pour une matrice, cette notation correspond aux  $j$  premières colonnes et rangées de celle-ci. Soit  $\rho_j(b) = \rho(X_{|j}, b)$  l'intensité  $j$ -dimensionnelle de la caractéristique d'Hadwiger, pour  $1 \leq j \leq N$ . Pour  $j = 0$ , on écrit  $\rho_0(b) = P\{X \geq b\}$ . Pour une matrice  $M$  de dimension  $n \times n$ , on écrira  $\text{detr}_j(M)$  comme la somme des déterminants de toutes les sous-matrices principales de  $M$  de taille  $j \times j$ ,  $j = 1, \dots, N$ . On écrit aussi  $\text{detr}_0(M) = 1$ . Enfin, soit  $s_j = 2\pi^{j/2}/\Gamma(j/2)$  l'aire de la  $(j-1)$ -sphère de rayon 1 dans  $\mathbb{R}^j$ . On peut alors définir une mesure  $j$ -dimensionnelle  $\mu_j(C)$  de l'ensemble  $C$  :

$$\mu_j(C) = \frac{1}{s_{N-j}} \int_{\partial C} \text{detr}_{N-1-j}\{c(C)\} dt, \quad j = 0, \dots, N-1, \quad (19)$$

où  $c(C)$  est la matrice de courbure de  $\partial C$ , de taille  $(N-1) \times (N-1)$ . On écrit aussi  $\mu_N(C) = |C|$ .

Finalement, on obtient la formulation suivante pour l'espérance de la caractéristique d'Hadwiger d'un champ stationnaire et isotrope :

$$E[\chi(A_b)] = \sum_{j=0}^N \mu_j(C) \rho_j(b). \quad (20)$$

Le terme  $j = N$  dans cette somme correspond simplement au terme volumique  $\chi_V(A_b)$ . Les autres termes proviennent donc tous de la contribution de la frontière  $\partial C$ . Le développement complet de cette dernière expression est présenté dans [19].

La condition d'isotropie n'est pas nécessaire en toute rigueur. En fait, il suffit que le champ soit isotrope en  $N-1$  dimensions lorsque l'une des variables est fixée [30]. Cette subtilité peut paraître futile puisqu'elle ne se manifeste que

lorsque l'une (et une seule) des dimensions du champ brise l'isotropie. En fait, c'est précisément ce qu'il advient lorsque l'on projette un champ isotrope de dimension  $N$  dans un espace de dimension  $N + 1$  par sa transformée en ondelettes continue.

### 2.1.5. Caractéristique d'Hadwiger d'un champ multi-résolution

Définissons un champ multi-résolution  $Y(t, s)$  de la manière suivante : soit  $W(t)$  un champ stationnaire isotrope quelconque (il n'est pas nécessaire que ses dérivées existent) défini sur un domaine  $C_W$ . Soit  $Y(t, s)$  le champ défini par :

$$Y(t, s) = \langle W, \psi_s \rangle = \int \psi_s(t - h)W(h)dh \quad (21)$$

où  $\psi_s(t) = e^{Ns/2}\psi(e^st)$ , et  $s \in [s_1, s_2]$ . Nous verrons plus loin (section 3.2.1) que le champ  $Y$  est analogue à une transformée en ondelette continue, avec  $\psi(\cdot)$  l'ondelette associée et en considérant l'échelle  $a = e^{-s}$ . On voit facilement que pour une valeur de  $s$  fixée et une fonction  $\psi(\cdot)$  isotrope, le champ  $Y$  est stationnaire et isotrope en  $t$  puisqu'il ne constitue qu'une convolution du champ avec un noyau isotrope. Cependant, une inspection rapide nous montre que ce champ n'est pas conjointement stationnaire en  $(t, s)$ .

C'est ici que le commentaire amené dans la section précédente à propos de l'isotropie en  $N - 1$  dimensions devient utile. Le champ  $Y(t, s)$  respecte cette condition d'isotropie "faible", ce qui permet d'appliquer les résultats obtenus précédemment pour un champ stationnaire et complètement isotrope.

Un avantage fondamental de cette approche de type ondelette et qui n'est pas souligné ni dans le papier de Siegmund et Worsley [29] ni dans Worsley [30] réside dans l'assouplissement des conditions de régularité du champ "spatial"  $W(t)$ . En effet, même un champ qui ne serait continu en aucun point de son domaine mais dont la convolution avec une ondelette respecte les conditions de régularité de Adler [13] permettra le calcul de sa caractéristique d'Hadwiger.

Nous présentons maintenant un exemple canonique de champ stochastique discontinu presque sûrement partout sur son support, le *bruit blanc*. Ce champ

stationnaire est défini par sa fonction de distribution gaussienne de moyenne nulle et de variance unitaire en tout point de son support. La réalisation en chaque point est complètement indépendante, i.e. la fonction de corrélation spatiale est donnée par une distribution en delta de Dirac. Soit

$$\Lambda = e^{-2s} \int \dot{\psi} \dot{\psi}' dt, \quad (22)$$

$$\kappa = \int [t' \dot{\psi} + (N/2)\psi]^2 dt, \quad (23)$$

où  $\dot{\psi} = \partial\psi/\partial t$ . La notation  $(\cdot)'$  représente la transposée, lorsque la dimension de l'espace  $N > 1$ .

Nous avons montré dans la section précédente que le calcul de l'espérance de la caractéristique d'Hadwiger pour un champ stationnaire et faiblement isotrope se résume au calcul de ses intensités  $j$ -dimensionnelles, par l'expression  $E[\chi(A_b)] = \sum_{j=0}^N \mu_j(C) \rho_j(b)$ . Pour un champ gaussien respectant ces conditions, on a :

$$\rho_j(b) = \frac{e^{-b^2/2}}{(2\pi)^{\frac{j+1}{2}}} \left\{ \frac{\lambda_1^{j/2} + \lambda_2^{j/2}}{2} \text{He}_{j-1}(b) + \frac{\lambda_1^{j/2} - \lambda_2^{j/2}}{j} \sqrt{\frac{\kappa}{2\pi}} G_j(b, \kappa) \right\}, \quad (24)$$

où  $G_j(b, \kappa)$  est un polynôme de degré  $j$  en  $b$  et de degré  $\lfloor j/2 \rfloor$  en  $1/\kappa$ , donné par

$$G_j(b, \kappa) = \sum_{m=0}^{\lfloor j/2 \rfloor} \frac{(-1)^m j!}{2^m m! (j-2m)!} \left( \frac{1}{\kappa} \right)^m \frac{\text{He}_{j-2m}(b)}{1-2m}, \quad (25)$$

où  $\lambda_1$  et  $\lambda_2$  sont les variances du champ  $Y$  aux valeurs limites d'échelle  $s_1$  et  $s_2$  respectivement, et où  $\text{He}_j(\cdot)$  sont les polynômes d'Hermite d'ordre  $j$ .

On a alors tous les éléments nécessaires pour décrire la statistique d'excédence de ce type de champ. Une généralisation de ce résultat ainsi qu'un exemple d'application générique seront présentés dans le premier chapitre.

## 2.2. PRÉSENTATION DE L'ARTICLE

Dans ce chapitre nous reprenons les travaux de Siegmund et Worsley à propos de la caractéristique d'Euler de champs stochastiques gaussiens multi-échelle [29]. Tout le travail de Siegmund et Worsley porte sur le cas particulier d'un champ multi-échelle généré par la convolution d'un bruit blanc avec un noyau lissant, avec un paramètre continu supplémentaire conditionnant la taille du noyau, i.e. l'échelle. Rappelons que cette approche a pour but de permettre la détection d'un signal (dont la forme serait proche de celle du noyau) présent dans un bruit blanc lorsque la taille du signal n'est pas connue à l'avance.

Or en toute généralité la décorrélation du bruit sous-jacent n'est pas toujours vérifiée. On s'intéresse alors à un test qui permettrait d'obtenir le même résultat pour des bruits dits "colorés", i.e. dont la fonction de corrélation n'est pas le delta de Dirac. On s'attend à ce qu'une approche semblable à celle développée par Siegmund et Worsley permette malgré tout d'obtenir un test basé sur la caractéristique d'Euler. Une telle approche fait nécessairement appel à l'inversion d'une matrice de covariance qui ne possède plus certaines propriétés simples qui permettait, pour un bruit non-corrélé, un calcul d'une complexité raisonnable. La première partie de l'article qui suit démontre qu'il est possible de récupérer ces propriétés et énonce les conditions nécessaires pour que l'espérance de la caractéristique d'Euler demeure une quantité aisément calculable.

En pratique cette méthodologie est appliquée à la détection de signaux multi-dimensionnels, tels que l'activité cérébrale en Imagerie par Résonance Magnétique fonctionnelle (IRMf). On travaille alors avec des données volumiques (3D), et l'ajout d'un paramètre continu d'échelle projette le problème en une inspection d'un espace 4-dimensionnel. En terme numérique l'implémentation est donc très lourde. Nous présentons dans la deuxième partie de l'article une optimisation numérique qui réduit grandement le coût calculatoire supplémentaire relatif à l'ajout d'une dimension d'échelle.

## Maxima of the Continuous Wavelet Transform of correlated data

PHILIPPE ST-JEAN, *Université de Montréal*

JEAN-MARC LINA, *Université de Montréal*

BERNARD GOULARD, *Université de Montréal*

### Abstract

The study of continuous wavelet transform (CWT) of signals through the behavior of its local maxima is a well-developed field that already led to useful applications in signal and image analysis. Meanwhile, the study of level upcrossings of random field is based on expected values of random quantities related to local maxima of the field. Generalizing the notion of level upcrossings from one dimension to higher dimensional spaces leads to the problem of evaluating the expected value of the Euler characteristic of excursion sets on those fields. This has been done by Adler (1981) and further extended by Siegmund and Worsley (1995) who proposed an extension of the method to test for signals not only of unknown location but of unknown scale as well, using an approach quite similar to the CWT. Even for an “irregular” field which does not respect Adler’s condition, a proper use of the CWT leads to a representation where the field becomes regular. We first show that this allows to apply Adler’s method to a more general family of irregular random fields as, for instance, a fractional Brownian motion. Then, we introduce a fast implementation based on the discrete dyadic wavelet decomposition that allows to perform the analysis with fewer operations than the method originally proposed by Siegmund and Worsley. Finally, we apply this method in order to detect a sharp but continuous signal in a background noise.

### 2.3. INTRODUCTION

The study of the behaviour of the extremas of the wavelet representation of signals or functions are at the cross-road of most of the theoretical and practical issues of the wavelet applications. At the origin, we have the important result due to Jaffard [31] that relates the pointwise Lipschitz regularity of functions with the decay of the high amplitude of the wavelet coefficients in the cone of influence of each singularity. Simultaneously, Mallat and Zhong [8] obtained an algorithm that iteratively approximates a 2d signal only from the wavelet modulus maxima throughout scales. To some extent, this encapsulation of the signal through wavelet extremas is the basic of the “zero-tree coding” algorithm advocated by Shapiro

[32] for compression. The notion of persistency of wavelet maxima across scale is also at the origin of the thermodynamical description of fractals [33] and motivated the more recent hierarchical Markovian modelisation of images [34]. Finally, the thresholding and shrinkage techniques used in the wavelet domain for estimating deterministic signal corrupted with noise also rely on the large amplitude coefficients [35].

The present work exposes an other point of view borrowed from the study of the topological properties associated with random fields in high-dimensional space. More precisely, the field under consideration is related to a wavelet representation of some signal  $f$  in the time-scale space  $(t, s)$

$$W_\psi[f](t, s) = e^{-\frac{s}{2}} \int \psi\left(\frac{t' - t}{e^s}\right) f(t') dt' \quad (26)$$

while the value of interest is the topological characterization (*i.e.* the Euler index) of excursion sets defined by the position of the wavelets modulus larger than some variable threshold. Indeed, it can be shown that the study of level upcrossings of random field, *i.e.* random wavelet coefficients, is based on expected values of random quantities that are related to local maxima of the field. The ultimate motivation of this analysis is the detection of unknown signals in a homogeneous gaussian random field. For instance, figures (2.8) and (2.9) both show the same realization of a fractional Brownian motion, although some deterministic signal was added in the latter; in this case the goal is to obtain a test for the presence of signals in such background noise.

In the first part of this paper, we present some general results on exceedence statistics for multi-dimensional fields that were obtained mostly by Adler [13]. A more complete and more specific presentation is found in Worsley [28], section 2.4.1 being essentially an overview of what is found in this later paper. Siegmund and Worsley [29] investigated the detection of a signal in a smooth zero-mean gaussian field by choosing an appropriate scale dependent smoothing kernel. This can be interpreted as a wavelet-type projection of an irregular one-dimensional

field (the white noise process) to a space of higher dimension where the new representation is smooth. We will explain in more details the possible advantages of this interpretation. First, we extend the previous results of Siegmund and Worsley[29] and we propose a method to compute the Euler characteristic for homogeneous gaussian fields with arbitrary spatial correlation function. An example for a long memory process is given by applying this method to fractional brownian motion. Since the representation of an irregular random field can only be complete if the scale parameter goes down to 0, the Euler characteristic for that representation will generally diverge. However, one can still learn about the field by studying the behavior of the Euler characteristic as the scale tends toward 0.

In Section 2.5, we consider the numerical implementation of the methods. It is shown that dyadic wavelet representation of the signal is sufficient for computing the Euler characteristics : the topological index can be well estimated on a lattice rather than on the full time-scale plane.

Finally, Section 2.6 is concerned with the detection of signal in a background of noise with known spatial correlations.

## 2.4. EXCEEDENCE STATISTICS OF MULTI-DIMENSIONAL FIELDS

### 2.4.1. Smooth homogeneous fields

Let us state the following problem for a random field in one dimension, which is trivially resolved. Let  $X(t)$ ,  $0 \leq t \leq T$  be a one-dimensional, homogeneous, almost surely continuous Gaussian field : how many times, on average, will this field exceed a given threshold  $b$ ?

This problem can also be stated as : what is the mean number of distinct intervals over which the field value is superior to  $b$ ? It appears that this number can be written as the expected value of the number of points for which the field takes the value  $X(t) = b$  while the first derivative with respect to  $t$  is positive. Any of these points effectively accounts for one interval. It is also possible that the

field exceeds  $b$  at  $t = 0$ , adding an extra interval, and this should be counted in as well. Hence in this simple case, the distribution of the bivariate field  $(X, dX/dt)$  is sufficient to obtain the desired expected value.

This idea can be generalized to multi-dimensional fields, although some difficulties will appear. Let us illustrate with an example in two dimensions. The field  $X(\mathbf{t})$  is now defined on a region  $C \in \mathbb{R}^2$ . A realization of the field can now be seen as a set of smooth mountains, valleys and lakes on the region  $C$  (this new field is also a.s. continuous). For a given large value of  $b$ , this realization of the field  $X$  might exceed  $b$  only on a small closed region of  $C$ , the summit of the highest mountain, so to say. In this case, we certainly hope that the generalization would lead to consider this peak as one bi-dimensional “interval”. The same thing should happen for smaller values of  $b$ , when other peaks will count in as well. Yet what should be done when  $b$  is small enough that a “lake” comes in, i.e. when the field exceeds  $b$  over an annular subset of  $C$ ? In order to answer that question, we first list appropriate properties for the quantity we are trying to define.

As in the one-dimensional case, we would like to be able to express this quantity in terms of the joint distribution of the field and its derivatives, and hopefully to reduce it to a count of single points (a “ponctual” representation) with specific qualities (as  $X = b, dX/dt > 0$  for the 1D case). Also, it is readily seen that the number of intervals exceeding  $b$  (in 1D) is invariant under any smooth transformation of the domain  $[0, T]$ , and we would like to bring this feature to the general case too, including rotations. We shall now present some definitions that we will need to introduce the quantity we are looking for.

Let  $X(\mathbf{t})$ ,  $t = (t_1, \dots, t_N) \in \mathbb{R}^N$  be a stationary random field and  $C$  a compact subset of  $\mathbb{R}^N$ . We define the *excursion set*  $A_b$  of  $X(t)$  as the set of points of  $C$  where  $X(t)$  exceeds the threshold  $b$ , i.e.

$$A_b = \{t \in C; X(t) \geq b\}. \tag{27}$$

$A_b$  is of course a subset of  $\mathbb{R}^N$ .

Let  $B$  be a compact subset of  $\mathbb{R}^N$  such that its intersection with any hyperplane of dimension  $k$ ,  $k = 1, \dots, N$  is connected. We will call  $B$  a *basic*. The union of a finite number of basics is called a *basic complex* (provided that all of their intersections are also basics, c.f. [13]).

Under certain regularity conditions that can be found in [13] (chap. 3) (essentially, we ask for  $X$  to be homogeneous, a.s. continuous and twice differentiable), the excursion set  $A_b$  will a.s. be a basic complex.

Finally, we define a quantity  $\chi(A)$  which we will refer to as the *Hadwiger characteristic* of any basic complex  $A$ , and which corresponds precisely to the multi-dimensional generalization of the number of exceedence intervals that we are looking for.  $\chi(A)$  is simply the number of disjoint intervals of  $A$  (individual points also counts as intervals) for dimension  $N = 1$ . In  $N > 1$  dimensions,  $\chi(A)$  is recursively defined as :

$$\chi(A) = \sum_u [\chi(A \cap \mathcal{E}_u) - \chi(A \cap \mathcal{E}_u^-)], \tag{28}$$

where  $\mathcal{E}_u = \{t \in C; t_N = u\}$  and

$$\chi(A \cap \mathcal{E}_u^-) = \lim_{v \uparrow u} \chi(A \cap \mathcal{E}_v). \tag{29}$$

Our goal was to obtain a characteristic that was both invariant under smooth transformations of coordinates (which is the case for this topological Hadwiger characteristic) and that could be obtained by counting individual points (a “punctual” representation). Looking into definition (28), one realizes that  $\chi(A)$  is obtained from a sum of characteristics of lower dimension objects, since  $A \cap \mathcal{E}_u$  which is the intersection of  $A$  with a lower dimensional object ( $\mathcal{E}_u$  is an hyperplan of dimension  $N - 1$ ) is also of lower dimension ( $N - 1$ ). Where do the terms in (28) come from, i.e. how can  $\chi(A \cap \mathcal{E}_u) - \chi(A \cap \mathcal{E}_u^-)$  be non-zero? If  $A$  is two-dimensional, then  $A \cap \mathcal{E}_u$  is one dimensional, i.e. it is a set of intervals (and

individual points). Now in order for  $A \cap \mathcal{E}_u$  to contain fewer or more intervals than  $A \cap \mathcal{E}_u^-$ , there has to be a frontier of  $A$  tangent to the hyperplane  $\mathcal{E}_u$  (in this case a line), so that a part of  $A$  that did not intersect with  $\mathcal{E}_u^-$  does with  $\mathcal{E}_u$ . It is worth noting that the Hadwiger characteristic is equal to the Euler characteristic for any basic complexes, and in this regard we will refer to the later in the rest of this paper.

Back to the random field  $X$  and its excursion set  $A_b$ , we are looking for points where the frontier of  $A_b$  is tangent to a hyperplane  $\mathcal{E}_u$ . This will happen at a point where :

$$X = b, \tag{30}$$

$$dX/dt_i = 0, i = 1..N - 1 \tag{31}$$

$$dX/dt_N > 0. \tag{32}$$

The  $X = b$  condition is obvious, since we want the point to be on the frontier of  $A_b$ . For  $i = 1..N - 1$ , the  $dX/dt_i = 0$  condition simply expresses the fact that  $A_b$  is tangent to  $\mathcal{E}_u$ . Finally,  $dX/dt_N$  must be greater than zero because we took the limit from below in (29). Note that the characteristic will (a.s.) change by no more than  $\pm 1$ , the sign of the change being controlled by the parity of the number of negative eigenvalues of the second derivative matrix of  $X$ . Note that this parity corresponds to the sign of the local curvature of  $X$ , a quantity that one expects to see in the evaluation of a topological characteristic. Hence we have the ponctual representation we are looking for, which now depends not only on the joint distribution of  $(X, dX/dt)$ , but on the larger  $(X, dX/dt, d^2X/dt^2)$ . It is also possible that points found on the frontier of the domain  $C$  contribute to the characteristic, similarly to the point  $t = 0$  in the one-dimensional case.

Expressions for the expected value of the Euler characteristic in the general case of  $N$ -dimensional random fields can be found in [30]. In the following, we have decided to focus on the special case of one-dimensional homogenous fields

with arbitrary spatial correlation, as the extension to higher dimension can be achieved following the procedure presented in [30].

We end this section with a comment on the homogeneity condition on  $X$ . This condition can be weakened to homogeneity in  $N - 1$  dimensions for fixed value of the last coordinate [30]. This subtlety may seem futile, but it will become very important as we will introduce a continuous wavelet transform random field in the next section, which often respects homogeneity in its first  $N - 1$  spatial dimension for fixed value of the extra scale dimension.

#### 2.4.2. Homogeneous fields with arbitrary spatial correlation

In Siegmund and Worsley [29], the problem of detecting a signal in a smooth zero-mean gaussian field is studied. The multiresolution field is generated through the convolution of a white-noise process with a variable width smoothing kernel (a Gaussian-shaped kernel, for instance). Results from Adler and Worsley are both based on the assumption that the random field is somehow “smooth” ([13]). However, even though the white noise field is far from being smooth, a proper choice of the smoothing kernel can assure that the CWT field will. This can be interpreted as a projection of an irregular one-dimensional field (the white noise process) to a space of higher dimension where the new representation is smooth.

In 1+1 dimensions, let us consider the following random field

$$X(t, s) = C(s) \int \psi[e^s(h - t)]dZ(h) \quad (33)$$

$$= C(s) \cdot e^{-s/2} \cdot W_\psi[Z'](-s, t) \quad (34)$$

where  $Z$  is a Gaussian white noise process ;  $s$  represents the log of the usual scale dimension (i.e.  $s = -\log a$ ) and  $C(s)$  is a normalization factor (that depends only on scale) which will be defined later. Obviously, one falls back onto the usual CWT definition by imposing that  $C(s) = e^{s/2}$ .  $\psi(t)$  is our smoothing kernel ; at this point, we do not impose vanishing moments for  $\psi(t)$ , but we demand that it is symmetric (or anti-symmetric).

The new field  $X(t, s)$  is no longer homogeneous, although it is with respect to  $t$  for fixed value of  $s$ , and that is all what we need here.  $Z(t)$  is our one-dimensional underlying process, with the following properties :

$$E|Z(t_1) - Z(t_2)|^2 = \rho(t_1, t_2) \quad \text{and} \quad E[Z(t)] = 0 \quad (35)$$

and we assume that  $\rho$  is a *symmetrical* function of  $\Delta t = |t_1 - t_2|$  only. Actually, it can take the more general form of a sum of functions of either  $\Delta t$  or only one of the two variables at a time, since the later will disappear using symmetry arguments. The main condition for the integral in (33) to converge is that  $\rho$  should be finite.

Let us start with the simple case of a one-dimensional white noise that is projected to the two-dimensional scale-space through a CWT in the following manner :

$$X(t, s) = W_\psi[Z'](-s, t), \quad (36)$$

where we have imposed  $C(s) = e^{s/2}$  in (34). In this special case, we have  $\rho(\Delta h) = \delta(\Delta h)$  where  $\delta$  is the Dirac measure.

We shall bear in mind that the goal is to find an expression for the expected value of the Euler characteristic  $\chi(A_b(X))$  of the excursion set  $A_b(X)$  of the field  $X(t, s)$  (above the threshold  $b$ ), since we chose this quantity to represent the generalization of the exceedence statistics for random fields of dimension greater than 1. We write  $X_s$ ,  $X_t$  and  $X_{tt}$  for the partial derivatives of the field with respect to the scale and the spatial dimensions. Consider  $\phi_1(x_t, x)$  as the joint pdf of  $(X_t, X)$ , and  $\phi(x)$  as the pdf of  $X$  (with  $\Phi(x)$  the cumulative distribution). Assume that  $T$  is the (finite) length of the interval over which the field is defined, and  $[s_1, s_2]$  is the range of interest in the scale dimension. Siegmund and Worsley [29] showed that the expected value of the Euler characteristic is given by the

sum of the following terms :

$$E[\chi_V(A_b)] = -T \int_{s_1}^{s_2} E(X_s^+ X_{tt} | X_t = 0, X = b) \phi_1(0, b) ds, \quad (37)$$

$$E[\chi_E(A_b)] = \int_{s_1}^{s_2} E(X_s^+ | X = b) \phi(b) ds, \quad (38)$$

$$E[\chi_B(A_b)] = TE[X_t^+] \phi(b) + [1 - \Phi(b)], \quad (39)$$

where the “V,E,B” subscripts stand for volumetric, edge and base contributions to the characteristic (see [29]). The sum of those 3 terms can be explicitly computed and written as

$$\begin{aligned} E[\chi(A_b)] &= T \cdot (e^{s_1} - e^{s_2})(\lambda \kappa)^{1/2} b \phi(b) / 2\pi \\ &+ (T/2) \cdot (e^{s_1} + e^{s_2})(\lambda)^{1/2} \phi(b) / (2\pi)^{1/2} \\ &+ (s_2 - s_1) \kappa^{1/2} \phi(b) / (2\pi)^{1/2} \\ &+ [1 - \Phi(b)], \end{aligned} \quad (40)$$

where the  $\lambda$  and  $\kappa$  parameters are defined as the following covariances :

$$\lambda = e^{2s} E[X, X_{tt}] = -e^{2s} E[X_t, X_t], \quad (41)$$

$$\kappa = E[X, X_s]. \quad (42)$$

Hence the variance matrix of the joint distribution of  $(X, X_s, X_t, X_{tt})$  is a critical object in this context.

We now come back to our task of generalizing this method to irregular one-dimensional fields. For instance, this idea can be exploited to apply the method to non-smooth gaussian fields with arbitrary spatial correlation function. Remember that the continuous wavelet transform is a complete (redundant) representation of a field in the limit of infinitely small scales. In this limit, however, there is no reason why the Euler characteristic should be finite. Still, the generalization gives the exact expected value of the characteristic for any scale interval that does not

contain  $a = 0$  (or  $s = -\log a = \infty$ ), as well as the behavior of the characteristic as the scale  $a$  tends toward 0.

Let us first compute the scale-spatial covariance of  $X(t_1, s_1)$  and  $X(t_2, s_2)$ , which will be useful later on; we have

$$\begin{aligned} & Cov[X(t_1, s_1), X(t_2, s_2)] \\ &= C(s_1)C(s_2) \int \overline{\psi[e^{s_1}(h_1 - t_1)]} \psi[e^{s_2}(h_2 - t_2)] dh_1 dh_2 Cov[W(h_1), W(h_2)] \end{aligned} \quad (43)$$

This can be rewritten as ( $\psi_{h_1}$  means derivative with respect to  $h_1$ ) :

$$\begin{aligned} & C(s_1)C(s_2) \int \int \rho(|h_1 - h_2|) \overline{\psi_{h_1}[e^{s_1}(h_1 - t_1)]} \psi_{h_2}[e^{s_2}(h_2 - t_2)] dh_1 dh_2 \\ &= C(s_1)C(s_2) \int \int \rho(|h_1 - (h_2 - \Delta t)|) \overline{\psi_{h_1}[e^{s_1}h_1]} \psi_{h_2}[e^{s_2}h_2] dh_1 dh_2 \\ &= C(s_1)C(s_2) \mathcal{F}^{-1} \left[ \hat{\rho}(\omega) \cdot (-i\omega \hat{\psi}(e^{-s_1}\omega)) \cdot i\omega \hat{\psi}(e^{-s_2}\omega) \right] (\Delta t) \end{aligned} \quad (44)$$

with

$$\mathcal{F}^{-1}[\hat{f}](t) = \frac{1}{2\pi} \int \hat{f}(\omega) \cdot e^{i\omega t} d\omega. \quad (45)$$

In order to get an expression for the Euler characteristic, we need to compute the variance matrix of the following joint distribution of  $X$  and some of its derivatives :  $Var(X, X_s, X_t, X_{tt})$ . This can be done by taking derivatives of the scale-spatial covariance that we have just computed, and then evaluating it at  $\Delta t = 0$  and  $s_1 = s_2 = s$ . For example, the  $Cov(X, X_t)$  term of the variance matrix is given by :

$$Cov[X, X_t] = \left\{ \frac{\partial}{\partial t_1} Cov[X(t_1, s_1), X(t_2, s_2)] \right\}_{\Delta t=0, s_1=s_2} \quad (46)$$

Expressed in the Fourier domain, this is written as the zero-frequency integral, i.e. :

$$Cov[X, X_t] = C^2(s) \int -(i\omega)^3 \hat{\rho}(\omega) |\hat{\psi}(e^{-s}\omega)|^2 d\omega \quad (47)$$

Since  $\rho$  is a real symmetrical function of its argument, as well as our filter  $\psi$ , the previous integral is null. Similar symmetry arguments imply that  $Cov[X_t, X_s]$  and  $Cov[X_t, X_{tt}]$  terms of the variance matrix are also null.

Therefore, the most general form of the matrix is the following :

This matrix is quite similar to the one obtained by Siegmund and Worlsey in the special case of  $W(t)$  being a white noise, although there is now an extra  $\zeta$  term representing the local covariance of the scale-space field  $X$  with its scale derivative. This is inconvenient since the derivation of the final expression for the Euler characteristic contains an intermediate step (37) where the following expected value must be obtained :  $E[X_s^+ X_{tt} | X = b]$ . Evaluation of this last term necessitates an inversion of the covariance matrix of  $(X, X_s, X_{tt})$ , and the extra  $\zeta$  term leads to an awkward form of the final characteristic formula.

We will now show that by imposing  $\alpha = Var[X] = 1$  through a proper choice of the normalisation factor  $C(s)$ , we automatically get  $\zeta = Cov[X, X_s] = 0$ , hence falling back on a manageable expression for the Euler characteristic. Under this condition, the derivation of other parameters of the matrix ( $\kappa, \lambda_1, \lambda_2$ ) will also be easily obtained. An explicit expression for the  $\alpha = Var[X]$  term is given by :

$$\begin{aligned} Var[X] &= \{Cov[X(t_1, s_1), X(t_2, s_2)]\}_{\Delta t=0, s_1=s_2} \\ &= C^2(s) \int \omega^2 \hat{\rho}(\omega) \hat{\psi}^2(e^{-s}\omega) d\omega \end{aligned} \tag{48}$$

Defining  $C(s)$  as :

$$C(s) = \left( \int \omega^2 \hat{\rho}(\omega) |\hat{\psi}(e^{-s}\omega)|^2 d\omega \right)^{-1/2}, \tag{49}$$

we get  $\alpha = 1$ . We are now ready to state the following result :

PROPOSITION 2.4.1. For  $X(t, s)$  a scale-space random field defined as in (33), the covariance of  $X(t, s)$  and  $X_s(t, s) = \partial X / \partial s$  is null everywhere if and only if

$$C(s) = K \left( \int \omega^2 \hat{\rho}(\omega) |\hat{\psi}(e^{-s}\omega)|^2 d\omega \right)^{-1/2}, \tag{50}$$

where  $K$  is some constant.

PROOF : Let us compute  $\zeta = Cov[X, X_s]$  with  $C(s)$  as above.  $\zeta$  can then be written as :

$$Cov[X, X_s] = \left\{ \frac{\partial}{\partial s_1} Cov[X(t_1, s_1), X(t_2, s_2)] \right\}_{\Delta t=0, s_1=s_2} \quad (51)$$

The r.h.s. of this last equation is simply half the derivative along  $s$  of the variance of  $X(t, s)$ , since

$$\begin{aligned} \frac{\partial}{\partial s_1} \{Cov[X(t_1, s_1), X(t_2, s_2)]\}_{\Delta t=0, s_1=s_2} &= \\ \frac{\partial}{\partial s_2} \{Cov[X(t_1, s_1), X(t_2, s_2)]\}_{\Delta t=0, s_1=s_2} & \end{aligned} \quad (52)$$

(from symmetry over  $s_1$  and  $s_2$ ) and

$$\begin{aligned} \frac{d}{ds} Cov[X(t, s), X(t, s)] &= \\ & \frac{\partial}{\partial s_1} \{Cov[X(t_1, s_1), X(t_2, s_2)]\}_{\Delta t=0, s_1=s_2=s} \\ & + \frac{\partial}{\partial s_2} \{Cov[X(t_1, s_1), X(t_2, s_2)]\}_{\Delta t=0, s_1=s_2=s}; \end{aligned} \quad (53)$$

hence  $Cov[X, X_s] = 0$  if and only if the variance of  $X(t, s)$  does not depend on  $s$ , which is ensured by (50).

