Legendre polynomial-based robust Fourier transformation and its use in reduction to the pole of magnetic data

Daniel Oduro Boatey Nuamah¹ · Mihály Dobróka² · Péter Vass² ·
Mátyás Krisztián Baracza³

Received: 31 May 2021 / Accepted: 3 August 2021 / Published online: 16 August 2021
© The Author(s) 2021

Abstract

A new inversion based Fourier transformation technique named as Legendre-Polynomials Least-Squares Fourier Transformation (L-LSQ-FT) and Legendre-Polynomials Iteratively Reweighted Least-Squares Fourier Transformation (L-IRLS-FT) are presented. The introduced L-LSQ-FT algorithm establishes an overdetermined inverse problem from the Fourier transform. The spectrum was approximated by a series expansion limited to a finite number of terms, and the solution of inverse problem, which gives the values of series expansion coefficients, was obtained by the LSQ method. Practically, results from the least square method are responsive to data outliers, thus scattered large errors and the estimated model values may be far from reality. A definitely better option is attained by introducing Steiner's Most Frequent Value method. By combining the IRLS algorithm with Cauchy-Steiner weights, the Fourier transformation process was robustified to give the L-IRLS-FT method. In both cases Legendre polynomials were applied as basis functions. Thus the approximation of the continuous Fourier spectra is given by a finite series of Legendre polynomials and their coefficients. The series expansion coefficients were obtained as a solution to an overdetermined non-linear inverse problem. The traditional DFT and the L-IRLS-FT were tested numerically using synthetic datasets as well as field magnetic data. The resulting images clearly show the reduced sensitivity of the newly developed L-IRLS-FT methods to outliers and scattered noise compared to the traditional DFT. Conclusively, the newly developed L-IRLS-FT can be considered to be a better alternative to the traditional DFT.

Keywords Fourier transformation · Legendre polynomials · Series expansion · Data processing

✉ Péter Vass
gfvassp@uni-miskolc.hu

¹ Department of Geological Engineering, Faculty of Civil and Geo-Engineering, College of Engineering, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

² Department of Geophysics, University of Miskolc, Miskolc-Egyetemváros 3515, Miskolc, Hungary

³ Research Institute of Applied Earth Sciences, University of Miskolc, Miskolc-Egyetemváros 3515, Miskolc, Hungary

1 Introduction

An important part of geophysical studies is to conclude the inner earth from data collected at or near the surface of the earth. The measured data is somehow connected to one or more interesting properties of the earth. These physical properties can be adequately estimated by solving an inverse problem. After modelling the subsurface volume under investigation, an inverse problem can be resolved to find the probable values of the model parameters. In geophysical inversion, the objective is to find an earth model defined by a set of physical values that matches observational data. Data acquisition techniques are undergoing permanent improvements over the years, hence, more cutting-edge data processing methods are required. The objective of this study was to develop new data processing algorithms by applying inverse theory to Fourier transformation for geophysical applications. In signal processing, a change of data from the time domain to frequency domain is a common practice. Commonly, the Discrete Fourier Transformation (DFT) algorithm is used to establish the discrete frequency components of a distinctly sampled time-domain dataset. However, the noise sensitivity of the processing method must be taken into account as the noise recorded in the time domain is adequately transformed into the frequency domain. In order to provide significant reduction of the impact of data noise and outliers in data processing, a more robust method known as the Steiner Iteratively Reweighted Least Square Fourier Transformation (S-IRLS-FT) was developed by Dobróka et al. (2015). The algorithm was evaluated on several levels including an application on magnetic data set (Dobróka et al. 2017) to prove its noise reduction capability. By combining the IRLS algorithm with Cauchy-Steiner weights the Fourier transform was handled as a robust inverse problem and the Fourier spectrum discretized utilizing series expansion. These series expansion based inversion methods have been successfully applied in processing gravity data (Dobróka and Völgyesi 2010), borehole geophysical data (Szabó 2004; Dobróka et al. 2016) as well as Induced polarization data (Turai et al. 2010).

2 Theoretical basis of the IRLS-FT method

The developed algorithm uses series expansion based discretization of the Fourier spectrum with Legendre polynomials as basis functions of discretization, and the solution of an inverse problem provides the estimated values of expansion coefficients. The robustified process uses a variant of the IRLS method, which is relied on the Steiner weight instead of the Cauchy one to facilitate the internal iterative recalculation of weights.

