DETECTION OF ABRUPT CHANGES IN SOUND SIGNALS WITH THE HELP OF WAVELET TRANSFORMS

A. GROSSMANN*, M.HOLSCHNEIDER*
R. KRONLAND - MARTINET+, J. MORLET++

**Centre de Physique Théorique, Section 2, CNRS-Luminy*
Case 907, F-13288 Marseille CEDEX 9
FRANCE

+CNRS , LMA,
31 Chemin Joseph Aiguier
F-13009 Marseille CEDEX
FRANCE

++ORIC , 370 bis, Rue Napoleon Bonaparte
F- 92500 Rueil-Malmaison
FRANCE

INTRODUCTION

So-called "wavelet" or ("ondelette") methods have been used recently with success in a variety of fields; (see e.g. (1) to (6)).

The present paper, which follows a related one (3) is mainly concerned with sound signals. The "wavelet" techniques are here applied to a discussion of questions like: "When did "something new" appear in the signal?" and: "How does the new component behave in a neigbourhood of its arrival time?" . Roughly speaking, a convenient answer to the first question is given by the lines of constant phase of the wavelet transform, which tend to converge towards the points of interest; the second question is answered by the characteristic behaviour of the modulus of the wavelet transform.

Our discussion will start with some basic theory; we continue with the analysis of computer-generated signals with known threshhold behaviour,in the presence of computer-generated noise. We end with some speech and scratchy music.

This paper was written with the aim of illustrating some applications of wavelet transforms. We don't feel competent to discuss in general the topic of abrupt changes in signals, and refer the reader to (7).

INVERSE PROBLEMS
ISBN 0-12-014581-2

WAVELET TRANSFORMS

Scale-covariant Representations

Consider a fairly arbitrary real-valued function s(t), which will be called a signal; for the sake of definiteness, the argument t will be called time. Let us introduce an additional parameter, the positive "scale parameter" a. We want to associate to s(t) a function S(t,a) which will allow us to disentangle the properties of s(t) corresponding to different scale ranges.

If a is small, (i.e. if the point (t,a) lies close to the edge of the (t,a) half-plane), then, in a sense to be made precise, the function S(t,a) describes the local behaviour of s, around t, while S(t,a) (a large) describes large-scale properties.

Let us discuss for a moment the properties of this time-and-scale half-plane H. One should always keep in mind that the boundary (a = 0) does not belong to H; zero is not a reasonable scale. There are clearly two natural families of transformations on H. The first one consists of the shifts parallel to the boundary: a-->a and t--->t - t_o. The second family consists of re-scalings: fix t_o, and consider the family

$$(t - t_o) \dashrightarrow \lambda (t - t_o), \qquad a \dashrightarrow \lambda a \qquad (\lambda > 0)$$

The natural surface element on H is a^{-2} dt da. It does not change under shifts and re-scalings. We now list a set of requirements to impose on a corespondence s(t) -->S(t,a) so that it can be confortably used as a description of signal on H.

(i) The correspondence s --> S should be linear.(Between suitable function spaces)

(ii) The correspondence should not depend on the choice of the origin of time. If S(t,a) is the transform of s(t), then, for any b, S(t-b,a) should be the transform of s(t-b).

(iii) If s(t) has as transform S(t,a), then, for any $\lambda > 0$,the re-scaled signal $\lambda^{1/2}s(\lambda t)$ (which has the same energy as s) has as transform the function $S(\lambda t,\lambda a)$.

(iv) There should be an "energy conservation law": this is crucial in order to ensure that there is no loss of information in going from s to S and that the signal s can be easily reconstructed from its transform.The requirement is that there should exist a constant C, independent of the choice of s(t), such that, for all square integrable signals s(t), one has

$$\int s(t)^2 dt = C \iint |S(t,a)|^2 \, a^{-2} da \, dt$$

where S(t,a) is the transform of s(t).

An explicit construction of such correspondences s-->S is provided by so-called wavelet transforms which will be described in the following section.

Notations :

We allow the signal s(t) to be a generalized function (distribution). The inner product

$$(g,f) = \int g^*(t)\, f(t)\, dt$$

(where the star denotes complex conjugation), will be naturally extended when possible. For convenience in calculations, we introduce the following families of unitary operators:

(i) Shifts: $(T^b f)(t) = f(t-b)$ (b real)

(ii) Unitary dilations: $(D^a f)(t) = a^{-1/2} f(t/a)$ $(a > 0)$

(iii) Multiplications by exponential $(E^b f)(t) = e^{ibt} f(t)$ (b real)

(iv) Fourier transform: $(Ff)(\omega) = (2\pi)^{-1/2} \int e^{-i\omega t} f(t)\, dt$

The commutation properties between these operators are easily calculated and are constantly used. The most important for us is the relation $D^a T^b = T^{ab} D^a$.

Wavelets:

The wavelets that we consider here are complex-valued functions such that

(i) $(g,g) < \infty$,

(ii). The Fourier transform $g\^(\omega)$ is real and vanishes for $\omega \leq 0$, and finally

(iii) $\int |g\^(\omega)|^2 (1/\omega)\, d\omega < \infty$

The last condition, (g should have no zero-frequency component) is crucial for the"energy conservation law".