This completes the proof. Therefore, one can say that the normalisation of the variance of the random scale-spatial field decorrelates the field from its scale derivative.

Let us now compute the  $\kappa$  term, which represents the variance of the  $X_s$  field. We define the following quantity :

$$q(s_1, s_2) = \int \omega^2 \hat{\rho}(\omega) \hat{\psi}(e^{-s_1}\omega) \hat{\psi}(e^{-s_2}\omega) d\omega \quad (54)$$

So that we have  $C(s) = (q(s, s))^{-1/2}$ . Then :

$$\kappa = Var[X_s] = \left\{ \frac{\partial^2}{\partial s_1 \partial s_2} Cov[X(t_1, s_1), X(t_2, s_2)] \right\}_{\Delta t=0, s_1=s_2}$$

$$\begin{aligned}
 &= \left\{ \frac{\partial}{\partial s_2} \left[ C(s_2) \cdot \left( \frac{dC(s_1)}{ds_1} q(s_1, s_2) + C(s_1) \frac{\partial q(s_1, s_2)}{\partial s_1} \right) \right] \right\}_{s_1=s_2} \\
 &= \left( \frac{dC(s)}{ds} \right)^2 \cdot q(s, s) \\
 &\quad + 2 \frac{dC(s)}{ds} \cdot C(s) \cdot q_1(s, s) + C^2(s) \cdot q_{11}(s, s) \tag{55}
 \end{aligned}$$

where  $q_1$  and  $q_{11}$  are defined as :

$$\begin{aligned}
 q_1(s, s) &= \left\{ \frac{\partial}{\partial s} q(s, s') \right\}_{s=s'} \\
 q_{11}(s, s) &= \left\{ \frac{\partial^2}{\partial s \partial s'} q(s, s') \right\}_{s=s'} \tag{56}
 \end{aligned}$$

But from (13), the expression for  $\kappa$  simplifies to :

$$\begin{aligned}
 \kappa &= \left( \frac{-q_1(s, s)}{q^{3/2}(s, s)} \right)^2 \cdot q(s, s) + 2 \left( \frac{-q_1(s, s)}{q^{3/2}(s, s)} \right) \cdot \frac{q(s, s)}{(q(s, s))^{1/2}} + \frac{q_{11}(s, s)}{q(s, s)} \\
 &= \frac{q_1^2}{q^2} - 2 \frac{q_1^2}{q^2} + \frac{q_{11}}{q} \\
 &= \frac{q_{11}}{q} - \frac{q_1^2}{q^2} \tag{57}
 \end{aligned}$$

Similar calculation leads to the following form for  $\lambda_1$  and  $\lambda_2$  :

$$\begin{aligned}
 \lambda_1 &= \frac{q_1}{q} \\
 \lambda_2 &= \frac{q_{11}}{q} - \frac{q_1^2}{q^2} \tag{58}
 \end{aligned}$$

This result extends Siegmund and Worsley's result [29] for a white noise process, with the same adequate choice of normalisation factor. Coming back to equation (37), the only difference between the general case and the white noise process is the possibly non-trivial dependence of the covariances upon the scale

variable  $s$ , obtained from the derivation above. In any event, the integrals can always be computed numerically.

Furthermore, the variance matrix of  $(X, X_t, X_s, X_{tt})$  is all that we need to extend this method to other fields such as  $\chi^2$ , for which the expected value of the Euler characteristic has already been obtained [30]. We will now turn to an example that illustrates our method.

### 2.4.3. Fractional brownian motion

We are now looking at the following field :

$$X(t, s) = C(s) \int \psi[e^s(h - t)]dB_H(h) \tag{59}$$

The correlation function for fractional brownian motion is given by :

$$EB_H(t_1)B_H(t_2) = \frac{1}{2} (|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H}) \tag{60}$$

where  $H$  lies between 0 and 1. One might object that this process is not stationary, which is true. However, the two extra terms which depend solely on  $t_1$  and  $t_2$  disappear in all the covariances by symmetry arguments. Typical realizations of fBm for different values of the parameter  $H$  are shown on figure 2.10. Note the increasing regularity of the signal with  $H$  ; for  $H < 1/2$ , increments are anti-correlated, while they are positively correlated for  $H > 1/2$ . Assuming a Gaussian kernel for  $\psi$ , one gets that the variance of  $X$  is equal to :

$$Var[X(t, s)] = C^2(s) \cdot e^{-2sH} H2^{2H}\Gamma(H + 1/2) \tag{61}$$

So that  $C(s)$  is set to :

$$C(s) = \frac{e^{sH}}{\sqrt{H2^{2H}\Gamma(H + 1/2)}} \tag{62}$$

Finally, the matrix elements are computed :

$$\kappa(H) = 1 - H$$

$$\lambda_1(H) = \lambda_2(H) = 1 - H$$

$$\epsilon(H) = (H - 2)(H - 1) \quad (63)$$

Some precaution must be taken regarding the regularity conditions on  $X(t, s)$ . For a Gaussian field, the finiteness of all third derivatives of  $X$  is a sufficient condition to ensure that realizations of  $X$  have almost surely continuous derivatives up to second order ([29]), which is required for the previous results to be valid. In the white noise case studied by Siegmund and Worsley, this condition is equivalent to demand that the integral of the product of any pair of third derivatives of  $\psi(h)$  times  $h^6$  is finite. For instance, this is verified when  $\psi(h)$  is a Gaussian kernel.

In our case, the regularity condition is slightly different : the double integral of any pair of fourth derivative of  $\psi(h_1)$  and  $\psi(h_2)$  times  $h_1^3 \cdot h_2^3 \cdot |h_1 - h_2|^{2H}$  must be finite. Again, this is verified when  $\psi(h)$  is a Gaussian kernel, and  $0 < H < 1$ .

It must be noted that for the case  $H = 1/2$ , we fall back on Siegmund and Worsley's result for a white noise process. This is easily explained by the fact that our field  $X(t, s)$  is defined as a translated integral of a smoothing kernel over a stochastic measure, which is the same as convolving the "derivative" of the initial random field with a smoothing kernel. In the case of a classical brownian motion ( $H = 1/2$ ), this generalized derivative is simply the white noise process itself. This results from the fact that taking twice the derivative of  $\rho$  in this case yields a delta function. Furthermore, the duality of the convolution process also allows to see this calculation as the convolution of the derivative of the kernel with the random field itself. We have already used this fact in the derivation of the general method in the previous section. The case for which we do not have  $H = 1/2$  is of more interest, since in this case the increments of the process are not independent (or the derivative is not a white noise). We produced numerical simulations of fBm for some values of  $H$ , computed their Euler characteristic for each realization and average them together (fig.2.11). They completely agree with the theoretical result that we derived above.

Our result also holds for fBm with  $H > 1$  (i.e. processes that a.s. possess derivatives) although in this case it is necessary for the kernel to possess sufficient null moments, hence to be a wavelet. If this condition is not respected, the  $q$ -integrals in the derivation of the formula might diverge.

## 2.5. DYADIC WAVELET REPRESENTATION

We now turn to some numerical aspects that we have to consider in the implementation of the method we introduced in the previous section. It is known that the continuous wavelet transform is a dramatically redundant representation of a field, and that there exists a minimal representation in the same scale-space field that contains all the information, i.e. a set of points from which it is possible to reconstruct the whole scale-space field through a reproducing kernel. In the context of evaluating the Euler characteristic of such a field, one might wonder if it is possible to exploit this minimal representation in order to reduce computational work. The answer to that question is yes, and we will show how this can be done in this section.

### 2.5.1. Numerical evaluation of the Euler characteristic for a field defined on a lattice

For a 1-dimensional signal, the extra scale dimension yields a 2-dimensional scale-space field on which we would like to evaluate the Euler characteristic of excursion set above a threshold. Adler [13] gives an approximation to the characteristic when the field is “sampled”, i.e. when it is defined on a lattice. Even in the case of a so-called continuous wavelet transform, the numerical implementation can only result in a finite resolution representation of the field. Adler’s approximation consists essentially in counting the number of points, lines and faces of the lattice whose values are above the given threshold. A line is considered “above” when both of its extremities are, and a face when the four points (in the case of a rectangular lattice) satisfy the condition. The number of points and

faces are added together, and the number of lines is subtracted from the sum. In higher dimensional spaces the generalization is similar, except that cubes, hypercubes etc... are counted as well. In the limit of an infinitely small lattice cell, the approximate characteristic tends toward the exact one.

Of course, the implementation is not done by testing on all points, lines and faces at every threshold. As we scan threshold values, an element (point, line or face) will contribute to a change in the Euler characteristic only when the threshold hits the element value. It is easy to see that there are twice as many lines as there are points or faces, if we neglect for now the edges of the field. A "site" will be defined as an ensemble of one point, two lines and a face, which can be chosen for instance as the left superior corner for the point, and the left and superior edges for the lines. Now for each site, the contribution of the site to the characteristic will be zero for all threshold values below the minimum value of all elements in the site (which is always the face, since it depends on the three others). Hence one can start at the face value, and then test for the two lines and the corner until the maximum of these three elements is reached, above which the contribution falls back to zero. In doing so, the right and inferior edges of the field are not taken into account, so that two additional loops are needed to include the contribution of the points and lines that form these edges.

### 2.5.2. How does the scale-space field look like ?

It is obvious that the scale-space field has different regularity properties along the spatial dimensions depending on the scale at which one is looking. Assuming that the wavelet used is  $C^\infty$ , one might find (on average, at a given scale) twice as many level crossings as are expected on a scale twice larger, for an irregular one-dimensional spatial field. This means that the probability of a strong non-linear behavior between two points in the lattice is always smaller for points representing larger scales in the scale-space field, as the resolution of the lattice tends to infinity (spacing between the sites tends toward zero). This suggests that

the resolution of the lattice could be adapted so that it would be looser for large scales than for small ones. This corresponds precisely to one aspect of the minimal representation theorem of dyadic wavelets, i.e. that the number of coefficients necessary to describe a spatial field as a scale-space field at a given scale is twice the number of coefficients needed at a scale twice larger, for orthonormal families of wavelets.

Therefore we propose to replace the canonical rectangular grid by the scale-adapted grid shown on figure (2.12). For a one dimensional signal sampled over  $N$  points (we assume that  $N$  is a power of 2 for simplicity), the scale space field sampled on a rectangular grid at every power of two along the scale dimension and for  $N$  points along the spatial one (at each scale) yields a total number of  $N(\log_2(N) + 1)$  points over which wavelet coefficients must be calculated, and  $N(\log_2(N))$  sites over which the Euler characteristic is evaluated. Obtaining the Euler characteristic on a dyadic grid constructed with only  $2N$  sites will be shorter by a factor of  $(\log_2(N))/2$ . In  $D$  dimensions, this factor becomes  $\frac{2^D-1}{2}(\log_2(N))$ .

Yet in practice it is preferable to impose some slight spatial-wise oversampling of the dyadic tiling, say by a factor of four (along the dimension  $t$  here), or else to oversample “voices” between “octaves” in the scale dimension (in practice by a factor of five to ten) in order to assure that no maxima of the field  $X(t, s)$  would be completely missed. This still results in a good gain in speed for signals of reasonable length.

## 2.6. RESULTS

The following question naturally arises at this point : Is it possible to test for the presence of a signal in a background of gaussian noise with spatial correlation function  $\rho$ ? We show some examples in the particular case of a fBm noise.

Our null hypothesis is that no signal is present ; we can then determine to which value of  $b$  corresponds a given level of confidence for the expected value of  $\chi(A_b)$ , say  $\{b|\chi(A_b) = 0.05\}$ . If the maximum value of the CWT of a signal

exceeds that value, then we can conclude that a signal is present (with 95% confidence). Hence the power of such a test is simply given by :

$$P\{X_{\max} > b\} \tag{64}$$

where we have

$$X(t, s) = C(s) \int \psi[e^s(h - t)]dZ(h) \tag{65}$$

and

$$dZ(t) = C(s_0)\xi f[e^{s_0}(t - t_0)]dt + dB_H(t). \tag{66}$$

Hence our signal is of the form  $f(t)$ , centered at  $t = t_0$ , and contracted by a factor  $e^{s_0}$ . The extra parameter  $\xi$  controls the amplitude of the signal. In the special case of  $f(t) = \psi(t)$  (i.e. when the wavelet matches the form of the signal to be detected), Siegmund and Worsley [29] showed that the power of the test can be approximated by :

$$P\{X_{\max} > b\} \simeq 1 - \Phi(b - \xi) + \phi(b - \xi)/\xi. \tag{67}$$

We decided to choose  $\xi$  such that the power of the test would be  $P\{X_{\max} > b\} = 0.5$ . This means that for signals of that strength, the test will identify the presence of the signal one time out of two on average. Figures 2.8, 2.9 and 2.13 show a typical example for which the test is successful in detecting such signal. Even though the signal is strong enough to significantly alter the value of the noise (compare figures 2.8 and 2.9 around  $x = 320$ ), it is not straightforward to declare the presence of the signal by a simple inspection of figure 2.9.

We can also test the validity of the proposed dyadic tiling approximation by investigating the global maximum statistics of the scale-space field. For 500 realizations of a classical brownian motion ( $h = 1/2$ ) with 256 points, the average underestimating error of evaluating the global maxima from a scale-adapted tiling (instead of a regular tiling) was of the order of  $3.52 \times 10^{-3}$ , whereas the average

global maxima for these realizations was 2.51. Bearing in mind that the scale-adapted tiling is 2.25 times faster in this case, the underestimation is a little price to pay. In figure 2.11, the average Euler characteristic has been computed with both regular and dyadic tiling of the CWT space; difference between the two methods is almost undetectable. The 2.25 factor may seem like not such a great gain, but this factor grows like  $\log_2(N)$  as the length of the sample does.

## 2.7. DISCUSSION

We come back to the simulation shown on figure (2.8) and figure (2.9); these two represent the same realization of an fBm, except for a small signal added by hand on the second (around  $x = 320$ ). The result of the procedure is shown on figure(2.13). It is to be noted that this algorithm not only detects the presence of a signal, but also gives a good approximation of its position in scale-space. This means that both spatial and scale information is obtained about the signal, since it is the global maximum of a given realization of the field  $X(t, s)$  which will get the Euler characteristic to exceed its 95% confidence threshold. Hence by keeping track of the global maximum of the field we now have a true detection algorithm for signals in any noisy (stationary) gaussian background.

The preceding arguments may lead to the false conclusion that only the global maximum of the field is of interest. From figure 2.11, one can appreciate the fact that for a relatively small number of realizations of a given gaussian process the expected value of the Euler characteristic for *all* values of  $b$  can help determine whether the model is adequate or not.

Indeed, the method we proposed strongly depends on a priori knowledge of the spatial correlation function  $\rho$ . Yet the method may seem trivial if one considers the following argument : if the spatial correlation function  $\rho$  is known a priori, why not simply whiten the noise in the first place and then apply Siegmund and Worsley's result ?

As a matter of fact, the whitening process can be seen as a modification of a simple wavelet (for which the white-noise calculation is easy to carry out, e.g. gaussian) obtained by dividing its Fourier transform by  $\sqrt{\hat{\rho}}$ . However there is no guarantee that this new wavelet will be stable, since  $\sqrt{\hat{\rho}}$  may be close or equal to zero at some point in the Fourier domain. Even if the wavelet is well-defined, its form will probably not match the one of the signal that is to be detected ; we know from the Matched Filter Theorem [36] that this leads to non-optimal signal detection.

It is worth noting that our method can also be applied to some non-stationary fields, provided that their CWT is. This is the case for fractional Brownian motion, and for all types of noise which have stationary increments.

## 2.8. CONCLUSION

We have proposed a geometrical approach to study exceedence statistics of irregular random fields based on the one developed by Adler and Worsley for smooth fields. We have shown that by projecting such a field to a higher dimensional space given by the CWT, it is possible to obtain a smooth representation for which Adler's method can be applied. We have also introduced the implementation of a fast algorithm to perform that analysis.

The obvious next step is to generalize the method to N-dimensional irregular Gaussian fields, but this task is quite straightforward since all the work of Adler and Worsley upon which this method is based was developed for any number of dimensions.

Up to this point we have also limited our work to homogeneous (or stationary) gaussian random fields only. For this particular family of random fields, we know that the spatial correlation function  $\rho(\Delta x)$  (the so-called correlogram) fully parametrizes the field, assuming that the mean is null at any point  $x$  ( $E[f(x)] = 0 \forall x$ ). This is obviously not the case for arbitrary random fields.

Our goal was to generalize Adler's method to irregular random fields, using a smooth representation (CWT in our case) to do so. Moving further, one can ask whether it is possible to work with fields which are well-defined in this smooth representation although hard to track when projected back to the standard space. Also, one can look into other smooth representations (similar to the CWT) of irregular fields. In particular, some random models in the wavelet domain (such as cascades model introduced in the study of fully-developed turbulence) with non-gaussian statistics leads to so-called multi-fractal signals [15]. These are good examples of signals for which the function  $\rho(\Delta x)$  is insufficient for a complete description of the field. Multi-fractals are partially characterized by a singularity spectrum, an object that shares many similarities with the Euler characteristic. The exact relationship between these two objects as well as a possible test for exceedence statistics of multi-fractals is the subject of ongoing work.

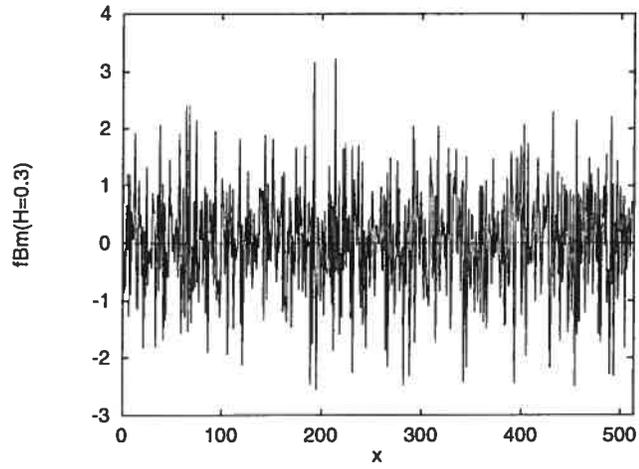


FIG. 2.8. A realization of a fBm noise for  $H = 0.3$ .

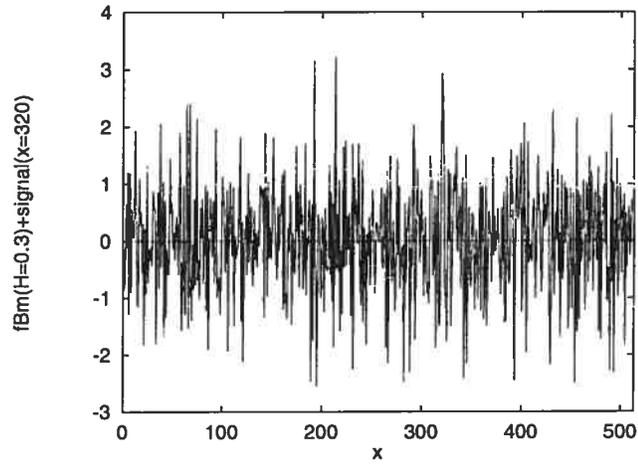


FIG. 2.9. The same realization of a fBm noise for  $H = 0.3$ , with a signal of the form  $xe^{-x^2}$  added at  $x = 320$ .

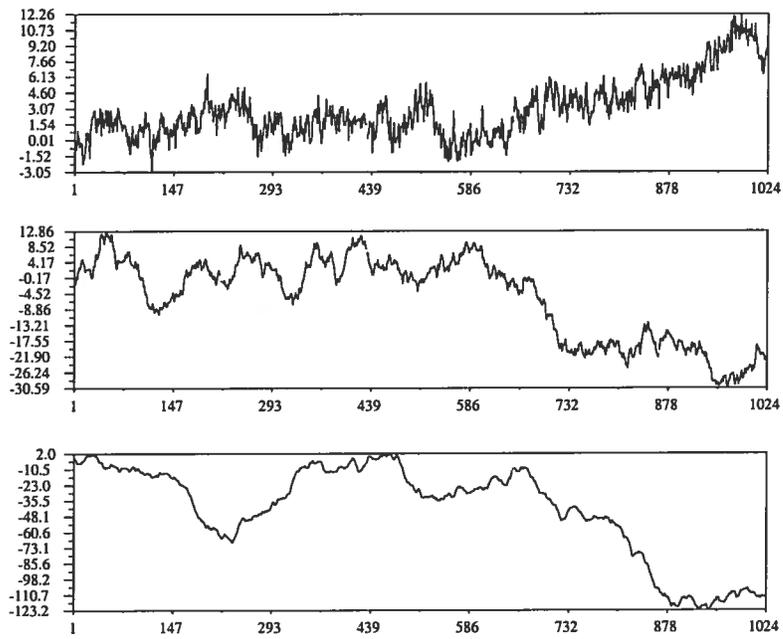


FIG. 2.10. Three realizations of fBm processes with  $H = 0.2$  (top),  $H = 0.5$  (classical Brownian motion, middle) and  $H = 0.8$  (bottom).

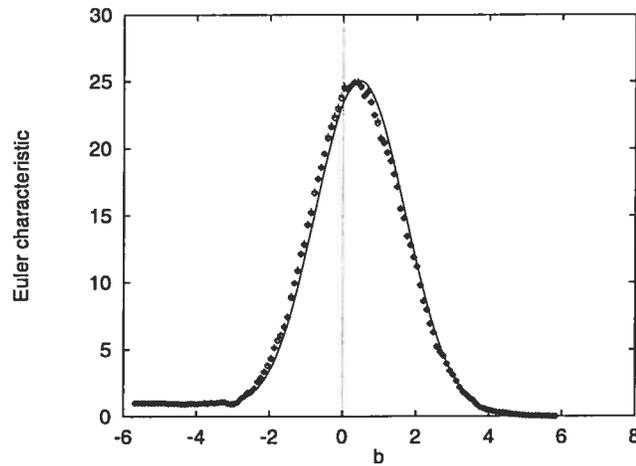


FIG. 2.11. Euler characteristic of a fBm with  $h = 0.7$ . The solid curve was computed from equation (37), while the circle points were obtained from an average of 50 fBm continuous wavelet transform simulations (with scale-normalized variance) from which the Euler characteristic was estimated. The cross have been computed on the discrete lattice described in Section 2.1. The simulations are in very good agreement with the theoretical expected value. For large values of  $b$ , the characteristic goes down towards zero as expected, since no regions of  $X(t, s)$  for which the field should take very high values. Similarly, for large negative values of  $b$ , the whole field should be above that value over all its domain, which has an Euler characteristic of one.

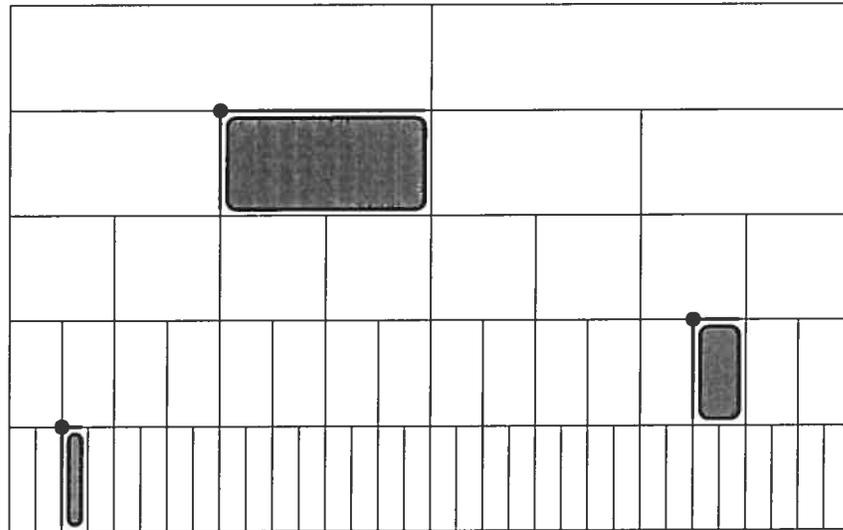


FIG. 2.12. Scale-adapted tiling with examples of individual sites representing one point, two lines and a face.

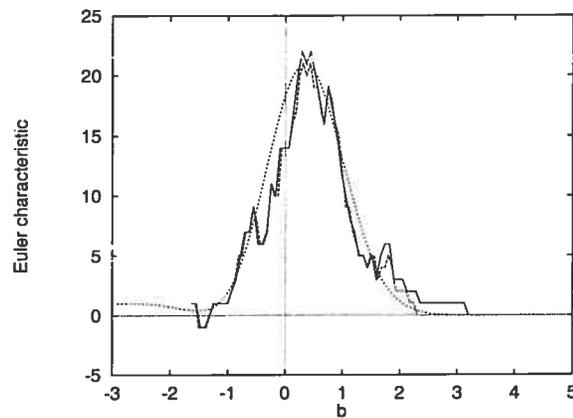


FIG. 2.13. Evaluation of the Euler characteristic for the noise only and for noise+signal. The theoretical value is plotted as well, with the 95% confidence threshold at 2.74, i.e.  $\chi(2.74) = 0.05$ . The noise+signal support goes beyond that value so that the signal is detected by the algorithm.

## Chapitre 3

---

### MODÈLES GRAPHIQUES MULTI-ÉCHELLE STATIONNAIRES

### 3.1. REVUE DE LITTÉRATURE : PHÉNOMÈNE D'INTERMITTENCE EN TURBULENCE, APPROCHE MULTI-FRACTALE ET MODÈLE EN CASCADES

#### 3.1.1. Équation de Navier-Stokes et symétries

Le comportement d'un flot turbulent (incompressible) est décrit par la vitesse  $v$  du fluide soumise à l'équation de Navier-Stokes (en trois dimensions) :

$$\partial_t \mathbf{v} + \mathbf{v} \cdot \nabla \mathbf{v} = -\nabla p + \nu \nabla^2 \mathbf{v}, \quad (68)$$

$$\nabla \cdot \mathbf{v} = 0. \quad (69)$$

Cette équation est non-linéaire. À ce jour, on n'est pas parvenu à confirmer ou infirmer l'existence d'une solution à cette équation sauf pour certaines conditions initiales et des conditions aux frontières particulières. On peut cependant décrire certaines caractéristiques du comportement des flots turbulents en étudiant les symétries du système. Nous suivons de près la présentation de Uriel Frisch dans cette section [37].

Considérons par exemple un flot uniforme dans la direction  $\hat{x}$  incidant sur un obstacle cylindrique (figure 3.1). On observe les symétries suivantes : invariance sous translation spatiale en  $\hat{z}$ , translation temporelle, parité en  $\hat{y}$ . On observe aussi une symétrie (imparfaite) sur une parité en  $\hat{x}$  de la forme :

$$(x, y, z) \rightarrow (-x, y, z), \quad (u, v, w) \rightarrow (u, -v, -w), \quad (70)$$

où  $(u, v, w)$  représentent les composantes de la vitesse  $v$ . Les trois premières symétries sont présentes dans l'équation de Navier-Stokes. La quatrième n'est vérifiée qu'en l'absence du terme non-linéaire. On observe effectivement une légère brisure de cette symétrie, légère parce que la vitesse caractéristique de ce flot est petite devant le terme de viscosité  $\nu$ . Ceci nous mène à définir un paramètre de contrôle

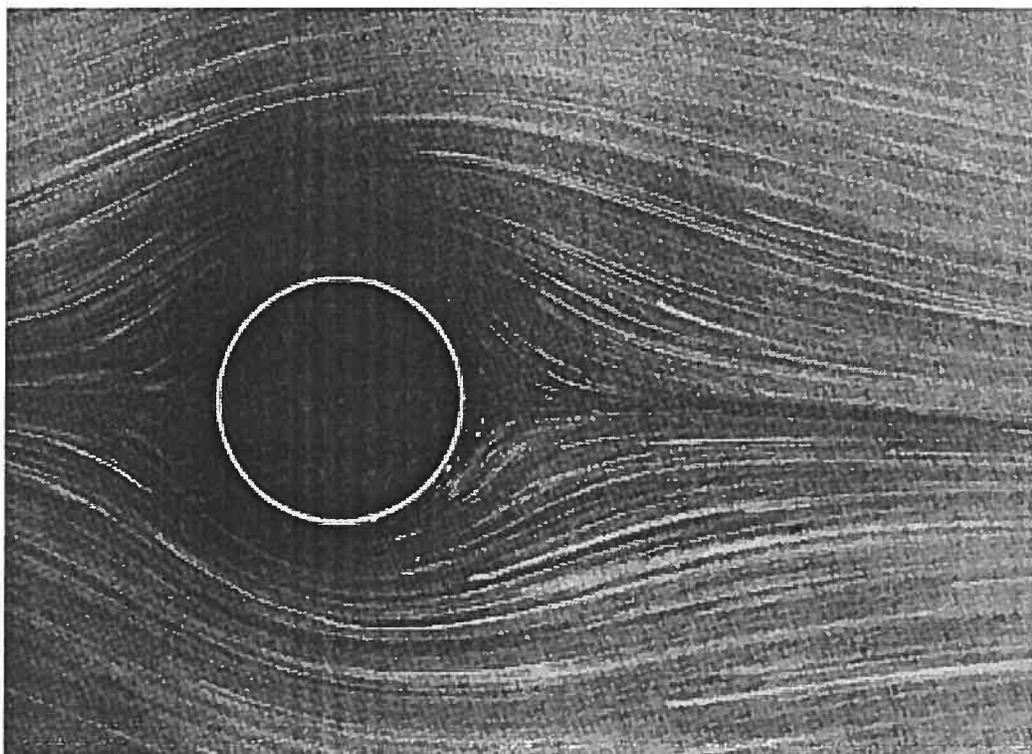


FIG. 3.1. Flot laminaire à bas Reynolds,  $R = 1.54$ . Le flot moyen va de gauche à droite. Photo S. Taneda.

du système, le nombre de Reynolds  $R$  :

$$R = \frac{LV}{\nu}, \quad (71)$$

où  $L$  et  $V$  sont les échelle et vitesse caractéristiques du système, et  $\nu$  la viscosité cinématique du fluide. En augmentant la vitesse ou en réduisant la viscosité (i.e. en augmentant  $R$ ), la brisure de la quatrième symétrie deviendra de plus en plus manifeste (figure 3.2). En augmentant progressivement  $R$  ( $R \sim 50$ ), les autres symétries seront également brisées. Dans ce cas on parle de brisure *spontanée*, puisque la symétrie est en accord à la fois avec les équations du mouvement et avec les conditions limite. Un phénomène intéressant se produit lorsque l'on augmente encore le nombre de Reynolds jusqu'à des valeurs largement plus élevées,  $R \sim 1000$ . On obtient alors un comportement aux allures chaotiques (figure 3.3). En

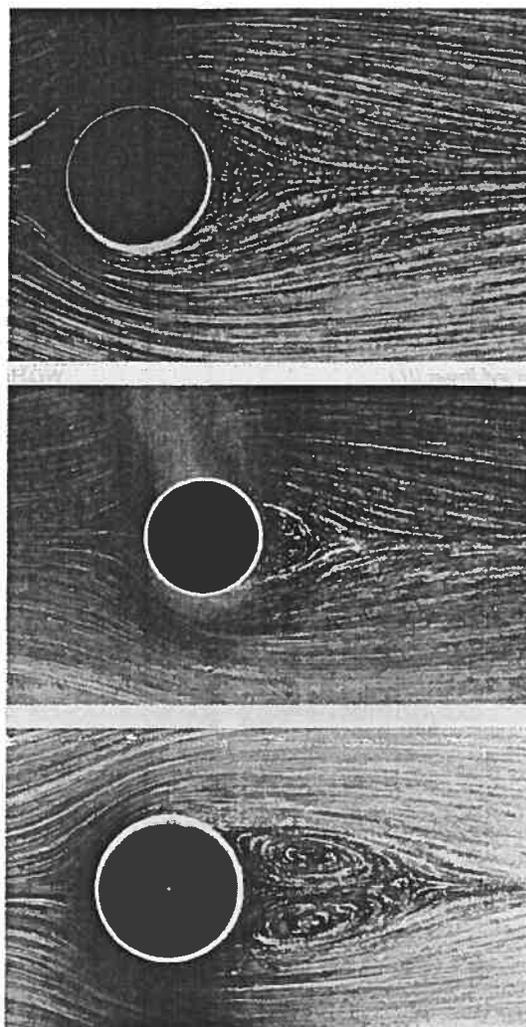


FIG. 3.2. Flot à des valeurs du nombre de Reynolds progressivement plus élevées ( $R = 9.6, 13.1, 26$  de haut en bas). Photo S. Taneda.

particulier si on utilise un grillage plutôt qu'un seul cylindre comme obstacle, on observe à une certaine distance du grillage le phénomène de *turbulence isotrope homogène*. Cette isotropie et cette homogénéité doivent être entendues dans un sens statistique. Vu comme un champ stochastique, le champ de vitesse  $v$  est considéré isotrope et stationnaire *au sens statistique* si sa distribution l'est. Il en va de même pour toute autre quantité physique du système turbulent.