For a quicker and easier solution of an inverse problem, we normally apply the simple least-squares method, which is very effective when data noise follows Gaussian distribution. In situations where the data errors are not consistent and sparsely distributed such as outliers, the calculated model value may be highly questionable, which creates a limiting factor to the application of the least-squares method since geophysical datasets often contain outliers (Nuamah and Dobroka 2019). Several theoretically accepted methods have been formulated over the years to adequately process data outliers in order to achieve statistical robustness. The Least Absolute Deviation (LAD) as a robust method minimizes the L1 norm of the deviation vector characterizing the misfit function between the observed and predicted data. Continual practical affirm investigations that inversion with minimization of the L1 norm provides good estimates when fewer large errors contaminate the

data (Dobroka et al. 2017). An easy approach is to use the Cauchy criterion which adopts a random-distributed data noise. To properly account for the variations in data noise during inversion, each data must contribute to the solution based on its error margins and this is achieved by weighting the data. By applying Cauchy weights, Cauchy inversion can be used as a robust optimization method (Amundsen 1991). Although a very efficient procedure is obtained by integrating the IRLS algorithm with Cauchy weights, it is a little challenging since the scale parameter of the weights has to be a-priori known. This challenge was resolved by the Most Frequent Value (MFV) method developed by Steiner (1988, 1997) which derives the scale parameters from the real statistics of the data set. Dobróka et al. 1991 first emphasized the possibility of inserting calculated MFV-weights based on Steiner's method into an IRLS algorithm to produce a very efficient and robust method. A successful application of the MFV method in processing borehole geophysical data was reported by Szűcs et al. (2006). The so-called Cauchy-Steiner weights were also useful in robust tomography algorithms by Szegedi and Dobróka (2014).

2.1 Legendre polynomials as basis functions

Legendre polynomials form a system of complete orthogonal polynomials, which has numerous applications in science and engineering fields of study. The orthogonality can be expressed in the following way if $P_n(x)$ denotes a Legendre polynomial of degree ' n '

$$\int_{-1}^1 P_m(x)P_n(x)dx = 0 \quad \text{if } n \neq m \quad (1)$$

Another distinguished property of the system is its definite parity, which is demonstrated by the relationship below

$$P_n(-x) = (-1)^n P_n(x) \quad (2)$$

These properties make it convenient when Legendre polynomials are used in series expansion to approximate a function in the interval $(-1, 1)$. Also, the Legendre differential equation and the orthogonality property are independent of scale. The Legendre differential equation is given as

$$(1-x^2)\frac{d^2y}{dx^2} - \frac{2xdy}{dx} + n(n+1)y = 0 \quad (3)$$

where $n > 0$, $|x| < 1$, or equivalently

$$\frac{d}{dx} \left[(1-x^2) \frac{dy}{dx} \right] + n(n+1)y = 0 \quad (4)$$

The solution of the above equations gives the Legendre functions of order ' n ' with a general solution expressed as

$$Y = AP_n(x) + BQ_n(x) \quad (5)$$

where $P_n(x)$ and $Q_n(x)$ are Legendre functions of the first and second kind of order ' n '. For $n=1, 2, 3, \dots$, $P_n(x)$ functions are referred to as Legendre polynomials and are given by Rodrigue's formula

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n \quad (6)$$

as well as in the recursive expression

$$(n+1)P'_{n+1}(x) = (2n+1)xP_n(x) - nP'_{n-1}(x) \quad (7)$$

3 The L-LSQ-FT and the L-IRLS-FT algorithms

As measured geophysical data always contain noise, the ability of a processing method to reduce or eliminate data noise is an advantageous feature. In order to add noise rejection capability to the Fourier transformation, we introduce the Legendre-Polynomials Least-Squares Fourier Transformation (L-LSQ-FT) and the Legendre-Polynomials Iteratively Reweighted Least-Squares Fourier Transformation (L-IRLS-FT).

