acta physica slovaca vol. 46 No. 4, 557 – 565 August 1996

INTRODUCTION TO WAVELETS AND APPLICATIONS TO CORRELATION MEASUREMENTS[1]

P. Lipa[2], M. Greiner[3], P. Carruthers[4]
Institut für Hochenergiephysik der Österreichischen Akademie der Wissenschaften, Nikolsdorfergasse 18, A-1050 Vienna, Austria

Received 16 April 1996 , accepted 28 May 1996

We discuss discrete orthogonal wavelet transforms and some applications to statistical and multiparticle physics.

1. Motivation

In multiparticle physics useful tools, such as factorial moments, cumulants, correlation integrals, void probabilities and combinants are currently explored and to some extent applied to the analysis of data. These correlation measures elucidate many interesting features of higher order correlations but share one common drawback: they rely on local hadron multiplicities and, hence, are not infrared stable. More precisely, the above mentioned measures depend crucially on the *number* of particles, so that it makes a big difference if a certain amount of energy is carried by one particle or is distributed over few nearby particles, say, a cluster.

For the design of infrared stable observables it is desireable to abstract from particles and find ways to characterize a "cluster" independently of the number of particles it contains. In other words, we seek strategies to smooth over the discrete nature of clusters in a fair and simple way. Wavelets are natural candidates for such strategies.

Orthogonal wavelets define a *multiresolution representation* of a signal, e.g. a collection of particles (points) in phase space. In a sequence of "local smoothing" and "differentiation" operations, the signal is decomposed into contributions from clusters, composed from clusters at smaller scales, which are in turn built from clusters at again smaller scales, and so on. Thus, wavelets provide a strategy to select clusters living only

[1] Invited lecture given by PL at the Triangle Meeting: School and Workshop on Heavy Ion Collisons, Bratislava, Slovakia, 4-9 September 1995

[2] E-mail address: lipa@hephy.oeaw.ac.at
[3] Present address: Institut für Theoretische Physik, Technische Universität, D-01062 Dresden, Germany; E-mail address: greiner@ptprs23.phy.tu-dresden.de
[4] Present address: Department of Physics, University of Arizona, Tuscon AZ-85721, USA; E-mail address: carruthers@ccit.arizona.edu

at a specific scale. This property raises hopes to tame the infrared problem ubiquitous in conventional correlation studies.

Unfortunately the language used to describe the wavelet transform is often involved and repelling for physicists. As a consequence, this community has hardly discovered the power of wavelets yet. For this reason we would like to give a simple introduction to the basic ideas tied to the wavelet transformation and the related concept of a multiresolution analysis. Additionally, we point out some applications.

2. Multiresolution decompositions and wavelets

Multiresolution decompositions chop up a signal into (not necessarily) mutually orthogonal contributions from nested sequence of scales. This chopping can be done by very fast algorithms which *implicitly* define and use wavelets. We explain the basic ideas by means of the familiar concept of histograms, giving rise to the simplest member of the wavelet family, the so-called Haar wavelet, known since beginning of this century.

Let us approximate an arbitrary one-dimensional function $\epsilon(x)$ in terms of a sequence of histograms, each having $2^j (j = 0, \ldots, J)$ bins; see fig.1 as a guideline. The binsize 2^{-j} defines the *scale* or *resolution* of a particular approximation. In principle the resolution j is allowed to go to $\pm\infty$, but in all practical applications we assume that the function $\epsilon(x)$ is known up to a finest scale J, which may be dictated by the resolution of the measurement device.

Fix a particular bin k of a histogram at scale j with bincontent ϵ_k; obviously any structure or fluctuation narrower than the binsize 2^{-j} is smoothed out. At the next finer scale $j+1$ the same structure is resolved by *two* bins and in most cases the first one will differ from ϵ_k by an amount $\tilde{\epsilon}_k$; then the second one deviates by $-\tilde{\epsilon}_k$ from the average. In this way, an arbitrary fluctuation within the bin k is captured in the *difference information* $\tilde{\epsilon}_k$ up to a resolution $2^{-(j+1)}$. Obviously this procedure can be iterated to finer and finer scales, dissecting an arbitrary fluctuation into independent contributions from different scales $0 \le j \le J$. If such a fluctuation at a certain scale j is large (i.e. has a significant $\tilde{\epsilon}_k$) we shall generously refer to it as a "cluster living at scale j".