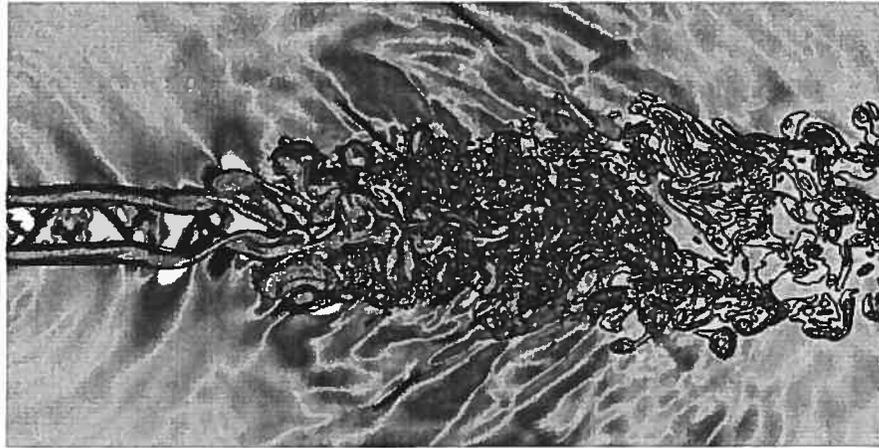


FIG. 3.3. Flot turbulent.

Ce rétablissement statistique des symétries du système, pour des flots à haut Reynolds, constitue la base du modèle de Kolmogorov [1].

### 3.1.2. Le modèle de Kolmogorov (K41)

Il n'est pas surprenant de voir brisées les symétries de l'équation de Navier-Stokes sous certaines conditions expérimentales, puisque les conditions limite, elles, ne respectent jamais parfaitement ces symétries. Dans l'exemple présenté dans la section précédente, la présence du cylindre brise toutes les symétries sous translation spatiale. Frisch exprime cette idée ainsi : “ *[...] the mechanisms responsible for the generation of the turbulent flow are usually not consistent with most of the possible symmetries[...]* ”

Cependant, pour de grandes valeurs du nombre de Reynolds, on observe que les symétries de Navier-Stokes sont rétablies dans un sens statistique. C'est cette observation qui a conduit Kolmogorov à formuler un certain nombre d'hypothèses concernant les flots turbulents.

Plutôt que d'énoncer ces hypothèses telles qu'elles ont été présentées dans les travaux originaux de Kolmogorov, nous reprenons la formulation de Frisch qui

“corrige” certaines lacunes du modèle original, mais qui mène aux mêmes conclusions ([37], chap. 6). La première hypothèse exprime l’idée de rétablissement des symétries :

**H1** : *In the limit of infinite Reynolds numbers, all the possible symmetries of the Navier-Stokes equation, usually broken by the mechanisms producing the turbulent flow, are restored in a statistical sense at small scales and away from boundaries.*

On entend par “petites échelles” les phénomènes se produisant à des échelles beaucoup plus petites que l’échelle caractéristique (nommée échelle “intégrale”) à laquelle est produite la turbulence, i.e. le diamètre du cylindre dans l’exemple ci-haut.

Formellement, on définit le rétablissement statistique des symétries en terme d’homogénéité des incréments de vitesse  $\delta\mathbf{v}_l$  :

$$\delta\mathbf{v}(\mathbf{r}, l) \equiv \mathbf{v}(\mathbf{r} + l) - \mathbf{v}(\mathbf{r}). \quad (72)$$

La deuxième hypothèse concerne spécifiquement la symétrie d’échelle du système :

**H2** : *Under the same assumptions as in H1, the turbulent flow is self-similar at small scales, i.e. it possesses a unique scaling exponent  $h$ .*

Cette deuxième hypothèse implique donc que

$$\delta\mathbf{v}(\mathbf{r}, \lambda l) \stackrel{\text{loi}}{=} \lambda^h \delta\mathbf{v}(\mathbf{r}, l), \quad \lambda > 0, \quad (73)$$

tant que  $l$  et  $\lambda l$  demeurent petits devant l’échelle intégrale.

On s’intéresse aussi au comportement de la dissipation d’énergie par unité de volume, que l’on note  $\varepsilon$ . La troisième hypothèse s’énonce en terme de cette quantité :

**H3** : *Under the same assumptions as in H1, the turbulent flow has a finite nonvanishing mean rate of dissipation  $\varepsilon$  per unit mass.*

Cette dernière hypothèse exprime le fait que la dissipation d’énergie locale demeure finie même lorsque la viscosité  $\nu$  tend vers zéro.

Dans ses travaux originels, Kolmogorov avait plutôt avancé que les propriétés statistiques à petite échelle étaient uniquement déterminées par l'échelle  $l$  et la dissipation  $\varepsilon$ . Partant de cette hypothèse, un simple argument d'analyse dimensionnelle permet alors d'obtenir l'unique valeur de  $h$  possible. En effet, la valeur moyenne du carré des incréments de vitesse,  $\langle(\delta\mathbf{v}(l))^2\rangle$ <sup>1</sup> est de dimension  $[L]^2[T]^{-2}$  alors que la dissipation  $\varepsilon$  s'exprime en  $[L]^2[T]^{-3}$ . On a donc :

$$\langle(\delta\mathbf{v}(l))^2\rangle = C\varepsilon^{2/3}l^{2/3}, \quad (74)$$

où  $C$  est une constante universelle sans dimension. Puisque le moment d'ordre 2 doit se comporter comme  $l^{2h}$ , on obtient pour  $h$  la valeur  $1/3$ . Mais cette dernière hypothèse a été invalidée par Landau en 1944 [38].

Cependant, partant des hypothèses H1, H2 et H3, on obtient de manière exacte la relation suivante pour la composante de la vitesse parallèle à la vitesse moyenne du flot (voir [37], sec. 6.2) :

$$\langle(\delta v_{\parallel}(l))^3\rangle = -\frac{4}{5}\varepsilon l, \quad (75)$$

appelée *loi des quatre cinquièmes*. Cette loi lie le moment d'ordre 3 de la distribution des incréments de vitesse à la dissipation d'énergie ; on en conclut également que  $h = 1/3$ .

En reprenant l'argument d'analyse dimensionnelle, on obtient aussi le comportement des moments d'ordre supérieurs, à travers la fonction de structure  $S_p(l)$  définie par

$$S_p(l) \equiv \langle(\delta v_{\parallel}(l))^p\rangle. \quad (76)$$

Pour respecter H2, on doit avoir :

$$S_p(l) = C_p\varepsilon^{p/3}l^{p/3}, \quad (77)$$

---

<sup>1</sup>On utilise  $\langle \cdot \rangle$  pour l'espérance d'une quantité dans la suite de ce chapitre afin d'alléger la notation et afin de se conformer à l'usage dans le domaine.

où  $C_p$  est une constante sans dimension indépendante de  $l$ , mais qui peut dépendre de  $p$ , et qui peut également dépendre des paramètres expérimentaux.  $C_3$ , cependant, doit être égal à  $-4/5$  de manière universelle, i.e. indépendamment de toute valeur des paramètres expérimentaux. C'est justement la non-universalité des  $C_p$  (pour  $p \neq 3$ ) qui est à la base de la correction apportée par Landau.

### 3.1.3. Intermittence

La théorie originale de Kolmogorov prévoit donc un comportement dans les échelles de la fonction de structure en loi de puissance :

$$S_p(l) \sim l^{p/3}. \quad (78)$$

C'est effectivement ce que l'on a longtemps observé expérimentalement, du moins pour de petites valeurs de  $p$ , i.e.  $p < 4$ . Il est difficile de mesurer ce comportement pour de grandes valeurs de  $p$ , puisque l'on cherche alors à mesurer des moments d'ordre élevé de la distribution de vitesse, influencés par les événements rares qui contrôlent la décroissance de la fonction de distribution. Il faut donc en pratique obtenir une statistique avec de grands échantillons, et s'assurer que cette statistique ne soit pas biaisée par une erreur systématique. Ce n'est que récemment que l'on est parvenu à mesurer avec une précision satisfaisante le comportement en échelle de  $S_p(l)$  pour des valeurs de  $p$  allant jusqu'à  $p = 30$  [39]. Les résultats de cette expérience font ressortir des corrections significatives au comportement prévu par Kolmogorov. On observe toujours un comportement en loi de puissance de la fonction de structure  $S_p(l)$  selon  $l$

$$S_p(l) \sim l^{\zeta_p}. \quad (79)$$

D'après ces résultats la dépendance en  $p$  de l'exposant  $\zeta_p$  ne serait pas linéaire, contredisant ainsi la prédiction de K41.

Parmi les modèles plausibles expliquant ce phénomène, le modèle multi-fractal introduit surtout par Frisch et Parisi propose une reformulation des hypothèses

H1, H2 et H3 ([37], chap. 8). Avant de passer à cette reformulation, nous présentons une interprétation géométrique de ce modèle.

Cette interprétation fait appel à une représentation hiérarchique du système turbulent. On considère que le mouvement du fluide turbulent peut être vu comme un ensemble de tourbillons de différentes tailles, où toutes les tailles sont présentes. On considère également que les plus petits tourbillons se “nourrissent” de l’énergie des grands, i.e. qu’il existe une cascade de transfert d’énergie allant des grandes échelles vers les petites ; cela correspond exactement à l’arrangement expérimental typique, où l’énergie est introduite aux grandes échelles (e.g. par un flot uniforme frappant un obstacle, à grande échelle) et où l’on observe tout de même des structures à toutes les échelles plus petites. Cette image en cascade reprend alors le principe sous-jacent de l’hypothèse H2 ci-haut concernant l’invariance sous changement d’échelle.

On peut alors considérer une vitesse typique à l’échelle  $l$  donnée par

$$v_l \sim \sqrt{\langle \delta v_{\parallel}^2(l) \rangle} \quad (80)$$

ainsi qu’un temps caractéristique à l’échelle  $l$ , appelé *temps de retournement* :

$$t_l \sim \frac{l}{v_l} \quad (81)$$

qui correspond au temps typique que met une structure de taille  $l$  à subir une déformation, due aux mouvements (relatifs) de ces composantes.

Le transfert d’énergie partant des structures de taille  $l$  vers les structures plus fines devrait alors se comporter comme :

$$\varepsilon \sim \frac{v_l^2}{t_l} \sim \frac{v_l^3}{l}, \quad (82)$$

et l’hypothèse H3 dicte que cette quantité ne dépend pas de l’échelle. On retrouve alors

$$v_l^2 \sim \varepsilon^{2/3} l^{2/3}, \quad (83)$$

i.e.  $h = 1/3$ . Mais le phénomène d'intermittence nous apprend que ce comportement n'est pas respecté.

Le modèle multi-fractal propose alors la correction suivante au modèle original : plutôt que de limiter l'invariance d'échelle à un seul exposant critique  $h$ , on considère plusieurs exposants critiques simultanément. Soulignons tout de suite que rien dans l'équation de Navier-Stokes n'indiquait qu'un seul exposant critique devait être considéré. En fait, l'équation de Navier-Stokes est invariante sous un nombre infini de transformations d'échelle.

Revenons d'abord à la valeur  $h = 1/3$ . Dans le modèle en cascade, on a fait l'hypothèse implicite que les structures de taille  $l$  remplissaient l'espace tri-dimensionnel. Si on imagine plutôt que l'énergie transmise d'une (grande) échelle vers une échelle plus fine se concentre en réalité dans une portion limitée de l'espace, on verra apparaître une valeur autre que  $1/3$  pour  $h$ . En effet, en considérant que l'énergie ne se transmet que dans une portion  $p_l$  de l'espace, avec  $p_l \sim l^{3-D}$ , on aura alors la relation suivante :

$$\varepsilon \sim \frac{v_l^3}{l} \cdot l^{3-D}. \quad (84)$$

On nomme *dimension fractale* la valeur  $D$ . La raison pour laquelle on utilise le terme *dimension* s'illustre ainsi : la probabilité qu'une boule de rayon  $l$  choisie au hasard dans un volume entre en contact avec une structure de dimension  $D$  (au sens classique) décroît comme  $l^{3-D}$ . Par exemple, le voisinage tri-dimensionnel d'une structure filamenteuse ( $D = 1$ ) est une "saucisse" dont la taille décroît comme  $l^{3-1}$ . La dimension fractale généralise cette idée à des valeurs non-entières de  $D$ . La valeur  $3 - D$  est alors appelée *co-dimension* de la structure étudiée. Nous présentons dans la section 3.2 une définition plus formelle de la dimension fractale.

On obtient évidemment le même résultat que précédemment lorsque  $D = 3$ . Cependant, si  $D \neq 3$ , on aura plutôt  $h = 1/3 - (3 - D)/3$  et, dans ce cas, la

fonction de structure devient

$$\zeta_p = p/3 + (3 - D)(1 - p/3). \quad (85)$$

On obtient toujours un comportement linéaire de  $\zeta_p$ , mais avec une pente différente de  $1/3$  cette fois. Conséquemment, si la cascade s'organise sur un sous-espace de  $\mathbb{R}^3$  de dimension fractale  $D$ , on devra observer ce type de comportement pour la fonction de structure.

Rien n'empêche la présence simultanée de plusieurs exposants d'échelle de valeurs différentes. D'autant qu'un tel processus, appelé processus *multi-fractal*, mènera à une fonction de structure possédant le caractère non-linéaire observé expérimentalement.

Formellement, il faut alors remplacer l'hypothèse  $H\bar{2}$  par la suivante :

**Hmf** : *Under the same assumptions as in H1, the turbulent flow is assumed to possess a range of scaling exponents  $I = (h_{\min}, h_{\max})$ . For each  $h$  in this range, there is a set  $\mathcal{I}_h \subset \mathbb{R}^3$  of fractal dimension  $D(h)$ , such that, as  $l \rightarrow 0$ ,*

$$\delta v_l(\mathbf{r}) \sim l^h, \quad \mathbf{r} \in \mathcal{I}_h. \quad (86)$$

*The exponents  $h_{\min}$  and  $h_{\max}$  and the function  $D(h)$  are postulated to be universal, i.e. independent of the mechanism of production of the flow ([37], chap. 8).*

On cherche à obtenir une expression pour le comportement dans les échelles de la fonction de structure  $S_p(l)$ . On peut l'écrire comme une intégrale sur le domaine des valeurs de  $h$  :

$$S_p(l) = \langle \delta v_l^p \rangle \sim \int_I l^{ph+3-D(h)} d\mu(h), \quad (87)$$

où  $I$  est l'intervalle des valeurs possibles de  $h$  et  $\mu(h)$  une mesure correspondant aux poids des différents exposants, dont l'expression formelle n'est pas pertinente pour l'établissement du comportement en échelle. On voit que le terme en  $l^{ph}$  provient de l'expression (86) tandis que le terme en  $l^{3-D(h)}$  correspond à la probabilité de se trouver à une distance proportionnelle à  $l$  du support de dimension  $D(h)$

des régions de comportement  $l^h$ . On s'intéresse au comportement de la fonction de structure à la limite des petites échelles, i.e.  $l \rightarrow 0$ ; dans ce cas pour chaque valeur de  $p$  l'intégrant sera complètement dominé par l'exposant minimal, i.e.

$$\lim_{l \rightarrow 0} \frac{\log S_p(l)}{\log l} = \zeta_p, \quad (88)$$

$$\zeta_p = \inf_h [ph + 3 - D(h)]. \quad (89)$$

La relation (89) correspond simplement à une *transformée de Legendre*. On a ainsi une relation permettant de lier le spectre des dimensions fractales  $D(h)$  à l'exposant d'échelle des moments des incréments de vitesse  $\zeta_p$ .

Si  $D(h)$  est une fonction concave (i.e. qu'elle possède une dérivée non-croissante partout), alors la transformée de Legendre est inversible et on peut écrire :

$$D(h) = \inf_p (ph + 3 - \zeta_p). \quad (90)$$

### 3.1.3.1. Dissipation d'énergie multi-fractale

En observant le phénomène d'intermittence à partir des données expérimentales sur les incréments de vitesse, on s'interroge naturellement sur des effets intermittents possibles sur l'autre quantité fondamentale, la dissipation d'énergie.

Plutôt que de se limiter à  $\varepsilon$ , qui était définie comme la valeur moyenne de la dissipation d'énergie par unité de volume, on considère une moyenne locale de la dissipation, introduite par Kolmogorov et Obukhov en 1962 [40, 41] :

$$\varepsilon_l(\mathbf{r}) = \frac{1}{(4/3)\pi l^3} \int_{|\mathbf{r}' - \mathbf{r}| < l} \frac{1}{2} \nu \sum_{ij} [\partial_j v_i(\mathbf{r}') + \partial_i v_j(\mathbf{r}')]^2 d^3 r'. \quad (91)$$

Cette définition locale suit la notion de multi-fractalité qui prévoit des comportements en échelle non-uniforme sur l'ensemble du support de la mesure, nécessitant une description locale.

Tout comme pour les incréments de vitesse, on prévoit un comportement de la dissipation locale dans les échelles donné par

$$\varepsilon_l(\mathbf{r}) \sim l^{\alpha-1}, \quad \forall \mathbf{r} \in \mathcal{D}_\alpha \subset \mathbb{R}^3; \quad \dim \mathcal{D}_\alpha = F(\alpha) \quad (92)$$

lorsque  $l \rightarrow 0$ . L'exposant  $\alpha-1$  plutôt que  $\alpha$  apparaît pour des raisons historiques.

Suivant le raisonnement pour les incréments de vitesse, on définit l'équivalent de la fonction de structure pour la dissipation locale, que l'on note  $\tau(q)$  :

$$\langle \varepsilon_l^q \rangle \sim l^{\tau(q)}. \quad (93)$$

Expérimentalement, on a mesuré une forme non-linéaire pour  $\tau(q)$ , indiquant à nouveau que le comportement local du flot turbulent n'est pas décrit correctement par K41 [42].

La théorie K41 permettait cependant de lier le comportement des incréments de vitesse à celui de la dissipation locale, à travers la loi du 4/5. L'argument original de Kolmogorov en était un d'analyse dimensionnelle, qui prévoyait que les incréments de vitesse devaient se comporter comme  $(\varepsilon l)^{1/3}$ .

Cette relation demeure en fait valable si on ajoute l'hypothèse qui suit. On considère d'abord les incréments de vitesse *sans dimension* :

$$\delta \tilde{\mathbf{v}}(\mathbf{l}) = \frac{\mathbf{v}(\mathbf{r} + \mathbf{l}) - \mathbf{v}(\mathbf{r})}{(\varepsilon l)^{1/3}}. \quad (94)$$

Kolmogorov propose que sous certaines conditions  $\tilde{\mathbf{v}}(\mathbf{l})$  et  $\varepsilon_l$  soient statistiquement indépendants, et que  $\tilde{\mathbf{v}}(\mathbf{l})$  possède une fonction de distribution universelle pour de grandes valeurs du nombre de Reynolds. C'est cette dernière hypothèse à laquelle on se réfère en tant que "refined similarity hypothesis". Elle implique que le comportement dans les échelles du moment d'ordre  $p$  des incréments de vitesse (contrôlé par  $\zeta_p$ ) soit le même que celui de la quantité  $(\varepsilon l)^{1/3}$ , ainsi

$$\delta \mathbf{v}(\mathbf{l})^p \sim \delta \tilde{\mathbf{v}}(\mathbf{l})^p \cdot (\varepsilon l)^{p/3} \quad (95)$$

ou encore

$$\zeta_p = p/3 + \tau(p/3). \quad (96)$$

Cette dernière expression décrit alors le lien fonctionnel entre les deux exposants.

### 3.1.4. Modèles en cascades

Il est intéressant de modéliser le comportement multi-fractal du flot à l'aide d'un modèle statistique simple. La méthode la plus simple sera de construire un modèle en cascades. Ce modèle pourra alors décrire le comportement statistique (dans les échelles) des incréments de vitesse, ou encore de la dissipation locale.

On construit un tel modèle en considérant d'abord un cube de taille  $l_0^3$ . On s'intéresse ici au cas d'une cascade sur la dissipation. On pose alors une valeur de dissipation  $\varepsilon$  uniforme sur tout le cube, positive et non-aléatoire. Ceci constitue l'étape  $n = 0$  de la construction.

Pour l'étape  $n = 1$ , on subdivise le cube en huit sous-cubes de taille  $(l_0/2)^3$ , et on obtient la valeur de dissipation locale sur chaque sous-cube en multipliant la valeur originale  $\varepsilon$  par une variable aléatoire  $W$  qui possède les propriétés suivantes :

$$W \geq 0, \quad \langle W \rangle = 1, \quad \langle W^q \rangle < \infty, \quad \forall q > 0. \quad (97)$$

Pour chaque sous-cube, la variable aléatoire  $W$  est indépendante de celle utilisée pour les sept autres sous-cubes. À l'étape  $n$ , on aura donc  $2^{3n}$  cubes de taille  $(l_0/2^n)^3$ . La valeur de la dissipation dans chacun de ces cubes sera donnée par

$$\varepsilon_l = \varepsilon W_1 W_2 \dots W_n, \quad (98)$$

où encore une fois tous les  $W_i$  sont indépendants et identiquement distribués. Le processus est répété *ad infinitum*.

Puisque la moyenne des  $W$  est égale à 1, la moyenne d'ensemble des  $\varepsilon_l$  sera égale à  $\varepsilon$ . Cependant les valeurs individuelles varieront d'un cube à l'autre, et en particulier la somme des dissipations sur huit sous-cubes ne correspondra pas à la valeur de la dissipation du cube formé par leur union, i.e. :

$$\varepsilon_l \neq \frac{1}{8} \sum_{j=1}^8 \varepsilon_{l/2}^{(j)}. \quad (99)$$

La cascade n'est donc pas conservative.

Pour les moments d'ordre  $q$  de la dissipation locale, on obtient ainsi

$$\langle \varepsilon_l^q \rangle = \varepsilon^q \left( \frac{l}{l_0} \right)^{\tau_q}, \quad (100)$$

avec

$$\tau_q = \log_2 \langle W^q \rangle. \quad (101)$$

L'expression (96) donne alors pour  $\zeta_p$  :

$$\zeta_p = p/3 - \log_2 \langle W^{p/3} \rangle. \quad (102)$$

L'exemple le plus simple de distribution sur  $W$  a été introduit par Novikov et Stewart [43].  $W$  y est décrit par une loi binomiale,  $W = 1/\beta$  avec probabilité  $\beta$ , et  $W = 0$  sinon. On a :

$$\Pr(W = \beta) = 1/\beta, \quad (103)$$

$$\Pr(W = 0) = 1 - 1/\beta, \quad (104)$$

$$\langle W^q \rangle = \beta^{1-q}, \quad (105)$$

ce qui donne pour les exposants

$$\tau_q = -(1 - q) \log_2 \beta, \quad (106)$$

$$\zeta_p = p/3 - (1 - p/3) \log_2 \beta. \quad (107)$$

Puisqu'on a un comportement linéaire pour les exposants, on est en présence d'un modèle mono-fractal, i.e. pour lequel une seule valeur de  $h$  est observée. Nous verrons plus loin d'autres distributions sur  $W$  qui conduisent à des comportements non-linéaires pour  $\tau_q$  et  $\zeta_p$ .

### 3.2. FORMALISME MULTI-FRACTAL

Dans cette section nous reprenons plus formellement les concepts de multi-fractalité introduits dans les sections précédentes. Nous décrivons également la Méthode des Maxima de la Transformée en Ondelettes, une méthodologie employée pour estimer numériquement les spectres multi-fractals des mesures et signaux étudiés. Nous nous plaçons d'abord dans le contexte de mesures fractales. Tous les concepts développés dans cette première partie seront ensuite exportés vers l'analyse de signaux (ou fonctions) fractals.

Pour une mesure  $\mu$  quelconque définie sur un support  $\text{Supp}(\mu)$ , on définit un exposant de singularité local  $\alpha(x_0)$  en  $x_0$  ainsi :

$$\alpha(x_0) = \lim_{\epsilon \rightarrow 0^+} \frac{\ln \mu(B_{x_0}(\epsilon))}{\ln \epsilon} \quad (108)$$

où  $B_{x_0}$  est une boule centrée en  $x_0$  et de taille  $\epsilon$ .

Nous aimerions maintenant obtenir un spectre décrivant la distribution de ces comportements singuliers sur le support de la mesure. Cependant ces singularités locales étant définies sur des points, elles ne seront pas denses sur le support de la mesure, du moins dans le cas général. Quel poids accorder alors à une valeur de  $\alpha$  donnée ? Puisque, de façon générale, l'ensemble des points pour lesquels on a une valeur de  $\alpha$  donnée sera de dimension fractale, on associe à chacune de ces valeurs la dimension fractale de cet ensemble. Nous définissons ici ce concept de dimension fractale de manière rigoureuse, au moyen de la *dimension de Hausdorff* [44].

La définition de la dimension de Hausdorff repose sur une généralisation de la notion de longueur. Cette *mesure de Hausdorff* est définie pour un ensemble  $S$  quelconque dans un espace métrique [45] :

$$l_\delta(S) = \lim_{\epsilon \rightarrow 0} l_{\delta, \epsilon}(S), \quad (109)$$

où

$$l_{\delta,\epsilon}(S) = \inf_{K(\epsilon)} \sum_{B_i \in K(\epsilon)} |B_i|^\delta. \quad (110)$$

L'infimum est pris sur l'ensemble des recouvrements  $K(\epsilon)$  de l'ensemble  $S$  par des boules  $\{B_i\}$  de diamètre inférieur ou égal à  $\epsilon$ . On peut maintenant définir en toute rigueur la *dimension de Hausdorff de  $S$*  comme étant la seule valeur  $d_H(S)$  telle que

$$l_\delta(S) = 0, \quad \forall \delta > d_H(S), \quad (111)$$

$$l_\delta(S) = +\infty, \quad \forall \delta < d_H(S). \quad (112)$$

$d_H(S)$  est donc l'unique valeur pour laquelle cette mesure est finie et non-nulle.

Le spectre des singularités  $f(\alpha)$  associé à la mesure  $\mu$  est alors défini ainsi :

$$f(\alpha) = d_H(x_0 \in \text{Supp}(\mu) | \alpha(x_0) = \alpha). \quad (113)$$

Cette définition constitue une généralisation complète de la dimension fractale en ce sens qu'elle décrit la distribution exacte de toutes les singularités présentes sur le support de la mesure. Elle possède cependant un inconvénient majeur : en pratique, pour pouvoir l'évaluer, il faut avoir accès à une plage d'échelle très large, ce qui n'est pas toujours le cas. Souvent, la convergence du comportement singulier en un point de la mesure sera trop lente pour être accessible en résolution finie (ou dans un temps de calcul raisonnable). Nous présentons ici une approche différente qui permet de régulariser ce problème d'ordre numérique [15].

Soit notre mesure  $\mu$  définie dans  $\mathbb{R}$ . On considère un pavage  $P(\epsilon)$  de  $\mathbb{R}$  par des boules de tailles  $\epsilon$ . Soit  $N(\epsilon)$  le nombre de ces boules de mesure non-nulle, que l'on numérote ainsi :  $\{B_i(\epsilon)\}_{i=1..N(\epsilon)}$ . On définit alors

$$\mu_i(\epsilon) = \int_{B_i(\epsilon)} d\mu \quad (114)$$

Pour tout  $q \in \mathbb{R}$ , on définit la fonction de partition

$$Z(q, \epsilon) = \sum_{i=1}^{N(\epsilon)} \mu_i^q(\epsilon). \quad (115)$$

Finalement, on définit le spectre  $\tau(q)$  associé à la mesure  $\mu$  par le comportement dans les échelles de la fonction de partition  $Z(q, \epsilon)$  :

$$Z(q, \epsilon) \sim \epsilon^{\tau(q)} \quad (116)$$

Mais, tel qu'on l'a vu dans la section 3.1.3, le spectre  $\tau(q)$  et le spectre  $f(\alpha)$  sont liés par une transformation de Legendre :

$$\tau(q) = \min_{\alpha} (q\alpha - f(\alpha)) \quad (117)$$

$$f(\alpha) = \min_q (q\alpha - \tau(q)) \quad (118)$$

Il est important de noter que les relations précédentes ne permettent pas de récupérer exactement un spectre à partir de l'autre, du moins pas en toute généralité. La transformée de Legendre ne redonne en effet que l'enveloppe convexe du spectre ; un exemple simple se trouve dans [15], p. 94.

### 3.2.1. Formalisme multi-fractal pour les fonctions et signaux

Nous avons introduit plus haut le modèle en cascade pour la dissipation d'énergie du flot turbulent. Il est intéressant d'obtenir un modèle semblable pour la vitesse du flot. On rencontrera alors certains problèmes. D'abord, la dissipation d'énergie (locale) est aisément interprétée comme une mesure ; elle est en effet définie comme une valeur moyenne sur une boule de taille  $l$ . La vitesse du flot, quant à elle, s'apparente plutôt à un signal (continu) prenant une valeur en tout point de l'espace.

Aussi, les différences de vitesse prennent des valeurs positives ou négatives, ce qui constitue un problème pour les moments d'ordre fractionnaire. On pourra alors utiliser la valeur absolue. Mais même dans ce cas, les différences de vitesse

prendront des valeurs arbitrairement proches de zéro, causant des ennuis pour l'évaluation numérique de moments d'ordre négatifs [33]. Il est donc utile de redéfinir l'approche multi-fractale pour ce type d'objet.

Toutes les définitions obtenues précédemment l'ont été pour une mesure dans  $\mathbb{R}$ . Il faut maintenant les transposer dans le contexte de l'analyse de fonctions (qui peuvent être des signaux, des images, etc...). Le degré de singularité local d'une mesure est décrit par le comportement du recouvrement partiel de la mesure par des boules de taille décroissante. Pour une fonction, la formulation est différente. Localement, on s'intéressera plutôt à une valeur  $h(x)$  telle que

$$|f(x+l) - f(x)| \sim Cl^{h(x)}. \quad (119)$$

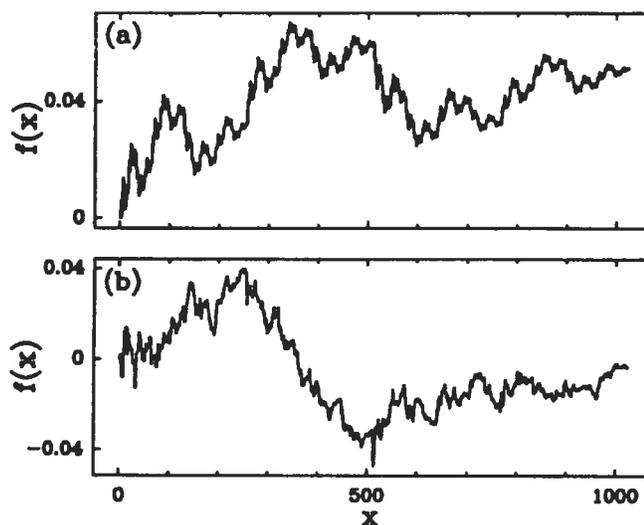


FIG. 3.4. Fonctions multi-fractales construites récursivement de manière a) déterministe ou b) aléatoire. (D'après [15].)