Wavelet transforms:

Choose a wavelet g satisfying the conditions (i), (ii), (iii) above.

For simplicity in notations, normalize g so that $2\pi\int|g^\wedge(\omega)|^2(1/\omega)d\omega=1$. The wavelet transform of the signal s with respect to the wavelet g is then the complex-valued function S defined as

$$S(t,a) = (T^tD^ag,s) = a^{-1/2}\int g^*((t'-t)/a)s(t')dt';$$

it can also be written in terms of the Fourier transforms of s and g as

$$S(t,a) = (E^{-t}D^{1/a}g^\wedge,s^\wedge) = a^{1/2}\int e^{ib\omega}\, g^\wedge(a\omega)\, s^\wedge(\omega)\, d\omega$$

One can check that the correspondence s --> S satisfies the requirements stated above. Given g, it is possible to characterize the functions S(t,a) that are wavelet transforms of some s. This involves "reproducing kernel" conditions which are described e.g. in (6).

Wavelet transforms of homogeneous functions

A function f is said to be homogeneous of degree α if it satisfies, for every $a > 0$, $f(at) = a^\alpha f(t)$ An example is a signal that vanishes for $t < 0$ and then starts as a (in general fractional) power. In terms of the unitary operator D^a, the homogeneity condition is written as

$$D^{1/a}f = a^{\alpha+(1/2)}\, f.$$

The above definition of homogeneity refers to the point $t = 0$. We shall say that a function f is homogeneous of degree α at the point $t = b$, if the function T^bf is homogeneous of degree α at $t=0$.

We shall need some special families of homogeneous functions. For $\mathrm{Re}\alpha>-1$, consider the function

$$u_+(\alpha,t) = (1/\Gamma(\alpha+1))\, t^\alpha \qquad \text{for } t > 0$$

$$= 0 \qquad \text{for } t \leq 0$$

Here $t^\alpha = e^{\alpha \ln t}$.

Considered as a family of generalized functions, $u_+(\alpha,.)$ depends analytically on α. Furthermore this family can be analytically continued from the half-plane $\mathrm{Re}\,\alpha > -1$ to the whole α plane. Explicit expressions for this continuation are given in Gel'fand and Shilov's book on generalized functions. In particular, one has,

$$u_+(-n,t) = \delta^{(n-1)}(t) \qquad (n = 1,2,...)$$

(derivative of Dirac's delta function). The function $u_+(\alpha,t)$ is homogeneous of order α.

Let s(t) be homogeneous of degree α at $t_o = 0$. Let S(b,a) be its wavelet transform. Then S(b,a) satisfies

$$S(b,a) = a^{\alpha + (1/2)}\, S(b/a,\, 1)$$

Indeed, one has

$$S(b,a) = (T^b D^a g,\, s) = (D^a\, T^{b/a},s) = (T^{b/a}g, D^{1/a}s)$$

$$= a^{\alpha+(1/2)}(T^{b/a}g,s) = S(b/a,1).$$

If s(t) is homogeneous at an arbitrary t_o, one has

$$(T^{b\text{-}to}\ D^a g,\, s) \;=\; (T^b\, D^a g, T^{to}\, s)$$

giving

$$S(b\text{-}t_o,a) = S((b\text{-}t_o)/a,\, (b\text{-}t_o)/a).$$

In order to simplify notations, we shall take $t_o = 0$.

Let us now discuss the behaviour of S(b,a) as a tends to 0, i.e. as the point (b,a) approaches the boundary of the scale half-plane. We consider only approaches along straight lines. One clearly has to distinguish several cases:

(i) The line intersects the point of homogeneity (b=0,a=0). Along this line one has (b/a) = const, so

$$S(b,a) = a^{\alpha+(1/2)}\, C$$

where C = S(const , 1). If $C \neq 0$, then

$$\ln |S(b,a)| = [\mathrm{Re}\ \alpha + (1/2)]\ \ln a + \ln|C|;$$

$$\arg(S(b,a)) = \mathrm{Im}\ \alpha\ \ \ln a + \arg(C)$$

Notice an immediate consequence: If α is real, then the phase of S(b,a) is constant along lines converging to the point of homogeneity. This feature will be seen to survive, to a good approximation, as a detector of abrupt changes in real-life signals.

(ii) The line does not intersect the point of homogeneity. The behaviour of S(b,a) is then governed not only by the factor $a^{\alpha+(1/2)}$ but also by the factor S(b/a, 1). As a tends to zero, b/a tends to infinity. If the product $g\^(\omega)s\^(\omega)$ is smooth enough (in particular, if the zero of $g\^(\omega)$ at $\omega=0$ is of high enough order) then this factor decreases rapidly as a tends to zero, giving rise to behaviour very different from that described under (i). This feature again survives in real-life situations.

Before looking at actual signals, we illustrate the above results with an example where the wavelet transform is given in closed form.

