A Multiresolution Method for Detecting Higher Order Discontinuities from Irregular Noisy Samples

M. Randrianarivony and G. Brunnett

Abstract. We consider the problem of detecting higher order discontinuities of functions that are represented by discrete noisy irregular data sets. First, we show briefly the relationship between a function having higher order discontinuity and a local maximum of the appropriate wavelet transform of the given exact data. Then, we describe an algorithm that allows to detect the discontinuity for the case of noisy sampling data. Furthermore, we analyze an approach to extend our method to irregular samples. A further application of a nonlinear minimization may improve the accuracy of the results considerably. Numerical tests are presented to support our method.

§1. Introduction

Let $\tilde{\boldsymbol{x}}_i = (x_i, y_i) \in \mathbb{R}^2$ (resp. $\tilde{\boldsymbol{x}}_i = (x_i, y_i, z_i) \in \mathbb{R}^3$) be data points which belong to an unknown curve $\mathcal{K} \subset \mathbb{R}^2$ (resp. surface $\mathcal{S} \subset \mathbb{R}^3$). We assume that $\mathcal{K}$ (resp. $\mathcal{S}$) is piecewise $\mathcal{C}^k$ with $k \geq 2$ and globally continuous. In geometric modeling, it is often desirable [12] to partition $\mathcal{K}$ (resp. $\mathcal{S}$) into sub-curves $\mathcal{K}_p$ (resp. sub-surfaces $\mathcal{S}_p$) which are subsequently easier to fit.

In practice the data points can be obtained from 3D scanners. So, they are often inexact, i.e. we observe $\boldsymbol{x}_i = \tilde{\boldsymbol{x}}_i + \boldsymbol{\eta}_i$, where $\boldsymbol{\eta}_i$ are noise and $\tilde{\boldsymbol{x}}_i$ are the actual points belonging to $\mathcal{K}$ or $\mathcal{S}$. The aim of this paper is to locate special features including corners or edges of $\mathcal{K}$ or $\mathcal{S}$. More specifically, we are interested in finding locations where $\mathcal{K}$ or $\mathcal{S}$ does not have C^k-continuity with $k = 1, 2$. The only information that we know is the discrete and inaccurate data $\boldsymbol{x}_i$. In this document, we restrict ourselves

Curve and Surface Fitting: Saint-Malo 2002
XXX (eds.), pp. 1–10.

ISBN XXX.

to curves which can be represented as functions, i.e., there is a function f with

$$(x_i, y_i) \in \mathcal{K}, \qquad \text{if} \quad y_i = f(x_i)\,. \tag{1}$$

Note that our goal in this paper is *not* to find $\mathcal{K}$ (resp. $\mathcal{S}$) as in [3]. That will be our ultimate goal in some future work. In [1,15,14], the authors use wavelets to detect discontinuities of a function f from *regular* samples, i.e. from

$$y_i = f\left(\frac{i}{N}\right) + \varepsilon_i\,, \qquad i = 0, 1, ..., N\,, \tag{2}$$

one tries to find σ with $f(\sigma^+) \neq f(\sigma^-)$. We intend to generalize that technique with respect to two points:

- Find higher order discontinuities, i.e. $f^{(k)}(\sigma^+) \neq f^{(k)}(\sigma^-)$, $k = 1, 2$.
- Because we often meet samples which are not regular in practice, we allow nonuniform samples, i.e.

$$y_i = f(x_i) + \varepsilon_i\,, \quad x_i \in]0, 1[\,, \quad i = 0, ..., N\,. \tag{3}$$

In Section 2 we recall some topics in multiresolution analysis, and present some characterization of C^k-discontinuities. The main results are in Sections 3 and 4 where we see first the effect of noise and its remedy, and second we present how to treat nonuniform samples.

§2. Multiresolution Scheme

Since we intend to solve this problem by a multiresolution method, we need to recall briefly some characteristics.