This procedure defines the simplest case of a multiresolution analysis [1]. The signal $\epsilon(x)$ can be represented by a sum of independent contributions with finer and finer detail information (the functions in the right column in fig. 1), each one capturing only those fluctuations that live between two adjacent scales.

Wavelet theory starts with expressing the above scheme in a more fancy way: the histogram at finest resolution scale J is now written in terms of 2^J box functions $\phi^H_{Jk}(x) = \phi^H(2^J x - k)$, each representing one bin with width 2^{-J} and position $k = 0, \ldots, 2^J - 1$. They are constructed from the unit box function, the so-called *scaling function*,

$$\phi^H(x) = \phi^H_{00}(x) = \begin{cases} 1 & \text{for } 0 \le x < 1 \\ 0 & \text{else} \end{cases} \tag{1}$$

by a dilation with contraction factor 2^{-J} and a translation (shift) by an integer k. The

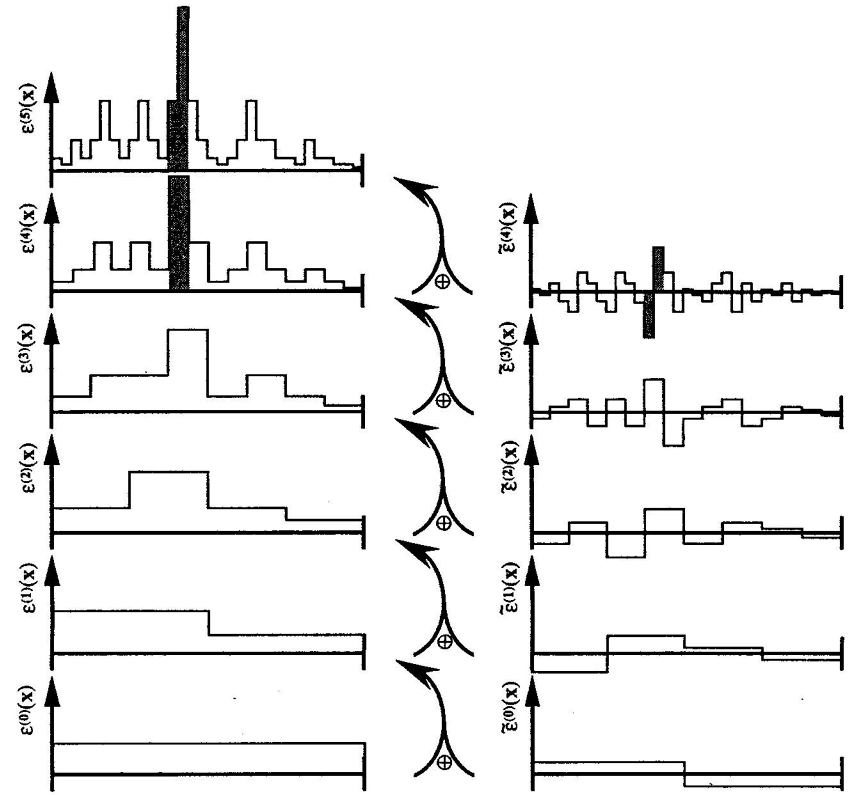

Fig. 1. A multiresolution decomposition of a random signal (upper left corner) at resolution $J = 5$. The left column of histograms shows a sequence of smoothed approximation while the right column represents the orthogonal Haar wavelet decomposition.

histogram of the function $\epsilon(x)$ at the finest scale J is then written as a series:

$$\epsilon(x) \to \epsilon^{(J)}(x) = \sum_k \epsilon^{(J)}_k \phi^H_{Jk}(x), \tag{2}$$

where $\epsilon^{(J)}_k = \int \epsilon(x)\phi^H_{Jk}(x)\mathrm{d}x$ is the content of the k-th bin at resolution J.

