Hilbert transform using a robust geostatistical method

O Marashly\textsuperscript{1,*}, M Dobroka\textsuperscript{2}

\textsuperscript{1}Ph.D. student, Geophysics Department, Miskolc University, H-3515 Miskolc Egyetemvaros, Hungary
\textsuperscript{2}MTA-EM Applied Geoscience Research Group, Miskolc University, H-3515 Miskolc Egyetemvaros, Hungary

*Corresponding author: gfomar@uni-miskolc.hu

Abstract. In this paper, we introduced an efficient inversion method for Hilbert transform calculation which can be able to eliminate the outlier noise. The Most Frequent Value method (MFV) developed by Steiner merged with an inversion-based Fourier transform to introduce a powerful Fourier transform. The Fourier transform process (IRLS-FT) ability to noise overthrow efficiency and refusal to outliers make it an applicable method in the field of seismic data processing. In the first part of the study, we introduced the Hilbert transform stand on an efficient inversion, after that as an example we obtain the absolute value of the analytical signal which can be used as an attribute gauge. The method depends on a dual inversion, first we obtain the Fourier spectrum of the time signal via inversion, after that, the spectrum calculated via transformation of Hilbert transform into time range using a efficient inversion. Steiner Weights is used later and calculated using the Iterative Reweighting Least Squares (IRLS) method (efficient inverse Fourier transform). Hermite functions in a series expansion are used to discretize the spectrum of the signal in time. These expansion coefficients are the unknowns in this case. The test procedure was made on a Ricker wavelet signal loaded with Cauchy distribution noise to test the new Hilbert transform. The method shows very good resistance to outlier noises better than the conventional (DFT) method.

1. Introduction

Seismic attributes play an important role in seismic data processing and interpretation. They can be physical attributes (further classified as pre-stack or post-stack ones having sub-classes of instantaneous or wavelet attributes) or geometrical attributes used to obtain certain information about dip, azimuth and discontinuity (further classifications are also exist). They are used for detecting fractures or faults and geological lineaments which are below the seismic resolution for providing important structural and stratigraphic information in three dimensions.

Amplitude, phase, and frequency are the main characteristics elements of seismic traces. Extracting instantaneous and local values of these parameters is a useful way of revealing hidden features. In addition to the basic attributes, weighted averaging of instantaneous seismic attributes is beneficial. This technique is most useful for reducing spikes and noise variations [1].

On another important side, it’s become an important tool to geoscientists for reservoir characterization and quality control. Since the paper published by Taner et al. [2], the subject area has expanded and significantly developed. Today, we can talk about a wide range of attributes, divide them into physical and geometric attributes, classify procedures according to whether they are applied before or after processing, interpreted on one or more seismic channels, etc. In developing and
applying measurement and data processing procedures, it is important to manage and improve the signal-to-noise relationship as much as possible.

Fourier transform is often a priority in the creation of attribute stations. Szegedi and Dobroka [3] proposed a robust Fourier transform process (IRLS-FT) built on an inversion basis. In this paper, we demonstrate that the method works effectively in suppressing outlier noise and can repair the signal-to-noise relationship by up to an order of magnitude. As a first step, we present the production of the Hilbert transform as part of a transformed robust/resistant inversion, which plays a decisive role in the definition of the complex channel. The reflection strength attribute was introduced in the article which represents the total energy strength of the trace and is independent of phase due to which it is more suitable to analyze the amplitude anomalies in detail.

2. The analytical signal

The starting point for calculating the basic attribute stations is the creation of an analytical signal (analytical or complex channel). The concept of analytical signal in data processing was introduced by the Nobel Prize laureate Hungarian physicist Dénes Gábor [4]. His ambition was to use the powerful mathematical tools of quantum mechanics in signal processing (using square-integrable complex functions as elements of the so-called Hilbert space). To do this he introduced the analytical signal

\[ s(t) = u(t) + j u_H(t) \]  

(1)

where \( u(t) \) is the time signal and

\[ u_H(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} u(\tau) \frac{d\tau}{\tau - t} \]  