Definition 1. [5] A wavelet is a function $\psi \in L^2(\mathbb{R})$ which satisfies

$$\int_{\mathbb{R}} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty\,. \tag{4}$$

The wavelet transform of $f \in L^2(\mathbb{R})$ with respect to ψ is

$$(W_\psi f)(s, u) := \frac{1}{\sqrt{|s|}} \int_{\mathbb{R}} f(x) \psi\left(\frac{x - u}{s}\right) dx\,, \qquad \forall s \in \mathbb{R} \setminus \{0\},\ \forall u \in \mathbb{R}\,.$$

The discrete wavelet transform (DWT) is given by

$$d_k^m := (W_\psi f)(2^m, k2^m) \quad m, k \in \mathbb{Z}\,. \tag{5}$$

Definition 2. [8,9] A sequence of nested linear spaces

$$\cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \cdots$$

is a multiresolution analysis if

(C1) $\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R})$,

(C2) $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$,

(C3) $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j-1}$,

(C4) there is $\phi \in L^2(\mathbb{R})$: such that $\{\phi(x - k) : k \in \mathbb{Z}\}$ forms an orthonormal basis of V_0.

Remark 3. In practice, a multiresolution analysis is used to create wavelets [5]. In most cases, the scaling function ϕ is not given explicitly. Only a discrete sequence h_k in a two-scale relation is provided (see [9,11]). And there is an extremely fast algorithm ([4,9]) to compute all the discrete wavelet coefficients in (5) with the help of h_k. In the subsequent discussion, we deal only with Daubechies wavelets [5].

Theorem 4. [Effect of noise] Let v_t be a Gaussian white noise which will describe subsequently the noise in the measurements. Let s_τ be any sequence tending to zero as $\tau \to 0$. Then for all $\delta \in]0,1[$,

$$P\left[\max_{x\in[0,1]} |W_\psi v(s_\tau, x)| < C_\tau |2\log s_\tau|^{1/2}\right] \to 1-\delta \quad \text{as} \quad \tau \to 0\,,$$

where C_τ tends to 1.

Proof: Let $\delta \in]0,1[$ be given. Define $\tilde{M} := \max_{x\in[0,1]} |W_\psi v(s_\tau, x)|$ and

$$B(s) = (2|\log s|)^{1/2} + \frac{1}{(2|\log s|)^{1/2}} \log A \quad \text{with} \quad \mathrm{A} := \frac{1}{2\pi}[\int (\psi'(\mathrm{u}))^2 \mathrm{du}]^{1/2}\,.$$

Since v_t is a Gaussian white noise, $W_\psi v$ is a stationary Gaussian process ([15]). Hence, Corollary A1 of [2] ensures that for all μ

$$P[|2\log s_\tau|^{1/2}(\tilde{M} - B(s_\tau)) < \mu] \to \exp[-2e^{-\mu}]\,.$$

By taking $\mu = -\log(-0.5\log(1-\delta))$, we have

$$P[|2\log s_\tau|^{1/2}(\tilde{M} - B(s_\tau)) < -\log(-0.5\log(1-\delta))] \to 1-\delta\,,$$

$$P[|2\log s_\tau|^{1/2}\tilde{M} < 2|\log s_\tau| + \log A - \log(-0.5\log(1-\delta))] \to 1-\delta\,.$$

By defining $S := \log A - \log(-0.5\log(1-\delta))$,

$$P[\max_{x\in[0,1]} |W_\psi v(s_\tau, x)| < (2|\log s_\tau|)^{1/2} + \frac{S}{(2|\log s_\tau|)^{1/2}}] \to 1-\delta\,.$$

The theorem is proved with $C_\tau := 1 + \frac{S}{|2\log s_\tau|}$. □

Characterization of C^k-discontinuities

Suppose we have the following model:

$$z(dx) = f(x)dx + \tau v(dx), \tag{6}$$

where τ is the noise amplitude.

Lemma 5. *Let $k = 1, 2$ and let ψ be a Daubechies wavelet of order $m \geq k+1$. Then there exists $\theta \in L_2(\mathbb{R})$ such that*

$$\psi = \theta^{(k)} \quad \text{and} \quad \int \theta(t)dt = 0. \tag{7}$$

Proof: This is a direct consequence of Theorem 6.2. of [9] and the fact that the Daubechies wavelet ψ has $m > k$ vanishing moments [13]. □