Explicit example

Denote by h_o the gaussian:

$$h_o(t) = \exp(-t^2/2).$$

The numerical examples in this paper were calculated with the wavelet g(t) having as Fourier transform

$$g^\wedge(\omega) = h_o(\omega - \Omega)$$

where $\Omega = \pi(2/\ln 2)^{1/2} \approx 5.33$. This is the wavelet introduced by one of us (J.M.) many years ago. Strictly speaking, one should add to this function appropriate counterterms in order to make it satisfy the conditions (ii) and (iii) above. In practice, however, these counter-terms are negligibly small.

The wavelet transform of of $u_+(\alpha,.)$ with respect to this wavelet can be easily obtained bu noticing that one of the known integral representations of parabolic cylinder functions D_ν is

$$(E^\Omega h_{o,} u_+(\alpha,.)) = \exp(-\,\Omega^2/4)\; D_{-\alpha-1}(i\Omega).$$

One obtains, writing $c = b/a$,

$$(T^b D^a E^\Omega h_o,\, u_+(\alpha,.))$$

$$= \exp(i\Omega c)\;\; a^{\alpha+(1/2)} \exp(-(\Omega - ic)^2/4)\; D_{-\alpha-1}(c+i\Omega)$$

Remarks on random signals and on noise

We now want to give some ideas about the behaviour of the wavelet transform of noise. Consider a stationary random process s(t); that is a set of random variables, indexed by the time t. It is quite natural to suppose that the average of s(t) is zero:

$$< s(t) > = 0$$

(the brackets "<,>" denote the average over all possible realisations of the random process s(t)). We will now consider the correlation function at two points:

$$\Phi(t_1 - t_2) = < s(t_1)\ s(t_2) >.$$

Since the process is stationary, Φ depends only on the time difference. For $t_1=t_2$ we obtain - up to an infinite normalisation constant - the variance of the random variable s(t). It is convenient to introduce the Fourier-transform $F\Phi$ of Φ, which is usualy called the characteristic function:

$$F\Phi(\omega) = (2\pi)^{-1/2} \int e^{-i\omega t}\, \Phi(t)\, dt$$

Now consider the wavelet transform S(b,a) of this proces. For each fixed point (b,a) on the half-plane, this defines a random variable, and so S(b,a) is a random process over the half-plane.One sees that, in the mean, S(b,a) vanishes as does s(t):

$$<S(b,a)> = <a^{-1/2} \int g^*((t'-t)/a) s(t') dt'>$$

$$= a^{-1/2} \int g^*((t'-t)/a) <s(t')> dt' = 0.$$

We shall now consider the correlation between the wavelet transform at two different points in the half-plane:

$$<S^*(b,a) S(b',a')>$$

$$= <(aa')^{-1/2} \int g^*((t-b)/a)\ g((t'-b')/a')\ s(t)\ s'(t')\ dtdt'>$$

$$= (aa')^{-1/2} \int g^*((t-b)/a)\ g((t'-b')/a')\ <s(t)\ s'(t')>\ dtdt'$$

$$= (aa')^{-1/2} \int g^*((t-b)/a)\ g((t'-b')/a')\ \Phi(t-t')\ dtdt'$$

We now replace $\Phi(t-t')$ by the inverse Fourier transform of the characteristic function $F\Phi$. Changing the order of integration we find:

$$<S^*(b,a) S(b',a')>$$

$$= (2\pi)^{-1/2} (aa')^{-1/2} \int g^*((t-b)/a)\ g((t'-b')/a')\ e^{i\omega(t-t')}\ F\Phi(\omega)\ d\omega\ dtdt'>$$

$$= (2\pi)^{1/2}\ (aa')^{1/2} \int (Fg)^*(a\omega)\ (Fg)(a'\omega)\ e^{i\omega(b-b')}\ F\Phi(\omega)\ d\omega$$

To understand better the significance of this equation we will consider the special case of a white noise. This is the extremal situation where the two random variables s(t) and s(t') are completely independent whenever $t \neq t'$:

$$\Phi(t-t') = \sigma^2\, \delta(t-t'); \quad F\Phi = \sigma^2 / \sqrt{2\pi}$$

Then the two-point correlation function is essentially given by the reproducing kernel:

$$<\bar{S}(b,a)S(b',a')> = \sigma^2 c_g \, p_g((b-b')/a', \ a/a').$$

where

$$p_g(u,v) = (T^u D^v g,g)\,(1/c_g), \quad \text{with } c_g = 2\pi\int|g(\omega)|^2\,(1/\omega)\,d\omega.$$

So, since p_g is localized around (0,1) say on $(\pm\Delta, 1\pm\Delta)$, we see that the wavelet transform at two points (b,a), $(b+\Delta b, \ a+\Delta a)$ become essentially independent whenever

$$\Delta b \,/\, a < \Delta, \text{ and } \Delta a \,/\, a < \Delta$$

In Fig.1 a we show the modulus of the wavelet transform of a realisation of a white noise. One remarks the different black regions indicating coherent behaviour of the transform. As expected, the size $\Delta \ln(a)$ - note the logarithmic scale - is essentially constant, whereas the size in the b - direction Δb depends on a.