3.1 Basic formulae for 1D case

In a one-dimensional case of physical applications, the Fourier transform (often abbreviated by FT) is well-defined as

$$U(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u(t) e^{-j\omega t} dt, \quad (8)$$

where u represents a quantity related to a finite energy phenomenon and it is dependent on time denoted by t , ω is the angular frequency and j symbolizes the imaginary unit. The frequency spectrum $U(\omega)$ is a complex-valued continuous function representing the Fourier transform of a real-valued time function. Observably, a geophysical quantity measured in the time domain is adequately represented in the frequency domain by the Fourier transform. A backward transition from the frequency domain to the time domain is also possible and is provided by the inverse Fourier transform.

$$u(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} U(\omega) e^{j\omega t} d\omega \quad (9)$$

Expressing the Fourier transform as an inverse problem required the frequency spectrum $U(\omega)$ to be defined by a discrete parametric model. To realize this condition, we assumed that the parameter $U(\omega)$ can be estimated with satisfactory accuracy if a finite series expansion is used as the approximation of the frequency spectrum

$$U(\omega) \cong \sum_{i=1}^M B_i \Psi_i(\omega), \quad (10)$$

With the expression B_i being a complex-valued expansion coefficient and Ψ_i denoting a member of a properly chosen set of real-valued basis functions. The number of terms is limited to M . The theoretic value of a *continuous signal* in time domain at time moment t_k can be given by the inverse Fourier transform as follows

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

Inserting the expression given in Eq. (9) one finds that

$$u_k^{theor} \cong \frac{1}{\sqrt{2\pi}} \int_a^b \left(\sum_{i=1}^M B_i \Psi_i(\omega) \right) e^{j\omega t_k} d\omega = \sum_{i=1}^M B_i \frac{1}{\sqrt{2\pi}} \int_a^b \Psi_i(\omega) e^{j\omega t_k} d\omega, \quad (12)$$

where interval $[a, b]$ is the domain of the selected set of basis functions. It is clear that the accuracy of this approximation is highly dependent on how the domain of basis functions correlates to the bandwidth of the *continuous signal*.

In practice, not a continuous but a *discrete signal* with finite number of samples is usually available for the estimation of frequency spectrum. Therefore, the frequency range is limited by the Nyquist-Shannon sampling theorem. So that this limited frequency range can be preserved, it must be mapped on the interval of basis functions by means of a linear scaling operation.

Introducing the notation ω' for the scaled angular frequency, one can write that

$$G_{k,i} = \frac{1}{\sqrt{2\pi}} \int_a^b \Psi_i(\omega') e^{j\omega' t_k} d\omega', \quad (13)$$

where $G_{k,i}$ constitutes the elements of the Jacobian matrix of the size N -by- M . Here, N denotes the number of time domain samples. The model was parameterized by introducing the Legendre polynomials as basis functions to give

$$G_{k,n} = \frac{1}{\sqrt{2\pi}} \int_{-1}^1 P_n(\omega') e^{j\omega' t_k} d\omega' \quad (14)$$

In fact, Eq. (14) corresponds to the inverse Fourier transform of the Legendre polynomial of degree n . Thus, we can symbolize the general element of the Jacobian matrix in the following way

$$G_{k,n} = \mathcal{F}_k^{-1} \{P_n(\omega')\} \quad (15)$$

The basic idea of introducing a new inversion-based Fourier Transformation method is to approximate the inverse Fourier transform of $P_n(\omega')$ by means of the inverse discrete Fourier transform (IDFT) procedure:

$$G_{k,n} = IDFT_k \{P_n(\omega')\}. \quad (16)$$

The values of the $P_n(\omega')$ functions are accurate (noise-free), so the application of IDFT (or IFFT) is independent of the noise problem (of the data set). By using this procedure, the $G_{k,n}$ elements of the Jacobian matrix can be generated numerically. The theoretical value of the *discrete signal* at time moment t_k is

$$u^{theor}(t_k) = u_k^{theor} = \sum_{i=1}^M B_n G_{k,n} \quad (17)$$

and the k -th element of the data deviation vector is written as

$$e_k = u_k^{meas} - u_k^{theor} = u_k^{meas} - \sum_{n=1}^M B_n G_{kn}. \quad (18)$$

3.2 Basic formulae for 2D case

For a function $u(x, y)$, the 2D Fourier transform can be calculated by the following integral

$$U(\omega_x, \omega_y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u(x, y) e^{-j(\omega_x x + \omega_y y)} dx dy. \quad (19)$$