Theorem 6. *Let $k = 1, 2$ and suppose we have a* discontinuity *of $f^{(k)}$ at $\sigma \in]0, 1[$, i.e.*

$$f^{(k)}(\sigma^+) \neq f^{(k)}(\sigma^-). \tag{8}$$

Suppose also that $f \in C^{k+1}(]0, 1[\backslash\{\sigma\})$. Define

$$\sigma_\star := \arg \max_{0 \leq x \leq 1} |W_\psi z(s_\tau, x)|,$$

$$s_\tau := (\tau^2 |log\, \tau|^2)^{1/(2k+1)}. \tag{9}$$

Then

$$|\sigma_\star - \sigma| = O_p(s_\tau). \tag{10}$$

Remark 7. The idea behind this theorem is: the position $\sigma_\star$ where the absolute value of the wavelet transform has a local maximum is a good estimate of the unknown exact discontinuity σ (see Fig. 1 for an illustration). Furthermore, (10) specifies that its convergence speed is of order s. Note that this theorem is a generalization of the first part of Theorem 2 in [15]. A similar theorem can be found in [9].

Proof of Theorem 6:

Part 1 [Maximum location]: Let θ be the function defined in (7). Because ψ is a Daubechies wavelet, it has compact support [5] and therefore the construction of θ in Theorem 6.2 of [9] shows that θ also has compact support. Consequently, one can easily prove (see page 24 of [5]) from Lemma 5 that θ fulfills the admissibility condition (4). Multiple partial integrations yield

$$|W_\theta f^{(k)}(s, u)| = \frac{1}{s^k} |W_\psi f(s, u)|. \tag{11}$$

Now, we are able to apply the results in [15] to $g := f^{(k)}$. Let A be such that: $supp(\theta) \subset A$ and A symmetric, i.e. $r \in A \Leftrightarrow -r \in A$. Then (see the proof of Theorem 2 of [15])

$$|W_\theta g(s, x)| \leq K s^{3/2} \qquad \forall x \quad \text{such that} \quad |\sigma - x|/s \notin A.$$

$$\max \left\{ |W_\theta g(s, x)| \, : \, \frac{|\sigma - x|}{s} \in A \right\} \geq K s^{1/2}.$$

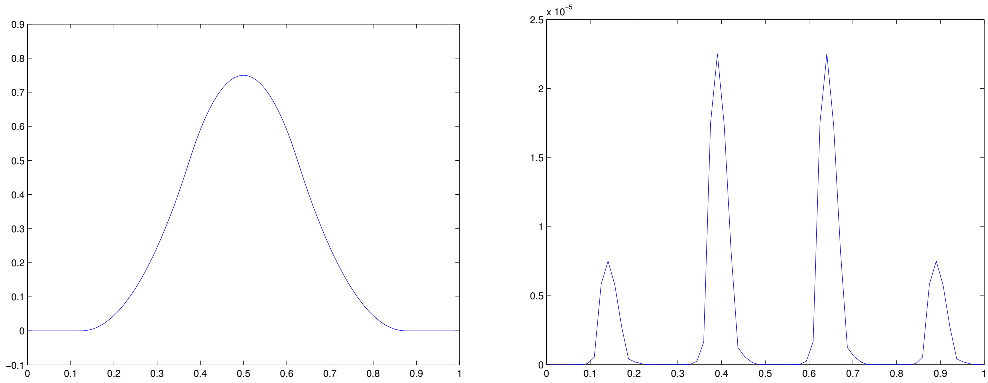

Fig. 1. Function with C^2 discontinuity and its wavelet transform.

These two last relations with (11) then yield

$$|W_\psi f(s,x)| \leq K s^{k+(3/2)} \qquad \forall x \quad \text{such that} \quad |\sigma - x|/s \notin A\,, \tag{12}$$

$$\max\left\{ |W_\psi f(s,x)| \,:\, \frac{|\sigma - x|}{s} \in A \right\} \geq K s^{k+(1/2)}\,. \tag{13}$$

Part 2 [Noise influence]: On the one hand, from (13) and Theorem 4, we have with probability tending to $1-\delta$

$$\max\{|W_\psi z(s_\tau, x)| : (\sigma - x)/s_\tau \in A\} \geq K s_\tau^{k+\frac{1}{2}} - 2\tau C_\tau |\log s_\tau|^{1/2}.$$

Since $C_\tau \to 1$, it is bounded by a constant, and so

$$\max\{|W_\psi z(s_\tau, x)| : (\sigma - x)/s_\tau \in A\} \geq K\tau(|\log\tau| - |\log s_\tau|^{1/2}).$$