The above outlined multiresolution analysis is formalized by the series expansion into contributions $\tilde{\epsilon}^{(j)}(x)$ from different scales (compare with Fig. 1),

$$\begin{aligned} \epsilon^{(J)}(x) &= \epsilon^{(0)}(x) + \textstyle\sum_j \tilde{\epsilon}^{(j)}(x) \\ &= \epsilon_{00}\phi^H_{00}(x) + \textstyle\sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} \tilde{\epsilon}_{jk}\psi^H_{jk}(x). \end{aligned} \tag{3}$$

with

$$\tilde{\epsilon}_{jk} = 2^j \int \epsilon(x)\psi^H_{jk}(x)\mathrm{d}x. \tag{4}$$

The wavelet amplitudes $\tilde{\epsilon}_{jk}$ are most conveniently represented in a one-dimensional histogram $(\epsilon_{00}, \tilde{\epsilon}_{00}, \tilde{\epsilon}_{10}, \tilde{\epsilon}_{11}, \tilde{\epsilon}_{20}, \ldots, \tilde{\epsilon}_{23}, \ldots, \tilde{\epsilon}_{J-1,0}, \ldots, \tilde{\epsilon}_{J-1,2^{J-1}-1})$ with 2^J bins.

What are the basis functions $\psi^H_{jk}(x)$, that make such a series expansion possible? Simply by inspection of fig. 1 one can deduce that they have to be dilated and translated copies of a unit difference function, $\psi^H_{jk}(x) = \psi^H(2^j x - k)$ with

$$\psi^H(x) = \psi^H_{00}(x) = \begin{cases} 1 & \text{for } 0 \leq x < 1/2 \\ -1 & \text{for } 1/2 \leq x < 1 \\ 0 & \text{else} \end{cases} \tag{5}$$

Note that the functions $\psi^H_{jk}(x)$ are orthogonal with respect to the shift index k and the dilation index j: $2^j 2^{j'} \int \psi^H_{jk}(x)\psi^H_{j'k'}(x)\mathrm{d}x = \delta_{jj'}, \delta_{kk'}$.

This orthogonal basis was introduced by Haar in the early 1900's, in todays language the $\psi^H_{jk}(x)$ are called "Haar wavelets". The major breakthrough in wavelet theory was achieved by Mallat and Daubechies [1, 2] in the late 80's, when they showed that the above multiresolution expansion can be based on more general functions with nicer mathematical properties than the Haar wavelets (which are discontinuous and, consequently, badly localized in Fourier space).

More specifically, Daubechies constructed several families of wavelets and associated scaling functions, which are orthonormal, have compact support and are at least continuous or even smoother (several times differentiable) [2]. They are solutions of the functional equations

$$\phi(x) = \sum_k c_k \phi(2x-k) \quad \text{and}$$

$$\psi(x) = \sum_k (-1)^k c_{1-k} \phi(2x-k) \ . \tag{6}$$

In principle the solution of these equations can be found numerically by iteration, once a finite set of coefficients c_k is given. However, for most choices of c_k the solutions will *diverge*.

The box function (1) and the Haar wavelet (5) represent the simplest convergent solution; they are determined by only two coefficients: $c_0 = c_1 = 1$ and all the others are zero.

The next simple, though highly nontrivial, convergent solution involves four nonvanishing coefficients

$$c_0 = \frac{1}{4}(1+\sqrt{3}), \quad c_1 = \frac{1}{4}(3+\sqrt{3}), \quad c_2 = \frac{1}{4}(3-\sqrt{3}), \quad c_3 = \frac{1}{4}(1-\sqrt{3}),$$

leading to the continuous and orthogonal Daubechies D4-wavelet. These and many more wavelets are visualized, e.g. in [2]. Their construction represents a beautiful and ingenious piece of applied mathematics. Generally one has a trade-off between compactness (length) and smoothness of a wavelet: the smoother the wavelet becomes (and, hence, the better it is localized in Fourier space) the broader the compact support has to be.