Cette valeur décrit bien le comportement de la fonction dans un voisinage de  $x$  en fonction de l'échelle, paramétrée ici par la variable  $l$ . Sur la figure 3.4, on trouve deux exemples de fonctions possédant des comportements locaux de ce type et pour lesquelles  $h(x)$  dépend du point  $x$  auquel on s'intéresse ; on parlera

alors de fonctions *multi-fractales*. On appelle exposant de Hurst local (en  $x$ ) la fonction  $h(x)$  qui respecte la condition exprimée dans l'équation (119). Cet exposant s'apparente à l'exposant de singularité  $\alpha$  que nous avons défini plus tôt ; on cherchera maintenant à écrire le spectre des singularités pour une fonction de la même façon :

$$D(h) = d_H(\{x, h(x) = h\}). \quad (120)$$

Cette définition de l'exposant de singularité ne permet cependant pas d'étudier des fonctions dont la dérivée à un certain ordre est singulière bien que la fonction elle-même ne le soit pas (en fait, la fonction est même dérivable dans ce cas). Une généralisation toute simple permet d'inclure ces comportements singuliers sur les dérivées et nous amène à définir un autre exposant, appelé exposant de Hölder, et qui correspond exactement à l'exposant de Hurst dans le cas où la fonction elle-même possède une singularité en un point. Si pour tout  $x$  dans un voisinage de  $x_0$  on a :

$$|f(x) - P_n(x - x_0)| \leq C|x - x_0|^h, \quad (121)$$

où  $C$  est une constante et  $P_n$  est un polynôme de degré  $n$ , alors on dira que la fonction est Lipschitz  $h$  en ce point  $x_0$ . Le plus grand  $h$  pour lequel cette inégalité est vérifiée se nomme exposant de Hölder local de  $f(x)$  en  $x_0$ .

Nous avons introduit cette définition plus complète en prétendant qu'elle permettait d'analyser le comportement auto-similaire de fonctions qui possédaient un certain nombre de dérivées, mais dont la dérivée à un certain ordre était singulière au point étudié. On dirait alors qu'une fonction dont la valeur de  $h$  se situe entre deux entiers  $n$  et  $n+1$  possède  $n$  dérivées, sa dérivée  $n$ -ième étant singulière, avec un exposant de Hölder de valeur  $h - n$  situé entre 0 et 1. Cette affirmation n'est pas vraie en toute généralité. Les conditions formelles pour que cela soit vérifié seront décrites en détails dans la section 3.2.1.3 ; nous y étudierons les

limites du formalisme multi-fractal en abordant un contre-exemple, les singularités oscillantes. Pour l'instant nous continuons notre inspection du comportement singulier d'une fonction sans nous en soucier.

Nous avons suggéré au chapitre 1 que l'analyse en ondelettes constituait un outil idoine pour l'étude du caractère auto-similaire de fonctions. C'est ici que cette méthode entre en scène, dans le contexte de l'évaluation du spectre des singularités d'une fonction. Nous reprenons ici la définition déjà énoncée au chapitre 1.

Nous avons alors défini la Transformée en Ondelettes (T.O.) continue en  $x_0$  d'une fonction  $f(\cdot)$  selon l'ondelette  $\psi(\cdot)$  à l'échelle  $a$  ainsi :

$$T_\psi[f](x_0, a) = \frac{1}{a} \int \psi\left(\frac{x - x_0}{a}\right) f(x) dx. \quad (122)$$

Rappelons que la T.O. peut-être interprétée comme une convolution locale d'une fonction  $f(\cdot)$  avec un noyau  $\psi(\cdot)$  dilaté par un facteur d'échelle  $a$ . On choisira pour  $\psi(\cdot)$  une fonction d'intégrale nulle ; nous reportons à nouveau la discussion concernant cette dernière condition à la section 3.2.1.1. Remarquons cependant que cette condition d'intégration fait en sorte que la fonction  $\psi(\cdot)$  doit osciller au moins une fois autour de zéro afin que son intégrale s'annule ; on s'attend alors à ce que la T.O. fasse ressortir le caractère d'oscillation locale de  $f(\cdot)$  à une échelle donnée.

Remarquons que l'équation (122) peut aussi s'écrire

$$T_\psi[f](x_0, a) = \int \psi(x) f(ax + x_0) dx. \quad (123)$$

Maintenant, considérant une fonction  $f(x)$  possédant un exposant de Hölder  $h(x_0)$  en  $x_0$ , on peut écrire son développement en puissance autour de ce point :

$$f(x) = c_0 + c_1(x - x_0) + \dots + c_k(x - x_0)^k + C|x - x_0|^{h(x_0)} \quad (124)$$

On introduit alors ce développement dans l'expression (122) :

$$T_\psi[f](x_0, a) = c_0 M_0 + c_1 M_1 a + \dots + c_k M_k a^k + C \int \psi(x) |ax|^{h(x_0)} dx. \quad (125)$$

où les  $M_i$  sont les moments d'ordre  $i$  de l'ondelette  $\psi$ . On peut imposer que cette ondelette possède au moins  $k$  moments nuls, i.e.  $M_i = 0, i = 0..k$ ; on aura alors pour  $T_\psi[f](x_0, a)$  :

$$T_\psi[f](x_0, a) = C \int \psi(x) |ax|^{h(x_0)} dx \quad (126)$$

$$\simeq a^{h(x_0)} C \cdot T_\psi[f](x_0, 1) \sim a^{h(x_0)}. \quad (127)$$

On voit donc que le comportement de la transformée en ondelettes sur la ligne ( $b = x_0, a$ ) est contrôlé par l'exposant de Hölder local de la fonction en  $x_0$ . Hélas, l'inverse n'est pas vrai : le comportement de la T.O. sur cette ligne ne permet pas d'obtenir la valeur de l'exposant de Hölder. Cependant, sous certaines conditions, le comportement de la T.O. dans un voisinage de cette ligne (dans un cône, en fait) nous permet d'extraire cet exposant. Nous verrons plus loin quelles sont exactement ces conditions, et ce qu'il advient lorsqu'elles ne sont pas respectées. Pour l'instant, en supposant qu'elles le sont, on a la relation suivante :

$$\max_{|x-x_0| \leq Ca} |T_\psi[f](x, a)| = O(a^{h(x_0)}). \quad (128)$$

Cette relation nous apprend que les valeurs maximales que prennent les coefficients de la T.O. à l'intérieur d'un cône dans le plan espace-échelle défini par  $|x - x_0| \leq Ca$  nous permettent d'extraire la valeur de l'exposant de Hölder local  $h(x_0)$  en étudiant le comportement de ces maxima lorsque l'échelle  $a$  tend vers 0. Ce résultat est obtenu à partir du théorème de Jaffard énoncé plus loin ; nous retrouvons dans l'expression (128) l'essence de ce que nous avons annoncé au chapitre 1 à propos de l'information portée par les maxima du module.

L'objectif est toujours de calculer un spectre multi-fractal pour une fonction en considérant l'exposant de Hölder comme valeur de singularité (ou de "fractalité") locale. Dans le contexte des mesures, nous avons considéré le comportement du recouvrement d'une mesure par des boules de rayons décroissants pour écrire notre fonction de partition

$$Z(q, \epsilon) = \sum_{i=1}^{N(\epsilon)} \mu_i^q(\epsilon). \quad (129)$$

Puisque pour une fonction c'est le comportement des maxima de la T.O. qui joue ce rôle, cette fonction de partition devient :

$$Z(q, a) = \sum_b |T_\psi[f](b, a)|^q. \quad (130)$$

où la somme est faite sur les valeurs de  $b$  telles que  $|T_\psi[f](b, a)|$  soit maximal, pour une valeur fixée de  $a$ .

En fait, pour des raisons d'implémentation numérique, cette somme sera limitée uniquement aux maxima qui font partie d'une chaîne, i.e. pour lesquels il existe des maxima voisins dans un petit voisinage en échelle. Cela évite que de petites valeurs maximales isolées viennent polluer numériquement la fonction de partition pour des valeurs de  $q$  négatives. En effet, une seule valeur proche de zéro dans cette somme fait diverger toute la fonction pour  $q$  négatif. Malgré cette précaution, il est toujours possible qu'un maximum faisant partie d'une ligne de maxima et qui change de signe à une échelle donnée prenne une valeur presque nulle. Pour éviter cette situation indésirable, on réécrit la fonction de partition en la régularisant :

$$Z(q, a) = \sum_{l \in \mathcal{L}(a)} \left( \sup_{a' \leq a} |T_\psi[f](b, a')| \right)^q. \quad (131)$$

Le supremum pris sur les échelles plus petites nous assure alors que de très petites valeurs ne puissent apparaître dans la somme. De manière analogue à l'expression (116), on écrit pour le comportement dans les échelles de la fonction de partition :

$$Z(q, a) \sim a^{\tau(q)}. \quad (132)$$

Il est important de noter que  $\tau(q)$  est ici défini en toute généralité et ne correspond pas uniquement à l'exposant d'échelle de la dissipation d'énergie dans un flot turbulent. En fait, dans le cadre de l'analyse multi-fractale de fonctions par la MMTO, il correspond plutôt à l'exposant  $\zeta(p)$  des différences de vitesse. Nous conservons toutefois la notation  $\tau(q)$  dans la prochaine section pour exprimer le fait que ces résultats ne se limitent pas au contexte de la turbulence.

### 3.2.1.1. *W-cascades*

Armé de ce nouvel outil, il est maintenant possible de définir un modèle de cascade pour une fonction, e.g. le signal de vitesse en turbulence. Les *W-cascades* sont en tout point analogues au modèle en cascade sur la mesure de dissipation, à ceci près que ce sont les coefficients de la transformée en ondelettes dyadique qui interviennent à chaque étape de la génération.

La définition de la transformée en ondelettes dyadique est similaire à celle de la transformée en ondelettes continue présentée ci-haut, hormis que la base de décomposition est orthonormale ; elle ne possède donc pas le caractère redondant de la transformée continue, qui plonge un signal uni-dimensionnel dans un espace à deux dimensions en ajoutant une dimension d'échelle.

En considérant que l'échelle intégrale constitue une taille maximale du signal étudié, on se contentera de travailler sur une base de l'espace des fonctions périodiques  $L^2_{\text{per}}([0, 1])$ [46]. On sait construire une base orthonormale pour cet espace à partir de deux fonctions  $\phi(x)$  et  $\psi(x)$ , en utilisant des dilatations et translations discrètes de  $\psi(x)$  :

$$\psi_{j,k} = 2^{j/2} \psi(2^j x - k), \quad j \geq 0, \quad 0 \leq k < 2^{j-1}. \quad (133)$$

Lorsque certaines conditions sont vérifiées par  $\phi(x)$  et  $\psi(x)$ , la famille de fonctions  $\{\phi(x), \psi_{j,k}(x)\}$  constitue une base orthonormale de  $L^2_{\text{per}}([0, 1])$ . Parmi ces conditions, la plus notable est celle de la moyenne nulle, i.e. la fonction  $\psi(x)$  doit

avoir son premier moment nul :

$$\int_0^1 \psi(x) dx = 0. \quad (134)$$

C'est cette condition qui vaut à  $\psi(x)$  le nom d'*ondelette*, puisqu'elle doit alors osciller au moins une fois autour de 0.

On définit alors les coefficients en ondelettes (dyadiques)  $\{c_\phi, c_{j,k}\}$  d'une fonction  $f(x)$  ainsi :

$$c_\phi = \langle f, \phi \rangle = \int_0^1 f(x) \cdot \phi(x) dx, \quad (135)$$

$$c_{j,k} = \langle f, \psi_{j,k} \rangle = 2^{j/2} \int_0^1 f(x) \cdot \psi_{j,k}(x) dx \quad (136)$$

Puisque la base en ondelettes est orthonormale, on saura reconstruire la fonction  $f(x)$  à partir de ses coefficients par l'expression suivante :

$$f(x) = c_\phi \phi(x) + \sum_{j \geq 0} 2^{-j/2} \sum_{0 \leq k < 2^j} c_{j,k} \psi_{j,k}(x). \quad (137)$$

On remarque que pour des valeurs de  $j$  successives, le paramètre  $k$  prend toujours le double du nombre de valeur qu'à l'échelle précédente (on a deux fois plus de coefficients associés à l'échelle  $j + 1$  qu'à l'échelle  $j$ ). Ceci est analogue à la structure du modèle en cascade défini plus haut pour la mesure de dissipation.

Une  $W$ -cascade sur le signal de vitesse sera alors définie par la généralisation triviale du modèle en cascade déjà décrit (pour la dissipation), en remplaçant simplement la dissipation locale sur une boule de taille  $l = 2^{-n}$  par le coefficient en ondelettes  $c_{j,k}$  du signal de vitesse à l'échelle  $j = n$ .

Cependant, un choix arbitraire concernant la distribution des variables aléatoires  $W$  ne conduira pas automatiquement à une fonction bornée puisque la somme dans l'expression (137) peut ne pas converger. Les conditions de convergence s'expriment en terme de l'exposant  $\tau(q)$  associé à la  $W$ -cascade et de son spectre de singularité.

Écrivons d'abord les probabilités suivantes :

$$q_j^\alpha = \Pr\{\max_k |c_{j,k}| > 2^{-j\alpha}\} \quad (138)$$

et

$$p_j^\alpha = \Pr\{|c_{j,k}| > 2^{-j\alpha}\}, \quad (139)$$

qui ne dépend pas de  $k$ . On remarque que

$$q_j^\alpha \leq 2^j p_j^\alpha \quad (140)$$

puisque la probabilité que la plus grande valeur de  $|c_{j,k}|$  soit bornée par  $2^{-j\alpha}$  est certes plus petite ou égale à la somme des probabilités individuelles. Mais puisque le coefficient  $c_{j,k}$  s'écrit comme le produit de  $j$  variables aléatoires  $W$ , on peut réécrire  $p_j^\alpha$  ainsi :

$$p_j^\alpha = \Pr\left\{\sum_{i=1}^j \log_2 |W_i| \geq -j\alpha\right\}. \quad (141)$$

Une borne supérieure s'obtient en faisant appel à la théorie des grandes déviations.

**PROPOSITION 3.2.1.** *Si  $E[\log_2 |W|] < +\infty$  alors  $\forall \alpha < -E[\log_2 |W|]$  et  $\forall \epsilon > 0$ ,  $\exists J > 0$  tel que  $\forall j > J$*

$$p_j^\alpha < e^{j\epsilon} \cdot 2^{j(F(\alpha)-1)}, \quad (142)$$

où  $F(\alpha)$  est définie comme la transformée de Legendre de  $\tau(q)$

$$F(\alpha) = \inf_q (q\alpha - \tau(q)), \quad (143)$$

et où  $\tau(q)$  est définie ainsi :

$$\tau(q) = -\log_2 E(|W|^q) - 1, \quad \forall q \in \mathbb{R}. \quad (144)$$

De plus,  $F(\alpha)$  est une fonction concave telle que  $F(-E[\log_2 |W|]) = 1$  et qui est croissante sur l'intervalle  $]-\infty, -E[\log_2 |W|]$  et décroissante sur  $[-E[\log_2 |W|], \infty[$ .

La preuve se trouve dans l'annexe B de [46] et fait elle-même appel à [47] en définissant la transformée de Cramér dans le contexte de la théorie des grandes déviations. Une analogie simple entre la transformée de Cramér et la notion d'entropie statistique est présentée dans la section 8.6.4 de [37].

De la proposition 3.2.1 on déduit celle qui suit :

**PROPOSITION 3.2.2.** *Pour une  $W$ -cascade telle que définie ci-haut, la convergence sera assurée dès que  $0 < -E[\log_2 |W|]$  et que*

$$\exists \eta > 0, F|_{[0, \eta]} < 0. \quad (145)$$

Nous donnons maintenant la preuve de la proposition 3.2.2. Posons  $\alpha = \eta/2$ , de sorte que  $0 < \alpha < -E[\log_2 |W|]$  et que  $F(\alpha) < 0$ . Utilisant la proposition 3.2.1 :

$$q_j^\alpha < e^{j\epsilon} \cdot 2^{jF(\alpha)} \quad (146)$$

et ce  $\forall \epsilon > 0$  et pour  $j$  suffisamment grand. Par conséquent

$$\sum_{j=1}^{\infty} q_j^\alpha < \infty. \quad (147)$$

Le fait que cette somme soit bornée implique, par le lemme de Borel-Cantelli, que la probabilité d'observer l'évènement associé à  $q_j^\alpha$  un nombre de fois infini est nulle. On a finalement :

$$\sum_{j>J} \sum_{0 \leq k < 2^j} 2^{-j} c_{j,k}^2 \leq \sum_{j>J} \max_k c_{j,k}^2 \leq \sum_{j>J} 2^{-2\alpha j} < \infty, \quad (148)$$

assurant la convergence de la cascade.

### 3.2.1.2. *Approche thermodynamique*

La fonction de partition que nous avons définie laisse croire qu'il est possible de donner une interprétation thermodynamique à notre méthode d'évaluation du spectre multi-fractal d'une fonction. Par exemple, puisque l'on a une augmentation du nombre de maxima lorsque l'on se déplace vers les petites échelles, et que

les résultats qui nous intéressent se trouvent justement à la limite des échelles infiniment petites, on peut penser que cette limite correspond à la limite thermodynamique d'un grand nombre de particules. On peut alors regarder l'inverse du logarithme de l'échelle comme un *volume*, i.e.  $V = -\ln a$ . On a alors

$$|T_\psi[f](b_i, a)| \sim e^{-h_i V}, \quad (149)$$

$$D(h) \sim V^{-1} \ln N(h, V) \quad (150)$$

où  $N(h, V)$  correspond au nombre de maxima de l'ordre de  $e^{-hV}$  et s'interprète comme la dégénérescence de l'état d'énergie  $h$ ; l'indice  $i$  identifie chaque micro-état (chaque maximum individuel).  $h_i$  est en effet vue comme l'énergie par unité de volume du micro-état  $i$  et  $D(h)$  s'interprète comme l'entropie par unité de volume de tous les états "d'énergie"  $h$ . On n'a toujours pas fait appel à la fonction de partition; c'est que cette première série d'analogies constitue une approche micro-canonique, i.e. où l'énergie totale  $h \cdot V$  est fixée.

En introduisant la fonction de partition  $Z(q, a)$ , on identifie  $|T_\psi[f](b, a)|^q \sim e^{-hVq}$  au poids de Boltzmann et  $\tau(q) = \log_a Z(q, a)$  n'est autre que l'énergie libre de Helmholtz à la température  $q^{-1}$ . Cette fois, on somme sur toutes les énergies possibles, et on retrouve l'approche canonique.

On voit alors qu'il est possible d'obtenir des valeurs moyennes de l'énergie  $\bar{h}(q)$  et de l'entropie  $\bar{D}(q)$  pour toute température  $q$  :

$$\bar{h}(q, a) = \frac{\sum \{|T_\psi[f](b, a)|^q \cdot [\log_a |T_\psi[f](b, a)|]\}}{Z(q, a)} \quad (151)$$

$$\bar{D}(q, a) = \frac{\sum \left\{ |T_\psi[f](b, a)|^q \cdot \left[ \log_a \left( \frac{|T_\psi[f](b, a)|^q}{Z(q, a)} \right) \right] \right\}}{Z(q, a)} \quad (152)$$

En éliminant  $q$  entre ces deux équations, on retombe sur le spectre de singularité  $D(h)$ . C'est d'ailleurs cette approche qui est utilisée en pratique pour calculer le spectre  $D(h)$  (figure 3.5) [33].

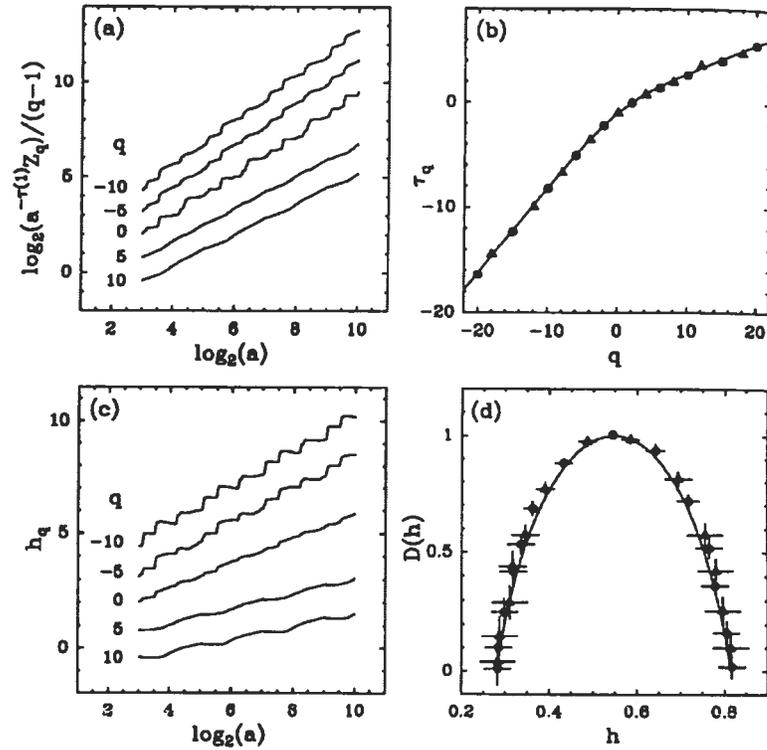


FIG. 3.5. Calcul de  $\tau(q)$  et de  $D(h)$  associés aux fonctions multifractales de la figure 3.4. a)  $\log_2(a^{-\tau(1)} Z(q, a)) / (q - 1)$  en fonction de  $\log_2(a)$ . b)  $\tau(q)$  en fonction de  $q$ . c)  $h(q, a)$  en fonction de  $\log_2(a)$ . d)  $D(h)$  en fonction de  $h$ . Dans b) et d) : (●) fonction déterministe; (△) fonction aléatoire; (-) courbe théorique. (D'après [15], fig.35.)

### 3.2.1.3. Singularités oscillantes

L'objectif initial est maintenant atteint, celui d'obtenir un formalisme multifractal pour les fonctions. Nous avons souligné en cours de route que le comportement des maxima de la T.O. ne nous permettait d'extraire l'exposant de Hölder local que sous certaines conditions. Dans cette section nous précisons ce commentaire.

Le théorème suivant exprime les conditions formelles concernant le comportement hölderien d'une fonction  $f(\cdot)$  [48] :

a) Soit  $f \in L^2(\mathbb{R})$  et Lipschitz  $\gamma$  en  $x_0$  avec  $\gamma \leq n$  où  $n$  représente le nombre de moments nuls de l'ondelette. Alors, sa T.O. vérifie :

$$|T_\psi[f](x, a)| = O(a^\gamma + |x - x_0|^\gamma). \quad (153)$$

b) Inversement, si on a que  $\gamma < n$  et que

$$\text{i) } \exists \nu > 0 \text{ tel que } |T_\psi[f](x, a)| = O(a^\nu) \text{ uniformément en } x, \quad (154)$$

$$\text{ii) } |T_\psi[f](x, a)| = O(a^\gamma + |x - x_0|^\gamma / |\ln |x - x_0||), \quad (155)$$

alors  $f$  est Lipschitz  $\gamma$  au point  $x_0$ .

Supposons pour l'instant que la condition (b-i) soit toujours vérifiée, et négligeons la correction logarithmique dans la condition (b-ii). Alors, la condition (a) devient nécessaire et suffisante. Elle nous apprend ceci : si tous les maxima de la T.O. liés à une singularité en  $x_0$  se trouvent à l'intérieur d'un cône défini par  $|x - x_0| \leq Ca$ , la condition (a) devient

$$\max_{|x - x_0| \leq Ca} |T_\psi[f](x, a)| = O(a^{h(x_0)}). \quad (156)$$

tel que nous l'avons écrit plus tôt. Cependant, il existe des comportements locaux pour lesquels aucun cône (même si la valeur de  $C$  est très grande) ne contient tous les maxima provoqués par la singularité en  $x_0$ . Nous appellerons *singularités oscillantes* les singularités pour lesquelles on trouve un tel comportement. Un exemple classique nous est donné par la fonction  $f(x) = \sin(1/x)$ . Pour cette fonction, l'analyse en T.O. nous donne des lignes de maxima qui convergent toutes vers des points autres que là où se trouve la singularité ! De manière plus générale, on observe ce type de comportement pour toute fonction de la forme :

$$f(x) = |x - x_0|^\gamma \sin\left(\frac{2\pi}{|x - x_0|^\beta}\right), \quad (157)$$

où l'on suppose  $\beta > 0$  et  $\gamma > -1$ . Ce type de fonction (figure 3.6) "détourne" la T.O. en quelque sorte, en présentant systématiquement des oscillations d'échelles

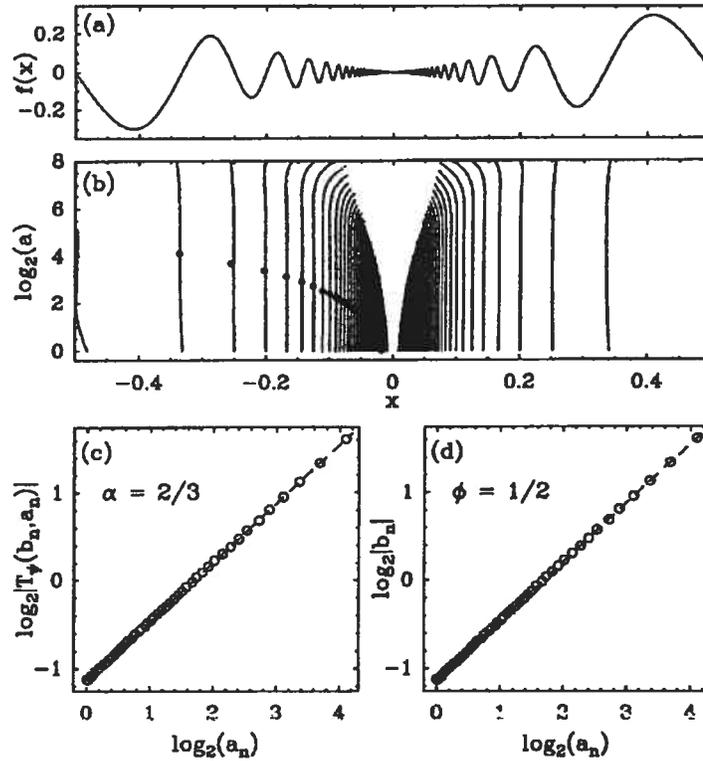


FIG. 3.6. a) Exemple d'une singularité oscillante :  $f(x) = |x|^\gamma \sin(2\pi/|x|^\beta)$  pour  $\gamma = 4/3$  et  $\beta = 1$ . b) Maxima du module de la T.O. de  $f(x)$ . Les points ( $\bullet$ ) marquent la position des maxima globaux de la T.O. sur chaque ligne de maxima. c)  $\log_2 |T_\psi(b_n, a_n)|$  vs  $\log_2 a_n$ . La pente de ce graphe redonne  $\alpha(x_0 = 0)$ . d)  $\log_2 b_n$  vs  $\log_2 a_n$ . Cette pente donne un estimé de  $\phi(x_0 = 0)$ . (D'après [49].)

de plus en plus fines lorsque l'on s'approche de la singularité, ne permettant donc pas à l'ondelette de les saisir. On utilise souvent le terme anglais "chirp" pour décrire ces singularités, en opposition à des singularités de type "cusp" pour lesquelles le comportement singulier est contrôlé par des maxima se trouvant à l'intérieur d'un cône. Les singularités oscillantes démontrent un comportement le long des lignes de maxima qui ne caractérise pas l'exposant de Hölder local. En fait, pour une fonction de la forme présentée dans l'équation (157), une évaluation

de cet exposant basé sur la méthode MMTO donne une valeur infinie, alors que le comportement hölderien local doit clairement être donné par  $\gamma$ .

Sur la figure 3.6, on voit bien que les lignes de maxima ne convergent pas vers le point  $x_0 = 0$  lorsque  $a \rightarrow 0$ . Cependant, si l'on s'intéresse à la position du maximum pour chaque ligne de maxima (i.e. le maximum global sur chaque ligne), on voit que cette série de points, lorsque l'on saute d'une ligne à l'autre, converge elle vers le point  $x_0$ . Notons par  $(b_n, a_n)$  cette séquence de points en choisissant arbitrairement un de ces points comme début de séquence,  $(b_0, a_0)$ . On peut montrer [49] que sur cette séquence de points, le module de la T.O. pour un signal de type "chirp" se comporte comme

$$|T_\psi[f](b_n, a_n)| \sim a_n^{\alpha(x_0)}, \quad (158)$$

où  $\alpha(x_0) = \gamma/(\beta+1)$ . Puisque  $\beta > 0$ , on voit que le nouvel exposant  $\alpha$  ainsi calculé ne correspond pas exactement à l'exposant de Hölder que l'on cherche, mais est en fait plus petit que ce dernier par un facteur  $(\beta + 1)$ . Pour retrouver l'exposant de Hölder, on doit donc parvenir à évaluer  $\beta$ . Cela devient possible si l'on étudie comment évoluent les positions  $b_n$  au fur et à mesure que l'on s'approche de  $x_0$ , i.e. lorsque  $n \rightarrow \infty$ . On sait que pour la famille de signaux à laquelle on s'intéresse ici, chaque ligne de maxima correspond à une oscillation du signal lui-même (il est facile de s'en convaincre en "promenant" une ondelette de petite taille sur le signal). On doit donc avoir que les abscisses  $b_n$  s'approchent de  $x_0$  de la manière suivante :

$$|b_n - x_0| \sim n^{-1/\beta}. \quad (159)$$

De plus, la séquence des échelles  $a_n$  doit être proportionnelle à la distance entre deux abscisses consécutives, puisque l'échelle à laquelle la corrélation entre l'ondelette et la fonction oscillante est maximale doit être proportionnelle à ce que

l'on pourrait appeler la "fréquence instantanée" de la fonction, qui elle est justement donnée par la différence entre ces deux abscisses. On peut alors écrire  $a_n \sim b_n - b_{n+1} \sim n^{-(\beta+1)/\beta}$ . Inversant cette dernière relation et remplaçant  $n$  par  $a_n^{-\beta/(\beta+1)}$  dans (159), on obtient

$$|b_n - x_0| \sim a_n^{\phi(x_0)} \quad (160)$$

lorsque  $n \rightarrow \infty$ , et où  $\phi(x_0) = 1/(\beta+1)$  (figure 3.6). On retrouve alors l'exposant de Hölder local  $h(x_0)$  à partir des valeurs de  $\alpha(x_0)$  et de  $\phi(x_0)$  :

$$h(x_0) = \alpha(x_0)/\phi(x_0) \quad (161)$$

Tout ce développement est pour l'instant limité à des fonctions de type "chirp". On peut montrer [49] qu'en définissant ces deux exposants de la même façon, mais en ajoutant la condition  $\phi = \max(1, \phi)$ , cette approche peut être généralisée pour une plus grande famille de fonctions. Entre autre, il est possible de construire récursivement des fonctions pour lesquelles on retrouve dans tout voisinage d'une singularité (oscillante ou pas) une infinité de singularités, oscillantes et non-oscillantes.