Since $\frac{s_\tau}{\tau} \to \infty$, we have $s_\tau \geq \tau$ for sufficiently small τ. Therefore,

$$\max\{|W_\psi z(s_\tau, x)| : (\sigma - x)/s_\tau \in A\} \geq K\tau(|\log\tau| - |\log\tau|^{1/2}) =: a(\tau). \tag{14}$$

On the other hand, we can use the same technique to see that for $(\sigma - x)/s_\tau \notin A$,

$$|W_\psi z(s_\tau, x)| \leq K\tau|\log\tau|^{1/2} =: b(\tau). \tag{15}$$

One can prove that $a(\tau)/b(\tau) \to \infty$. So for sufficiently small τ, $a(\tau) > b(\tau)$. Therefore, by combining (14) and (15), $\max_{x\in[0,1]} |W_\psi z(s_\tau, x)|$ must be obtained for $\sigma_\star$ with $\frac{\sigma - \sigma_\star}{s_\tau} \in A$, i.e. $\sigma_\star = \sigma + s_\tau r$ with $r \in A$. $\square$

Remark 8. The former estimates can be used to determine the threshold value if we want to use wavelet thresholding method like in [6].

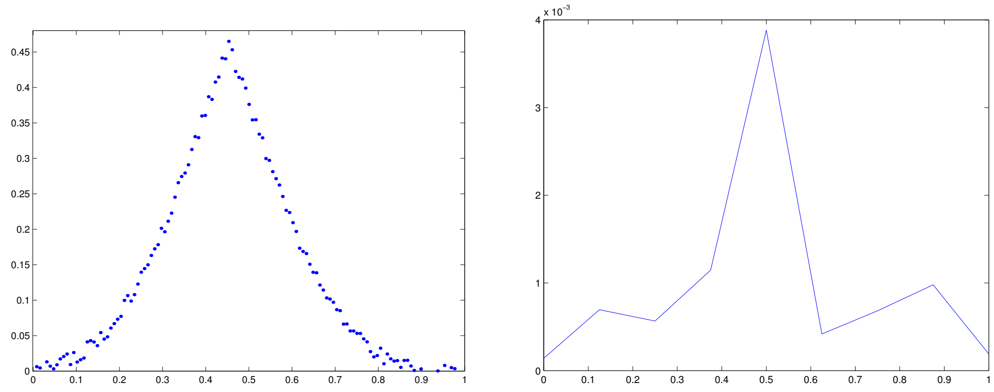

Fig. 2. Noisy data and its wavelet transform at scale $s = 2^{-3}$.

§3. Coping with Noise

Problem pertaining to noise

Let us consider first the regular case, i.e. $x_i = i/N$, $\forall\, i = 0, 1, ..., N$. In practice, the absolute value of the wavelet transform might have a lot of local maxima: the ones which are due to discontinuities and the ones which are due to noise. Let us consider for example the noisy data in Fig. 2. We compute the DWT at different scales. We can notice that results in small scale are more affected by noises than those in large scale (Fig. 2, 3). For example, one cannot recognize in the right plot of Fig. 3 which local maximum is actually corresponding to the desired discontinuity of the derivative. Furthermore, large scales $s = 2^m$ give inaccurate results because the DWT are only computed at $u = k2^m$ (see (5)). For example, in scale $s = 2^{-3}$, we can only have the DWT for $u = 0$, 0.125, 0.25, 0.375, 0.5, 0.625, 0.75, 0.875, 1. So if the exact discontinuity is $\sigma = 0.631$, then 0.625 is the best approximation of it. That means: if we choose small scales, we have trouble with the noise but if we choose large scales, then we lose accuracy. In the following discussion, we give a solution to overcome that problem.

Situation in the neighborhood of the discontinuity

We see from (12) that for those x which are far from σ, we have

$$|W_\psi f(s, x)| \le C s^{k+(3/2)} \to 0 \ . \tag{16}$$

That implies that the wavelet transform of z is very much affected by noise if x is far from σ. Indeed, (16) yields in that case

$$W_\psi z(s, x) = W_\psi f(s, x) + \tau W_\psi v(s, x) \approx \tau W_\psi v(s, x) \ .$$

That idea does not hold anymore for those x which are close to σ because we have a lower bound in (13), while we had an upper bound in (12).