The inverse transform is expressed by the formula below

$$u(x, y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(\omega_x, \omega_y) e^{j(\omega_x x + \omega_y y)} d\omega_x d\omega_y, \quad (20)$$

where $U(\omega_x, \omega_y)$ is the 2D spatial-frequency spectrum, x and y represent the spatial coordinates and ω_x, ω_y denote the spatial angular frequencies. The expression below was used in discretizing the continuous Fourier spectrum

$$U(\omega_x, \omega_y) = \sum_{n=1}^N \sum_{m=1}^M B_{n,m} \Psi_n(\omega_x) \Psi_m(\omega_y), \quad (21)$$

where $\Psi_n(\omega_x)$ and $\Psi_m(\omega_y)$ are frequency-dependent basis functions with the domain of interval $[a, b]$. The expansion coefficients are denoted by $B_{n,m}$, N and M are used for specifying the numbers of terms in the finite 2D series expansion. To develop the L-IRLS-FT algorithm in 2D, the same inversion procedures in the one-dimensional case was followed. The general form of the Jacobian matrix in the case of two-dimensional series expansion based inverse Fourier transform is given as

$$G_{nm}^{kl} = \frac{1}{\sqrt{2\pi}} \int_a^b \Psi_n(\omega_x) e^{j\omega_x x_k} d\omega_x \cdot \frac{1}{\sqrt{2\pi}} \int_a^b \Psi_m(\omega_y) e^{j\omega_y y_l} d\omega_y, \quad (22)$$

where G_{nm}^{kl} is an element of the Jacobian matrix. Similarly to the 1D case, the linear scaling of independent variables cannot be avoided when the angular frequency ranges coming from the Nyquist-Shannon sampling theorem and the domain of basis functions do not match. Thus, the notations ω'_x and ω'_y must be introduced for the scaled spatial angular frequencies.

Parameterization of the model is achieved by introducing the Legendre polynomials as basis functions to give,

$$G_{nm}^{kl} = \frac{1}{\sqrt{2\pi}} \int_{-1}^1 P_n(\omega'_x) e^{j\omega'_x x_k} d\omega'_x \cdot \frac{1}{\sqrt{2\pi}} \int_{-1}^1 P_m(\omega'_y) e^{j\omega'_y y_l} d\omega'_y \quad (23)$$

or in a more formal notation

$$G_{n,m}^{kl} = \mathcal{F}_k^{-1} \{P_n(\omega'_x)\} \cdot \mathcal{F}_l^{-1} \{P_m(\omega'_y)\} \quad (24)$$

We employed the frequently used inverse discrete Fourier transform (IDFT) to calculate the inverse Fourier transform of Legendre polynomials,

$$G_{n,m}^{kl} = IDFT \{P_n(\omega'_x)\} \cdot IDFT \{P_m(\omega'_y)\} \quad (25)$$

By using this procedure the $G_{n,m}^{kl}$ elements of the Jacobian matrix can be generated numerically. At this point, a new inversion method is to be defined. The theoretical value of the data at point (x_k, y_l) is given as

$$u^{theor}(x_k, y_l) = u_s^{theor} = \sum_{n=1}^N \sum_{m=1}^M B_{n,m} G_{n,m}^{kl}$$

Which can be simplified further to

$$u_s^{theor} = \sum_{i=1}^I B_i G_{s,i}$$

(where $i = n + (m - 1)N$, $s = k + (l - 1)N$)

The general element of the deviation vector can be given in the following form

$$e_s = u_s^{meas} - u_s^{theor} = u_s^{meas} - \sum_{i=1}^I B_i G_{s,i} \quad (26)$$

3.3 The L-LSQ-FT and the robust L-IRLS-FT methods

Using Eqs. (18) and (26) we can calculate the L_2 -norm, the misfit function as

$$E_2 = \sum_{k=1}^N e_k^2 = \sum_{k=1}^N (u_k^{meas} - u_k^{theor})^2 = \sum_{k=1}^N \left(u_k^{meas} - \sum_{n=1}^M B_n G_{kn} \right)^2.$$

The minimization of this function gives the normal equation of the L-LSQ-FT (Gaussian Least Squares) method

$$\mathbf{G}^T \mathbf{G} \vec{B} = \mathbf{G}^T \vec{u}^{meas}$$

resulting in the solution

$$\vec{B} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \vec{u}^{meas}.$$

The Weighted Least Squares procedure (WLSQ) can be a robust inversion method with significant noise reduction capability and reduced outlier sensitivity. This is because each weighted data contributes to the solution according to its error margin, hence obtaining a good result (Dobroka et al. 2012). The deviation vector was minimized using Cauchy weights, and additionally reformed to Cauchy-Steiner weights. This was important because the Gaussian Least Squares Method (LSQ) is applicable only when data noise follows a regular distribution, and the algorithm minimizes the L_2 -norm of the deviation vector between the theoretical and calculated data. Ideally, most geophysical measurements contain irregular noise with sparsely occurring outliers making the least-squares method (LSQ) less efficient for processing. The minimized weighted L_2 -norm of the deviation vector is given as