Donc, si une fonction possède un certain nombre de ces singularités dites oscillantes, la méthode MMTO qui mène au formalisme multi-fractal pour les fonctions n'est pas applicable en toute rigueur lorsque qu'une caractérisation en terme du comportement hölderien est essentielle.

### 3.2.2. Cascades et stationnarité

Dans tout ce qui a été présenté ci-haut, il n'a été nulle part fait mention de stationnarité du processus. Pourtant, l'invariance sous translation est une symétrie fondamentale de l'équation de Navier-Stokes et de nombreux autres systèmes. La stationnarisation du modèle en W-cascades (section 3.2.1.1) fait l'objet du travail présenté dans l'article qui suit. Pour l'instant nous souhaitons montrer de manière

simple que le processus en W-cascades tel que défini dans la section 3.2.1.1 n'est pas stationnaire. Nous conservons ici la même notation que celle introduite dans cette section.

On entend par stationnarité l'indépendance de la statistique des coefficients  $c_{j,k}$  de la transformée en ondelettes dyadique par rapport à  $k$  pour une valeur d'échelle  $j$  fixée. De manière plus formelle on définit deux types de stationnarité : la stationnarité *au sens faible* ou *au sens large*, où l'on se contente de vérifier que l'espérance du produit de deux coefficients ne dépend que de la distance entre ces coefficients ; la stationnarité *au sens strict*, où l'on exige que tous les moments joints de tout ordre impliquant tout nombre de coefficients soit invariant sous translation. La stationnarité au sens strict implique celle au sens large, mais l'inverse n'est pas vrai.

Dans le cas d'un modèle en W-cascades tel que défini dans la section 3.2.1.1, ni l'un ni l'autre de ces deux types de stationnarité n'est respectée. En effet, pour deux coefficients voisins  $c_{j,k}$  et  $c_{j,k+1}$  ou  $k$  est pair, ces coefficients partagent un parent direct, c'est-à-dire que l'on peut écrire pour ces deux coefficients

$$c_{j,k} = c_{j-1,k/2} * W_{j,k} \quad (162)$$

$$c_{j,k+1} = c_{j-1,k/2} * W_{j,k+1} \quad (163)$$

où  $W_{j,k}$  et  $W_{j,k+1}$  sont indépendants et identiquement distribués. Or lorsque  $k$  est impair ceci n'est plus vrai puisque les deux coefficients ne partagent plus le même parent direct. Dans le pire des cas, lorsque  $k = 2^{j-1} - 1$ , les deux coefficients ne partagent en fait aucun parent sauf à la racine de l'arbre. Ces deux coefficients sont alors parfaitement décorrélés, ce qui n'est certes pas le cas en toute généralité pour des coefficients partageant un parent direct. Ainsi on en conclut que le modèle en W-cascades ne respecte pas une des symétries fondamentales de l'équation de Navier-Stokes, à savoir l'invariance sous translation.

### 3.3. PRÉSENTATION DE L'ARTICLE

Nous avons présenté à la section 3.1.4 le modèle en cascade, introduit en phénoménologie de la turbulence pour générer une mesure de dissipation d'énergie manifestant le comportement multi-fractal prédit par l'hypothèse Hmf. Ce type de modèle peut également être repris pour générer des signaux de vitesse turbulents à travers les  $W$ -cascades présentées dans la section [46].

Il est alors intéressant d'étudier la structure de corrélation spatiale et temporelle de tels signaux. Cette approche nécessite généralement une hypothèse supplémentaire de stationnarité (ou d'homogénéité) afin qu'un estimateur de corrélation spatiale ou temporelle puisse être formellement défini. Malheureusement les modèles en cascades actuels *ne respectent pas cette condition supplémentaire*, tel qu'il l'est démontré dans l'article qui suit. Le problème jusqu'ici a été contourné en définissant un estimateur de corrélation "moyenné" [46] qui ne représente pas en toute généralité la quantité que l'on cherche véritablement à mesurer, à savoir la forme de la corrélation spatiale homogène ; des méthodes de stationnarisation approximative ont également été développées, qui encore une fois ne garantissent pas la validité de l'estimateur de corrélation spatiale [50].

Nous présentons dans l'article qui suit une méthode permettant de définir des modèles graphiques multi-échelle respectant la condition de stationnarité. Nous dérivons les conditions que doit respecter cette nouvelle construction et la forme générale prise par la fonction de corrélation spatiale. Dans le cas du modèle gaussien nous montrons que la stationnarité est assurée au sens strict, alors qu'elle l'est au sens large pour les statistiques non-gaussiennes. Dans ce dernier cas, nous dérivons des conditions assurant la stationnarité de tous les moments d'ordre inférieur ou égal à  $p$  pour  $p$  fini.

# Stationary Multiscale Graphical Models

PHILIPPE ST-JEAN, *Université de Montréal*

JEAN-MARC LINA, *Université de Montréal*

BERNARD GOULARD, *Université de Montréal*

## Abstract

We present a stationary model for cascading process defined on dyadic digraphs. This model is intended to describe self-similarity of certain physical processes, such as those encountered in the field of fully-developed turbulence. Those processes respect a differential equation which displays scale symmetry to a certain extent, such as the Navier-Stokes equation. However, the “cascade” model actually used to describe those self-similar processes omit to conform to another fundamental symmetry of the system : invariance under translation. Hence the cascade model generates non-stationary processes, whereas the non-stationarity appears in sole consequence of this oversimplifying model.

We propose to modify the standard cascade model by stating a sufficient condition that ensures stationarity of the process, while keeping the marginal distribution of the coefficients intact such that the scaling behavior of the process remains unchanged.

## 3.4. INTRODUCTION

We present a stationary model for cascading process defined on dyadic digraphs. The motivation for this paper originated from the work of Arnéodo et al. on cascade model in fully developed turbulence, which was introduced to describe multi-fractal behavior (intermittency) in wind tunnel velocity field signals [33, 46, 51]. The model allows to simulate velocity fields with parameter values

(such as the Reynold number) that cannot be attained experimentally. For instance there is interest in predicting the behavior when the Reynold number goes to infinity, requiring either the viscosity tending towards zero or the size of the tunnel to be arbitrarily large [52]. The model is defined as a Markov dyadic tree over an orthogonal wavelet decomposition of the velocity signal, and is referred to as a  $\mathcal{W}$ -cascade model [46].

The  $\mathcal{W}$ -cascade model is intended to describe self-similarity characteristics of processes. Those processes respect a differential equation which displays scale symmetry to a certain extent, as for the Navier-Stokes equation in the case of turbulence. However, the cascade model omits to conform to another fundamental symmetry of the system : invariance under translation. This invariance is typically verified experimentally [37]. However the  $\mathcal{W}$ -cascade generates non-stationary processes, whereas the non-stationarity appears in sole consequence of an oversimplified model.

Hence  $\mathcal{W}$ -cascade model on wavelet dyadic trees lacks the stationarity property, while the stochastic processes to be described are stationary. A problem arises when it comes to the estimation of the spatial correlation function on simulated data. This problem is usually put aside in a somewhat inelegant fashion by averaging correlation estimators along the spatial dimension, with little or no justification for doing so.

In a different context, efforts were put forward in order to generate “almost stationary” processes with a multi-scale graphical modelling approach, such as in [50] where a Kalman-filter prediction algorithm on multiscale graphs is studied. Their approach results in apparent stationarity of the process, meaning that the form of the covariance function “looks” stationary, i.e. fluctuations of the covariance function when changing the position where it is evaluated are small relative to pointwise variance. This was not a problem in [50] as their algorithm did not require stationarity. In the case of fully developed turbulence such an approach is invalid ; the problem is just hidden deeper, as higher order (n-point) moments are

generally not guaranteed to respect this apparent stationarity behavior. Furthermore, one might wonder what kind of behavior on the spatial covariance function (or on higher order moments) should be expected when strict stationarity of the model is imposed. The main goal of our work is to elaborate a model which is strictly stationary, while conserving the self-similarity properties of the previous models, those being enlisted by the behavior of pointwise statistics (i.e. on every single node) through scales. For instance, quantities such as the singularity spectrum of the process must be preserved [33, 53].

One important drawback of non-stationary models on dyadic multiscale graphs lies in the fact that the topology of the dyadic graph becomes visible in the process, since its spatial covariance function depends strongly on the position of the node in the tree. This artefactual structure is of course never observed on the processes one tries to model. This flaw becomes unacceptable when one is interested in modelling the precise form of the spatial correlation function of a wavelet cascade process, or the correlation between dyadic wavelet coefficients.

We propose to modify the standard cascade model by stating a sufficient condition that ensures stationarity of the process, while keeping the marginal distribution of the coefficients intact. We begin by presenting some non-stationary models which will be used as a starting point. Then we derive a sufficient condition in this context, assuming that the model is gaussian. We move on to prove that this condition is also sufficient in a much wider context than the gaussian case, i.e. that stationarity of higher order moments is ensured provided some extra conditions are imposed.

We show that our solution also brings increased versatility to the model, as it allows to control the shape of the spatial correlation function to a certain extent (something that was not possible with the previous models). Then we derive estimators for the extra parameters involved. Finally we show comparison between non-stationary models and our model on simulated data.

### 3.5. NEAREST-NEIGHBOR MODELS

We begin by describing the two simplest non-stationary models (which we hereafter refer to as nearest-neighbor models). The first one is defined on a tree structure such as the one depicted on figure (3.7), a single-rooted binary tree. This tree structure is commonly seen in discrete dyadic wavelet decomposition on orthonormal basis, whereas each node represents a wavelet coefficient. The coefficients are arranged in levels which represent scales. One readily sees that every node is the parent of exactly two child nodes. The cascade process is defined as follows : Let  $c_{0,0} = 0$  be the coefficient at the root (top) of the tree, and  $c_{j,k}$

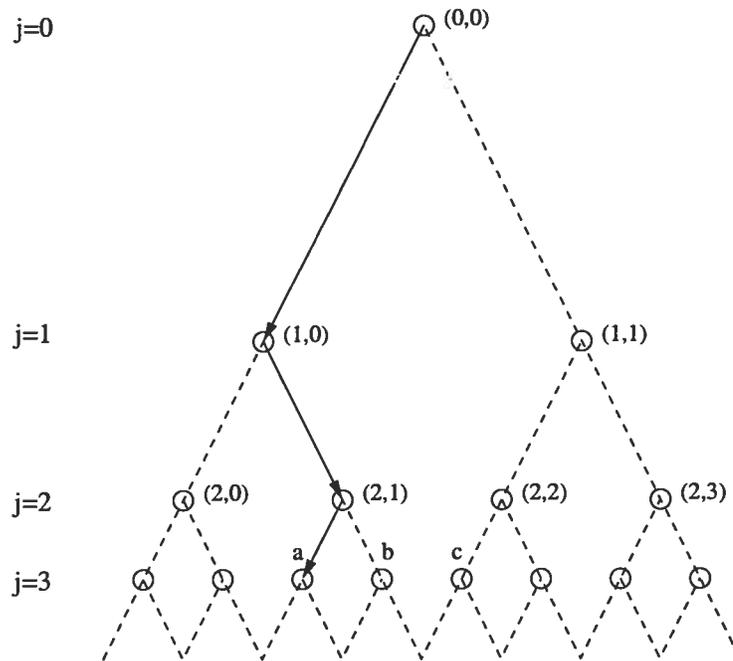


FIG. 3.7. The single parent model. The solid line indicates ancestors of coefficient  $a$ .

the  $k$ -th coefficient at level  $j$ . Then any  $c_{j,k}$  is generated by adding a  $W_{j,k}$  random variable to its parent coefficient, i.e. :

$$c_{j,2k} = c_{j-1,k} + W_{j,2k}$$

$$c_{j,2k+1} = c_{j-1,k} + W_{j,2k+1}.$$

We mention here that if  $W$ 's are iid gaussian variables, then every coefficient will also be gaussian, and that the whole tree of coefficients defines a multi-gaussian variable. We will assume that this is the case until we reach section 3.8. At any level  $j$  of the tree, it is possible to define a discrete spatial covariance function  $C$  as :

$$C^{(j)}(k, \Delta k) = E[c_{j,k} \cdot c_{j,k+\Delta k}],$$

with the additional assumption that the  $W$ 's are zero-mean. In this model,  $C$  is a function of  $j$ ,  $\Delta k$  and  $k$  as well. This means that the covariance of two coefficients on the same level depends not only on their relative distance ( $\Delta k$ ), but also on their position in the tree. For instance, coefficients  $a$  and  $b$  on figure(3.7) show stronger correlation than coefficients  $b$  and  $c$  as the former share an immediate common parent, although relative distance between  $a$  and  $b$  and between  $b$  and  $c$  are the same. The stationarity constraint we want to impose is for the covariance  $C$  to be a function of  $j$  and  $\Delta k$  solely.

We now present another dyadic model, depicted on figure (3.8). This model is no longer a tree, but still is an acyclic digraph with a single source. In this case, half of the coefficients have only one parent, while the others have two. The expression for those with one parent (the ones with even index  $k$ ) is the same as in the single-parent model above, while the odd-indexed coefficients are given by the second expression :

$$\begin{aligned} c_{j,2k} &= c_{j-1,k} + W_{j,2k} \\ c_{j,2k+1} &= \alpha \cdot c_{j-1,k} + \beta \cdot c_{j-1,k+1} + W_{j,2k+1} \end{aligned}$$

where  $\alpha$  and  $\beta$  can take any real value (typically  $1/2$  for both). Again, it is obvious that this model is not stationary in our sense, as the covariance for a pair of even-indexed coefficients will in the general case be different from a pair of odd-indexed ones (as the latter share a parent).

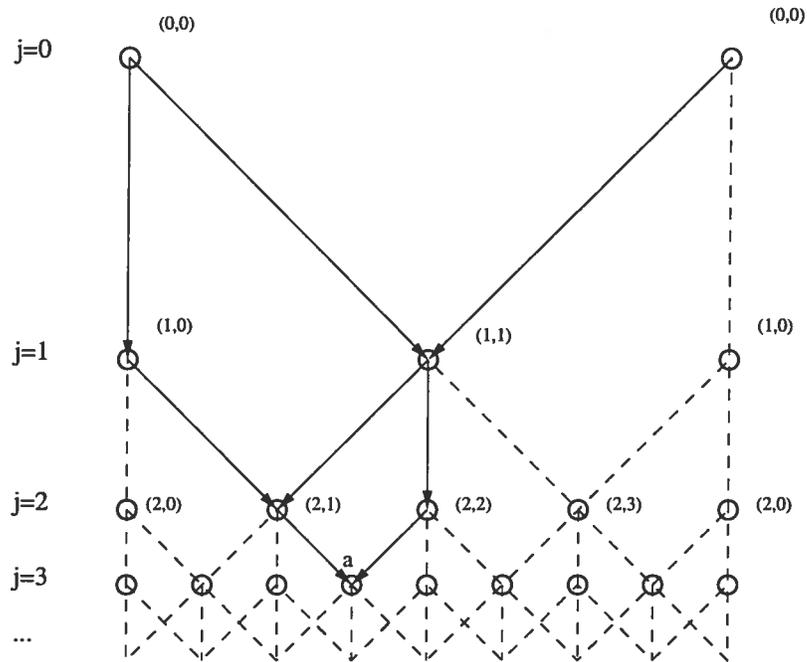


FIG. 3.8. The one-two parent model. Solid lines indicates all ancestors of coefficient  $a$ .

This last model (the “one-two” parent model) will be our starting point in building a stationary model, although we might as well have chosen the preceding one. The only difference lies in the computation of the stationarity constraints which happens to be somewhat simpler to perform using the one-two parent model.

### 3.6. STATIONARY MODEL ( $\alpha$ -MODEL)

In order to obtain stationarity, it is necessary to add complexity to the model. In the case of nearest-neighbor models, a child coefficient  $c_{j,k}$  can be considered as depending solely on the value taken by its direct parents  $\gamma(c_{j,k})$  (either one or two of them, as above). This is equivalent to a first-order markovian relationship along branches of the tree, of the type (for even-indexed) :

$$P(c_{j,k}|\Omega_{j-1}) = P(c_{j,k}|\gamma(c_{j,k}))$$

where  $\Omega_j$  represents all coefficients on levels  $0, 1, \dots, j$ . However,  $c_{j,k}$  will also be correlated to coefficients in  $\Omega_j$  other than its direct parents. By adding possible direct contribution from other ancestors in the model, one can recover all the properties of nearest-neighbor models (in terms of marginal distributions and parent-child covariances, for instance), while gaining stationarity. It is actually possible to do so by limiting the influence of direct contribution to the coefficients found at the preceding level only, i.e. by forcing the digraph to allow edges between nodes on adjacent levels only.

One simple way to achieve this is to impose translation invariance on the covariance of coefficients found on the same level, by exploiting symmetries of the tree structure at hand.

In the model we present, the coefficients are written as :

$$c_{j,k} = W_{j,k} + \sum_{k'=0}^{2^{j-1}-1} \alpha_{k,k'}^{(j)} c_{j-1,k'} \quad (164)$$

with  $c_{0,0}=0$ . The  $c_{j,k}$  now depend directly on all coefficients found on level  $j-1$ . The value of  $\alpha_{k,k'}^{(j)}$  for a given triplet  $(j, k, k')$  corresponds to the weight associated to the edge going from  $c_{j-1,k'}$  to  $c_{j,k}$ . We refer to such a model as an  $\alpha$ -model. A graphical representation is shown on figure 3.9.

We still consider the  $W_{j,k}$  to be identically distributed, although not necessarily independent from one another (yet independent from the  $c_{j-1}$ ).

We ask for  $\alpha_{k,k'}^{(j)}$  to depend only on some distance function of  $k$  and  $k'$ , i.e.

$$\alpha_{(k,k')}^{(j)} = \alpha_{d(k,k')}^{(j)}.$$

In our case, the obvious choice for  $d(k, k')$  is  $d(k, k') = |2k' - k|$ , which corresponds to the actual distance along the  $k$ -axis on fig(3.9). For simplicity we assume that  $\alpha$  is symmetric over its index (i.e.  $\alpha_{-i} = \alpha_i$ ), so that we can drop the absolute value in the distance function. This is done without loss of generality, as the derivation in the following section also holds for the anti-symmetric part of  $\alpha$ . Usage of non-symmetric form for  $\alpha$  will be discussed further in section 3.7.1.

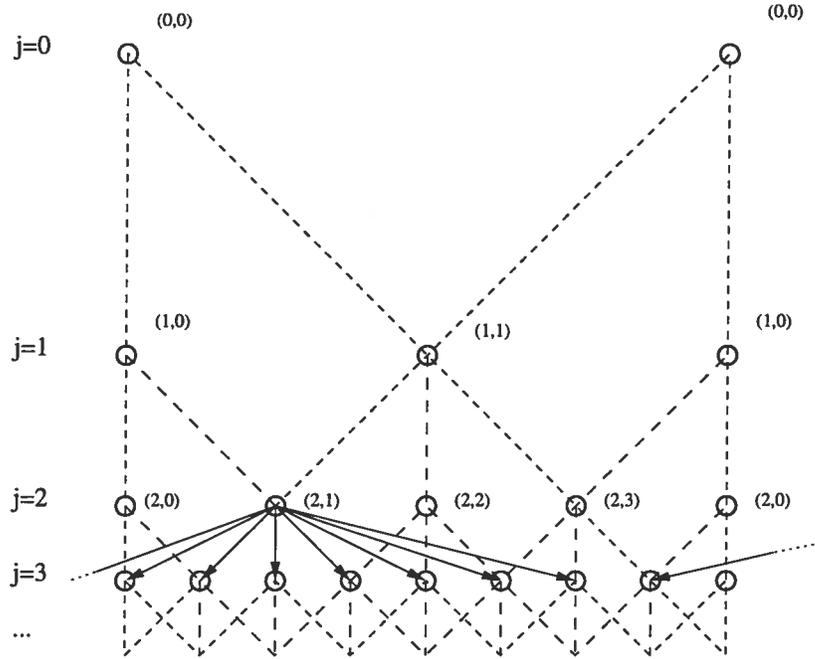


FIG. 3.9. The  $\alpha$  model.

### 3.7. STATIONARITY CONSTRAINTS

From now on we impose that the  $W_{j,k}$  are i.i.d.  $N(0, \sigma_W)$ , without loss of generality. At a given scale  $j$  we want the  $c_j$ 's to share the same marginal distribution ( $N(\mu(j), \sigma(j))$ ); the form of the pdf is guaranteed to be Gaussian since each coefficient is expressed as a linear combination of the  $W_{j,k}$ . The zero-mean condition on the  $W$ 's also implies that  $\mu(j) = 0, \forall j$ . This has no influence on stationarity, which depends solely on second order moments.

The task is to build a model for which coefficients found at the same level have covariances depending on their mutual distance only; in order to impose this condition, we proceed by induction. We assume that the inner-level covariance for level  $j-1$   $\text{Cov}[c_{j-1,k}, c_{j-1,k+\Delta k}]$  is a function of  $\Delta k$  (and  $j$ ) only, and demand that the inner-level covariance at level  $j$  respects the same condition.

This covariance is written as

$$C^{(j)}(k, \Delta k) = C_1^{(j)}(k, \Delta k) + \text{Cov}_W(\Delta k),$$

where

$$C_1^{(j)}(k, \Delta k) \equiv E \left[ \left( \sum_{k'=0}^{2^{j-1}-1} \alpha_{2k'-k}^{(j)} c_{j-1,k'} \right) \left( \sum_{k''=0}^{2^{j-1}-1} \alpha_{2k''-(k+\Delta k)}^{(j)} c_{j-1,k''} \right) \right], \quad (165)$$

and where

$$\text{Cov}_W(\Delta k) = E[W_{j,k} \cdot W_{j,k+\Delta k}]$$

is assumed to depend solely on  $\Delta k$ . Thus the process will be stationary when  $C_1^{(j)}(k, \Delta k)$  depends only on  $\Delta k$  for all  $j$ .

We can rewrite (165), using a simple change of variables (from  $k''$  to  $k' + k''$ ) :

$$C_1^{(j)}(k, \Delta k) = \sum_{k'=0}^{2^{j-1}-1} \sum_{k''=0}^{2^{j-1}-1} \alpha_{2k'-k}^{(j)} \alpha_{2(k''+k')-(k+\Delta k)}^{(j)} C^{(j-1)}(0, k''), \quad (166)$$

where we have used the induction hypothesis in the rightmost part of the last expression, as it corresponds exactly to the covariance at level  $j - 1$ .

As we try to make this expression independent of  $k$ , we shall first notice that replacing  $k$  by  $k + 2m$ ,  $m \in \mathbb{Z}$  does not change the expression, thanks to the periodicity of  $\alpha$ . However, replacing  $k$  by  $k + 1$  (or by  $k + 2m + 1$ , which is equivalent) modifies the expression. This means that imposing the expression to be invariant when changing  $k$  in  $k + 1$  is sufficient to obtain stationarity.

We will first consider the case where  $\Delta k$  takes an even value. The two  $\alpha$ 's in (166) then share parity on their index, either both even or both odd. Let  $\eta$  and  $\xi$  be even and odd sub-samplings of  $\alpha$ , respectively :

$$\begin{aligned} \eta_i^{(j)} &= \alpha_{2i}^{(j)} \\ \xi_i^{(j)} &= \alpha_{2i-1}^{(j)} \end{aligned}$$

For  $C_1$  to remain unchanged when either  $k = 0$  or  $k = 1$  is to impose

$$E[c_{j,0} c_{j,\Delta k}] = E[c_{j,1} c_{j,1+\Delta k}].$$

Since  $\Delta k = 2m$ ,  $m \in \mathbb{Z}$ , we write

$$\begin{aligned} \sum_{k'=0}^{2^{j-1}-1} \sum_{k''=0}^{2^{j-1}-1} \eta_{k'}^{(j)} \eta_{k''+k'-m}^{(j)} C^{(j-1)}[k''] = \\ \sum_{k'=0}^{2^{j-1}-1} \sum_{k''=0}^{2^{j-1}-1} \xi_{k'}^{(j)} \xi_{k''+k'-m}^{(j)} C^{(j-1)}[k''] \end{aligned}$$

where  $C^{(j-1)}[k''] = E[c_{j-1,k'} \cdot c_{j-1,k''+k'}]$  is even around  $k'' = 0$ .

The last equation is written in terms of circular convolutions :

$$\left( \eta^{(j)} \otimes \eta_-^{(j)} \otimes C^{(j-1)} \right) [m] = \left( \xi^{(j)} \otimes \xi_-^{(j)} \otimes C^{(j-1)} \right) [m], \quad (167)$$

where  $(\eta_-)_i = \eta_{-i}$ . Through Fourier transform, this becomes (we omit the  $j$ 's for simplicity) :

$$\hat{\eta}[l] \cdot \hat{\eta}[-l] \cdot \hat{C}[l] = \hat{\xi}[l] \cdot \hat{\xi}[-l] \cdot \hat{C}[l].$$

$\hat{C}$  and  $\hat{\eta}$  are both real and even, since  $C$  and  $\eta$  are ( $\eta$  is the even sub-sampling of  $\alpha$ , also even).  $\xi$  is also real, but not even. Hence  $\hat{\xi}$  can be written in polar form,  $\hat{\xi}[l] = \hat{A}_\xi[l] e^{i\theta_\xi[l]}$ , with  $\hat{A}_\xi$  and  $\theta_\xi$  even and odd, respectively.

Regardless of the form of  $\hat{C}$ , the constraint will be verified when

$$\hat{A}_\xi^2[l] = \hat{\eta}^2[l]. \quad (168)$$

This means that  $\hat{\eta}$  and  $\hat{\xi}$  are the same, up to a complex phase. The phase  $\theta_\xi$  will be such that  $\xi$  is even around  $1/2$ . We lost necessity of the condition in this last step, as possible zeros of  $\hat{C}$  could soften the condition on  $\hat{\eta}$  and  $\hat{\xi}$ .

Repeating the same exercise for  $\Delta k = 2m + 1$ ,  $m \in \mathbb{Z}$ , one falls on the following expression :

$$\left( \eta^{(j)} \otimes \xi_-^{(j)} \otimes C^{(j-1)} \right) [m] = \left( \xi^{(j)} \otimes \eta_-^{(j)} \otimes C^{(j-1)} \right) [m + 1],$$

which transforms to

$$\hat{\eta}[l] \cdot \hat{\xi}[-l] \cdot \hat{C}[l] = e^{\frac{2\pi il}{N}} \hat{\xi}[l] \cdot \hat{\eta}[-l] \cdot \hat{C}[l]$$

where  $N = 2^{j-1}$ . By simplifying both  $\hat{C}$  and  $\hat{\eta}$  in this last expression, all that remains is the same condition on the phase of  $\hat{\xi}$  as mentioned above. Hence no extra condition is to be imposed on  $\eta$  and  $\xi$ .

We conclude that the process will be stationary for any choice of even  $\hat{\eta}$ , provided that the modulus of  $\hat{\xi}$  is the same, and that its phase is expressed as

$$\begin{aligned}\hat{\xi}[l] &= \hat{\eta}[l] \cdot e^{i\theta_\xi[l]}, \\ \theta_\xi[l] &= \frac{-\pi l}{N}, \quad l = -N/2 + 1 \dots N/2.\end{aligned}\tag{169}$$

This means for instance that one can choose values for even coefficients in  $\alpha$ , and compute odd ones by performing a frequency shift in Fourier space.

### 3.7.1. Symmetry of $\alpha$

We have assumed so far that  $\alpha$  was symmetric over its index. The derivation above can be repeated with the anti-symmetric part of  $\alpha$ , with the same result. The only effect of removing the symmetry condition on  $\alpha$  will be that (169) now represents the frequency shift between  $\hat{\eta}$  and  $\hat{\xi}$  rather than the exact phase of the latter, as  $\hat{\eta}$  is now allowed a non-null complex phase.

Non-symmetric behavior of  $\alpha$  can be used in the modelling of past-future dependencies for time-based cascading processes, as is the case in modelling fluctuations of financial time series. See for instance [54, 53].

### 3.7.2. Interpretation of the “frequency shift” condition

We will now show that the “frequency shift” condition (169) tells us that the support of  $\hat{\alpha}$  is limited to half of its length, i.e.  $\alpha$  is a band-limited discrete function.

We begin by upsampling both  $\eta$  and  $\xi$  by a factor of two (by introducing zeros at every odd-indexed value), and then shifting  $\xi$  by one, such that the sum of the two leads to an expression for  $\alpha$  :

$$\alpha = \eta_{\uparrow 2} + \xi_{\uparrow 2, \text{shft}}\tag{170}$$

The upsampling of  $\eta$  and  $\xi$  implies that their Fourier transform is simply duplicated on the domain of  $\hat{\alpha}$ , which is appropriately twice as large. The shift on  $\xi$  introduces a frequency shift in the Fourier domain. Hence we have

$$\begin{aligned} \widehat{(\eta \uparrow 2)} &= \hat{\eta}, \\ \widehat{(\xi \uparrow 2, \text{shft})} &= \hat{\eta} \cdot e^{\frac{-2\pi il}{2^j}} \cdot \begin{cases} e^{\frac{-\pi il}{2^{j-1}}} & l = -(2^{j-2}) + 1 \dots 2^{j-2} \\ e^{\pi i - \frac{\pi il}{2^{j-1}}} & l = 2^{j-2} + 1 \dots 2^{j-1} \\ e^{-\pi i - \frac{\pi il}{2^{j-1}}} & l = -(2^{j-1}) + 1 \dots -(2^{j-2}) \end{cases}, \quad (171) \end{aligned}$$

where it is understood that  $\hat{\eta}$  is duplicated over the  $[-(2^{j-1}) + 1, 2^{j-1}]$  support of  $\hat{\alpha}$ . The phase factors in (171) cancel out and give

$$\widehat{(\xi \uparrow 2, \text{shft})} = \hat{\eta} \cdot \begin{cases} 1 & l = -(2^{j-2}) + 1 \dots 2^{j-2} \\ -1 & l = 2^{j-2} + 1 \dots 2^{j-1} \\ -1 & l = -(2^{j-1}) + 1 \dots -(2^{j-2}) \end{cases}$$

Thus the sum in (170) is expressed in the Fourier domain as

$$\hat{\alpha} = 2 \cdot \hat{\eta} \cdot \chi^{(j-1)}, \quad (172)$$

where  $\chi^{(j-1)}$  is the characteristic function over  $[-(2^{j-2}), 2^{j-2}]$ .

Hence the support of  $\hat{\alpha}$  is limited to half its domain, which makes  $\alpha$  a band-limited function. This implies that the support of  $\alpha$  itself is of full length (i.e.  $2^j$ ). An important consequence of this condition is that it is not possible to generate a stationary process with only one or two parents at each node, which in return explains why the nearest-neighbor models could not be stationary. For instance the one-two parent model, where  $\alpha$  is given as  $[0 \dots 0, 1/2, 1, 1/2, 0 \dots 0]$ , is not band-limited and thus could not lead to a stationary process.

An interpretation of this condition is found by looking at the set of nodes at scale  $j - 1$  as a band-limited process; since it only regroups  $2^{j-1}$  nodes, it cannot contain frequencies higher than that number. Consequently it cannot generate by itself higher frequencies unless some additional random variables (such as the  $W$ 's) are added. Trying to increase artificially the frequency content using solely

the nodes at scale  $j - 1$  always results in creating two different processes, one for odd and one for even nodes at scale  $j$  that are then simply interlaced, leading to the artefactual non-stationarities we are trying to avoid. This makes for an intuitive explanation of why the support of  $\alpha$  must be limited to half its length in the frequency domain.

### 3.8. STATIONARITY OF HIGHER ORDER MOMENTS

So far we have proven that it is possible to ensure strict-sense stationarity at least for Gaussian processes, as second-order moments completely describe the statistics in this case. The question arises of whether the same condition implies stationarity of higher order moments as well (i.e.  $\rho \geq 3$ ) for non-gaussian distributions, with the benefit of providing stronger stationarity for a much larger class of marginal distributions on the nodes. It is shown in this section that this is the case to a certain extent.