(2)

is the Hilbert-transform of the time signal. The analytical signal or complex trace was first discussed in details by Bracewell [5] and Taner [2] defining the most important seismic attributes: the reflection strength, the instantaneous phase and frequency, etc. The concept and its use is well-discussed in seismic literature. According to Eq. (2), the Hilbert transform is generated as a convolution of the time signal \( u(t) \) with function \( \frac{1}{\pi t} \). In the frequency domain, this relation can be written as

\[ F\{u_H(t)\} = F\{u(t)\}F \left\{ -\frac{1}{\pi t} \right\} \]  

(3)

where \( F \) denotes the Fourier transform. As \( F \left\{ -\frac{1}{\pi t} \right\} = -j \, \text{sgn}(\omega) \), introducing the notation

\[ U(\omega) = F\{u(t)\} \]  

(4)

one can write

\[ F\{u_H(t)\} = -j \, \text{sgn}(\omega)U(\omega) = U_H(\omega) \]  

(5)

giving the Hilbert transform as

\[ u_H(t) = F^{-1}\{U_H(\omega)\} \]  

(6)
It can be seen that when calculating the Hilbert transform the Fourier transform and its inverse are to be used. It is well known that the traditional DFT and IDFT algorithms are noise-sensitive in the case of non-Gaussian noise in the data set. To define a robust procedure an inversion-based Fourier transform method (IRLS-FT) was introduced by Dobroka et al. [6]. In order to use the noise-rejection capacity of the method, we can use IRLS-FT in calculating $U(\omega)$ as

$$U(\omega) = \mathcal{F}_{\text{IRLS}}[u(t)]$$

with IDFT in Eq. (6). A full inversion-based method can be produced when the $\mathcal{F}^{-1}_{\text{IRLS}}$ procedure is used also in Eq. (6).

3. Reflection strength attribute and its noise sensitivity

Once the analytical signal (6) is known, the attributes can be produced. The reflection strength (instantaneous amplitude, paving) is further examined as an example, which is the absolute value of the analytical signal

$$A(t) = \sqrt{u(t)^2 + v(t)^2}$$

The noise sensitivity of this attribute is illustrated using the Ricker wavelet [7] shown in Figure 1 (in blue). The time series shows a 10 Hz wave packet localized at 0.1 sec with a sampling interval of 0.005 sec in the [-1,1] (sec) domain. The data set (I) was generated by adding Gaussian noise (with a standard deviation of 0.0025 and zero mean) to the noise-free data. A data system (II) with outlier errors was created by generating noise following the Cauchy distribution of scale parameter $\epsilon=0.04$. Figures 1–3 show the reflection strength channel calculated by the conventional process for noise-free inputs laden with Gauss noise (I) and Cauchy noise (II). Hilbert's transformed functions were carried out using the discrete Fourier transform process (DFT) and its inverse (IDFT).

![Figure 1. The noiseless time-domain Ricker wavelet (blue) and noiseless reflection strength (red).](image-url)
Figure 2. The noisy Ricker wavelet (data set I) (blue) and noisy reflection strength (red). The discrete Fourier transform method (DFT) and its inverse (IDFT) were used.

Figure 3. The noisy Ricker wavelet (data set II) (blue) and noisy reflection strength (red). The discrete Fourier transform method (DFT) and its inverse (IDFT) were used.

It can be seen that the reflection strength gauge calculated based on the data system containing outliers II is particularly noisy. To characterize noise sensitivity, we introduce
distance in the data space, which is $d(\text{I}) = 0.0240$ for data set I and for the data set II the distance $d(\text{II}) = 0.0444$ (N being the number of the samples). Figure 3 justifies the development of a Hilbert transform process more resistant to outlier noise. Since Hilbert's transformed training is based on the traditional Fourier transform (DFT), it is obvious that the inversion-based Fourier transform procedure (IRLS-FT) utilizing Steiner weights [3] is used to solve the task.