Setting (b,a)=(b',a') we find that the average energy $|S|^2$of the white noise is everywhere the same and is given by:

$$|S|^2 = \sigma^2 \,||g||_2{}^2$$

Hovever, for correlated random processes, the local mean energy will depend on the length scale a. For a stationary process it always is independent of b.

EXAMPLES

The pictures below represent the wavelet transform of various signals, computed with repect to the gaussian wavelet shifted in frequency, which was described above. The signal is shown in the black strip above the transform. The scale parameter a, on a logarithmic scale, points downwards from the signal, so that the upper edge corresponds to the smallest chosen value of the parameter a. The logarithmic scale distorts the straight phase lines described in the text into a funnel-like shape.

Figure 1 is the wavelet transform of a computer-generated noise.

Figures 2 ,4 and 6 show the wavelet transform of the function that vanishes for $t < 0$ and is equal to $t^\alpha \, e^{-t}$ for $t \geq 0$. The values of α are, successively, 0, 1/2 and 2. In figures 3,5 and 7, a noise at 30% level has been added to the signal.

Figure 8 shows the beginning a speech signal. The total time is 32 ms. On the picture of modulus, one sees first a high-frequency component, followed by a part carrying lower frequencies. The two onsets can be seen in the phase picture 8b.

The pictures 9a and 9b are the wavelet transform of the music coming from an old record . The total length of time shown is 16 ms.The feature that can be seen on the left half of Fig 9a corresponds to a scratch.Notice the general similarity of of this peak with the academic examples. Again, the position of this perturbation can be inferred quite precisely from the lines of constant phase in Fig 9b.

ACKNOWLEDGEMENTS

This work was performed within the framework of the RCP 820 ("Ondelettes") of the CNRS. We have benefited from numerous discussions with our colleagues.

References

1. Lemarié, P.G. and Meyer,Y. (1986)."Ondelettes et Bases hilbertiennes." To appear in Revista Matematica Ibero-Americana, volume dedicated to A. Calderon.

2. Battle,G:(1986). "A Block Spin Construction of Ondelettes". Preprint, Mathematics Deparment, Cornell University.

3. Kronland-Martinet, R., Morlet J. and Grossmann, A. (1987). "Analysis of Sound Patterns through Wavelet Transforms". To appear in International Journal of Pattern Recognition and Artificial Intelligence, Special Issue on Expert Systems and Pattern Analysis.

4. Daubechies, I. and Paul, T.(1986) "Wavelets and Applications". To appear in the Proceedings of the 8th International Congress on Mathematical Physics, Marseille 1986.World Scientific, Singapore.

5. Goupillaud, P., Grossmann, A. and Morlet, J. (1984) Geoexploration 23, 85-102.

6. Grossmann, A. and Morlet, J. (1984). SIAM J. Math. An. 15, 723-736.

7. Basseville, M. and Benveniste, A. "Detection of Abrupt Changes in Signals and Dynamical Systems" . Springer, Berlin, 1986.

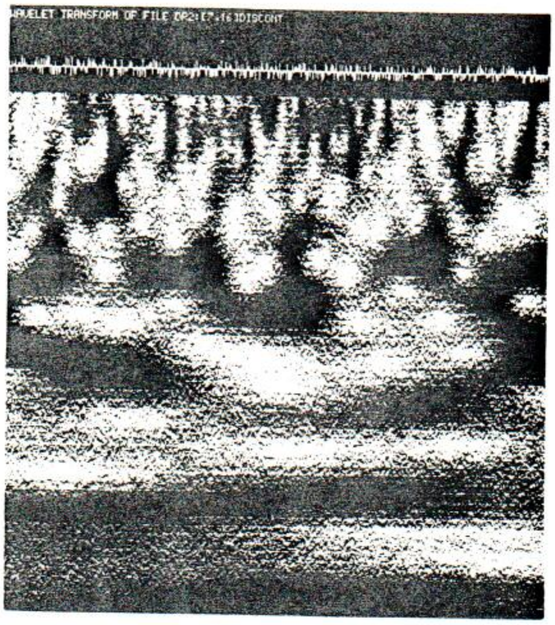

Fig 1.a

Modulus of the wavelet transform of a random signal. Notice the varying sizes of the black regions due to the dependence of the correlation length on the scale a.

