Geosciences and Engineering, Vol. 8, No. 13 (2020), pp. 9–19.

HILBERT TRANSFORM USING THE MOST FREQUENT VALUE METHOD

OMAR AL MARASHLY – MIHÁLY DOBRÓKA

Miskolc University
gfomar@uni-miskolc.hu, dobroka@uni-miskolc.hu

Abstract: In the study, we present a robust inversion method for calculating Hilbert transform, a process that also provides resistance to outlier noise. The inversion-based Fourier transform process combined with the Most Frequent Value method (MFV) developed by Steiner can effectively make the Fourier transform more robust. The resistance of the robust Fourier transform process (IRLS-FT) to outliers and its outstanding noise suppression capability justify the method being tried in the field of seismic data processing. As the first stage, we present the production of the Hilbert transform based on a robust inversion, and as an application example we calculate the absolute value of the analytical signal that can be produced as an attribute gauge (instantaneous amplitude). The new algorithm is based on a dual inversion: we determine the Fourier spectrum of the time signal (channel) by inversion, and the spectrum obtained by the transformation required for the Hilbert transform is transformed into the time range with a robust inversion. The latter operation is carried out using the Steiner weights calculated using the Iterative Reweighting Least Squares (IRLS) method (robust inverse Fourier transform based on inversion). To discretize the spectrum of the time signal, we use the scaled Hermite functions in a series expansion. The expansion coefficients are the unknowns in the inversion. The new Hilbert transform procedure was tested on a Ricker wavelet loaded with Cauchy post-distribution noise. The results show that the procedure has remarkable resistance to outlier noises and noise suppression an order of magnitude better than that calculated by the conventional (DFT) method.

Keywords: *Inversion-based Fourier Transform, robust inversion, robust Hilbert Transform*

1. INTRODUCTION

Attributes play an important role in seismic data processing and interpretation. They can be used to obtain certain information (physical or geometric parameters) that can’t be available before. Since the paper published by Taner et al. (1979), the subject area has expanded and significant development. 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 Dobróka (2014) 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.

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 (1946). 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

$$s(t) = u(t) + ju_H(t) \tag{1}$$

where

$$u_H(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} u(\tau)\,\frac{d\tau}{\tau - t} \tag{2}$$

is the Hilbert transform of the time signal. According to *Equation (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

$$\mathcal{F}\{u_H(t)\} = \mathcal{F}\{u(t)\}\,\mathcal{F}\left\{-\frac{1}{\pi t}\right\}, \tag{3}$$

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

$$U(\omega) = \mathcal{F}\{u(t)\} \tag{4}$$

one can write

$$\mathcal{F}\{u_H(t)\} = -j\,\mathrm{sgn}(\omega)U(\omega) = U_H(\omega)\,, \tag{5}$$

giving the Hilbert transform as

$$u_H(t) = \mathcal{F}^{-1}\{U_H(\omega)\}\,. \tag{6}$$

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 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 Dobróka et al. (2017). In order to use the noise-rejection capacity of the method we can use IRLS-FT in calculating $U(\omega)$ as

$$U(\omega) = \mathscr{F}_{IRLS}\{u(t)\}$$

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

$$A(t) = \sqrt{u(t)^2 + v(t)^2}. \tag{7}$$

The noise sensitivity of this attribute is illustrated using the Ricker wavelet (Ricker, 1953) 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 $\varepsilon = 0.04$. *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).

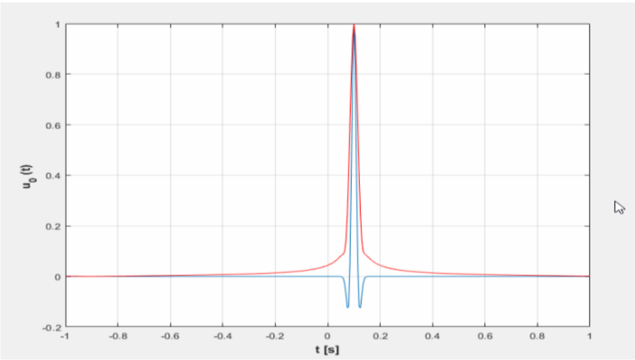

Figure 1. *The noiseless time domain Ricker wavelet (blue) and noiseless reflection strength (red)*