4. The inversion-based Fourier transform (IRLS-FT)

The Fourier transform links the time domain of the signal registration and the frequency domain of the signal test according to the following formulas:

$$ U(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u(t) e^{-j\omega t} \, dt , \quad u(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} U(\omega) e^{j\omega t} \, d\omega $$

The frequency spectrum $U(\omega)$ is a Fourier transform of the time signal $u(t)$, which is usually a continuous function of the complex value. When using series expansion based discretization, the spectrum is written in terms of a suitably chosen system of basis functions:

$$ U(\omega) = \sum_{n=1}^{M} B_n \Psi_n(\omega) $$

where $B_n$ denotes the complex expansion coefficients, $\Psi_n(\omega)$ denotes the $n$-th known basis function, and $M$ is the total number of the basis functions. If the Fourier transform is understood as an overdetermined inverse problem, we must first designate the direct problem, which is the inverse Fourier transform, defined in the case of the $k$-th measurement datum as

$$ u^{(elm)}(t_k) = u_k^{(elm)} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} U(\omega) e^{j\omega t_k} \, d\omega $$

Combining the above equations, the calculated (theoretical) data can be determined according to the following system of linear equations:

$$ u^{(elm)}(t_k) = u_k^{(elm)} = \sum_{n=1}^{M} B_n G_{kn} $$

Where

$$ G_{kn} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Psi_n(\omega) e^{j\omega t_k} \, d\omega = \mathcal{F}^{-1}_k(\Psi_n(\omega)) $$

is the Jacobi matrix. The deviation vector of measured and calculated data is given by

$$ e_k = u_k^{(measured)} - u_k^{(calculated)} = u_k^{(measured)} - \sum_{n=1}^{M} B_n G_{kn} $$
The expression contains the expansion coefficients, which are determined by minimizing some norm of the deviation vector. Knowing the expansion coefficients, the spectrum can be determined at any frequency as

\[ \mathcal{U}^{(estimated)}(\omega) = \sum_{n=1}^{M} B_n^{estimated} \Psi_n(\omega) \]  

(15)

5. The robust generation of Hilbert transform

To produce the Hilbert transform we need to know the \( \mathcal{U}(\omega) \) spectrum of the signal. To improve the signal-to-noise ratio, the Fourier transform is performed by the IRLS -FT method, and then the spectrum is multiplied by the function \(-j \text{ sgn}(\omega)\). We then return to the time domain by an inverse Fourier transform. This can be done by applying IDFT and also by using an inversion-based inverse Fourier transform. The latter procedure can be performed by robust inversion defined here. The starting point is the expression of the inverse Fourier transform

\[ \mathcal{U}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u(t)e^{-j\omega t} dt \]  

(16)

The direct problem is given by the formula of the Fourier transform, where the time function \( u(t) \) is discretized in the form of a series expansion

\[ u(t) = \sum_{n=1}^{M} B_n \Psi_n(t) \]  

(17)

After substitution, the following formula is obtained for the k-th sampling element of the spectrum

\[ \mathcal{U}_k(\omega_k) = \sum_{n=1}^{M} B_n \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Psi_n(t) e^{-j\omega_k t} dt = \sum_{n=1}^{M} B_n G_{kn} \]  

(18)

Where

\[ G_{kn} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Psi_n(t) e^{-j\omega_k t} dt = \mathcal{F}_k[\Psi_n(t)] \]  

(19)

is the Jacobi matrix, the elements of which can be thought of as Fourier transforms of the basis function system. The calculation of the complex integral in the formula can be avoided by choosing the basic functions of series expansion from the eigenfunctions of the Fourier transform.

6. Choosing basis function system - the Hermite functions

The Hermite functions have been chosen as basis functions because in the case of inverse problems it is advised to use complete, orthonormal function systems to reduce the number of unknown parameters and to improve the stability of the inversion procedure. The Hermite functions require proper scaling because in geophysical applications the frequency covers a wide range. The scaled Hermite polynomials can be calculated using the Rodriguez formula

\[ h_n(\omega, \alpha) = (-1)^n e^{\alpha \omega^2} \left( \frac{d}{d\omega} \right)^n e^{-\alpha \omega^2} \]  

(20)
and can be determined by the recursive formula [8]

\[
h_{n+1}(\omega, \alpha) = 2 \omega \alpha h_n(\omega, \alpha) - 2n \alpha h_{n-1}(\omega, \alpha)
\] (21)

where \(\alpha\) is a scaling factor. The first and second Hermite polynomials are

\[
h_0(\omega, \alpha) = 1
\]

(22)

\[
h_1(\omega, \alpha) = 2 \alpha \omega
\]

(23)