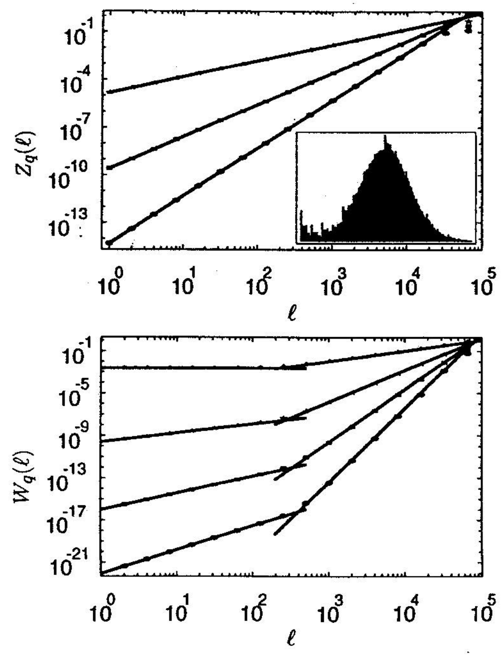

Fig. 2. Multifractal and Haar-wavelet analysis of a small selfsimilar noise distribution superimposed to a smooth Gaussian function (see inset). Plotted are the moments (a) $Z_q(\ell)$ and (b) $W_q(\ell)$ versus ℓ for $q = 1$ (triangles), 2 (stars), 3 (squares) and 4 (circles).

As a generalization of the multiresolution analysis presented for the Haar wavelet before, the equations (2)-(4) and (6) define a multiresolution analysis for any specific choice of wavelets. The resulting "histograms" at the various scales are *not* step functions, but acquire the degree of smoothness of the underlying wavelet.

In the literature many defferent wavelets with various additional properties have been constructed: biorthogonal wavelets, symmetric wavelets, wavelets on a closed interval, wavelet packets, ... For a good survey about these extensions we refer the interested reader to the books of Chui [3], Kaiser [4] and Meyer [5]; confer also the second edition of 'Numerical Recipes' [6].

3. Some applications in (statistical) physics

The overwhelming success of wavelets in signal analysis is founded in its ability to extract and analyze efficiently local details living on a hierarchy of scales. Hence, the wavelet transformation appears to be extremely attractive to study and describe complex reactions in general, especially those exhibiting texture and patterns involving multiple scales. In particular, the self-similarity aspect is eye-catching; hence, the description of hierarchically organized (stochastic) complex reactions should be facilitated tremendously by wavelet transforms.

In order to elucidate this last point we begin with a simple example [7]: Consider a smooth Gaussian function with a small selfsimilar (multifractal) noise added; see inset

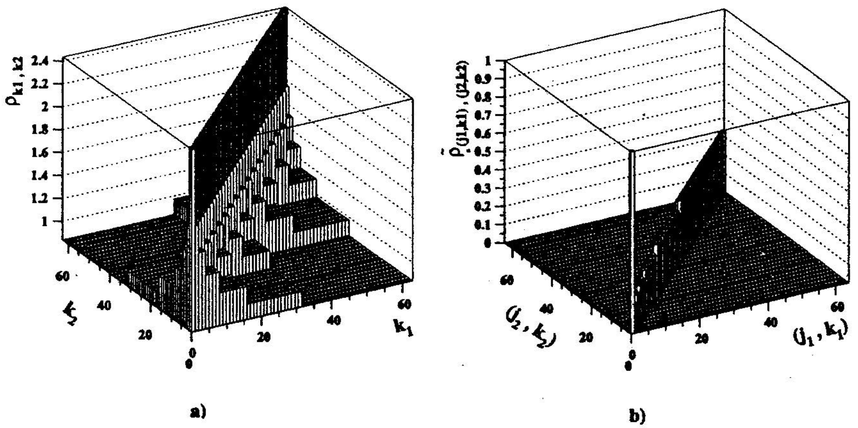

Fig. 3. (a) Spatial two-bin correlation density $\rho_{k_1k_2}$ and (b) the corresponding Haar-wavelet correlation density $\tilde{\rho}_{(jk_1)(jk_2)}$ for the p-model cascade.