Let us express this comment formally :

PROPOSITION 3.8.1. *For an  $\alpha$ -model as defined in (164), with  $\alpha$  respecting conditions (168, 169), then :*

$$E[c_{j,k} c_{j,k+\Delta k_1} \dots c_{j,k+\Delta k_{p-1}}]$$

does not depend on  $k$ ,  $\forall \Delta k_i \in \mathbb{Z}$ ,  $i \in [1 \dots p - 1]$ , provided that the expected value above exists and that the following condition holds at every scale :

$$\hat{\eta}^{(j)} \left[ \sum_i^{p-1} l_i \right] \cdot \prod_{i=1}^{p-1} \hat{\eta}^{(j)}[-l_i] = \tag{173}$$

$$\hat{\eta}^{(j)} \left[ \sum_i^{p-1} l_i \right] \cdot \prod_{i=1}^{p-1} \hat{\eta}^{(j)}[-l_i] \cdot \exp \left( i\theta \left[ \sum_i l_i \right] - i \sum_i \theta[l_i] \right),$$

$\forall l_i \in [0, 2^{j-1} - 1]$ ,  $i \in [1 \dots p - 1]$ .

The proof can be found in the Appendix.

Let us look at (173) and see what it implies. In the case of  $p = 2$ , the expression is trivially respected and no new condition is imposed, as expected. For  $p \geq 3$  we see that (173) will be respected as long as  $\theta[l]$  is a linear function of  $l$ . Although this is indeed the case for  $l \in [-2^{j-2} + 1, 2^{j-2} - 1]$  where  $\theta[l] = -\pi l / 2^{j-1}$ , the sum in the argument of  $\hat{\eta}$  in the LHS of (173) can go beyond this interval. As  $\theta[l]$  is reproduced periodically outside the interval, it is easily checked that an extra factor of  $-1$  will appear for certain combinations of values for the  $l_i$ . This is the case when  $l_1 = l_2 = 2^{j-2} - 1$  with  $p = 3$ , for instance. Hence a new condition must be imposed on  $\hat{\eta}$  for  $p \geq 3$ .

A sufficient condition for (173) is derived here. The extra  $-1$  factor appears only when  $\sum l_i$  goes out of the  $[-2^{j-2} + 1, 2^{j-2} - 1]$  interval. However we can force  $\hat{\eta}[l]$  to be zero on enough points such that the sum of indices on non-zero values never grows out of the interval. For instance with  $p = 3$ , limiting the support of  $\hat{\eta}$  to half of its length will achieve this goal, since the sum involves exactly two terms. For any order  $p$ , this will hold when limiting the support of  $\hat{\eta}$  to a  $1/(p-1)$  fraction of its length.

Hence in order to provide high-order moment stationarity of the model when the marginal distribution on the nodes is not gaussian, it is necessary to “start” the cascading process at a scale where there is enough nodes such that the sufficient condition derived above is not trivial, whereas  $\hat{\eta}[l]$  would be forced to be null everywhere except at  $l = 0$ .

This result brings greater generality to the model, as it allows to obtain stationary of higher order moments for multi-scale process with any non-gaussian multivariate distribution, provided these moments exist.

The proposed  $\alpha$ -model can be considered as a method of removing structures induced by the digraph in a multi-scale Markov model rather than being limited to provide wide-sense stationarity, as the stationarity constraints do not depend on the specific form of the distribution on the nodes.

### 3.9. FORM OF THE COVARIANCE FUNCTION

The stationarity constraint derived just above leads to some specific behavior of the covariance function, which we describe in this section. Assuming that we have  $\alpha$  respecting the stationarity condition (172), we write the covariance at level  $j$  as

$$C^{(j)}[\Delta k] = C_1^{(j)}[\Delta k] + C_W^{(j)}[\Delta k]. \quad (174)$$

For even values of  $\Delta k$ , we can rewrite (167) :

$$C_1^{(j)}[2m] = (\eta^{(j)} \otimes \eta^{(j)} \otimes C^{(j-1)}) [m]. \quad (175)$$

Whereas for odd  $\Delta k$ , (169) gives

$$C_1^{(j)}[2m + 1] = (\eta^{(j)} \otimes \xi^{(j)} \otimes C^{(j-1)}) [m].$$

Considering the last two expressions as null for odd and even values of  $\Delta k$  respectively, we see that adding them together gives the full expression for  $C_1$ . Using Fourier transform once again and following the same line as in section 3.7.2,

$$\hat{C}_1^{(j)}[l] = \hat{C}^{(j-1)}[l] \cdot \hat{\eta}^{(j)}[l] \cdot \left\{ \hat{\eta}^{(j)}[l] + e^{\frac{2\pi i l}{2^j}} \hat{\xi}^{(j)}[l] \right\}.$$

The length of the support of  $\hat{C}_1^{(j)}$  is  $2^j$ , while that of  $\hat{\eta}^{(j)}$ ,  $\hat{\xi}^{(j)}$  and of  $\hat{C}^{(j-1)}$  is  $2^{j-1}$ ; since those functions are all periodic, the latter are simply replicated twice on the support of  $\hat{C}_1^{(j)}$ .

By (174), we get a recursive form :

$$\hat{C}^{(j)}[l] = \hat{C}^{(j-1)}[l] \cdot \hat{\eta}^{(j)}[l] \cdot \left\{ \hat{\eta}^{(j)}[l] + e^{\frac{2\pi i l}{2^j}} \hat{\xi}^{(j)}[l] \right\} + \hat{C}_W^{(j)}[l]$$

Coming back to polar notation for  $\hat{\xi}$ , and using (168) :

$$\begin{aligned} \hat{C}^{(j)}[l] &= \hat{C}^{(j-1)}[l] \cdot (\hat{\eta}^{(j)}[l])^2 \cdot \left\{ 1 + e^{\frac{2\pi i l}{2^j}} \cdot e^{i\theta_\xi[l]} \right\} + \hat{C}_W^{(j)}[l] \\ &= \hat{C}^{(j-1)}[l] \cdot (\hat{\eta}^{(j)}[l])^2 \cdot 2\chi^{(j-1)} + \hat{C}_W^{(j)}[l]. \end{aligned} \quad (176)$$

Iterating this last expression, we get a formula for the covariance at any level :

$$\hat{C}^{(j)}[l] = \left[ \sum_{j'=1}^{j-1} 2^{j-j'} \cdot \hat{C}_W^{(j')} [l] \cdot \chi^{(j')} [l] \prod_{j''=j'+1}^j (\hat{\eta}^{(j'')} [l])^2 \right] + \hat{C}_W^{(j)} [l]$$

This last equation is quite general ; it describes the form of the covariance function for any stationary process based on our  $\alpha$ -model.

In order to generate a realization of one such process, it is necessary to choose (or estimate) both  $\hat{\eta}^{(j)} [l]$  and  $\hat{C}_W^{(j)} [l]$ . We start by a simple example to show what form of covariance function can be expected. It is readily seen that the following choice for  $\hat{\eta}$  is acceptable :

$$\hat{\eta}[l] = 1 \tag{177}$$

This is just a delta function for  $\eta$ .  $\alpha$ , however, will not be a delta function because of the contribution of  $\xi$ .  $\alpha$  will rather be a “low-pass” delta function. We also consider a delta function for  $C_W$  (i.e.  $W$ 's are i.i.d.), which is admissible as there is no constraint over  $C_W$ .

Equation (176) then simplifies to

$$\hat{C}^{(j)} [l] = \left[ \sum_{j'=1}^{j-1} 2^{j-j'} \chi^{(j')} [l] \right] + 1,$$

and we obtain a “staircase” form for the Fourier Transform of the covariance function  $\hat{C}$ . (fig. (3.10)). We see that the covariance function behaves like  $\log_2(\hat{C}^{(j)} [l]) \sim j - \log_2[l + 1]$ , i.e.  $\hat{C}^{(j)} [l] \sim \frac{j}{l+1}$ . Of course, a function with a smoother decrease for  $\hat{\eta}$  will result in a smoother  $\hat{C}$  too.

### 3.10. ESTIMATORS FOR $\hat{\eta}$ AND $\hat{C}_W$

We shall now describe how to estimate parameters of the model given a realization of the process. Estimators for  $\hat{\eta}$  and  $\hat{C}_W$  are sufficient to completely characterize the  $\alpha$ -model under the gaussian assumption ; this is understood from the fact that an expression for pairwise covariances of all wavelet coefficients can

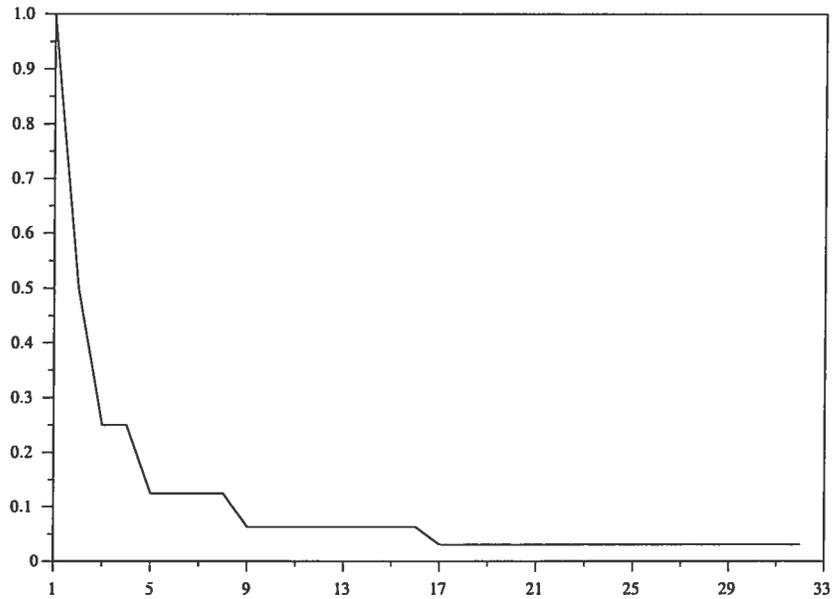


FIG. 3.10.  $\hat{C}[l]$  for  $\hat{\eta}$  defined in (177).

be derived from  $\hat{\eta}$  and  $\hat{C}_W$ , regardless of whether the coefficients belong to the same level or not.

Indeed, in order to obtain a simple estimator for  $\hat{\eta}$  we begin by computing the covariance function between coefficients found on two adjacent levels :

$$C^{(j-1,j)}[\Delta k] \equiv \text{Cov}(c_{j-1,k'}, c_{j,k}),$$

where  $\Delta k = 2k' - k$  still. We will only need this covariance for even values of  $\Delta k$ , i.e. the even subsampling  $C_e^{(j-1,j)}[\Delta k]$ . Then we have :

$$\begin{aligned} C_e^{(j-1,j)}[\Delta k] &= \sum_{k'=0}^{2^{j-1}-1} \alpha_{2k'-\Delta k}^{(j)} C^{(j-1)}[k'] \\ &= (\eta^{(j)} \otimes C^{(j-1)}) \left[ \frac{\Delta k}{2} \right] \end{aligned}$$

Considering that  $C_e^{(j-1,j)}$  is null for odd values of  $\Delta k$ , we have  $\hat{C}_e^{(j-1,j)}[l] = \hat{\eta}^{(j)}[l] \cdot \hat{C}^{(j-1)}[l]$ . Hence we obtain an estimator for  $\hat{\eta}^{(j)}[l]$  :

$$\hat{\eta}^{(j)}[l] = \frac{\hat{C}_e^{(j-1,j)}[l]}{\hat{C}^{(j-1)}[l]} \quad (178)$$

For  $\hat{C}_W$ , we use equation (175) to write :

$$\hat{C}_W^{(j)}[l] = \hat{C}^{(j)}[l] - \hat{C}^{(j-1)}[l] \cdot (\hat{\eta}^{(j)}[l])^2 \cdot 2\chi_{[-(2^{j-2}), 2^{j-2}]}$$

It is possible to get a simpler expression for the even sub-sampling of  $\hat{C}_W$  :

$$\begin{aligned} \hat{C}_{W,e}[l] &= \hat{C}_e^{(j)}[l] - \hat{C}_e^{(j-1)}[l] \cdot (\hat{\eta}^{(j)}[l])^2 \\ &= \hat{C}_e^{(j)}[l] - \hat{C}_e^{(j-1,j)}[l] \cdot \hat{\eta}^{(j)}[l] \end{aligned} \quad (179)$$

where we used (178) in order to obtain (179).

### 3.11. SIMULATIONS AND ANALYSIS

In this section, we compare the behavior of non-stationary models for cascading process with our proposed stationary model. To that end, we have generated multiple realizations of the process for each model, estimated correlations between nodes (i.e. pairs of nodes at distance  $\Delta k$  from each other, both nodes being found at the same scale) at the finest scale  $j = J$  and plotted their values as a function of the position  $k$  of one of the node in the graph.

We have generated three series of process realizations, one with the single-parent model, one with the one-two parent model and the other using the stationary  $\alpha$ -model. For the latter we chose the simple case of  $\hat{\alpha}$  being a truncated gaussian window and  $C_W$  a delta function. This was done over 8 scales ( $j = 1 \dots 8$ ), and the correlations are estimated on nodes found at the last scale  $j = 8$ . Figure (3.11) shows estimation of the correlation of the  $c_{j,k}$ 's as a function of position  $k$  and inter-nodal distance  $\Delta k$  for the single-parent model. One very well sees the dyadic structure of the tree. This process is highly non-stationary; for instance certain nodes are almost completely uncorrelated ( $\rho \simeq 0$ ) while most are strongly

correlated ( $\rho > 0.6$ ) when  $\Delta k = 1$ . All figures are subsampled by a factor of 4 for clarity.

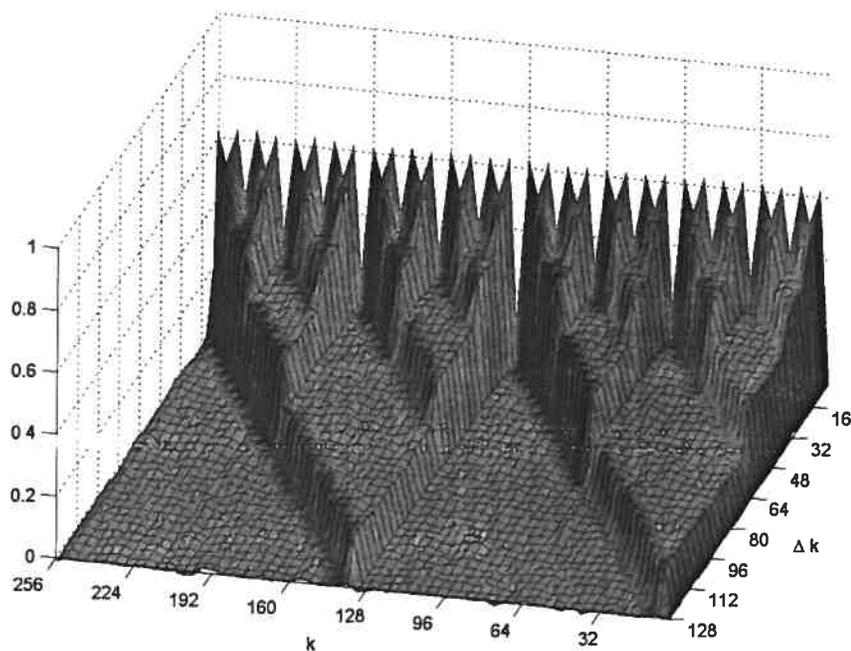


FIG. 3.11. Correlations between nodes as a function of absolute position of the node ( $k$ ) and inter-nodal distance ( $\Delta k$ ) for the single-parent non-stationary model ( $j=8$ ,  $\sigma_W = 1$ ). Blocky non-stationary artifacts are obvious.

Figure (3.12) shows the same plot for the one-two parent model this time. As this model is closer to the stationary model in terms of number of parent nodes involved in the graph, non-stationarities do not appear as strikingly as for the single-parent model. Still the dyadic structure is obvious and the process is non-stationary.

Figure (3.13) shows the same plot for the  $\alpha$ -model. Elimination of the artefactual dyadic structure is obvious when compared to figure (3.11).

All simulations involved 20000 realizations of graphs.

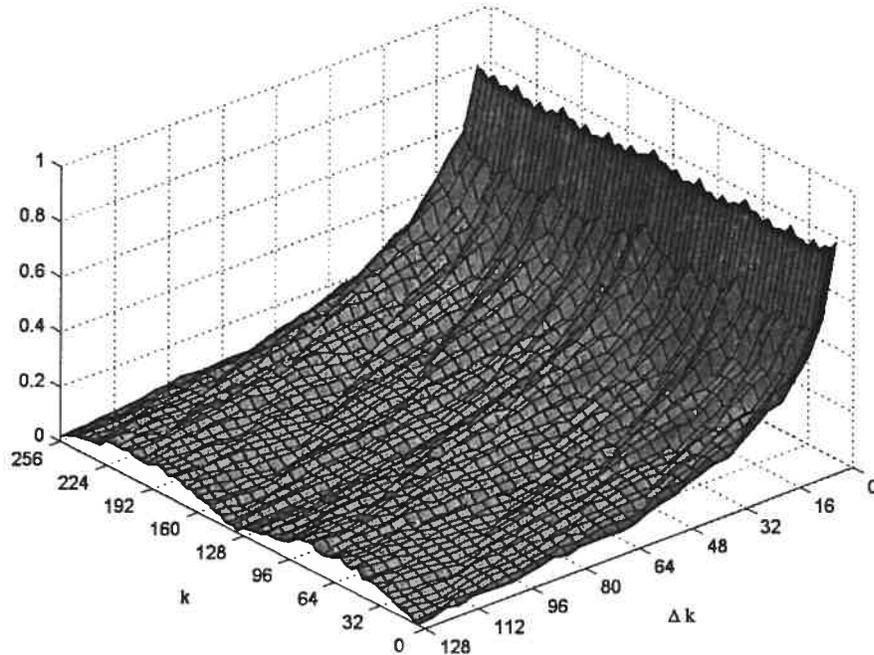


FIG. 3.12. Correlations between nodes as a function of absolute position of the node ( $k$ ) and inter-nodal distance ( $\Delta k$ ) for the two-parent non-stationary model ( $j=8$ ,  $\sigma_W = 1$ ). Artifacts are subtler than the ones on figure (3.11), yet still visible at any inter-nodal distance.

### 3.12. DISCUSSION AND CONCLUSIONS

A method has been proposed to ensure stationarity of multi-scale graphical models processes. The method relies on the necessary independence of second-order statistics over spatial translation. In the case of a multivariate gaussian joint distribution on the set of nodes, the condition is also sufficient and thus leads to strict-sense stationarity of the process rather than approximate stationarity as was the case in [50].

The construction of the model led us to an expression for the spatial covariance on the nodes of the graph given the parameter values of the model. Hence

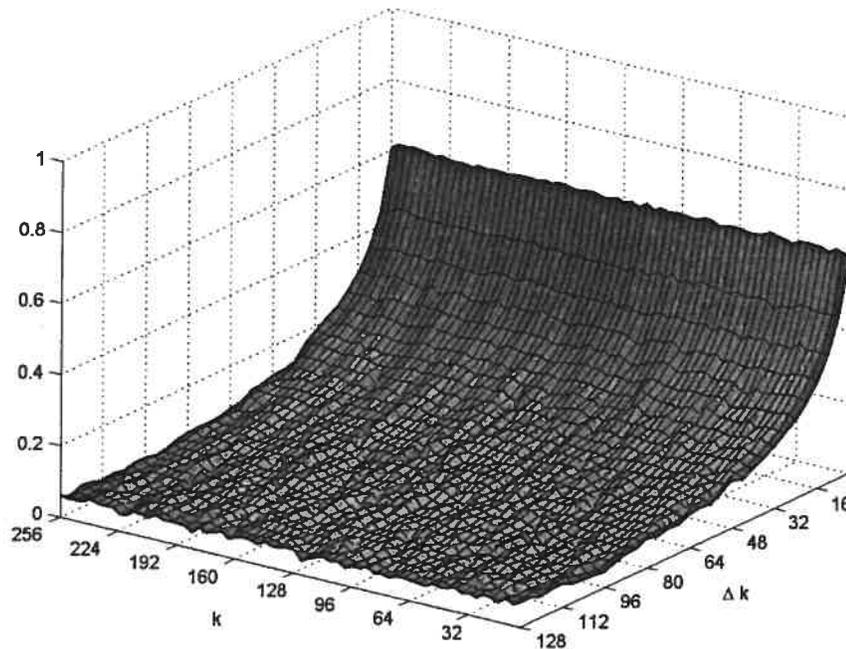


FIG. 3.13. Correlations for the stationary  $\alpha$ -model as a function of absolute position of the node ( $k$ ) and inter-nodal distance ( $\Delta k$ ). ( $j=8, \sigma_W = 1$ )

we have constructed a whole family of processes for which a set of parameters (the  $\alpha$ 's) can be estimated.

We have also shown that it is possible to ensure stationarity of higher order moments with additional conditions imposed on the model. The question whether a modification of the model could lead to strict sense stationarity for any marginal distribution on the nodes is the subject of ongoing work. For instance it can be shown that a dyadic model starting with a number of nodes (on the first level) which is not a power of two can be made strictly stationary regardless of the marginal distribution on the nodes; however the stationarity conditions in this case involve a strong delocalization of the spatial covariance function.

Regarding the problem of fully developed turbulence cascades, the proposed method allows not only to model scaling behavior on the marginal distribution

on the nodes, but also to model the form of the spatial covariance function. For instance a  $\log^2(\Delta k)$  behavior for the shape of the spatial covariance has been observed with great universality over different Reynolds number and experimental setup [52]. The method proposed here can be used to explore cascading processes that would explain the  $\log^2(\Delta k)$  behavior.

### 3.13. APPENDIX

PROOF OF PROPOSITION 3.8.1. Again we will proceed by induction. The  $p$ -order moments of the joint distribution of level  $j$  nodes can be expressed as :

$$C_1(c_{j,k}, c_{j,k+\Delta k_1}, \dots, c_{j,k+\Delta k_{p-1}}) = \sum_{\{K\}} \alpha_{2k_1-k}^{(j)} \alpha_{2(k_2+k_1)-(k+\Delta k_1)}^{(j)} \dots \alpha_{2(k_p+k_1)-(k+\Delta k_{p-1})}^{(j)} \cdot E[c_{j-1,k_1} \cdot c_{j-1,k_2+k_1} \cdot \dots \cdot c_{j-1,k_p+k_1}] \quad (180)$$

where  $C_1$  is defined as in (165), and  $K = \{k_i | i = 1 \dots p, k_i = 0 \dots 2^{j-1} - 1\}$ . The induction hypothesis allows us to write the expectation in the rightmost part of (180) as  $C^{(j-1,p)}(k_2, \dots, k_p)$ . We want this expression to be equal for  $k = 0$  and  $k = 1$  regardless of the parity of the  $\Delta k_i$ . First consider the case where all  $\Delta k_i$  are even. We write  $m_i = \Delta k_i / 2$ . For  $k = 0$ , we have

$$C_1 = \sum_{\{K\}} \eta_{k_1}^{(j)} \eta_{k_2+k_1-m_1}^{(j)} \dots \eta_{k_p+k_1-m_{p-1}}^{(j)} C^{(j-1,p)}(k_2, \dots, k_p)$$

Let us compute the Fourier transform along the  $m_i$  :

$$\hat{C}_1 = \sum_{\{M\}} e^{-2\pi i l \cdot m / 2^j} \cdot \sum_{\{K\}} \eta_{k_1}^{(j)} \eta_{k_2+k_1-m_1}^{(j)} \dots \eta_{k_p+k_1-m_{p-1}}^{(j)} C^{(j-1,p)}(k_2, \dots, k_p)$$

Changing all  $m_i$  to  $m_i + (k_i + k_1)$  :

$$\hat{C}_1 = \sum_{\{M\}} \sum_{\{K\}} e^{-2\pi i l \cdot m / 2^j} \cdot e^{-2\pi i l_1 (k_1+k_2) / 2^j} \cdot \dots \cdot e^{-2\pi i l_{p-1} (k_1+k_p) / 2^j} \cdot \eta_{k_1}^{(j)} \eta_{-m_1}^{(j)} \dots \eta_{-m_{p-1}}^{(j)} C^{(j-1,p)}(k_2, \dots, k_p)$$

leading to

$$\hat{C}_1 = \hat{\eta}^{(j)}\left[\sum_{i=1}^{p-1} l_i\right] \cdot \prod_{i=1}^{p-1} \hat{\eta}^{(j)}[-l_i] \cdot \hat{C}^{(j-1,p)}[l_1, \dots, l_{p-1}]$$

For  $k = 1$ , the exact same expression as above holds where all  $\eta$  are replaced with  $\xi$ . The second-order condition imposed  $\hat{\xi}[l] = \hat{\eta}[l] \cdot e^{i\theta[l]}$ . Hence for  $k = 1$ ,  $\hat{C}_1$  becomes

$$\hat{C}_1 = \hat{\eta}^{(j)}\left[\sum_i^{p-1} l_i\right] \cdot e^{i\theta[\sum_i l_i]} \cdot \prod_{i=1}^{p-1} \hat{\eta}^{(j)}[-l_i] \cdot e^{-i\theta[l_i]} \cdot \hat{C}^{(j-1,p)}[l_1, \dots, l_{p-1}] \quad (181)$$

Even though  $\theta[l] = -\pi l/2^j$  is a linear function of  $l$  over the interval  $[-2^{j-2} + 1, 2^{j-2} - 1]$ , the sum over  $i$  of the  $l_i$  can go out of this interval for  $p \geq 3$ . Thus (181) imposes a new condition on  $\hat{\eta}$  and  $\hat{\xi}$ . This condition is expressed as

$$\hat{\eta}^{(j)}\left[\sum_i^{p-1} l_i\right] \cdot \prod_{i=1}^{p-1} \hat{\eta}^{(j)}[-l_i] = \hat{\eta}^{(j)}\left[\sum_i^{p-1} l_i\right] \cdot e^{i\theta[\sum_i l_i]} \prod_{i=1}^{p-1} \hat{\eta}^{(j)}[-l_i] \cdot e^{-i\theta[l_i]}, \quad (182)$$

$\forall l_i \in [0, 2^{j-1} - 1], i = 1 \dots p - 1$ .

Let us now consider the case where one of the  $\Delta k_i$  is odd while the others remain even. Assume without loss of generality that  $\Delta k_1$  is the odd one. For  $k = 0$ , the effect is simply to change the  $\hat{\eta}[-l_1]$  into  $\hat{\xi}[-l_1]$ , such that in this case

$$\hat{C}_1 = \hat{C}_{1,\text{all even}} \cdot e^{i\theta[l_1]} \quad (183)$$

Conversely for  $k = 1$ , we have  $\hat{\xi}_-[l_1]$  changing into  $\hat{\eta}_-[l_1] \cdot e^{-2\pi i l_1/2^j}$ ; the “missing”  $\hat{\xi}_-[l_1]$  implies the absence of the  $e^{i\theta[l_1]}$  term in (181).  $\hat{C}_1$  becomes

$$\hat{C}_1 = \hat{C}_{1,\text{all even}} \cdot e^{-2\pi i l_1/2^j} \cdot e^{-i\theta[l_1]} \quad (184)$$

Replacing  $\theta[l] = -\pi l/2^j$  in this last expression proves that both (183) and (184) are the same, provided the condition (182) for the case where all  $\Delta k_i$  were even is respected. Hence we find no new condition on  $\alpha$  when considering one  $\Delta k_i$  to be odd.

The same exercise can be repeated for all combinations of even and odd  $\Delta k_i$ . Hence condition (182), along with (168, 169) is sufficient to provide stationarity of order  $p$  moments of the joint distribution of the nodes.

□

## Chapitre 4

---

# UNE INTERPRÉTATION DU MODÈLE DE SHE-LEVEQUE BASÉE SUR LA STATISTIQUE D'ORDRE

#### 4.1. REVUE DE LITTÉRATURE : LE MODÈLE DE SHE-LEVEQUE

L'approche multi-fractale permet de modéliser le comportement des incréments de vitesse et de la dissipation locale d'un flot turbulent, mais elle ne spécifie pas quelle doit être la forme fonctionnelle des exposants  $\zeta_p$  et  $\tau_p$ .<sup>1</sup> Un certain nombre de modèles multi-fractals ont été proposés au courant des dernières années, modèles qui correspondent plus ou moins bien aux valeurs expérimentales obtenues pour  $\zeta_p$  lorsque  $p \geq 4$ .

La majorité de ces modèles font appel à des paramètres ajustables dont le sens physique n'est pas établi clairement. En 1994, She et Leveque ont présenté un modèle de la dynamique turbulente qui ne fait appel à aucun paramètre ajustable (ou du moins pas directement) et qui prévoit un comportement de  $\zeta_p$  en très bon accord avec les données expérimentales [16]. Ce modèle fait appel au concept de *filaments de vorticit *, ph nom ne observ  aussi bien sur les donn es exp rimentales [55] que dans les simulations num riques. Les travaux de Vincent et Meneguzzi [56], qui d montrent la pr sence de filaments sous simulation num rique, ont confirm  que ce ph nom ne  tait bel et bien un effet de la dynamique de Navier-Stokes. Le mod le de She-Leveque cherche   expliquer l'ensemble du comportement non-lin aire des exposants  $\zeta_p$  et  $\tau_p$    partir de la pr sence de ces filaments.

Nous pr sentons ici ce mod le, avec la mise en garde suivante : bien que le mod le ne fasse appel   aucun param tre ajustable arbitraire il repose malgr  tout sur un ensemble d'hypoth ses, et plus particuli rement sur un postulat dont la justification en termes physiques n'est pas clairement  tablie. *L'article qui suit la pr sente section propose une justification statistique de ce postulat.* Cette justification statistique fait appel   certaines notions de statistique d'ordre qui sont  galement pr sent es dans cette section,   la suite du mod le de She-Leveque.

---

<sup>1</sup>Dans cette section on utilisera la notation  $\tau_p$  ou  $\tau(p)$  plut t que  $\tau(q)$  pour respecter la notation adopt e par She et Leveque

Le modèle de She-Leveque repose sur une caractérisation de la dissipation locale  $\varepsilon_l$  à travers une hiérarchie de *structures de fluctuation*  $\varepsilon_l^{(p)}$  définies par le rapport de moments successifs de la dissipation (voir l'expression 93), i.e.

$$\varepsilon_l^{(p)} = \frac{\langle \varepsilon_l^{p+1} \rangle}{\langle \varepsilon_l^p \rangle}. \quad (185)$$

Les deux extrémités de cette hiérarchie,  $\varepsilon_l^{(0)}$  et  $\varepsilon_l^{(\infty)}$ , sont associées respectivement à la dissipation moyenne  $\bar{\varepsilon}$  (celle introduite par Kolmogorov dans K41, que l'on note ici avec une barre pour éviter toute confusion) et aux structures filamentaires. Lorsque  $l \rightarrow 0$ , la dissipation moyenne demeure constante. Cependant, le comportement des structures filamentaires démontre une dépendance dans les échelles. C'est cette différence de comportement en échelle entre les différentes structures qui permet d'expliquer l'apparition du spectre multi-fractal associé à la forme non-linéaire des exposants  $\zeta_p$  et  $\tau_p$ . Les structures  $\varepsilon_l^{(p)}$  aux valeurs intermédiaires ( $0 < p < \infty$ ) correspondent à des "éléments de fluide" (ou structures) n'ayant pas tout à fait atteint la cohérence des structures filamentaires.