The scaled Hermite polynomials fulfil the condition of orthogonality

\[
\int_{-\infty}^{\infty} e^{-\alpha \omega^2} \cdot h_n(\omega, \alpha) \cdot h_m(\omega, \alpha) d\omega = \frac{\pi}{\sqrt{\alpha}} (2\alpha)^n n! \delta_{nm}, \delta_{nm} = \begin{cases} 0, & n \neq m \\ 1, & n = m \end{cases}
\]

(24)

In the formula, \(\delta_{nm}\) denotes the Kronecker-delta symbol. The scaled Hermite function can be defined as

\[
H_n(\omega, \alpha) = \frac{e^{-\frac{\alpha \omega^2}{2}} \cdot h_n(\omega, \alpha)}{\sqrt{\frac{\pi}{\alpha}} n! (2\alpha)^n}
\]

(25)

where \(h_n(\omega, \alpha)\) denotes the \(n\)-th scaled Hermite polynomial. Thus the introduced Hermite functions are orthonormal

\[
\int_{-\infty}^{\infty} H_n(\omega, \alpha) \cdot H_m(\omega, \alpha) d\omega = \delta_{nm}
\]

(26)

To get a fast and simple formula for the calculation of Jacobi’s matrix a special feature of the Hermite functions was used. As we can see earlier, \(G_{k,n}\) is the inverse Fourier transform of the basis function. This is the main reason for selecting the Hermite functions because their non-scaled versions are eigenfunctions of the Fourier transform [9]

\[
F \{H_n^{(0)}(t)\} = (-j)^n H_n^{(0)}(\omega)
\]

(27)

and can be written for the inverse Fourier transform

\[
F^{-1}\{H_n^{(0)}(\omega)\} = (j)^n H_n^{(0)}(t)
\]

(28)

Jacobi’s matrix can be expressed by the scaled Hermite functions \(H_n(\omega, \alpha)\) as

\[
G_{k,n} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} H_n(\omega, \alpha) \cdot e^{j\omega t_k} d\omega = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sqrt{\alpha} H_n^{(0)}(\omega) \cdot e^{j\omega t_k} d\omega
\]

(29)

Introducing the following notations

\[
\omega t = \omega^\prime t^\prime, \omega^\prime = \sqrt{\alpha} \omega, t^\prime = \frac{t}{\sqrt{\alpha}}
\]

(30)
it can be computed easily that

\[ G_{k,n} = \frac{1}{\sqrt{\alpha}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} H_n^{(0)}(\omega') \cdot e^{j\omega' t} d\omega' = \frac{1}{\sqrt{\alpha}} F^{-1}\{H_n^{(0)}(\omega')\} \]  

(31)

Jacobi’s matrix can be written again based on Eq. (15)

\[ G_{k,n} = \frac{1}{\sqrt{\alpha}} (j)^n H_n^{(0)}(t') = \frac{1}{\sqrt{\alpha}} (j)^n H_n^{(0)}(\frac{t'}{\sqrt{\alpha}}) \]  

(32)

This is a very important result because Jacobi’s matrix can be derived quickly without integration. The discretized form of the spectrum can be written according to Eq. (3) using the scaled Hermite function system, where the expansion coefficients \( B_n \) are defined (including the expression of the Jacobi’s matrix) in the frame of the over-determined inverse problem. This inverse problem is highly over-determined because the number of measurement data is much higher than that of the parameters \( N>M \). In the case of the Least Squares method (LSQ), the \( L_2 \) norm of the deviation vector is minimized

\[ E_2 = \sum_{k=1}^{N} e_k^2 = \sum_{k=1}^{N} (u_k^{(measured)} - u_k^{(theor)})^2 = \sum_{k=1}^{N} (u_k^{(measured)} - \sum_{n=0}^{M} B_n \cdot G_{k,n})^2 = \min \]  

(33)

The well-known normal equation can be derived from the above condition in the following form

\[ G^T \hat{B} = G^T u^{measured} \]  

(34)

After this, we can estimate the complex series expansion coefficients

\[ \hat{B} = (G^T G)^{-1} G^T u^{measured} \]  

(35)

With this, the real and imaginary part of the Fourier spectrum can be calculated at any frequency by Eq. (3). The LSQ method gives optimal results in the case of a Gaussian distributed data set.