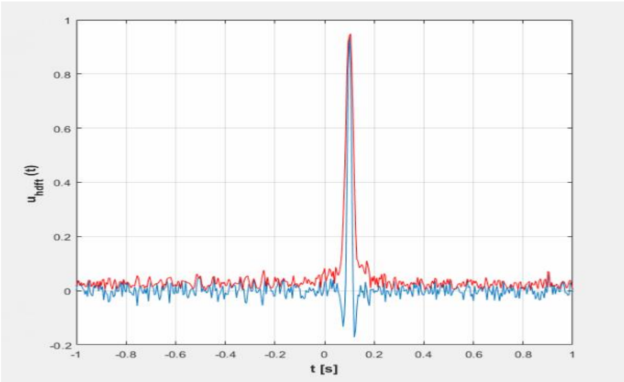

Figure 2. *The noisy Ricker wavelet (data set I) (blue) and noisy reflection strength (red). The discrete Fourier transform method (DFT) and its inverse (IDFT) were used.*

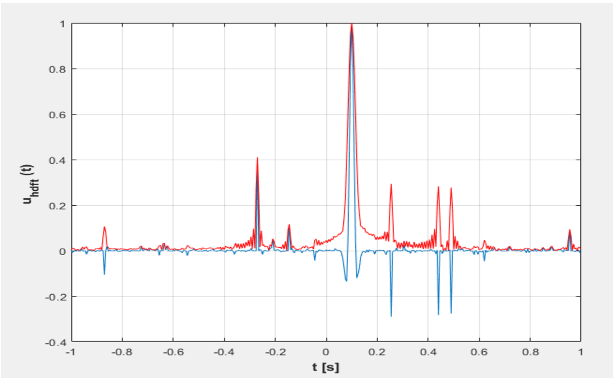

Figure 3. *The noisy Ricker wavelet (data set II) (blue) and noisy reflection strength (red). The discrete Fourier transform method (DFT) and its inverse (IDFT) were used.*

It can be seen that the reflection strength gauge calculated on the basis of the data system containing outliers II is particularly noisy. To characterize noise sensitivity, we introduce

$$d = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(A^{(noisy)}(t_k) - A^{(Noise-free)}(t_k))^2} \tag{8}$$

distance in the data space, which is d(I) = 0.0240 for data set I and for the data set II the distance d(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 (Szegedi and Dobróka, 2014) 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:

$$U(\omega) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} u(t)e^{-jwt}dt\,,\; u(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} U(\omega)e^{jwt}d\omega\,. \qquad (9)$$

The frequency spectrum *U(ω)* 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:

$$U(\omega) = \sum_{n=1}^{M} B_n \Psi_n(\omega)\,, \qquad (10)$$

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

$$u^{(elm)}(t_k) = u_k^{(elm)} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} U(\omega)e^{j\omega t_k}d\omega. \qquad (11)$$

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

$$u^{(elm)}(t_k) = u_k^{(elm)} = \sum_{n=1}^{M} B_n G_{kn}, \qquad (12)$$

where

$$G_{kn} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \Psi_n(\omega)e^{j\omega t_k}d\omega = \mathcal{F}_k^{-1}\{\Psi_n(\omega)\} \qquad (13)$$

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

$$e_k = u_k^{(measured)} - u_k^{(calculated)} = u_k^{(measured)} - \sum_{n=1}^{M} B_n G_{kn}. \qquad (14)$$

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

$$U^{(estimated)}(\omega) = \sum_{n=1}^{M} B_n^{estimated}\Psi_n(\omega). \qquad (15)$$

On an overly noisy data system, we demonstrated that the inversion-based Fourier transform achieves an order of magnitude better signal-to-noise ratio than the traditional DFT process.