$$E_w = \sum_{k=1}^N w_k e_k^2 \quad (27)$$

Applying Steiner's Most Frequent Value method (MFV), the scale parameter ϵ^2 was determined from the data set in an internal iteration loop. The so-called Cauchy-Steiner weights are given by

$$w_k = \frac{\epsilon^2}{\epsilon^2 + e_k^2}, \quad (28)$$

where ϵ^2 the Steiner's scale factor called dihesion is determined iteratively.

That results in a non-linear inverse problem occasioned by an estimated misfit function which is non-quadratic due to the introduction of the Cauchy-Steiner weights (because e_k contains the unknown expansion coefficients). Practically, this is solvable by applying the method of the Iteratively Reweighted Least Squares (Scales et al. 1988) which linearizes the procedure. In a stepwise order of solution, a 0 -th order solution $\vec{B}^{(0)}$ is derived by using the non-weighted LSQ method and the weights are recalculated as

$$w_k^{(0)} = \frac{\epsilon^2}{\epsilon^2 + (e_k^{(0)})^2}$$

with $e_k^{(0)} = u_k^{\text{measured}} - u_k^{(0)}$, where $u_k^{(0)} = \sum_{i=1}^M B_i^{(0)} G_{ki}$ and the expansion coefficients are given by the LSQ method. In the first iteration, the misfit function

$$E_w^{(1)} = \sum_{k=1}^N w_k^{(0)} e_k^{(1)2}$$

is minimized resulting in the linear set of normal equations

$$\mathbf{G}^T \mathbf{W}^{(0)} \mathbf{G} \vec{B}^{(1)} = \mathbf{G}^T \mathbf{W}^{(0)} \vec{u}^{\text{measured}}$$

The minimization of the new misfit function

$$E_w^{(2)} = \sum_{k=1}^N w_k^{(1)} e_k^{(2)2}$$

gives $\vec{B}^{(2)}$ which serves again for the calculation of $w_k^{(2)}$. This procedure is repeated giving the typical j -th iteration step

$$\mathbf{G}^T \mathbf{W}^{(j-1)} \mathbf{G} \vec{B}^{(j)} = \mathbf{G}^T \mathbf{W}^{(j-1)} \vec{u}^{measured} \quad (29)$$

with the $\mathbf{W}^{(j-1)}$ weighting matrix

$$W_{kk}^{(j-1)} = w_k^{(j-1)} \quad (30)$$

Every level of these iterations encloses an internal loop for the calculation of the Steiner's scale parameter which is repetitive until an accurate stop condition is met.

3.3.1 Numerical testing in 1D

A time-domain signal (Fig. 1) was created to test the noise reduction capability of the newly developed methods, L-LSQ-FT, L-IRLS-FT and the traditional DFT in one dimension. The noiseless time function of the test data is given by the expression below

$$u(t) = \kappa t^\eta e^{-\lambda t} \sin(\mu t + \phi) \quad (31)$$

where the Greek letters represent the parameters of the signal. Specified fixed values for the signal parameters as follows: $\kappa \approx 738.91$, $\eta = 2$, $\lambda = 20$, $\mu = 40\pi$, $\phi = \pi/4$.

The noise-free waveform was sampled at regular intervals of 0.005 (sec) measurement points ranging over the time interval of $[-1, 1]$ and processed using the traditional DFT method to give both the real and imaginary parts of the noise-free Fourier spectrum (Fig. 2). The L-LSQ-FT spectrum (Fig. 3) was also calculated using Legendre polynomials of the (maximal) order of $M=300$. For arithmetic purposes, the Fourier spectrum was calculated on data set converted between the interval $[-1, 1]$ in both x and y coordinates giving a suitable scale in the wavenumber domain. The output transformed spectra were similar for both the traditional DFT and the L-LSQ-FT. This demonstrates the applicability of L-LSQ-FT in processing noise-free data. Following the successful application of both methods to the noise-free signal, Gaussian and Cauchy

Fig. 1 Calculated noise-free waveform

Fig. 2 Processed DFT spectrum of the noise-free data

Fig. 3 Processed L-LSQ-FT spectrum of the noise-free waveform

noise were introduced into the noise-free signal (Fig. 1) for processing. In geophysical applications, the Gaussian noise distribution is occasionally encountered in data processing. The Gaussian noisy signal with 0 mean and 0.01 variance is given in Fig. 4.