To make the Fourier transform more robust an Iteratively Reweighted Least Squares (IRLS [5]) method using Cauchy weights was implemented. To achieve the optimal values of the unknown parameters \( B_n \) the following weighted norm is minimized

\[ E_w = \sum_{k=1}^{N} w_k e_k^2 \]  

(36)

where the weights are defined as

\[ w_k = \frac{e_k^2}{\epsilon^2 + e_k^2} \]  

(37)

and the \( k \)-th element of the deviation vector is

\[ e_k = u_k^{measured} - u_k^{theor} \]  

(38)
The scale parameter $\varepsilon$ of the Cauchy distribution is not a priori given because the data residuals change from iteration to iteration [10]. The weighted norm gives reliable results for inverse problems even if the measured data set contains outliers. The $j$-th iteration step of the normal equation is

$$G^T W^{(j-1)} G B^{(j)} = G^T W^{(j-1)} u^{measured} \quad (39)$$

This iteration is repeated until a proper stop criterion is met. Finally, the Fourier spectrum can be calculated at any frequency by using Eq. (10).

7. Numerical tests

In Figures 1, 2, and 3, examples were shown of noise-free and the absolute value of the analytical signal (reflection strength) calculated with a Ricker wavelet laden with Gaussian and Cauchy distribution noise. In Figures 4 and 5, the absolute value of the analytical signal (the reflection strength, as the first seismic attribute) generated by the robust inversion-based Fourier transforms of the same input signals is illustrated. Of course, our method for noise-free input provides the same result as the Fourier transform using the traditional DFT process. In the case of input data system I (Gauss noise), the inversion-based Hilbert transform in Figure 4 shows that it is less noisy than the data in Figure 2 produced by the traditional method. The slight improvement is characterized by a distance of $d(I)=0.0183$ in the data space. However, significant improvement is reflected in the processing of data system II (Cauchy noise) using the robust inversion method defined with Steiner weights as shown in Figure 5. Here, compared to Figure 3, we can see an almost complete suppression of the effect of the outlier data, which is caused by the distance in the data space $d(II) =0.0033$. There is nearly an order of magnitude difference between the data space distances obtained by the traditional ($d(II) =0.0444$) and inversion-based ($d(II) =0.0033$) Fourier transform processes.

![Figure 4](image_url)

**Figure 4.** Results after applying the new algorithm on the noisy signal using Gaussian noise distribution (data set I).
The above results confirm a significant improvement in signal/noise when using a robust inversion-based Fourier transform and an inverse Fourier transform built on a robust inversion basis to produce the Hilbert transform.

8. Conclusions
The new IRLS inversion-based Fourier transform method has been demonstrated to achieve high noise resistance to excessively noisy data. In this paper, we used this procedure in the calculation of the Hilbert transform. The algorithm presented is based on a dual inversion, on the one hand, we apply IRLS-FT to the Fourier transform and, on the other hand, after the production of the transformed spectrum of Hilbert, the inverse Fourier transform is also calculated using an inversion-based robust/resistant process. Although dual inversions require a significant calculation time compared to traditional DFT/IDFT transform, the numerical example presented shows that we can achieve a high degree of improvement with the inversion-based Hilbert transform process. Given the current capacity and speed of computers, it is likely that in some practical cases the extra computing time will be tolerable.

9. Acknowledgements
The research was carried out in project No. K-135323 is supported by the National Research, Development, and Innovation Office (NKFIH).

References
[1] Barnes, A. E. (2000). Geophysics, 65(1), 275-285
[2] Gábor, D. 1946. J. Inst. Elect. Eng., 93, pp. 429-441.
[3] Ricker, N. 1953. Geophysics, 18, pp. 10–40
[4] Gröbner, W, Hoffreiter, N. 1958. SpringerVerlag, Wien und Innsbruck.
[5] Steiner, F. 1997. Akadémiai Kiadó, Budapest
[6] Szegedi, H., Dobróka, M. 2014. Acta Geodaetica et Geophysica, 49, pp. 95–104
[7] Dobroka M, Szegedi H., Vass P. 2017. INTECH 2017
[8] Taner, M. T., Koehler, F. and Sheriff, R. E. 1979. Geophysics, 44(6), pp. 1041-1063
[9] Vaidyanathan, P. P., 2008. *IETE Journal of Education*, **49** (2), pp. 51-58.
[10] Bracewell, R. N., 1965, New York, McGraw-Hill, 268-71.