of Fig. 2a. In a conventional multifractal analysis this distribution is first expanded according to (2) into (Haar-) scaling functions ϕ_{jk} belonging to scale j; then the scaling behaviour of the partition function

$$Z_q(\ell = 2^{-j}) = \sum_k \left|\epsilon_k^{(j)}\right|^q \sim \ell^{\tau(q)} \quad (7)$$

in studied. Since the amplitudes $\epsilon_k^{(j)}$ reflect absolute values, this scaling approach is quite insensitive to the small selfsimilar noise as the large Gaussian background function dominates. Hence, $\tau(q) = \tau_{Gaussian}(q) = q - 1$ only reveals the trivial scaling properties of the smooth Gaussian; see Fig. 2a. On the orther hand, the (Haar-) wavelet expansion (3) focuses on differences between neighbouring parts of the distribution. As a consequence, the smooth Gaussian background drops out for small scales ℓ and the wavelet amplitudes $\tilde{\epsilon}_{jk}$ are only determined by the selfsimilar noise. Hence, the wavelet partition function

$$W_q(\ell = 2^{-j}) = \sum_k |\tilde{\epsilon}_{jk}|^q \sim \ell^{\beta(q)} \quad (8)$$

leads to scaling exponents $\beta(q) = \tau_{noise}(q)$ for small ℓ; for large ℓ the Gaussian behaviour takes over again and we find $\beta(q) = 2q - 1$. See Fig. 2b. Here the wavelet approach to multifractality reveals clearly the two different scaling regimes whereas the conventional multifractal approach is *incapable* to detect and isolate the fluctuations at small scales!

Another, more dynamical, example: So-called weight curdling models have been proposed as phenomenological descriptions of intermittent spatial fluctuations in fully developed turbulence [8, 9, 10]. Some of these models have also been used to simulate fluctuations in e^+e^- - and hadron-hadron collisions [11, 12]. Fig. 3a shows the spatial two-bin correlation density $\rho_{k_1k_2} = \langle \epsilon_{Jk_1}, \epsilon_{Jk_2} \rangle$ of the so-called p-model; it has been calculated analytically in ref. [13]. The power-law rise towards the diagonal is a