5. THE ROBUST GENERATION OF HILBERT TRANSFORM

To produce the Hilbert transform we need to know the $U(\omega)$ 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 sgn(ω). 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

$$U(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u(t) e^{-j\omega t} dt \,. \tag{16}$$

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

$$u(t) = \sum_{n=1}^{M} B_n \Psi_n(t). \tag{17}$$

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

$$U_k(\omega_k) = \sum_{n=1}^{M} B_n \, \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Psi_n\,(t) e^{-j\omega_k t} dt = \sum_{n=1}^{M} B_n G_{kn}, \tag{18}$$

where

$$G_{kn} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Psi_n\,(t) e^{-j\omega_k t} dt = \mathcal{F}_k\{\Psi_n(t)\} \tag{19}$$

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

$$h_n(\omega, \alpha) = (-1)^n e^{\alpha\omega^2} \left(\frac{d}{d\omega} \right)^n e^{-\alpha\omega^2} \tag{20}$$

and can be determined by the recursive formula (Gröbner and Hoffreiter, 1958)

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

where α is a scaling factor. The first and second Hermite polynomials are

$$h_0(\omega,\alpha)=1 \tag{22}$$

$$h_1(\omega,\alpha)=2\alpha\omega. \tag{23}$$

The scaled Hermite polynomials fulfil the condition of orthogonality

$$\int_{-\infty}^{\infty} e^{-\alpha\omega^2} \cdot h_n(\omega,\alpha)\cdot h_m(\omega,\alpha)\, d\omega = \sqrt{\frac{\pi}{\alpha}}(2\alpha)^n n!\,\delta_{nm}\,,\ \ \delta_{nm} = \begin{cases} 0, n \neq m \\ 1, n = m \end{cases}. \tag{24}$$

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

$$H_n(\omega,\alpha) = \frac{e^{-\frac{\alpha\omega^2}{2}} \cdot h_n(\omega,\alpha)}{\sqrt{\sqrt{\frac{\pi}{\alpha}}\, n!(2\alpha)^n}}\,, \tag{25}$$

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

$$\int_{-\infty}^{\infty} H_n(\omega,\alpha)\cdot H_m(\omega,\alpha)\, d\omega = \delta_{nm}\,. \tag{26}$$

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 (Vaidyanathan, 2008)

$$\mathcal{F}\left\{H_n^{(0)}(t)\right\} = (-j)^n H_n^{(0)}(\omega) \tag{27}$$

and can be written for the inverse Fourier transform

$$\mathcal{F}^{-1}\left\{H_n^{0}(\omega)\right\} = (j)^n H_n^{(0)}(t)\,. \tag{28}$$

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

$$G_{k,n} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} H_n(\omega,\alpha)\cdot e^{j\omega t_k}\, d\omega = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sqrt[4]{\alpha}\, H_n^{(0)}(\omega')\cdot e^{j\omega t_k}\, d\omega\,. \tag{29}$$

Introducing the following notations

$$\omega t = \omega' t', \; \omega' = \sqrt{\alpha}\,\omega, \; t' = \frac{t}{\sqrt{\alpha}} \tag{30}$$

it can be computed easily that

$$G_{k,n} = \frac{1}{\sqrt[4]{\alpha}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} H_n^{(0)}(\omega') \cdot e^{j\omega' t_k'} \, d\omega' = \frac{1}{\sqrt[4]{\alpha}} \, \mathcal{F}^{-1}\left\{ H_n^{(0)}(\omega') \right\}. \tag{31}$$

Jacobi's matrix can be written again based on *Equation (15)*

$$G_{k,n} = \frac{1}{\sqrt[4]{\alpha}} (j)^n H_n^{(0)}(t') = \frac{1}{\sqrt[4]{\alpha}} (j)^n H_n^{(0)}\left(\frac{t}{\sqrt{\alpha}} \right). \tag{32}$$

This is a very important result because the Jacobi matrix can be derived quickly without integration. The discretized form of the spectrum can be written according to *Equation (3)* using the scaled Hermite function system, where the expansion coefficients B_n are defined (including the expression of the Jacobi 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

$$E_2 = \sum_{k=1}^{N} e_k^2 = \sum_{k=1}^{N} (u_k^{(measured)} - u_k^{(theor)})^2 = \sum_{k=1}^{N} (u_k^{(measured)} - \sum_{n=0}^{M} B_n \cdot G_{k,n})^2 = \min. \tag{33}$$

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

$$\mathbf{G}^T \mathbf{G} \vec{B} = \mathbf{G}^T \vec{u}^{\,measured}. \tag{34}$$

After this, we can estimate the complex series expansion coefficients

$$\vec{B} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \vec{u}^{\,measured}. \tag{35}$$