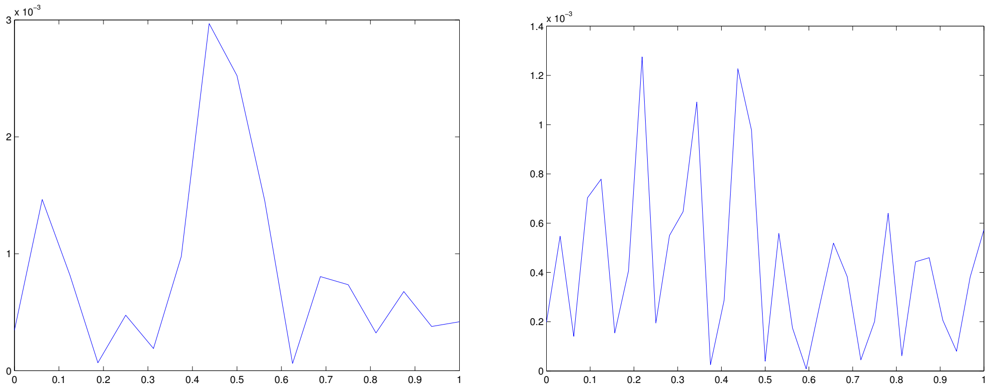

Fig. 3. Wavelet transforms at scales $s = 2^{-4}$ and $s = 2^{-5}$.

Remedy to reduce the influence of noise

We can infer from the previous fact that those x which are close to the C^k-discontinuity σ are less sensitive to noise than those which are far from σ. It is therefore of interest to work in a neighborhood of σ in order to reduce the influence of the noise and we can then use the following algorithm.

Algorithm

$I_0 :=]0,1[$ and $k := 0$

FOR $s = 2^{-c}$, 2^{-c-1},..., finest scale DO:

- Search in I_k for the maximum M_{k+1} of the wavelet transf. at scale s
- Define $I_{k+1} :=]M_{k+1} - s, M_{k+1} + s[$
- k=k+1.

END

Output= M_k.

In the loop, we begin by a scale 2^{-c} in which we are sure that the noise has very small influence.

Illustration

Let us consider again the data set in Fig. 2 which has only one corner at $\sigma = 29 \times 2^{-6} = 0.453125$. Note that there is no integer k with $\sigma = k2^{-m}$ and $m \leq 5$. Therefore we cannot expect to find σ if we only use scales larger than 2^{-6}. We begin by the scale $s = 2^{-3}$ where the noise has almost no influence (see Fig. 2). We apply the previous algorithm to get the sequence of nested intervals: $I_0 =]0,1[$, $I_1 =]0.375, 0.625[$, $I_2 =]0.375, 0.5[$, $I_3 =]0.40625, 0.46875[$ which all contain σ.

§4. Nonuniform Samples

Problem with non-uniformity and remedy

In the previous section we proposed a method that performs well for regular samples. However, samples are often non-equidistant in practice. The problem with applying the former techniques directly is that the DWT in (5) is only given at regular samples $k2^m$, so we must regularize the computations if we want to use the fast algorithm for computing the DWT. Let us consider the problem (3) again and let M be a function having the following properties

(R1) $M(\frac{i}{N}) = x_i$,

(R2) $M \in C^k$.

By defining $g := f \circ M$, we obtain from (3)

$$y_i = g\left(\frac{i}{N}\right) + \varepsilon_i \ . \tag{17}$$

Lemma 9. $k = 1, 2$. *Suppose (R1) and (R2) are fulfilled. Let* $\tilde{\sigma} \in (0, 1)$ *with* $M'(\tilde{\sigma}) \neq 0$. *The following properties are equivalent*

- $g^{(k)}$ *has discontinuity at* $\tilde{\sigma}$,
- $f^{(k)}$ *has discontinuity at* $\sigma := M(\tilde{\sigma})$.

Proof: Obvious □

In order to find the C^k-discontinuity from the nonuniform model (3), we first apply the former method to the regular samples $\{\frac{i}{N}\}$ of the uniform (17) and we have the output $\tilde{\sigma}_\star$. The final output for problem (3) is then $\sigma_\star := M(\tilde{\sigma}_\star)$. For all $\alpha > 0$, denote by $n(\sigma, \alpha)$ the number of samples having abscissae in $]\sigma - \alpha, \sigma + \alpha[$ and we define

$$\text{density}(\alpha) := \frac{n(\sigma, \alpha)}{\alpha} .$$

Let us investigate the dependence of the error $|\sigma - \sigma_\star|$ with the uniformity. We take the same number of samples $N = 128$, and change only the uniformity of the samples. In the second column of Table 1 it is shown that the error decreases as the density increases.

density(0.06)	$\lvert\sigma - \sigma_\star\rvert$	After minimization
266.66	4.0e-03	3.1903e-05
233.33	1.0e-02	4.8542e-05
166.66	3.0e-02	7.7056e-05
66.66	5.0e-02	3.1625e-05

Tab. 1. Density of the data, error, and error after minimization.