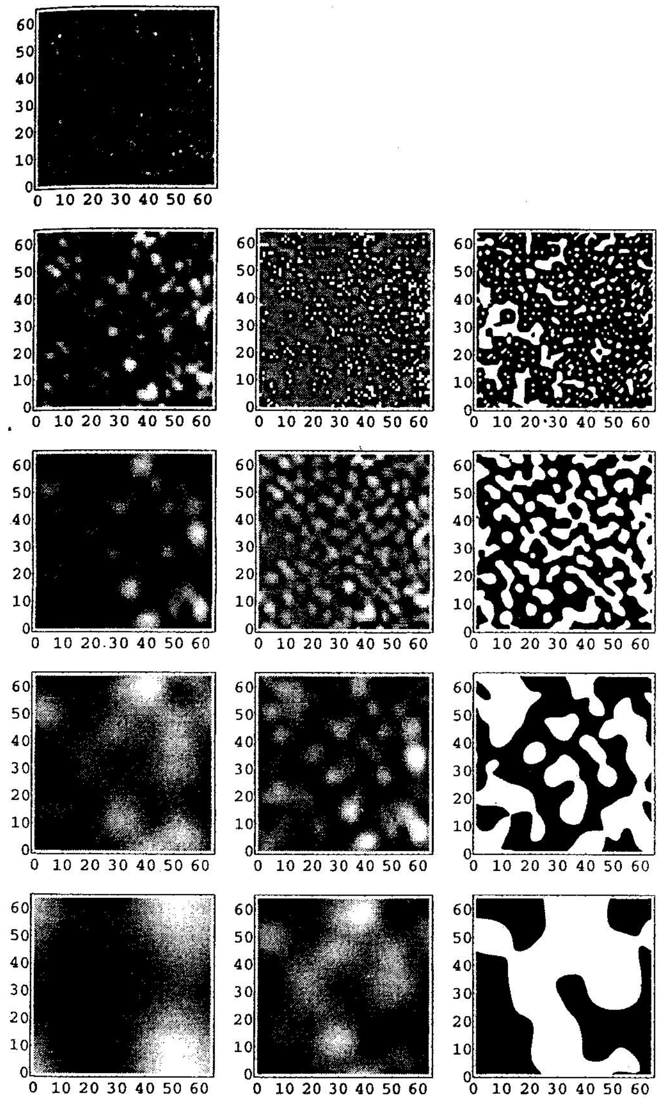

Fig. 4. Multiresolution analysis of a α-model realization in two dimensions with respect to the smooth compact Daubechies D12 wavelet. Left column: sequence of smoothing operations. Middle column: difference between two adjacent smoothed scales. Right column: multi-scale cluster boundaries of middle column emphasized by reduction to two gray values (white/black) for positive/negative regions in the difference information.

clear indication of the selfsimilarity of the hierarchical p-model cascade: the closer two bins are together, the more they share a common (cascade) history and the stronger they are correlated. This representation of the correlation density is based on the monoscale expansion (2); it is not an optimal choice since it does not take into account the hierarchical organisation of the process. Here, the wavelet representation (3) is the better choice. In fact it turns out that the second order wavelet correlation density $\tilde{\rho}_{(j_1k_1)(j_2k_2)} = \langle \tilde{\epsilon}_{j_1k_1} \tilde{\epsilon}_{j_2k_2} \rangle$ is completely "compressed" into the diagonal (see Fig. 3b); the Haar-wavelet amplitudes $\tilde{\epsilon}_{j_1k_1}$ have been ordered according to

$$\begin{aligned} \vec{\tilde{\epsilon}} &= \left(\epsilon_0^{(0)}, \tilde{\epsilon}_{00}, \tilde{\epsilon}_{10}, \tilde{\epsilon}_{11}, \tilde{\epsilon}_{20}, \ldots, \tilde{\epsilon}_{J-1,2^{J-1}-1} \right) \\ &\equiv \left(\tilde{\epsilon}_0, \tilde{\epsilon}_1, \tilde{\epsilon}_2, \tilde{\epsilon}_3, \tilde{\epsilon}_4, \ldots, \tilde{\epsilon}_{2^J-1} \right). \end{aligned} \quad (9)$$

All off-diagonal elements vanish and the staircase behaviour of the diagonal elements reveals the same power-law behaviour (indicating selfsimilarity) as found in the conventional correlation density $\rho_{k_1k_2}$ across the main diagonal. Definitely we have benefited here from the compression power of the wavelet transformation, which makes it such a powerful tool for signal transmission.

There is still more to gain: In analogy to the analyzing power of wavelets in signal analysis ("what frequencies come at what time?") higher order wavelet correlations provide direct information on how substructures living inside larger structures are organized [14]; in particle physics we would speak of correlations of subject within jets whereas in turbulence or astronomy it would be subclusters inside clusters (coherent structures). In this sense higher order wavelet correlations are extremely sensitive to the dynamics of the complex reaction under investigation.

4. Multiscale clustering in two-dimensional branching processes

To illustrate and visualize multiscale clusters in more detail, we consider hierarchical branching processes in two dimensions [14]. As a representative we take the two-dimensional α–model since it is easily generalized to two and higher dimensions. One possible realization of the two-dimensional α–model is shown in the upper left corner of Fig. 4. We have used a gray scale to indicate the population of regions in between large energy densities (white) and small energy densities (black). We observe that certain regions clump into clusters of various sizes while others are more or less void.