On the other hand, random noise which does not follow a regular distribution across a survey area often occur. This type of noise is mostly introduced into survey data from external sources such as data acquisition or survey designs and equipment limitations. They are inherent in geophysical data and are not related to the subsurface body of interest. Random noise reduction is a valid step to increase the signal to noise ratio in geophysical applications with several methods developed over the years to achieve this purpose (Liu et al. 2006; Al-Dossary and Marfurt 2007; Liu et al. 2009). This includes

Fig. 4 The generated noisy signal with Gaussian noise

the development of filters using various forms of transforms such as Wavelet Transform (Deighan and Watts 1997), S-Transform (Askari and Siahkooli 2008) and Fourier Transform (Dobroka et al. 2012). Failure to adequately suppress random noise affects the quality of processed data and interpretation.

Random noise following Cauchy distribution was added to the noise-free waveform to produce a noisy signal (Fig. 5) for processing. To demonstrate the noise reduction capability of the two methods, the Gaussian noisy signal (Fig. 4) was processed with the traditional DFT and the L-LSQ-FT methods. The resultant transformed spectra in the real and imaginary form are shown in Figs. 6 and 7 for the DFT and L-LSQ-FT respectively. The output signals show a considerable suppression of Gaussian noise by the L-LSQ-FT method compared to the traditional DFT method. Although a comparison between the real and imaginary spectrum as produced from the traditional DFT and the L-LSQ-FT does not

Fig. 5 The generated noisy signal with Cauchy noise

Fig. 6 Processed DFT spectrum of the Gaussian noisy signal ($D=1.03 \cdot 10^{-2}$)

Fig. 7 Processed L-LSQ-FT spectrum of the Gaussian noisy signal ($D=8.2 \cdot 10^{-3}$)

show much improvement in output Fourier spectra in both methods, the L-LSQ-FT algorithm was able to reject a substantial amount of the introduced Gaussian noise.

For quantitative characterization of the results, we introduce the RMS distance between (a) and (b) data sets (for example noisy and noiseless) in the time domain (data distance)

$$d = \sqrt{\frac{1}{N} \sum_{k=1}^N (u^{(a)}(t_k) - u^{(b)}(t_k))^2}, \quad (32)$$

as well as the frequency domain (model or spectra distance)

$$D = \sqrt{\frac{1}{N} \sum_{k=1}^N \left\{ (\operatorname{Re}[U^{(a)}(f_k) - U^{(b)}(f_k)])^2 + (\operatorname{Im}[U^{(a)}(f_k) - U^{(b)}(f_k)])^2 \right\}}. \quad (33)$$

In the case of the Gaussian noise, the distance between the noisy and noiseless data sets, $d=0.1032$. The model or spectra distance between the DFT spectrum of the noisy (contaminated with Gaussian noise) and the noiseless data sets gave $D=1.03 \cdot 10^{-2}$. There is some improvement using the L-LSQ-FT method characterized by the spectra distance between the noiseless and the noisy spectra: $D=8.2 \cdot 10^{-3}$. Similarly, the DFT gave a spectra distance $D=4.16 \cdot 10^{-2}$ for spectrum produced from the noisy Cauchy signal whilst the L-LSQ-FT gave $D=2.43 \cdot 10^{-2}$. From the analyses, a higher noise reduction capability was exhibited by the L-LSQ-FT method compared to the traditional DFT method. The results demonstrate the outlier and random noise sensitivity of the DFT and to some extent, the Least Squares Methods, hence there was the need to define a more robust method for outliers and random noise suppression. We, therefore, introduced the L-IRLS-FT Method. To test the noise reduction capability of the L-IRLS-FT, the Cauchy noisy signal (Fig. 5) was processed with the traditional 1D-DFT and L-IRLS-FT algorithms. The results are shown

Fig. 8 Processed DFT spectrum of the Cauchy noisy signal ($D=4.16 \cdot 10^{-2}$)

Fig. 9 Processed L-IRLS-FT spectrum of the Cauchy noisy signal ($D=1.32 \cdot 10^{-2}$)

in Figs. 8 and 9 respectively. The comparison of output signals proves that the newly developed L-IRLS-FT algorithm was effective in reducing the Cauchy noise component of the noisy signal.