With this, the real and imaginary part of the Fourier spectrum can be calculated at any frequency by *Equation (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

$$E_w = \sum_{k=1}^{N} w_k e_k^2 \tag{36}$$

where the weighs are defined as

$$w_k = \frac{\varepsilon^2}{\varepsilon^2 + e_k^2}, \tag{37}$$

and the *k*-th element of the deviation vector is

$$e_k = u_k^{measured} - u_k^{theor}. \tag{38}$$

The scale parameter ε of the Cauchy distribution is not a priori given because the data residuals change from iteration to iteration (Steiner, 1997). 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

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

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

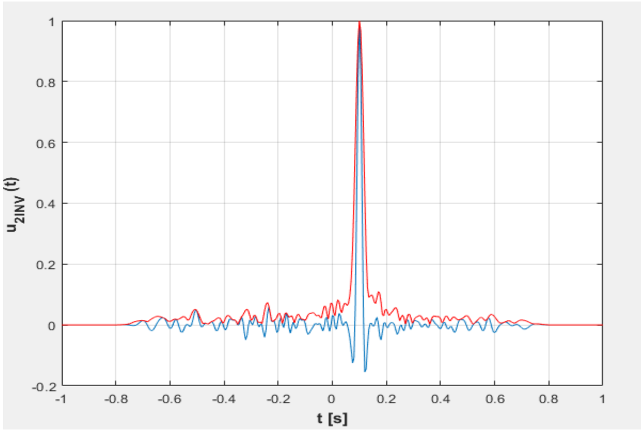

Figure 4. *Results after applying the new algorithm on the noisy signal using Gaussian noise distribution (data set I)*

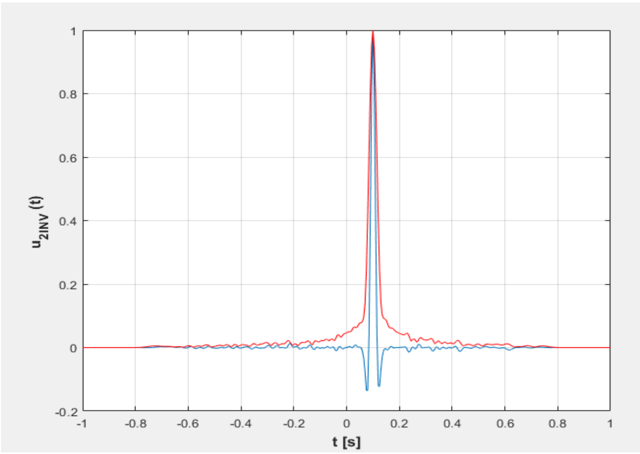

Figure 5. *Result after applying the new algorithm on the noisy signal using Cauchy distribution noise (data set II), we can see that the result is very effective on those type of noises*

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.

ACKNOWLEDGEMENTS

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

REFERENCES

[1] Gábor, D. (1946). Theory of Communication, part I. *J. Inst. Elect. Eng.*, 93, pp. 429–441.

[2] Ricker, N. (1953). The form and laws of propagation of seismic wavelets. *Geophysics*, 18, pp. 10–40.

[3] Gröbner, W., Hoffreiter, N. (1958). *Integraltafel. Zweiter Teil. Bestimmte Integrale*. SpringerVerlag, Wien–Innsbruck.

[4] Steiner, F. (1997). *Optimum Methods in Statistics.* Akadémiai Kiadó, Budapest.

[5] Szegedi, H., Dobróka, M. (2014). On the use of Steiner's weights in inversion-based Fourier transformation: Robustification of a previously published algorithm. *Acta Geodaetica et Geophysica,* 49, pp. 95–104.

[6] Dobróka, M., Szegedi, H., Vass, P. (2017). Inversion-based Fourier transform as a new tool for noise rejection. *INTECH 2017,* DOI:10.5772/66338.

[7] Taner, M. T., Koehler, F., Sheriff, R. E. (1979). Complex seismic trace analysis. *Geophysics*, 44 (6), pp. 1041–1063

[8] Vaidyanathan, P. P. (2008). Eigenfunctions of the Fourier Transform. *IETE Journal of Education*, 49 (2), pp. 51–58.