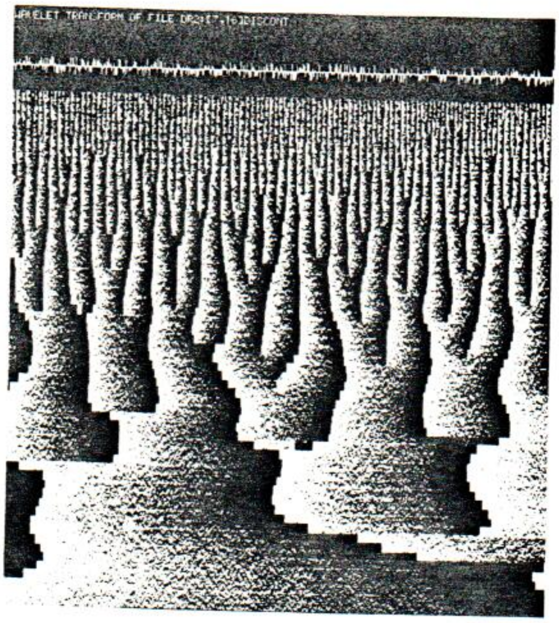

Fig 1.b

Here we represent the associated phase picture of fig 1.a.

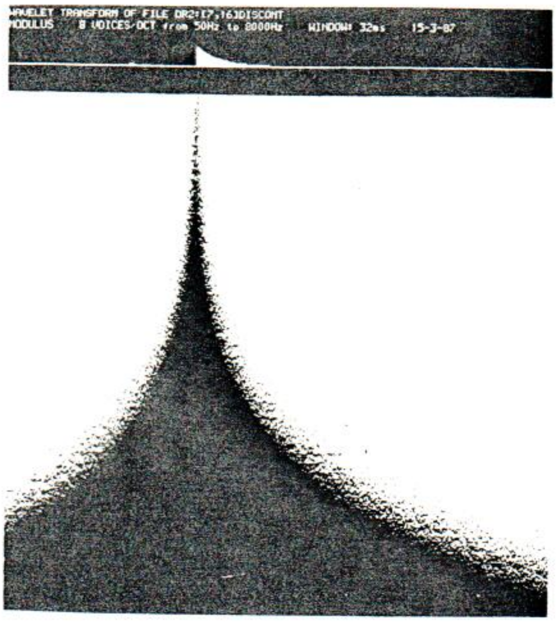

Fig 2.a

This picture represents the modulus of the wavelet transform of an "onset" signal. The exponent is 0.

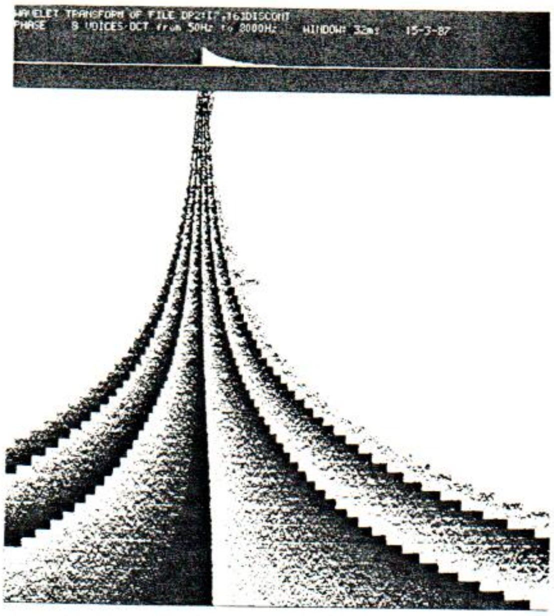

Fig 2.b

Here we show the phase picture associated to fig 2.a. Notice the convergence of the lines of constant phase to- wards the point on the border of the halfplane, where the singularity is situa- ted.

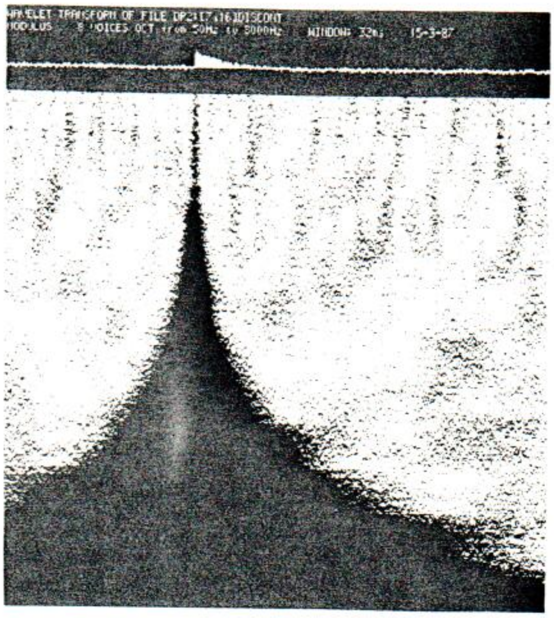

Fig 3.a

This picture shows the modulus of the wavelet transform of the same signal as in fig 2.a with 30% noise added. Notice that the characteristic feature of fig 2.a is preserved.