To analytically examine the results, we applied the RMS distance between two data sets (for example noisy and noiseless) in the frequency domain as well as the model or spectra distance as indicated by Eqs. (32) and (33). The DFT gave a spectra distance $D=4.16 \cdot 10^{-2}$ for spectrum produced from the noisy Cauchy signal whilst the L-IRLS-FT gave a spectra distance of $D=1.32 \cdot 10^{-2}$. From the above analyses, the L-IRLS-FT method showed a high noise reduction capability when both regular and irregular noise was added to the noise-free waveform for processing. The outcome completely validates the outlier and random noise sensitivity of the traditional DFT method.

3.3.2 Numerical testing in 2D

The newly developed inversion-based Fourier transforms, 2D L-LSQ-FT and 2D L-IRLS-FT were tested on a noise-free data set. A rectangular test surface of size $[-1, 1]$ units in both x and y directions was created (Fig. 10). An anomaly ($u=0.7$) in the centre of size $[-0.2, 0.2]$ units in both directions was created with a noise-free regular background. The generated data were sampled at an interval of $dx=dy=0.02$ units so the total quantity of data becomes $N=101 \cdot 101$. We then applied the 2D DFT and the 2D L-LSQ-FT algorithms to calculate the 2D Fourier spectrum of the noise-free discrete data set.

Figures 11 and 12 show the absolute value (amplitude spectrum) produced by the traditional 2D DFT and the 2D L-LSQ-FT respectively. The 2D L-LSQ-FT spectrum was estimated from Legendre polynomials of the (maximal) order of $M=45$. For scaling reasons, the data were initially changed between the interval $[-1, 1]$ in both x and y coordinates before the Fourier spectra were calculated. Visual comparison of the two spectra shows

Fig. 10 Generated noise-free 2D test surface

Fig. 11 Processed 2D-DFT amplitude spectrum of the noise-free data with space frequencies (f_x, f_y) on the horizontal and vertical axes

approximately the same output, indicating the effectiveness of both methods in processing a noise-free dataset.

After successfully applying both methods to the noise-free test surface, data noise following Gaussian and Cauchy distribution were introduced into the noise-free test surface for processing. The Gaussian noisy data is shown in Fig. 13 with a much rougher surface

Fig. 12 Processed 2D L-LSQ-FT amplitude spectrum of the noise-free data with space frequencies (f_x, f_y) on the horizontal and vertical axes

Fig. 13 Generated 2D noisy signal with Gaussian noise

area. Random noise following Cauchy distribution was added to the test surface to produce outlier data with spikes (Fig. 14).

To demonstrate the noise reduction capability of the two methods, the Gaussian and Cauchy noisy data were processed with the traditional 2D DFT and the 2D L-IRLS-FT

Fig. 14 Generated 2D noisy signal with Cauchy noise

methods. To quantitatively describe the results, we introduce the RMS distance between (a) and (b) data in the space domain as

$$d_{RMS} = \sqrt{\frac{1}{N} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} [u^{(a)}(x_i, y_j) - u^{(b)}(x_i, y_j)]^2} \quad (34)$$

And further in the space domain and the model distance as

$$D_{RMS} = \left[\frac{1}{N} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} (\text{Re} [U^{(a)}(\omega_{xi}, \omega_{yj})] - \text{Re} [U^{(b)}(\omega_{xi}, \omega_{yj})])^2 + \frac{1}{N} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} (\text{Im} [U^{(a)}(\omega_{xi}, \omega_{yj})] - \text{Im} [U^{(b)}(\omega_{xi}, \omega_{yj})])^2 \right]^{\frac{1}{2}}, \quad (35)$$

(N_x, N_y and $N = N_x N_y$ are relevant numbers of data or space-frequency points in the 2D test area).