Treatment of data which are less dense

It may happen that there are very few data or no data at all next to the discontinuity. That means that the density is very low. So we expect a very bad approximation if we use only the formerly described method. In that case, we take some few samples x_i, $i \in \mathcal{N}$, in a neighborhood of the approximation, and then we determine the coefficients p_i such that the curve in equation (18) best fits those data:

$$C(x) = \begin{cases} p_1 \left(\frac{x - p_2}{p_3 - p_2} \right), & x \leq p_3, \\ p_1 \left(\frac{x - p_4}{p_3 - p_4} \right), & x > p_3. \end{cases} \tag{18}$$

This means that we solve the minimization problem

$$\min_{p_1, \dots, p_4} \sum_{i \in \mathcal{N}} \| C(x_i) - y_i \|^2, \tag{19}$$

which can be solved with the iterative algorithm of Levenberg-Marquardt or Gauss-Newton [10] in which we use the result of the former algorithm as an initial guess. It is very fast in practice because we take only a very few samples. Note that the function C in (18) is a piecewise linear function with C^1-discontinuity at p_3. If we are interested in finding a C^2-discontinuity, we should replace (18) by a function which is piecewise C^2. The last column of Table 1 shows that applying the minimization (19) after performing the former algorithm really improves the accuracy of the results.

References

1. Antoniadis, A., Wavelet estimators for change-point regression model, in Centre de recherches mathématiques proceedings and lecture notes, 1999, 335–346.
2. Bickel, P. J., and M. Rosenblatt, On some global measures of the deviations of density function estimates, Ann. Statist. **1** (1973), 1071–1095.
3. Chen, H. Y., and H. Pottmann, Approximation by ruled surfaces, J. Comput. Appl. Math. **102** (1999), 143–153.
4. Cohen, A., I. Daubechies, and P. Vial, Wavelets on the interval and fast wavelet transforms, Appl. Comput. Harmonic Anal. **1** (1993), 54–81.
5. Daubechies, I., *Ten Lectures On Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM, Philadelphia, 1992.

6. Donoho, D., and I. Johnstone, Ideal spatial adaptation by wavelet shrinkage, Biometrika **81** No. 3 (1994), 425–455.
7. Eubank, R., *Nonparametric Regression And Spline Smoothing*, Marcel Dekker, New York, 1999.
8. Mallat, S., A theory for multiresolution signal decomposition: the wavelet representation, IEEE Trans. Pattern Anal. and Machine Intelligence **11** (1989), 674–693.
9. Mallat, S., *A Wavelet Tour Of Signal Processing*, Academic Press, San Diego, 1998.
10. Nocedal, J., and S. Wright, *Numerical Optimization*, Springer Series in Operations Research, New York, 1999.
11. Riedel, A., The wavelet transform on Sobolev spaces and its approximation properties, Numer. Math. **58** (1991), 875–894.
12. Varady, T., P. Benko, and G. Kos, Reverse engineering regular objects: simple segmentation and surface fitting procedures, Int. J. Shape Model. **4** (1998), 127–141.
13. Vidakovic, B., *Statistical Modeling By Wavelets*, John Wiley & Sons, inc., New York, 1999.
14. Wang, Y., Change curve estimation via wavelets, J. Amer. Statist. Assoc. **93** (1998), 163–172.
15. Wang, Y., Jump and sharp cusp detection by wavelets, Biometrika **82** (1995), 385–397.

M. Randrianarivony
Faculty of Computer Science
Technical Univeristy of Chemnitz
Strasse der Nationen 62
D-09107 Chemnitz
Germany
`maharavo@informatik.tu-chemnitz.de`
http://www.tu-chemnitz.de/~ranh

G. Brunnett
Same postal address
`brunnett@informatik.tu-chemnitz.de`
http://www.informatik.tu-chemnitz.de/~gdv