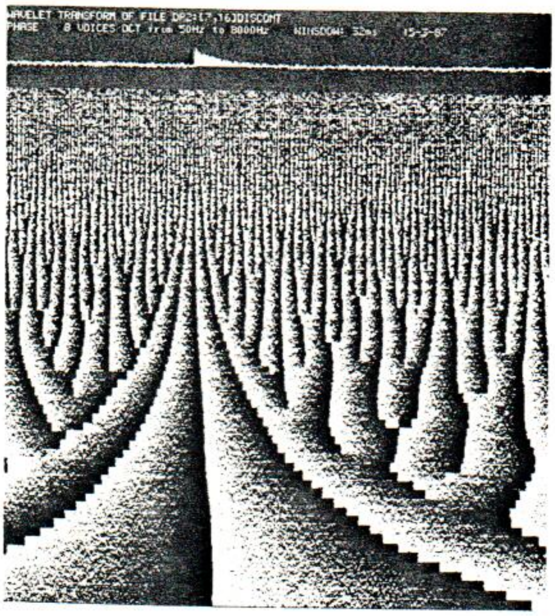

Fig 3.b

The associated phase picture of fig 3.a. Notice that the position of the onset can still be localised with the help of the lines of constant phase con- verging towards this point.

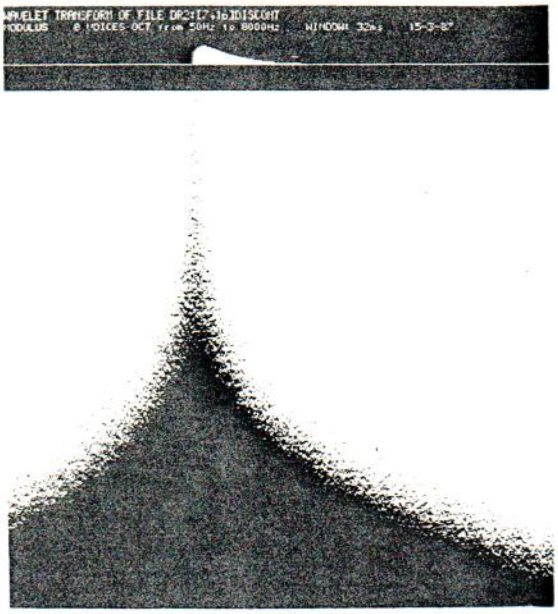

Fig 4.a

This picture represents the modulus of the wavelet transform of an "onset" signal. The exponent is 1/2. Notice that there is less energy in the high frequencies than in fig2.a.

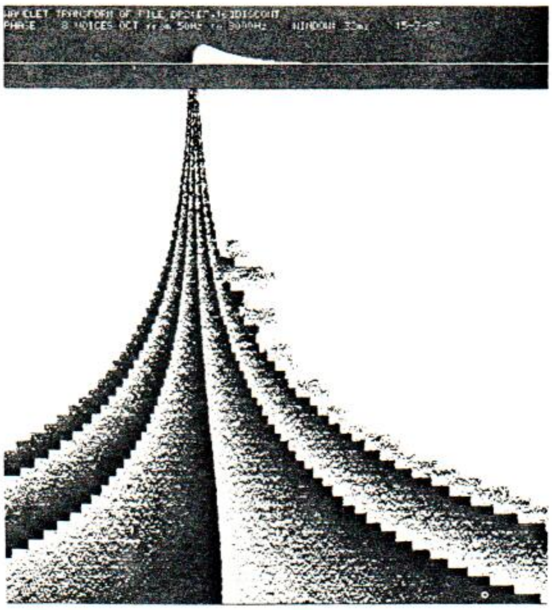

Fig 4.b

Here we show the phase picture associated to fig 4.a. The lines of constant phase converge towards the point on the border of the halfplane, where the singularity is situated.

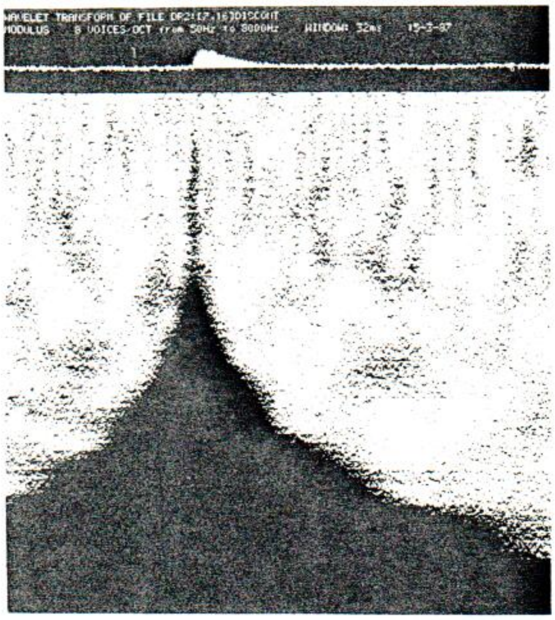

Fig 5.a

This picture shows the modulus of the wavelet transform of the same signal as in fig 4.a with 30% noise added. Notice that the characteristic feature of fig 4.a is preserved.