On justifie l'utilisation des structures de fluctuation  $\varepsilon_l^{(p)}$  ainsi : il a été observé de manière expérimentale que la présence de régions à dissipation locale élevée indique un degré de cohérence spatiale également plus élevé que lorsque la dissipation locale est faible. En fait, lorsque la dissipation locale est élevée, le fluide se déplace en un mouvement quasi-bidimensionnel autour de l'axe du filament. Puisque la quantité  $\varepsilon_l^{(p)}$  est proportionnelle à la présence de dissipation élevée, elle fournit un outil de différentiation de ces structures physiques plus ou moins bien organisées spatialement, i.e. avec un degré de cohérence indiqué par la valeur de  $p$ . Notons que le comportement dans les échelles de ces structures est intimement lié à l'exposant  $\tau_p$  à travers la relation suivante :

$$\varepsilon_l^{(p)} \sim \frac{l^{\tau_{p+1}}}{l^{\tau_p}} \sim l^{\tau_{p+1} - \tau_p}. \quad (186)$$

Le comportement dans les échelles de  $\varepsilon_l^{(p)}$  permet donc de récupérer la forme fonctionnelle de  $\tau_p$ , à une constante près. On s'intéresse ainsi au comportement dans les échelles de ces structures.

Tel que mentionné ci-haut, la structure  $\varepsilon_l^{(0)}$  correspond à la dissipation moyenne et ne dépend pas des échelles, d'après H3. Cependant,  $\varepsilon_l^{(\infty)}$  représente l'intensité des structures les plus intermittentes ; le modèle de She-Leveque fait l'hypothèse qu'il s'agit ici de structures filamentaires. Leur comportement dans les échelles est déduit d'un argument d'analyse dimensionnelle. La valeur  $\varepsilon_l^{(\infty)}$  prend la même dimension que la dissipation, i.e. une énergie par unité de temps :

$$\varepsilon_l^{(\infty)} \sim \frac{\delta E^\infty}{t_l}, \quad (187)$$

où  $t_l$  correspond toujours au temps de retournement des tourbillons de taille  $l$ . Par un raisonnement identique à celui de la section 3.1.3, on a :

$$t_l \sim (\bar{\varepsilon})^{-1/3} l^{2/3}. \quad (188)$$

La plus grande quantité d'énergie cinétique prête à être dissipée est de l'ordre de  $\delta v_0^2$ , i.e. la vitesse caractéristique aux grandes échelles ; une telle quantité d'énergie ne sera dissipée que suffisamment près d'une quasi-discontinuité du champ de vitesse, maximisant ainsi la dissipation. Mais cette image correspond justement à la structure filamentaire qui est formée d'un mouvement quasi bi-dimensionnel autour d'une structure unidimensionnelle dans laquelle on observe cette quasi-discontinuité du champ de vitesse. Le fluide en rotation autour du centre entraîne donc progressivement d'autres éléments de fluide moins cohérents, et dissipe l'essentiel de leur énergie au centre. L'énergie dissipée au centre de la structure filamentaire ne dépend donc pas de l'échelle, elle est simplement proportionnelle à l'énergie injectée aux grandes échelles. La dépendance dans les échelles de  $\varepsilon_l^{(\infty)}$  ne proviendrait que du temps de retournement, et on obtient :

$$\varepsilon_l^{(\infty)} \sim l^{-2/3}. \quad (189)$$

#### 4.1.1. La loi d'échelle universelle de She-Leveque

Le comportement dans les échelles des structures filamentaires permet de caractériser la forme fonctionnelle de  $\tau_p$ . En effet, pour  $p \rightarrow \infty$  on doit avoir :

$$\tau_{p+1} - \tau_p \rightarrow -2/3, \quad (190)$$

c'est-à-dire que

$$\tau_p \rightarrow -2p/3 + C_0. \quad (191)$$

$C_0$  correspond exactement à la co-dimension des structures les plus intermittentes, et puisque ce sont les filaments unidimensionnels on a alors  $C_0 = 2$ . On peut écrire pour  $\tau_p$  :

$$\tau_p = -2p/3 + 2 + f(p), \quad (192)$$

où  $f(p) \rightarrow 0$  lorsque  $p \rightarrow \infty$ .

Pour des valeurs de  $p$  intermédiaires, on considère que les structures associées ne sont pas parvenues à engendrer suffisamment de cohérence pour être filamenteuses ; elles ont alors une forme géométrique ressemblant à un ruban. Ces structures résultent d'une dynamique qui tend à former des filaments, mais dont la formation est contrainte par le désordre ambiant.

She et Leveque proposent l'hypothèse suivante : spatialement, les structures d'ordre  $p$  et d'intensité  $\varepsilon^{(p)}$  sont généralement entourées de structures d'ordre  $p - 1$  et d'intensité légèrement plus faible. Il en va de même dans la dimension temporelle, où la formation de structures d'ordre  $p$  est précédée de la formation de structures d'ordre  $p - 1$ . Partant de cette dynamique de formation des structures, She et Leveque suggère qu'il existe une relation liant le comportement dans les échelles des structures d'ordre  $p$  à celui des structures d'ordre  $p - 1$ . She et Leveque *postulent* que cette relation prend la forme suivante :

$$\varepsilon_l^{(p+1)} = A_p \varepsilon_l^{(p)\beta} \varepsilon_l^{(\infty)1-\beta}, \quad 0 < \beta < 1, \quad (193)$$

de sorte que  $\varepsilon_l^{(p)} \rightarrow \varepsilon_l^{(\infty)}$  lorsque  $p \rightarrow \infty$ .  $A_p$  ne dépend pas de l'échelle mais peut dépendre de  $p$  et de la configuration expérimentale. C'est cette dernière expression que l'on nomme *loi d'échelle universelle de She-Leveque*. Dans l'article qui suit cette section on expliquera comment l'on peut déduire la forme fonctionnelle de  $\tau_p$  à partir de cette loi d'échelle, qui est donnée par

$$\tau_p = -2p/3 + 2[1 - (2/3)^p]. \quad (194)$$

Auparavant, nous nous attardons tel qu'annoncé à certaines notions fondamentales de statistique de maxima qui n'ont pas été abordées au chapitre 2.1.

#### 4.1.2. Distributions asymptotiques de maxima

Tel qu'annoncé dans le chapitre 1, le corps du présent chapitre propose une interprétation du modèle de She-Leveque permettant de déduire leur loi d'échelle universelle plutôt que de la postuler, grâce à une contrainte statistique qui n'avait pas été observée par She et Leveque dans leurs travaux originaux.

Cette contrainte statistique fait appel à certaines notions de *statistique d'ordre*; on entend librement par statistique d'ordre toute statistique obtenue par ordonnancement d'un échantillon. Le maximum, le minimum, la médiane, les quantiles et l'écart maximal (la différence entre le maximum et le minimum) constituent des statistiques d'ordre. Ce type de statistique s'avère utile dans plusieurs contextes : la médiane et l'écart maximal, par exemple, peuvent servir d'estimateurs simples pour la moyenne et la variance. Certains estimateurs de tendance centrale accordent plus ou moins d'importance aux extrêmes, comme c'est le cas pour la moyenne tronquée qui rejette les valeurs maximale et minimale d'un échantillon en les considérant hors-norme. Cet estimateur entre aussi dans la grande famille des statistiques d'ordre.

Dans le chapitre 2, nous avons déjà abordé une statistique d'ordre puisque nous nous sommes attardés à la statistique de maxima (ou de minima) locaux d'un champ stochastique. Il convient de marquer ici la distinction entre maxima

locaux et maximum global d'un champ défini sur un domaine donné. Bien que le maximum local semble analogue au maximum global en ce qu'il constitue la valeur la plus élevée du champ dans une région donnée, il ne possède généralement pas les mêmes propriétés statistiques. En effet un maximum local, par définition, peut n'être maximal que dans une région de taille infinitésimale incluant le maximum local lui-même. Dans le cas d'un champ possédant certaines propriétés de régularité, i.e. l'existence de dérivées de premier et second ordre comme on l'a supposé pour tous les résultats obtenus au chapitre 2, cette petite région entourant le maximum local pourra ne contenir que des points où la valeur du champ est très fortement corrélée à celle du maximum local. Dans le cas d'un maximum global, si le domaine est suffisamment grand par rapport à l'échelle caractéristique de corrélation spatiale du champ, on pourra considérer que le maximum global correspond à la valeur maximale d'un grand nombre de variables aléatoires indépendantes. Cette distinction est primordiale.

Dans ce chapitre, nous nous intéressons spécifiquement à la notion de maximum global. Nous utiliserons librement cette idée de "région suffisamment grande" sans la formaliser d'avantage afin de pouvoir faire appel à des résultats connus concernant la statistique de valeur maximale pour un grand échantillon tiré de variables aléatoires indépendantes et identiquement distribuées. Une justification plus formelle de cette approche se trouve dans [13, 57, 58, 59].

La question qui nous intéresse est donc la suivante : quelle est la distribution asymptotique de la valeur maximale d'un échantillon de taille  $N$  tiré de variables aléatoires indépendantes et identiquement distribuées, lorsque  $N \rightarrow \infty$ ? Ce problème est difficile. Les contributions majeures nous viennent de Dodd [60], von Mises [61, 62], Fréchet [63], Fisher et Tippett [64], de Finetti [65], Gumbel [66], Gnedenko [67] et de Haan [68]. La réponse à cette question figure parmi les grands résultats classiques en statistique : si une distribution limite existe, i.e. qu'elle est bien définie lorsque proprement normalisée, alors elle doit prendre l'une des trois formes asymptotiques suivantes pour sa fonction de répartition (i.e. la primitive

de sa fonction de distribution) :

$$\Lambda_1(x) = \exp(-x^{-\alpha}), \quad x > 0; \quad (195)$$

$$\Lambda_2(x) = \exp(-(-x)^\alpha), \quad x \leq 0; \quad (196)$$

$$\Lambda_3(x) = \exp(-e^{-x}), \quad -\infty < x < \infty, \quad (197)$$

où  $\alpha > 0$ , et où  $x$  admet un changement de variable impliquant une translation et un facteur d'échelle, comme c'est le cas pour la loi normale, par exemple.

Nous présentons ici un argument intuitif expliquant le résultat ci-haut afin d'en favoriser l'appréciation.

On peut considérer que la valeur maximale dans un échantillon de taille  $N = mn$  est identique à la valeur maximale de  $n$  maxima obtenus en segmentant l'échantillon complet en  $n$  sous-ensembles regroupant  $m$  valeurs chacun. Lorsque  $m \rightarrow \infty$ , on aura à la fois l'échantillon complet et n'importe quel sous-ensemble qui devront converger vers la même distribution limite  $\Lambda(x)$ . Notons que  $n$  demeure fini. Remarquons aussi que la fonction de répartition du maximum de  $n$  valeurs (possédant elles-mêmes une fonction de répartition  $F(x) \equiv \Pr(X \leq x)$ ) est donnée par  $F(x)^n$ , puisque chacune des  $n$  valeurs doit être plus petite que  $x$  afin que le maximum le soit aussi. Cela n'est vrai que pour des variables indépendantes, comme c'est le cas ici. La distribution asymptotique, si elle existe, doit alors respecter la condition suivante :

$$\Lambda(a_n x + b_n)^n = \Lambda(x). \quad (198)$$

Cette dernière condition est en tout point analogue au théorème central limite pour une somme de variables aléatoires. En effet dans ce cas les paramètres  $a_n$  et  $b_n$  sont liés aux valeurs de moyenne et de variance de la loi asymptotique normale.

Les solutions à cette équation dépendent essentiellement des différentes valeurs pouvant être prise par  $a_n$ . Dans le cas où  $a_n \neq 1$ , on remarque que  $a_n x + b_n = x$  lorsque  $x = x_0 = b_n / (1 - a_n)$ . À cette valeur, on a alors :

$$\Lambda(x_0)^n = \Lambda(x_0), \quad (199)$$

ce qui implique que  $\Lambda(x_0)^n$  vaut 0 ou 1. Si la distribution asymptotique existe,  $x_0$  n'est qu'une constante qui peut être fixée à 0 sans perte de généralité. Dans ce cas pour  $a_n > 0$  on a :

$$\Lambda(x) = 0 \text{ si } x \leq 0, \Lambda^n(a_n x) = \Lambda(x) \text{ si } x > 0. \quad (200)$$

Pour  $a_n < 0$ , on a de manière semblable :

$$\Lambda(a_n x)^n = \Lambda(x) \text{ si } x \leq 0, \Lambda(x) = 0 \text{ si } x > 0. \quad (201)$$

Puis dans le cas où  $a_n = 1$ , on a simplement :

$$\Lambda(x + b_n)^n = \Lambda(x). \quad (202)$$

Les solutions à ces trois cas de figures correspondent précisément aux trois expressions énoncées ci haut (195).

Les conditions formelles édictant laquelle de ces trois solutions apparaît dépendamment de la distribution sous-jacente des variables aléatoires dont l'échantillon est obtenu seront énoncées dans la section 4.5.

## 4.2. PRÉSENTATION DE L'ARTICLE

Les deux chapitres précédents présentent des sujets sans relation apparente, le premier se concentrant sur la notion de statistique d'excédence et d'extréma, le deuxième concernant l'approche multi-fractale en turbulence pleinement développée. C'est dans ce troisième chapitre que les éléments présentés précédemment seront mis à l'œuvre conjointement.

Dans la section 4.1 nous avons présenté le modèle de She-Leveque en turbulence pleinement développée. Une forme fonctionnelle pour l'exposant d'échelle de la dissipation,  $\tau(p)$ , y est obtenue à travers la formulation d'un postulat concernant une relation d'échelle universelle liant les structures de fluctuation  $\varepsilon_l^{(p)}$  telles que définies par She et Leveque [16]. Or ce postulat constitue une hypothèse supplémentaire du modèle et apparaît comme un ajout *ad hoc* guidé par les résultats expérimentaux. Bien que la formulation de ce postulat soit partiellement justifiée par une explication physique concernant un processus d'entraînement (*entrainment process* dans la suite du chapitre) devant lier les structures d'ordre successifs, une infinité de formes fonctionnelles auraient tout aussi bien pu être valables d'après cet argument.

L'objet de ce chapitre est de revoir en détail le processus d'entraînement présenté par She et Leveque et de montrer que ce processus impose des contraintes statistiques au modèle qui n'ont pas été remarquées dans leurs travaux originaux, ni par la suite d'ailleurs. Ce processus d'entraînement est à la base de l'argumentation de She et Leveque qui les a amené à postuler leur loi d'échelle universelle. Nous montrerons que le processus d'entraînement est analogue à une transformation par remplacement maximal ; cette transformation fait intervenir des notions de statistique de maxima. Elle impose une contrainte forte sur la distribution asymptotique de l'ordre des structures de fluctuation. Un résultat classique en statistique d'ordre indique qu'il n'existe que trois distributions asymptotiques pour la valeur maximale d'un échantillon lorsque le nombre d'éléments dans l'échantillon tend vers l'infini [69]. L'une de ces trois distributions asymptotiques joue un rôle

prépondérant de par ces contraintes moins fortes sur la distribution sous-jacente, et c'est de cette distribution qu'il est possible de déduire la loi universelle de She-Leveque, qui mène à une prédiction sur la forme fonctionnelle des exposants  $\tau(p)$  et  $\zeta(p)$ .

Nous présentons aussi les formes fonctionnelles obtenues pour les deux autres distributions asymptotiques, dont un cas particulier est en partie analogue à un modèle en cascade log-normale. Un modèle semblable a été postulé en 1962 par Kolmogorov et Obukhov, quoi que dans notre contexte certaines difficultés associées à leur modèle sont évitées.

# An Interpretation of the She-Leveque Model based on Order Statistics

PHILIPPE ST-JEAN, *Université de Montréal*

## Abstract

We present an interpretation of the She-Leveque model in fully developed turbulence based on order statistics. Turbulent behavior at large values of the Reynolds number is often studied through the scaling behavior of moments of the distribution of the velocity differences and of the energy dissipation. The present interpretation leads to a derivation of the scaling exponents  $\zeta_p$  and  $\tau_p$  of these moments, without any postulate about a universal relation over the fluctuation structures such as the one used by She and Leveque. The interpretation is based on the fact that the hierarchy of fluctuation structures imposes constraints on the distribution of the order  $p$  of the structures. Hence the order  $p$  itself is seen as a random variable. As proposed by She and Leveque the hierarchy of the structures is such that the structure of largest order affects locally lower order structures through an entrainment process. This phenomenon leads to the Fisher-Tippett law, one of three asymptotic distributions for the extreme value of a random sample of  $n$  as  $n \rightarrow \infty$ .

## 4.3. INTRODUCTION

For the last sixty years or so the phenomenological study of fully developed turbulence has evolved from the celebrated work of Kolmogorov [1] to the recent general acceptance that significant discrepancy from his early work exists, the effect known as intermittency. This effect only manifests itself when rare outlier

events occur, forcing the need for gigantic statistics in order to be established properly. The study of scaling exponents (of high-order moments) of the velocity increments constitutes the obvious path in quantifying the presumed discrepancies. The scaling behavior of the longitudinal velocity increments,  $\delta v_l = v(x+l) - v(x)$ , is expressed as

$$\langle |\delta v_l|^p \rangle \sim l^{\zeta(p)}. \quad (203)$$

The Kolmogorov refined similarity hypothesis [41] links the behavior of  $\zeta(p)$  to the scaling behavior of the energy dissipation over a ball of size  $l$ ,  $\varepsilon_l$ , through the following relation :

$$\langle |\delta v_l|^p \rangle \sim \langle \varepsilon_l^{p/3} \rangle l^{p/3}. \quad (204)$$

Defining  $\tau(p)$  as the scaling exponent of  $\varepsilon_l$ , the relation becomes

$$\zeta(p) = p/3 + \tau(p/3). \quad (205)$$

The original work of Kolmogorov [1] (later referred to as K41) led to the prediction that  $\varepsilon_l$  was independent of  $l$ , with the obvious consequences that  $\tau(p) = 0$  and  $\zeta(p) = p/3$ . It is the observed discrepancies from this behavior of the scaling exponents that is referred to as *intermittency* [70, 56]. The term indicates the fact that non-linear corrections to  $\zeta(p)$  originate from intermittent behavior of the energy dissipation.

Over the past thirty years a number of phenomenological models have been proposed in order to account for non-linear corrections to  $\zeta(p)$  [37, 42]. Perhaps the most “popular” model was proposed ten years ago by She and Leveque [16]. It won notoriety not only because it best fits experimental data for large values of  $p$  [39], but mostly because it relies on an interesting physical picture that leads in a somehow natural fashion to an explicit form for the scaling exponents, namely  $\tau(p) = -2p/3 + 2[1 - (2/3)^p]$ . Dubrulle [71] put this model in the context of log-infinitely divisible cascade models [71, 72, 73, 74] and realized that this form of  $\tau_p$  corresponds to a log-Poisson distribution along the steps of the cascade. This

was presented independently by She and Waymire [74]. However the She-Leveque model relies on a *ad hoc* postulate of this explicit form. This postulate can be interpreted as the symptom of a missing link between a valid physical model (which includes constraints that have not been recognized yet) and a large set of experimental results.

The intent of this study is to provide this missing link through a statistical argument from which the explicit form for the scaling exponents can be derived rather than postulated.

In section 4.4 we give a brief presentation of the She-Leveque model. Section 4.5 investigates the underlying statistical constraint imposed by the She-Leveque model that leads to the explicit form for  $\tau(p)$  and  $\zeta(p)$ . In section 4.6 a more general result is obtained from fewer hypothesis than the She-Leveque model proposed.

#### 4.4. THE SHE-LEVEQUE MODEL

The model proposed by She and Leveque [16] characterizes the energy dissipation field  $\varepsilon_l$  using a hierarchy of fluctuation structures  $\varepsilon_l^{(p)}$  defined by the ratio of successive moments of  $\varepsilon_l$ , i.e.

$$\varepsilon_l^{(p)} = \frac{\langle \varepsilon_l^{p+1} \rangle}{\langle \varepsilon_l^p \rangle} \quad (206)$$

The two extreme structures  $\varepsilon_l^{(0)}$  and  $\varepsilon_l^{(\infty)}$  correspond respectively to the mean fluctuation structure  $\bar{\varepsilon}$  and to filaments [75], in view of the two-fluid model [76]. Intermediate values of  $p$  thus represent more or less organized fluid structures, perhaps ribbon or sheetlike.

It is argued in the She-Leveque model that the filamentary structures  $\varepsilon_l^{(\infty)}$  scale like  $l^{-2/3}$ . From the definition of  $\tau(p)$ , this amounts in stating that as  $p \rightarrow \infty$ ,

$$\tau(p+1) - \tau(p) \rightarrow -2/3, \quad (207)$$

hence  $\tau(p) \rightarrow -2p/3 + C$ . Using Legendre transform the constant  $C$  is readily seen as the codimension [37] of the high intensity structures (filaments), hence  $C = 3 - 1 = 2$ .

Perhaps the most important part of the model lies in the assumed interaction between structures of different order. She-Leveque propose that the fluctuation structures  $\varepsilon_l^{(p)}$  form a hierarchy in the sense that high order structures entrain lower order ones nearby. This phenomenon is explained in greater detail in the next section.

From the supposed hierarchy of those structures, they postulate the following relation between structures of adjacent order :

$$\varepsilon_l^{(p+1)} = A_p \varepsilon_l^{(p)\beta} \varepsilon_l^{(\infty)1-\beta} \quad (208)$$

From (206,208), we get the following expression for  $\tau_p$ , the scaling exponent of  $\langle \varepsilon_l^p \rangle$  :

$$\tau_{p+2} - (1 - \beta)\tau_{p+1} + \beta\tau_p + \frac{2}{3}(1 - \beta) = 0. \quad (209)$$

The most intermittent structures  $\varepsilon_l^{(\infty)}$  scales like  $l^{-2/3}$ , which in returns implies that  $\tau_p$  must be of the form  $\tau_p = -2/3p + 2 + f(p)$ , where  $f(\infty) = 0$ . We can then rewrite (209) as

$$f(p+2) - (1 + \beta)f(p+1) + \beta f(p) = 0. \quad (210)$$

This last expression has  $f(p) = \alpha\beta^p$  as its only non-trivial solution, and boundary conditions on  $\tau_p$  (namely  $\tau_0 = \tau_1 = 0$ ) leads to

$$\tau_p = -\frac{2}{3}p + 2 \left[ 1 - \left( \frac{2}{3} \right)^p \right]. \quad (211)$$

Although the derivation of the form of  $\tau_p$  starting with (208) is straightforward, the latter was postulated in a somewhat *ad hoc* fashion. Bear in mind that

this form of  $\tau_p$  corresponds to a log-Poisson cascade [71, 74]. However the form of  $\tau_p$  was obtained by She and Leveque without requiring the *cascade through scales* picture.

In the next section we propose a statistical interpretation of the She-Leveque model which shows that the underlying physics of the model forces the universal scaling behavior without having to formulate a postulate regarding its explicit form, given by (208). Still no explicit formulation of the cascade picture is required, staying in the same line as the She-Leveque model.

#### 4.5. STATISTICAL BEHAVIOR OF THE HIERARCHICAL DYNAMICS

One fundamental aspect of the She-Leveque model lies in their proposed hierarchical dynamics of the fluctuation structures  $\varepsilon_l^{(p)}$ . She and Leveque observe that the hierarchy of structures is the result of an entrainment process whereas high-order structures “constantly entrain surrounding, less ordered fluid elements”. Hence the dynamics that they propose is one of the *survival of the fittest* over the fluctuation structures  $\varepsilon_l^{(p)}$ .

We begin by relating the distribution of the order  $p$  of the structures to their intensities. For this we point out that a structure of order  $p$  represents a region where the fluid motion demonstrates a certain level of coherence [77] such that it contributes significantly to the intensity  $\varepsilon_l^{(p)}$  as defined by the ratio of moments  $p + 1$  and  $p$  of the dissipation.

As presented in [16], the intensity  $\varepsilon_l^{(p)}$  of a structure of order  $p$  has the same dimensions as the dissipation, i.e. energy over time. Hence this intensity is proportional to the quantity of energy available for dissipation through the structure. The order  $p$  quantifies the coherence of the structure. The maximum energy available, dissipated through structures for which  $p \rightarrow \infty$ , is bounded and given by the energy typical of large scale velocity fluctuations. Consequently the probability of presence of structures of order smaller or equal to  $p_0$  at any point  $(x, t)$  at

scale  $l$  is equal to the proportion of energy available to these structures, i.e. :

$$\Pr(p_l(x, t) \leq p_0) = \frac{\varepsilon_l^{(p_0)}}{\varepsilon_l^{(\infty)}}. \quad (212)$$

This is in line with the fact that  $\varepsilon_l^{(p)}$  is by definition a bounded, non-decreasing sequence of intensities over  $p$  [16]. In measure-theoretic terms,  $\varepsilon_l^{(p)}$  defines a *mass* associated to the variable  $p$ .

We now come back to the *survival of the fittest* dynamics. The model includes an explicit statement on the behavior of this entrainment process : since large order structures quickly absorb low order ones nearby, it is very unlikely that any structure of a given order  $p$  can be found spatially close to a structure of order much larger than  $p$ . Similarly in time it is very unlikely that the order  $p$  of a structure takes a small value when there was a large order structure around an instant before.

Hence the probability of having a  $p$ -order structure at a given point  $(x, t)$  in space and time is proportional to the probability of the maximum structure order found in a spatial and timewise neighborhood (of  $(x, t)$ ) being close to  $p$ . This imposes strong constraints on the possible statistical distribution of the structures.

Precisely it means that the following can be written about the distribution :

$$\Pr(p_l(x, t) = p_0) \sim \Pr\left(\max_{(x', t') \in B(x, t)} p_l(x', t') = p_0\right) \quad (213)$$

where  $\{(x', t') \in B(x, t)\}$  is some neighborhood around  $(x, t)$ . Looking at  $p_l(x, t)$  as a random field, the constraint then tells us that the pointwise probability distribution function (pdf) of  $p_l(x, t)$  must be invariant when replaced by the pdf of its maximum value in a neighborhood of any point  $(x, t)$ . Again it means that some time before there could not have been structures of order much higher than  $p$  around. However it allows for the entrainment process. This is in perfect agreement with the She-Leveque model. Note that random injection of energy alters this process ; still if a steady-state is reached, then it should obey the

constraint described above. Let us now see what actual form of the pdf of  $p$  can be expected.

It is a well-known result in order statistics that there are only three forms of asymptotic distributions that respect the later condition when the number of elements amongst which the maximum is taken is arbitrarily large. Indeed this is just the case here. Not because it can be assumed that the neighborhood is of a very large extent spatially (which would require that a very large number of structures can affect a given structure at a fixed time), but simply because the effect of taking maxima values as time goes by leads to the asymptotic distribution, i.e. the steady-state. Formally, we have the following result about these distributions :

**PROPOSITION 4.5.1.** *The cumulative distribution function (cdf) of the extremal value in a neighborhood of  $(x_0, t_0)$  must approach (under some mild conditions expressed later in this section) one of the three form of asymptotic distribution for the largest value in a random sample of  $N$  as  $N \rightarrow \infty$  [69], namely*

$$\Lambda_1(p) = \exp(-p^{-\alpha}), \quad p > 0; \quad (214)$$

$$\Lambda_2(p) = \exp(-(-p)^\alpha), \quad p \leq 0; \quad (215)$$

$$\Lambda_3(p) = \exp(-e^{-p}), \quad -\infty < p < \infty, \quad (216)$$

where  $\alpha > 0$ , and  $p$  is defined up to translation and scaling parameters.

We give here a short intuitive derivation of this result for self-containment sake, but refer the reader to [69] for a thorough discussion, and to [67] for a formal proof. We follow closely the discussion given in [69], which was originally introduced by Fisher and Tippett [64].

Consider taking the maximum value in a sample of  $N = mn$ ; this amounts in taking maximum values in  $n$  different subset (each containing  $m$  samples), then taking the maximum of those  $n$  values. Obviously we need both distributions to converge to the same form  $\Lambda(p)$  as  $m \rightarrow \infty$ . Since  $n$  remains finite, we use

the fact that the cdf of the maximum value of  $n$  i.i.d. random variables with cdf  $F_p(p_0) \equiv \Pr(p \leq p_0)$  is given by  $(F_p(p_0))^n$ , as each of the  $n$  values have to be smaller than  $p_0$  for the maximum to be too. Therefore the asymptotic distribution  $\Lambda(p)$  must be such that

$$\Lambda^n(a_n p + b_n) = \Lambda(p). \quad (217)$$

These asymptotic distributions play the same role for order statistics as the normal distribution for the sum of random variables, i.e. it is in most aspects analogous to the central limit theorem. Here  $a_n > 0$  and  $b_n$  are suitably chosen normalisation constants, just like the sum of  $n$  i.i.d. variables must be centered, and normalised by  $n$  in order to converge to a standard normal distribution. It is now easily checked that the three forms of  $\Lambda(p)$  presented above are the only solutions to (217).

When some minimal conditions on the underlying distribution are respected, the extremal value will converge to one of the three forms above; there exist known conditions on the underlying distribution (from which the maximum is taken) stating which one of the three forms is the asymptote, although in *most* cases it converges to the third form as conditions for convergence to the first two forms are more restrictive. These conditions were obtained by Gnedenko [67] and are given here in terms of the cdf  $P(x)$  of the underlying distribution.

PROPOSITION 4.5.2.  *$P(x)$  belongs to the domain of attraction of  $\Lambda_1$  if and only if*

$$\lim_{x \rightarrow \infty} \frac{1 - P(x)}{1 - P(kx)} = k^\alpha, \quad (218)$$

for every  $k > 0$ .

PROPOSITION 4.5.3.  *$P(x)$  belongs to the domain of attraction of  $\Lambda_2$  if and only if there exists an  $x_0$  such that*

$$P(x_0) = 1 \quad \text{and} \quad P(x_0 - \epsilon) < 1 \quad (219)$$

for every  $\epsilon > 0$ ,  
and if and only if

$$\lim_{x \rightarrow 0^-} \frac{1 - P(kx + x_0)}{1 - P(x + x_0)} = k^\alpha, \quad (220)$$

for every  $k > 0$ .

In essence, these conditions are equivalent to a power-law behavior of the tail of the cdf. We note that distributions which fall into the  $\Lambda_1$  domain are unlimited on the right, while they are limited for  $\Lambda_2$ . Distributions belonging to the  $\Lambda_3$  domain can occur in both cases; the formal conditions for  $\Lambda_3$  are more intricate, but a sufficient condition in the case where  $P(x)$  is unlimited on the right is given here.

PROPOSITION 4.5.4.  *$P(x)$  belongs to the domain of attraction of  $\Lambda_3$  if it is less than 1 for every finite  $x$ , it is twice differentiable for at least every  $x$  greater than some  $x'$  and it respects :*

$$\lim_{x \rightarrow \infty} \frac{d}{dx} \left[ \frac{1 - P(x)}{P'(x)} \right] = 0. \quad (221)$$

It is trivially checked that any exponential form belongs to the domain of attraction of  $\Lambda_3$ .

Figure 4.1 shows the evolution of the distribution of the maximum of  $N$  gaussian variables with mean  $\mu = 0$  and variance  $\sigma^2 = 1$  as the number of variables  $N$  grows. The gaussian distribution is one example of an underlying distribution that has  $\Lambda_3$  as its asymptote for the maximum value in a sample.

$\Lambda_1(p)$  and  $\Lambda_2(p)$  are known as Weibull and Fréchet distributions, respectively; their stretched-exponential behavior is reminiscent of the pdf for velocity increments at small scales [37]. We begin by considering only the third form  $\Lambda_3(p)$ , and will come back to the two others later on.

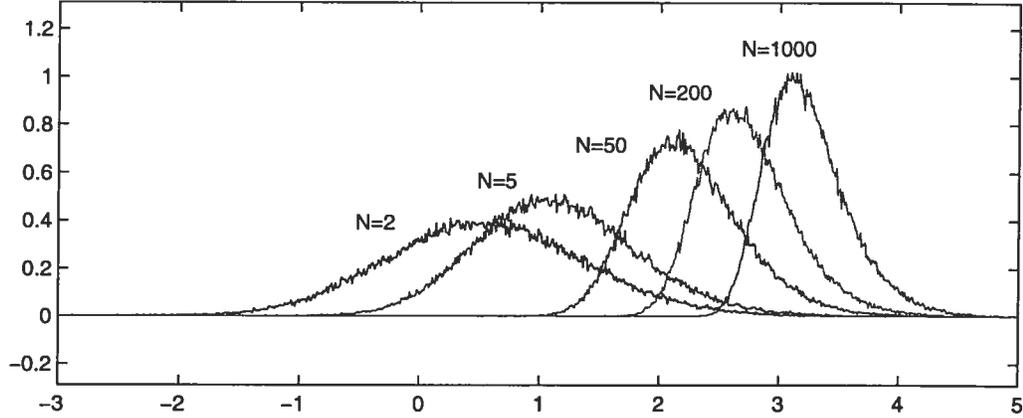


FIG. 4.1. Histograms of the maximum value of  $N$  gaussian variables for  $N = 2, 5, 50, 200, 1000$ . For  $N = 2$  the pdf is almost gaussian, as expected, while it tends to the asymmetrical  $\Lambda_3$  form for higher values of  $N$ . Each curve involves 100,000 realizations.