To quantify this picture, we explicitly perform a multiresolution decomposition similar to Fig. 1: First the energy densities at the finest resultion scale $J = 6$ are smoothed with the D12–scaling function on the next rougher scale $j = 5$. These averaged energy densities are depicted in Fig. 4 as the second figure from the top of the left column. Some detail about the subclustering occurring between the involved scales is obviously lost; this lost information can be recovered by keeping the difference between the resolutions $j = 5$ and $j = 6$, illustrated in the top of the middle column. These smoothing and differentiation operations are iterated through scales $j = 4, 3$ down to 2.

The left column of Fig. 4 shows density plots of a sequence of D12–smoothed approximations to the original configuration, whereas the middle column represents density plots of the wavelet transform, which form a sequence of mutually orthogonal details. In order to provide a better picture of the subclustering aspect of the wavelet transform, the details at various scales are exhibited again in the right column, but with two gray–values only: black for regions with negative values of the detail function, indicating

local voids, and white for regions where the detail function is positive, signaling the appearance of clusters at the various scales.

The borders between white and black regions in the right column of Fig. 4 are the zero–crossings of the wavelet transform. These serve as a natural definition for multiscale cluster boundaries. Fast algorithms exist [15] to detect these boundaries.

5. Outlook

All these observations let us hope, that the successive "smoothing" operations of the multiresolution representation, if defined analogously to conventional cluster- or jetfinding algorithms, might facilitate comparisons of experimental cluster correlation studies at hadron level with theoretical calculations of parton shower models. We expect that the "clusters" arising in the wavelet transform do not depend on the details of the hadronisation process at small Q^2, at least at large and intermediate scales, thus taming the infrared problem inherent in conventional correlation studies. Of course, these hopes and expectations remain to be verified by further numerical analyses of more realistic cascade models.

Acknowledgements We thank J. Kantelhardt and E. Roman for providing figures 3. P.L. warmly thanks the Austrian Academy of Science for support with an APART (Austrian Programme for Advanced Research and Technology) fellowship. M.G. gratefully acknowleges support from GSI, DFG and BMBF. This work was supported in parts by the US Department of Energy, grant no. DE-FG02-88ER40456.

References

[1] S. Mallat: *IEEE Trans. Pattern Anal. and Machine Intell.* **11** (1989) 674
[2] I. Daubechies: *Comm. Pure Appl. Math.* **41** (1988) 909
"Ten Lectures on Wavelets", Society for Industrial and Applied Mathematics (SIAM), Philadelphia (1992);
[3] C.K. Chui: *An Introduction to Wavelets*, Academic Press, Boston (1992);
Wavelets - *A Tutorial in Theory and Applications*, Academic Press, Boston (1992);
[4] G. Kaiser: *A Friendly Guide to Wavelets*, Birkhäuser, Boston (1994);
[5] Y. Meyer: Wavelets - *Algorithms and Applications*, SIAM, Philadelphia (1993);
[6] W. Press, S. Teukolsky, W. Vetterling, B. Flannery: *Numerical Recipes*, Cambridge University Press, Cambridge (1992);
[7] J. Kantelhardt, M. Greiner, E. Roman: *Physica* **A 220** (1995) 219
[8] B. Mandelbrot: *J. Fluid Mech.* **62** (1974) 719
[9] U. Frisch, P.L. Sulem, M. Nelkin: *J. Fluid. Mech.* **87** (1978) 719
[10] C. Meneveau, K.R. Sreenivasan: *Phys. Rev. Lett.* **59** (1987) 1424
[11] A. Bialas, R. Peschanski: *Nucl. Phys.* **B 273** (1986) 703
Nucl. Phys. **B 308** (1988) 857;
[12] P. Lipa, B. Buschbeck: *Phys. Lett.* **B 223** (1989) 465
[13] M. Greiner, P. Lipa, P. Carruthers: *Phys. Rev.* **E 51** (1995) 1948
[14] M. Greiner, J. Giesemann, P. Lipa, P. Carruthers : *Z. Phys.* **C 69** (1996) 305
[15] S. Mallat: *IEEE Trans. Inf. Theory* **37** (1991) 1019
S. Mallat, S. Zhong:*IEEE Trans. Patt. Anal. Mach. Intel.***14** (1992) 710;