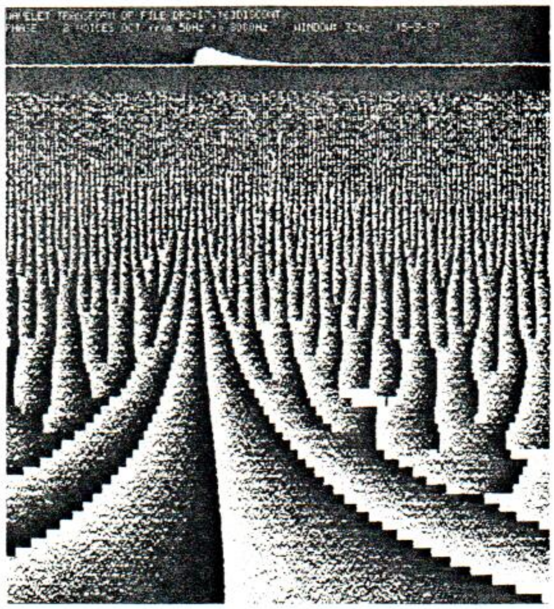

Fig 5.b

The associated phase picture of fig 5.a. Notice that the position of the onset can still be localised with the help of the lines of constant phase converging towards this point.

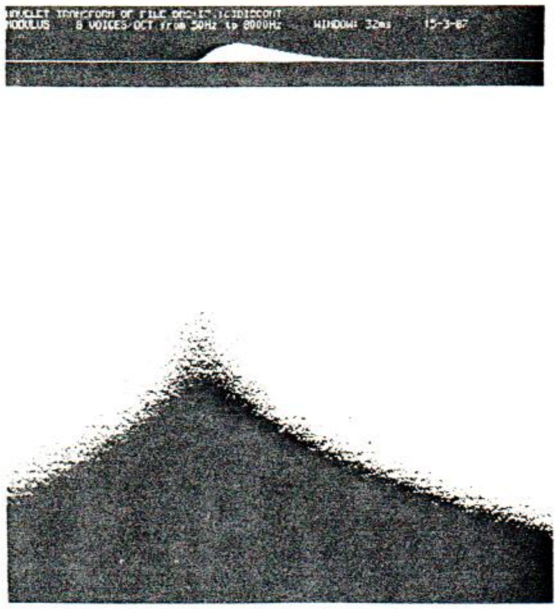

Fig 6.a

This picture represents the modulus of the wavelet transform of an "onset" signal. The exponent is 2. Notice that there is less energy in the high frequencies than in fig2.a and fig 4.a.

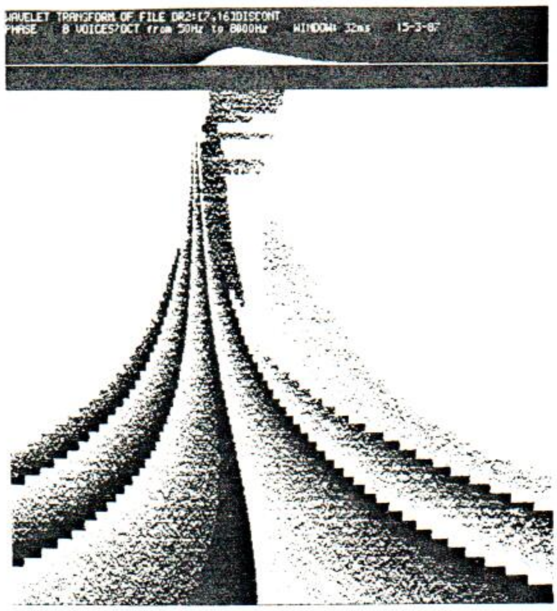

Fig 6.b

Here we show the phase picture associated to fig 6.a. Again the lines of constant phase converge towards the point on the border of the halfplane, where the singularity is situated.

Fig 7.a

This picture shows the modulus of the wavelet transform of the same signal as in fig 6.a with 30% noise added.

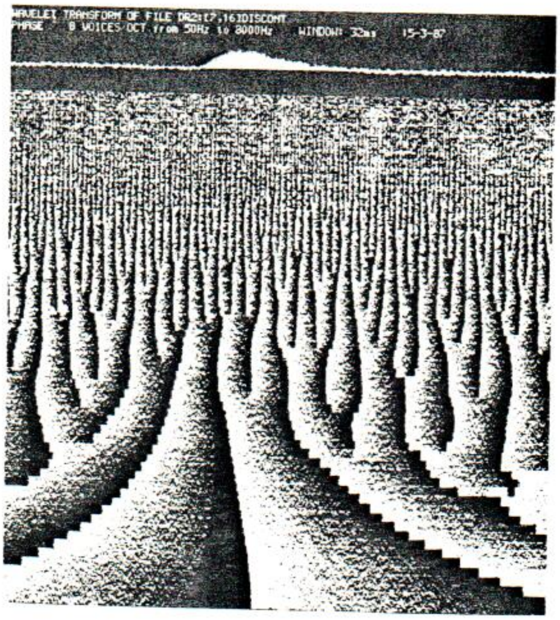

Fig 7.b

The associated phase picture of fig 7.a. Notice that the position of the onset can still be localised with the help of the lines of constant phase converging towards this point.