#### 4.5.1. Obtaining the She-Leveque universal scaling law from the Fisher-Tippett distribution

We now present a derivation of the She-Leveque universal scaling law from the constraint on the distribution of structures that we just derived.

$\Lambda_3(p)$  is known as the Gumbel distribution; it is a special case of the larger class of Fisher-Tippett distribution, with additional parameters for translation and scaling of its argument. Hence in this case the cdf of  $p$  is given by  $\Lambda_3(ap-b) = \exp(-e^{-(ap-b)})$ . The intensity of  $p$ -order structures can then be written as

$$\varepsilon_l^{(p)} = \varepsilon_l^{(\infty)} \cdot \exp(-e^{-(ap-b)}) \quad (222)$$

Note that  $a$  and  $b$  are allowed to depend on  $l$ .

We can now derive the following expression for  $\varepsilon_l^{(p+1)}$  with respect to  $\varepsilon_l^{(p)}$ :

$$\varepsilon_l^{(p+1)} = \varepsilon_l^{(\infty)} \cdot \exp(-e^{-(a(p+1)-b)}) \quad (223)$$

$$= \varepsilon_l^{(\infty)} \cdot \exp(-e^{-(ap-b)} e^{-a}) \quad (224)$$

$$= (\varepsilon_l^{(\infty)})^{1-e^{-a}} \cdot (\varepsilon_l^{(p)})^{e^{-a}} \quad (225)$$

But this is equivalent to (208) with  $\beta = e^{-a}$  and  $A_p = 1$ . Since  $\beta = 2/3$  in the context of the She-Leveque model [16], we get  $a = -\log(2/3)$ , independent of  $l$ . Of course  $A_p$  is also independent of  $l$ . Hence we have obtained (208) from a simple order statistics argument, rather than having to postulate it. Finally we also have that  $\varepsilon_l^{(0)}$  must not depend on  $l$ , such that

$$b = \log(-\log(l^{2/3})) = \log(2/3) + \log(-\log(l)). \quad (226)$$

So the dependence of  $\varepsilon_l^{(p)}$  over  $l$  comes from  $\varepsilon_l^{(\infty)}$  and  $b$ . Here  $l$  is non-dimensionalized by division by the integral scale  $l_0$ , such that  $l$  runs from 1 to 0.

It is worth emphasizing the fact that no new assumption have been introduced in the model; the asymptotic distribution was derived from a statistical interpretation of the She-Leveque model and its entrainment process. However it allows to get a much better understanding of the specific form of  $\tau(p)$ .

#### 4.6. GENERAL STUDY OF THE THREE ASYMPTOTIC FORMS

We have just showed that we can obtain the She-Leveque universal scaling law without having to postulate it, but rather from first principles based on order statistics and from some fundamental hypothesis forming the model. The main hypothesis was that there exists an entrainment process over the order  $p$  of the structures  $\varepsilon_l^{(p)}$ . There are however supplementary hypothesis such as the scaling behavior of  $\varepsilon_l^{(\infty)}$ , and the scaling constraints  $\tau(0) = \tau(1) = 0$ . We have also made the (unjustified) additional hypothesis that  $\Lambda_3$  was the proper distribution of the extreme. In this section we will obtain more general results relying solely on the entrainment process, and on the independence of  $\bar{\varepsilon} = \varepsilon_l^{(0)}$  over scale, for which there is strong experimental evidence.

In the She-Leveque model, (208) was used to derive the behavior of exponent  $\tau(p)$ ; the She-Leveque model did not provide any explicit expression for the intensity of structures  $\varepsilon_l^{(p)}$ , such as the one we just derived, namely (222). From the

latter we can write :

$$\tau(p+1) - \tau(p) = \lim_{l \rightarrow 0} \frac{\log(\varepsilon_l^{(p)}/\bar{\varepsilon})}{\log(l)} \quad (227)$$

$$= \lim_{l \rightarrow 0} \frac{\log(\varepsilon_l^{(\infty)}/\bar{\varepsilon}) + \log \Lambda_i(ap-b)}{\log(l)}, \quad (228)$$

where  $i = 1, 2, 3$ . Let us write

$$C_\infty \equiv \lim_{l \rightarrow 0} \frac{\log(\varepsilon_l^{(\infty)}/\bar{\varepsilon})}{\log(l)}, \quad (229)$$

with  $C_\infty = -2/3$  for the She-Leveque model. We will take a look at all three forms, starting from the last.

For the  $\Lambda_3$  form we have

$$\tau(p+1) - \tau(p) = C_\infty + \lim_{l \rightarrow 0} \left( \frac{e^{-(ap-b)}}{\log(l)} \right). \quad (230)$$

Using independence of  $\bar{\varepsilon}$  over scale we get

$$b = \log(C_\infty \cdot \log(l)). \quad (231)$$

Finally from the definition of  $C_\infty$ ,

$$\lim_{p \rightarrow \infty} \left( \lim_{l \rightarrow 0} \left( \frac{-e^{-(ap-b)}}{\log(l)} \right) \right) = 0. \quad (232)$$

A sufficient condition for this last equation to be true is to have  $a$  independent of  $l$ . In this case, we have

$$\tau(p+1) - \tau(p) = C_\infty \cdot (1 - e^{-ap}), \quad (233)$$

which leads to the following expression for  $\tau(p)$  :

$$\tau(p) = C_\infty \cdot \left( p - \frac{e^{-ap}}{e^{-a} - 1} \right) + C_0, \quad (234)$$

where  $C_0$  is again analogous to the codimension of highly dissipative structures. Assuming that  $\tau(1) = \tau(0) = 0$  imposes that  $a = -\log(C_\infty/C_0 + 1)$ . A final expression for  $\tau(p)$  is obtained :

$$\tau(p) = C_\infty \cdot p + C_0 \cdot \left[ 1 - \left( \frac{C_\infty}{C_0} + 1 \right)^p \right]. \quad (235)$$

We now come back to the two other asymptotic forms,  $\Lambda_1$  and  $\Lambda_2$ . The cdf for the second form is

$$\Lambda_2(ap - b) = \exp(-(-(ap - b))^\alpha), \quad (236)$$

where  $\alpha > 0$  and  $ap - b < 0$ .  $\Lambda_2(p)$  is equal to 1 for  $ap - b \geq 0$ . This means that there exists a structure of maximum order  $p_{\max} = b/a$  which saturates the dissipation. However as  $a$  and  $b$  can depend on  $l$ , the maximum attainable order could also depend on the scale.

We can write the following for  $\tau(p)$  in this case :

$$\tau(p+1) - \tau(p) = C_\infty + \lim_{l \rightarrow 0} \left( \frac{-(-(ap - b))^\alpha}{\log(l)} \right). \quad (237)$$

As all parameters may depend on  $l$ , this limit can take many forms. There is however one specific case of interest. When  $\alpha = 1$  (independent of  $l$ ), simple algebra leads to the following form for  $\tau(p)$  :

$$\tau(p) = \begin{cases} \frac{\mu}{2} \cdot (p - p^2), & p \leq -C_\infty/\mu, \\ C_\infty \cdot (p - 1) + f(p), & p > -C_\infty/\mu, \end{cases} \quad (238)$$

where  $f(p)$  is a periodic function of  $p$ . We recognize a form similar to the log-normal model [41, 40]. However, the quadratic behavior becomes linear (up to a periodic correction) for values of  $p > -C_\infty/\mu$ . We note that when  $\mu > -2/3(C_\infty + 3)$ , the quadratic portion of the velocity exponent  $\zeta(p)$  never reaches the maximum and  $\zeta(p)$  remains non-decreasing, avoiding the supersonic velocity problem of the true log-normal model [37]. In the case where  $C_\infty = -2/3$ , this is verified for all  $\mu > 0$ . However the periodic correction makes both  $\tau(p)$  and  $\zeta(p)$  non-concave. Bear in mind that this solution is obtained strictly from the entrainment process, without any direct hypothesis about the log-normal cascade.

Finally, let us look at the  $\Lambda_1(p)$  form. Its cdf is written as :

$$\Lambda_1(ap) = \exp(-(ap)^{-\alpha}), \quad (239)$$

this time with  $a > 0$  and  $p > 0$ . Note the absence of the  $b$  term, a consequence of the convergence constraint for this form [67]. Note also that  $p = 0$  is not allowed; yet as  $\varepsilon_l^{(0)}$  represents the mean fluctuation structure  $\bar{\varepsilon}$  which does not scale anyway, this shall not have any effect on the derivation of  $\tau(p)$ .

Following the same derivation as above :

$$\tau(p+1) - \tau(p) = C_\infty - \lim_{l \rightarrow 0} \frac{(ap)^{-\alpha}}{\log(l)}. \quad (240)$$

We need  $a$  to properly depend on  $l$ , else we fall back to K41. The only non-trivial case is

$$\tau(p+1) - \tau(p) = C_\infty + A \cdot p^{-\alpha}, \quad (241)$$

where  $A$  is a constant.

This last equation has the following solution :

$$\tau(p) = C_\infty \cdot p + C_0 + A \cdot \sum_{k=0}^{\infty} (k+p)^{-\alpha} \quad (242)$$

$$= C_\infty \cdot p + C_0 + A \cdot \zeta_H(\alpha, p), \quad (243)$$

where  $\zeta_H(\cdot, \cdot)$  is the Hurwitz Zeta function.<sup>2</sup>

Hence we have found a family of admissible solutions for the exponent  $\tau(p)$  based solely on the entrainment process, for all three asymptotes  $\Lambda_i$ ,  $i = 1, 2, 3$ . Although it might seem that the scaling behavior was imposed through proper scaling of parameters such as  $a$  and  $b$ , it can be verified that improper scaling of these parameters either leads to divergence or falls back onto K41.

#### 4.7. DISCUSSION AND CONCLUSION

We have presented a statistical interpretation of the She-Leveque model allowing the derivation of the *universal scaling law* that was postulated in [16]. This derivation relies on constraints that must be respected by the system in order to

---

<sup>2</sup>We add the unusual subscript in order to avoid confusion with the scaling exponent  $\zeta(p)$  of the velocity increments.

behave dynamically as proposed in the She-Leveque model. The constraints impose for the pdf of the structure order  $p$  to be one of three admissible asymptotic distributions for the maximum value in a neighborhood of a point. Out of those three valid distributions the Fisher-Tippett law is known to play a pre-eminent role because of its softer convergence conditions. The Fisher-Tippett law also happens to allow for the derivation of the conjectured scaling behavior of the energy dissipation as found in the work of She and Leveque.

It is also worth commenting with the Legendre transform perspective in mind. The Legendre transform allows to define singularity exponents  $h(p)$  given by

$$h(p) = \frac{\partial \tau(p)}{\partial p}. \quad (244)$$

Hence the scaling behavior of the fluctuation structures  $\varepsilon_l^{(\nu)}$  writes as

$$\tau(p+1) - \tau(p) = \int_p^{p+1} \frac{\partial \tau(p)}{\partial p} dp \quad (245)$$

$$= \int_p^{p+1} h(p) dp = \bar{h}(p). \quad (246)$$

She and Leveque argue that while for large value of  $p$  the distinction between  $h(p)$  and  $\bar{h}(p)$  is faint, it is not for small values of  $p$ ; this is in agreement with their picture of poorly organized structures for small values of  $p$ , whereas low order represents a mixture of fractal behaviors. Nevertheless, the derivation of the form of  $\tau(p)$  from an order statistics argument does not require this additional hypothesis. Assuming that the hierarchy can be formulated in terms of structures that scales like  $h(p)$  rather than  $\bar{h}(p)$ , the first order difference equation on  $\tau(p)$  is simply transformed into a trivial differential equation :

$$\frac{\partial \tau(p)}{\partial p} = e^{-(ap-b)} - 2/3, \quad (247)$$

still giving  $\tau(p) = -2p/3 + 2[1 - (2/3)^p]$  as a solution. For the  $\Lambda_1$  and  $\Lambda_2$  form however the solutions are different; in both cases they are expressed in term of power laws. The quasi-log-normal form (238) is then still valid, with the periodic correction eliminated. The goal here is not to argue whether the actual hierarchy

is formed of  $h(p)$  rather than  $\bar{h}(p)$ , but simply to point out that both pictures can lead to similar behavior for  $\tau(p)$ .

All results were obtained by considering the entrainment process proposed originally by She and Leveque, then describing its effect in terms of statistical constraints imposed on the intensity  $\varepsilon_l^{(p)}$  of fluctuation structures. Obviously the entrainment process itself is understood as an additional hypothesis on the phenomenology of fully developed turbulence. Ideally this hypothesis could be understood as the manifestation of a hidden symmetry of the Navier-Stokes equation. As presented above the entrainment process amounts to invariance of a solution under replacement by the maximum in a neighborhood. This transformation can be loosely interpreted (for a scalar field) as a local translation in the direction of the gradient. Local maxima, minima and saddle points appear as fixed points of this transformation. Perhaps an additional symmetry of the Navier-Stokes dynamics could be formulated in a similar fashion. This is the subject of ongoing work.

## CONCLUSION

---

Nous revoyons ici les résultats présentés dans les chapitres précédents. Au premier chapitre, nous avons repris les éléments principaux des travaux de Siegmund et Worsley à propos des statistiques d'excédence de champ multi-résolution. Nous avons généralisé leur résultat concernant la détection d'un signal de taille variable dans un bruit blanc à celle d'un signal noyé dans un bruit possédant une structure de corrélation spatiale non-triviale. Nous avons également indiqué comment il est possible de réduire significativement le temps de calcul nécessaire à cette détection en faisant appel à la résolution propre de chaque échelle entrant dans l'analyse.

Tous les résultats présentés dans ce chapitre dépendent cependant d'une hypothèse forte : le champ multi-échelle doit être gaussien. Worsley a déjà généralisé la méthode à certaines statistiques non-gaussiennes. Par exemple, il a obtenu un résultat semblable pour un champ en  $\chi^2$  [30]. Un champ multi-échelle en  $N + 1$  dimensions, provenant d'un champ de dimension  $N$  à laquelle s'ajoute la dimension d'échelle, sera gaussien si le champ sous-jacent l'est. Lorsque ça n'est pas le cas, le champ multi-échelle qui est obtenu par des combinaisons linéaires locales du champ sous-jacent pourra démontrer aux grandes échelles une statistique quasi-gaussienne, par le théorème central limite. Aux échelles plus fines, dépendamment de la structure de corrélation du champ sous-jacent, la statistique pourra s'éloigner d'une distribution gaussienne. On observe ce phénomène pour les incréments de vitesse d'un signal turbulent. Il serait certes intéressant de développer une méthodologie permettant d'appliquer l'approche topologique à ce type de champ.

L'approche multi-fractale en turbulence pleinement développée apparaît au deuxième chapitre. L'image d'un processus en cascade (pour la dissipation d'énergie) des grandes échelles vers les petites était déjà présente dans les travaux de Kolmogorov en 1941, et même avant [78]. Ce modèle en cascade a été formalisé pour les différences de vitesse en faisant appel, par exemple, aux  $W$ -cascades [46]. Pour tous ces modèles, la corrélation spatiale des éléments de la cascade a été négligée, et mène systématiquement à des processus ne respectant pas l'hypothèse de stationnarité. Nous avons présenté une approche permettant de stationnariser ces processus, au moins au sens large, tout en conservant leurs propriétés multi-fractale. Leur stationnarisation au sens strict nécessite de réduire le support fréquentiel de la fonction de corrélation spatiale inter-échelle de manière proportionnelle au moment d'ordre maximal qui doit être stationnarisé, au moins dans le cadre des modèles en cascades dyadiques. Nous avons également montré que les conditions de stationnarisation imposent des contraintes sur la décroissance de la fonction de corrélation spatiale pour ce type de processus.

Nous avons choisi pour ce faire de reprendre le modèle original de  $W$ -cascades sur une base multi-échelle *dyadique*. Cette approche dyadique, lorsqu'elle fut introduite, constituait la façon la plus simple de modéliser le processus en cascades, en particulier parce que cette représentation est à la fois complète et orthonormale. Les contraintes que nous avons obtenues et qui assurent la stationnarité au sens large du processus ne permettent que de s'approcher de la stationnarité au sens strict, et ce au prix des contraintes supplémentaires que nous venons de mentionner. Ces contraintes ne sont que des conséquences de la structure dyadique originale, tout comme les non-stationnarités du modèle original l'étaient. Il existe cependant d'autres schémas multi-échelle qui pourraient possiblement nous éviter ces ennuis. En particulier, un modèle en cascades dyadique possédant un nombre impair de racines peut être construit de manière stationnaire au sens strict. On rencontre cependant d'autres difficultés dans ce cas. Il serait toutefois

utile d'envisager d'autres décompositions multi-échelle avec en tête la propriété de stationnarité au sens strict.

Le troisième chapitre, qui contient le résultat le plus intéressant de la thèse, a repris certains éléments des deux chapitres précédents dans l'étude du modèle de She-Leveque en turbulence pleinement développée [16]. Faisant suite aux travaux de She et Leveque, nous avons revu pas-à-pas les éléments présentés dans leur article novateur de 1994 et identifié une contrainte statistique forte, sous-jacente à leur modèle et qui n'était pas exprimée jusqu'alors. Cette contrainte qui provient du processus d'entraînement exprimant la dynamique des structures de fluctuation  $\varepsilon_l^{(p)}$  permet de déduire plutôt que de postuler la relation entre les intensités de structures d'ordres successifs, menant ainsi à la forme fonctionnelle qui avait été postulée pour l'exposant  $\tau(p)$ .

De plus nous avons montré que la contrainte forte ainsi déduite ne conduit pas uniquement à la forme postulée par She et Leveque, mais qu'il existe plutôt trois distributions asymptotiques accessibles pour la répartition des ordres des structures. De ces trois distributions nous avons obtenus quelques formes possibles pour les exposants  $\tau(p)$  et  $\zeta(p)$ . Une de ces formes, à laquelle nous avons attribué le nom de quasi-log-normale, mérite une attention plus particulière. Cette forme est constitué d'une dépendance quadratique de l'exposant  $\tau(p)$  sur une partie de son domaine, puis d'une transition abrupte à une dépendance linéaire (ou quasi-linéaire puisqu'il peut apparaître un terme de correction périodique). Bien que le modèle purement log-normal ait suscité de l'intérêt dans la recherche d'un modèle d'intermittence, sa principale difficulté se situait dans la décroissance obligée du spectre  $\tau(p)$  à partir d'un certain point, puisque la dépendance quadratique concave atteint nécessairement un maximum. Cette décroissance implique l'apparition de vitesse supersonique dans le flot turbulent. Dans le cas du modèle quasi-log-normal, le passage à une dépendance linéaire croissante peut s'effectuer avant que ce maximum ne soit atteint, et évite ainsi le problème de vitesse supersonique.

Il est également intéressant de noter que les solutions envisagées sont obtenues à partir du simple processus d'entraînement, sans hypothèse directe sur le comportement multi-fractal du flot ou à travers un modèle en cascade dont les paramètres statistiques auraient été choisis *ad hoc* pour s'harmoniser aux données expérimentales. L'apparition d'un comportement multi-fractal se fait de manière plus naturelle, à travers l'évolution obligée des moments de la dissipation à travers les échelles.

Cette approche moins *paramétrisée* nous permet également de réduire les hypothèses à une seule, soit le processus d'entraînement. Ce processus d'entraînement, formulé de manière très simple, permet d'envisager un rapprochement fascinant avec les hypothèses originales de Kolmogorov [1]. Toute la grandeur de la théorie K41 réside en son économie d'hypothèses, des hypothèses à saveur très généraliste se résumant essentiellement au rétablissement (au sens statistique) de symétries de l'équation de Navier-Stokes pour de grandes valeurs du nombre de Reynolds, symétries brisées par les conditions aux frontières. Or le processus d'entraînement présenté dans le modèle de She-Leveque que nous avons réinterprété comme une contrainte statistique est alors décrit par l'invariance d'une représentation d'une solution (si elle existe) de l'équation de Navier-Stokes sous une transformation étonnante, celle qui remplace une valeur d'un champ scalaire par son maximum dans le voisinage de tout point. De manière informelle, on peut se représenter cette transformation comme étant analogue à une translation locale dans une direction parallèle au gradient du champ en tout point. Cette transformation a la particularité de dépendre, dans son expression formelle, de la solution même à partir de laquelle elle est effectuée, contrairement aux symétries habituelles telles que l'invariance sous translation globale, par exemple. On vérifie facilement que ce type de transformation locale a pour points fixes les optima locaux de la solution. Nous envisageons alors qu'une symétrie cachée de l'équation de Navier-Stokes puisse s'exprimer de manière semblable, expliquant par le fait de son rétablissement au sens statistique la forme universelle observée pour les

exposants d'échelle  $\tau(p)$  et  $\zeta(p)$ , sans pour autant contredire les autres symétries, en particulier la symétrie d'échelle qui admet, tel que décrit par Frish et Parisi [14], un spectre complet du comportement multi-fractal.

En 1941 Kolmogorov prédit le rétablissement au sens statistique de toutes les symétries de l'équation de Navier-Stokes ; bien que les symétries les plus étudiées, telle que l'invariance sous translation, réflexion, rotation et surtout l'invariance sous dilatation, soient effectivement vérifiées par l'équation de Navier-Stokes, elles ne sont certainement pas spécifiques à cette dernière. Pourquoi, alors, ne pas s'attarder plutôt à une symétrie qui serait vérifiée de manière plus spécifique par l'équation de Navier-Stokes et qui seule pourrait expliquer le comportement universel des exposants d'échelle observé expérimentalement ?

## BIBLIOGRAPHIE

---

- [1] A.N. Kolmogorov, C.R. Acad. Sci. USSR **30**, 301 (1941).
- [2] J. von Neumann, *Recent theories of turbulence* In : *Collected Works (1949-1963)* **6** (Pergamon Press, New York, 437-472, 1963).
- [3] S.A. Orszag, G.S. Patterson, In : *Statistical Models and Turbulence*, Lect. Notes in Physics **12** (Springer, Berlin, 127, 1972).
- [4] I. Daubechies, *Ten lectures on wavelet* (SIAM Philadelphia, 1992).
- [5] Y. Meyer, *Ondelettes* (Hermann, Paris, 1990).
- [6] S. Mallat, *A wavelet tour of signal processing*, (Academic Press, 1998).
- [7] C.J. Kicey, C.J. Lennard, *Fourier Analysis and Appl.***3** (1), 63 (1997).
- [8] S. Mallat, S. Zhong, *IEEE Trans. Patt. Recog. and Mach. Intell.* **14** (7), 710 (1992).
- [9] R. Carmona, In *Wavelets and statistics* (Springer-Verlag, Berlin, 96-108, 1995).
- [10] Z. Cvetkovic, M. Vetterli, *IEEE Trans. Signal Proc.* (March 1995).
- [11] K. Gröchenig, *NATO ASI 1991 On Probabilistic and Stochastic Methods in Analysis and Applications* (Kluwer, 1992).
- [12] Y. Meyer, *Wavelets : Algorithms and Applications* (SIAM 1993).
- [13] R.J. Adler, *The geometry of random fields* (Wiley, New York, 1981).
- [14] U. Frisch, G. Parisi, In : *Turubulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics*, edited by M. Ghil, R. Benzi, G. Parisi, 84 (North-Holland, Amsterdam, 1985).
- [15] J.-F. Muzy, *Annales de physique* **20**, 63 (1995).
- [16] Z.-S. She, E. Leveque, *Phys.Rev.Lett.* **72**, 336 (1994).
- [17] K.J. Worsley, *Chance* **9** (1), 27 (1996).

- [18] G.F. Smoot, C.L. Bennett, A. Kogut, E.L. Wright, J. Aymon, N.W. Boggess, E.S. Cheng, G. De Amici, S. Gulkis, M.G. Hauser, G. Hinshaw, P.D. Jackson, M. Jansen, E. Kaita, T. Kelsall, P. Keegstra, C. Lineweaver, K. Lowenstein, P. Lubin, J. Mather, S.S. Meyer, S.H. Moseley, T. Murdock, L. Rokke, R.F. Silverberg, L. Tenorio, R. Weiss, D.T. Wilkinson, *Astrophysical Journal* **396**, L1-L5 (1992).
- [19] K.J. Worsley, *Adv. Appl. Prob.* **27**, 943 (1995).
- [20] S. Torres, *Astrophysical Journal* **423**, L9-L12 (1994).
- [21] K.J. Friston, P. Jezzard, R. Turner, *Human Brain Mapping* **1**, 153 (1994).
- [22] K.J. Worsley, A.C. Evans, S. Marrett, P. Neelin, *J. Cereb. Blood Flow Metab.* **12**, 900 (1992).
- [23] H. Hadwiger, *Vorlesungen Über Inhalt, Oberfläche und Isoperimetrie* (Springer-Verlag, Berlin, 1957).
- [24] H. Hadwiger, *Math. Zeitschrift* **71**, 124 (1959).
- [25] L.A. Santaló, *Actualités Sci. Ind.* **1198** (1953).
- [26] L.A. Santaló, *Integral Geometry and Geometric Probability*, In : *Encyclopedia of Mathematics and its Applications Vol. 1* (Addison-Wesley, Reading, Mass. 1976);
- [27] D.E. Knuth, *Amer. Math Monthly* **99**, 403 (1992).
- [28] K.J. Worsley, *Annals of statistics* **23** (2), 640 (1995).
- [29] D.O. Siegmund, K.J. Worsley, *Annals of statistics* **23** (2), 608 (1995).
- [30] K.J. Worsley, *Adv. Appl. Prob.* **33** (4), 773 (2001).
- [31] S. Jaffard, *Publications Mathématiques* **35**, 155 (1991).
- [32] J.M. Shapiro, *IEEE Trans. Signal Proc.* **41** (12), 3445 (1993).
- [33] A. Arneodo, E. Bacry and J.F. Muzy, *Physica A* **213**, 232 (1995).
- [34] M.S. Crouse, R.D. Nowak and R.G. Baraniuk, *Wavelet-based statistical signal processing*, In : *Special issue on wavelets and filterbanks*, April 1998.
- [35] Donoho, J. Roy. Stat. Soc. **B39**, 301 (1995).
- [36] A. Rosenfeld, A.C. Kak, *Digital Picture Processing, Volume 2* (Academic Press, Orlando, 1982).

- [37] U. Frisch (1995), *Turbulence* (Cambridge University Press, 1995).
- [38] L.D. Landau, E.M. Lifshitz, *Fluid Mechanics*, 2nd edition (Pergamon Press, Oxford, 1987).
- [39] R. Benzi, S. Ciliberto, C. Baudet, F. Massaioli, R. Tripiccone, S. Succi, *Phy Re. E* **48**, 29 (1993).
- [40] A.M. Obukhov, *J. Fluid Mech.* **13**, 77 (1962).
- [41] A.N. Kolmogorov, *J. Fluid Mech.* **13**, 82 (1962).
- [42] C. Meneveau, K.R. Sreenivasan, *J. Fluid Mech.* **224**, 429 (1991).
- [43] E.A. Novikov, R.W. Stewart, *Izv. Akad. Nauk SSSR, Ser. Geofiz.*, 408 (1964).
- [44] F. Hausdorff, *Mathematische Annalen* **79**, 157 (1919).
- [45] A.S. Besicovitch, *Mathematische Annalen* **110**, 321 (1935).
- [46] A. Arnéodo, E. Bacry and J.F. Muzy, *J. Math. Phys.* **39**, 4142 (1998).
- [47] D.W. Strook, *An introduction to the theory of large deviations* (Springer-Verlag, Berlin, 1984).
- [48] S. Jaffard, *C.R. Acad. Sci. Paris* **308** I, 79 (1989).
- [49] A. Arnéodo, E. Bacry et J.-F. Muzy, *Phys Rev Lett* **74** (24), 4823 (1995).
- [50] H.-C. Huang and N. Cressie, In *Spatial Statistics : Methodological Aspects and Applications*, ed. M. Moore. Springer Lecture Notes in Statistics **159**, Springer-Verlag, New York, 83 (2001).
- [51] A. Arnéodo, S. Manneville and J.F. Muzy, *Eur. Phys. J. B* **1**, 129 (1998).
- [52] J. Delour, J.-F. Muzy and A. Arnéodo, *Eur. Phys. J. B* **23**, 243 (2001).
- [53] J.F. Muzy, J. Delour and E. Bacry, *Modelling fluctuations of financial time series : from cascade process to stochastic volatility model*, *Eur. Phys. J.* manuscript (2000).
- [54] A. Arnéodo, J.-F. Muzy and D. Sornette, *Eur. Phys. J. B* **2**, 277 (1998).
- [55] S. Douady, Y. Couder, M.E. Brachet, *Phys. Rev. Lett.* **67** (8), 983 (1991).
- [56] A. Vincent, M. Meneguzzi, *J. Fluid Mech.* **225**, 1 (1991).
- [57] H. Cramér, *Theor. Probability Appl.* **10**, 137 (1965).
- [58] V.A. Volkonskii, Yu. A. Rozanov, *Theory Prob. Appl.* **4**, 178 (1959).

- [59] V.A. Volkonskii, Yu. A. Rozanov, *Theory Prob. Appl.* **6**, 187 (1961).
- [60] E.L. Dodd, *Trans. Am. Math. Soc.* **25**, 523 (1923).
- [61] R. von Mises, *Selected papers of Richard von Mises 2* (AMS, Providence, 1964).
- [62] R. von Mises, *Rev. Math. Union Interbalkanique* **1**, 141 (1936).
- [63] M. Fréchet, *Ann. Soc. Polonaise Math.* **6**, 92 (1927).
- [64] R. A. Fisher, L. H. C. Tippett, *Proc. Camb. Phil. Soc.* **24**, 180 (1928).
- [65] B. de Finetti, *Metron* **9**, 127 (1932).
- [66] E.J. Gumbel, *Statistics of Extremes* (Columbia University Press, New York, 1958).
- [67] B. Gnedenko, *Ann. Math. Second Series* **44 (3)**, 423 (1943).
- [68] L. de Haan, *Mathematical Center Tracts* **32** (Mathematisch Centrum, Amsterdam, 1970).
- [69] H. A. David, *Order statistics*. (2nd edition, Wiley Series in Prob. and Math. Stat., Chapter 9, 1981)
- [70] F. Anselmet, Y. Gagne, E.J. Hopfinger, R.A. Antonia, *J. Fluid Mech.* **140**, 63 (1984).
- [71] B. Dubrulle, *Phys. Rev. Lett.* **73**, 959 (1994).
- [72] E.A. Novikov, *Phys. Rev. E* **50**, 3303 (1995).
- [73] B. Castaing, B. Dubrulle, *J. Phys. II France* **5**, 895 (1995).
- [74] Z.-S. She, E.C. Waymire, *Phys. Rev. Lett.* **74**, 262 (1995).
- [75] Z.-S. She, E. Jackson, S.A. Orszag, *Nature (London)* **344**, 226 (1990).
- [76] Z.-S. She, *Fluid Dyn. Res.* **8**, 143 (1991).
- [77] Z.-S. She, E. Jackson, S.A. Orszag, *Proc. R. Soc. London A* **434**, 101 (1991).
- [78] L.F. Richardson, *Weather Prediction by Numerical Process* (Cambridge University Press, Cambridge, 1922).