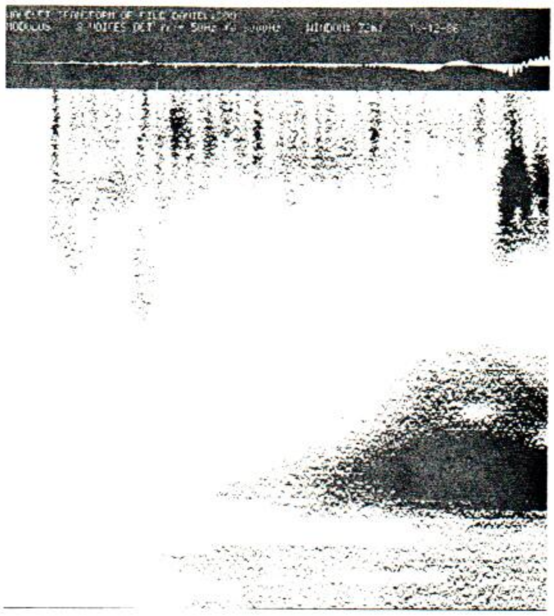

Fig 8.a

The modulus of the wavelet transform of a speach signal: the first 32 ms of "tik". The frequency range is from 50Hz to 8000Hz. Notice the two different characteristic features.

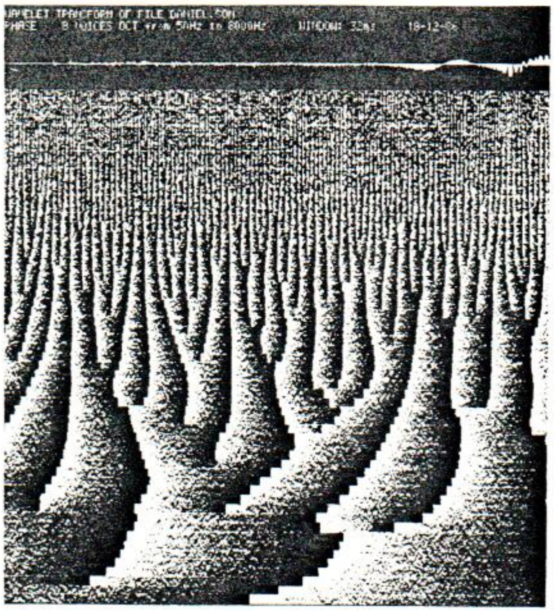

Fig 8.b

The phase picture associated to fig 8.a. Notice the convergence of the lines of constant phase towards the onsets of the two different features in fig 8.a.

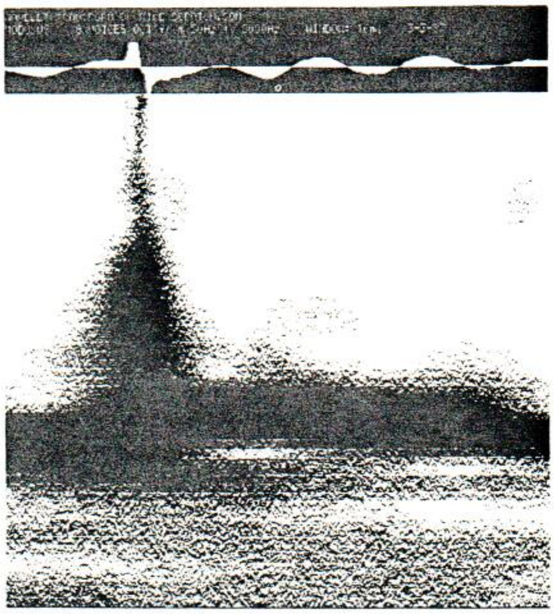

Fig 9.a

The modulus of the wavelet transform of a soundsignal taken from an old record. The scratch gives rise to a feature similar to the one of fig 2.a.

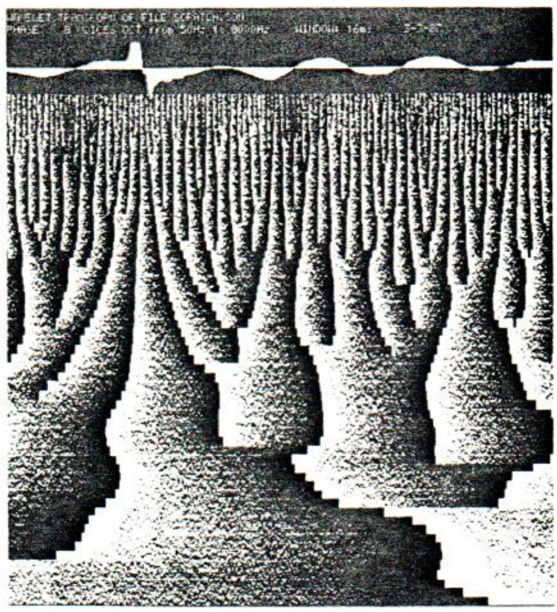

Fig 9.b

The modulus associated to fig 9.a. With the help of the lines of constant phase one can locate the scratch of the record.