The output processed Fourier spectrum for the 2D DFT and the 2D L-IRLS-FT for the Cauchy noisy data are shown in Figs. 15 and 16 below. It appears from the output spectra that the newly developed 2D L-IRLS-FT algorithm was effective in reducing the Cauchy noise component of the noisy data sets. The 2D L-IRLS-FT algorithm gave an almost noise-free spectrum demonstrating its effective noise reduction capability. For the processed Gaussian noisy dataset, the model or spectra distance between the DFT spectrum of the noisy (contaminated with Gaussian noise) and the noiseless DFT processed data sets was $D = 1.700 \cdot 10^{-3}$. There was sufficient improvement characterized by the spectra distance between the noiseless and the noisy (given by 2D L-IRLS-FT) spectra: $D = 1.231 \cdot 10^{-3}$. Likewise, the 2D DFT gave a spectra distance $D = 3.413 \cdot 10^{-3}$ for

Fig. 15 Processed 2D DFT spectrum of the Cauchy noisy Data with space frequencies (f_x, f_y) on the horizontal and vertical axes ($D=3.413 \times 10^{-3}$)

Fig. 16 Processed 2D L-IRLS-FT spectrum of the Cauchy noisy Data with space frequencies (f_x, f_y) on the horizontal and vertical axes ($D=6.759 \times 10^{-4}$)

spectrum produced from the noisy Cauchy data (Fig. 15) whilst the 2D L-IRLS-FT gave an improved spectra distance of $D=6.759 \times 10^{-4}$ (Fig. 16). Accordingly, the 2D L-IRLS-FT method showed a remarkable noise reduction capability when Cauchy noise was added to the test surface for processing. The results wholly validate the outlier and random noise sensitivity of the traditional 2D DFT method. Hence, we recommend the new method, the 2D L-IRLS-FT, when a robust and resistant solution is necessary for subduing the effects of random noise.

4 Application to reduction to the pole of magnetic data

The above tests prove the appreciable noise rejection capacity of the L-IRLS-FT method. This promising feature implies its application among others, in the field of reduction to the pole of magnetic data. As is well-known the shape and position of the maximum and minimum values of a magnetic anomaly are dependent on the inclination and declination angles. To eliminate this feature, the reduction to the magnetic pole is used. The result of this linear transformation is a magnetic anomaly of the same source assuming to be magnetized in the vertical direction (Kiss 1981). The transformation in the space-frequency (f_x, f_y) domain takes the form

$$T_P(f_x, f_y) = S_P(f_x, f_y) T(f_x, f_y) \quad (36)$$

where $T(f_x, f_y)$ is the Fourier transform of the total magnetic anomalies, $S_P(f_x, f_y)$ is the transfer function of the reduction to the pole operation

$$S_P(f_x, f_y) = (f_x^2 + f_y^2) / (N(f_x^2 + f_y^2)^{1/2} + j(Lf_x + Mf_y))(n(f_x^2 + f_y^2)^{1/2} + j(lf_x + mf_y)).$$

here L , M and l , m are direction cosines containing the declination and inclination angles (Kiss 1981). The total field reduced to pole is given by the inverse Fourier transformation

$$T_P(x, y) = \mathcal{F}_k^{-1} \{ T_P(f_x, f_y) \}. \quad (37)$$

The procedure contains the Fourier transform of the measured data, so in the case of noisy datasets containing outliers, it can be important to use FT algorithms with special noise rejection capacity. To demonstrate the efficiency of the L-IRLS-FT method (using Cauchy-Steiner weights) in the procedure of reduction to pole we defined a synthetic dataset calculated on a C-S shaped anomaly with uniform 1[A/m] magnetization in depth between (Al-Dossary and Marfurt 2007; Amundsen 1991) km. The magnetic field was calculated by using the method of Kuranatnam (1981). Figure 17a shows the noise-free magnetic field and Fig. 17b shows its pole reduced version calculated by the traditional DFT.

To simulate noisy dataset containing outliers 5% noise of Gaussian distribution was added to the noise-free data and a randomly selected 20% of the dataset was further

Fig. 17 The noise-free magnetic field (a) and its reduced to the pole version using the traditional DFT (b)

Fig. 18 The noisy magnetic field containing outliers (a) and its reduced to the pole version using the traditional DFT (b)

Fig. 19 The noisy magnetic field containing outliers (a) and its reduced to the pole version using the L-IRLS-FT method containing Cauchy-Steiner weights (b)

contaminated by an additional 20% noise. The noisy dataset and its pole reduced version calculated by the traditional DFT is shown in Fig. 18a and b, respectively. As it can be seen the pole reduction results are very noisy.

Using the L-IRLS-FT method in the reduction to pole procedure of the same noisy dataset the result is shown in Fig. 19b. It can be seen that the new L-IRLS-FT algorithm using Cauchy-Steiner weights has appreciable outlier rejection also in the reduction to pole procedure.

5 Conclusion

We propose a new, advanced and vigorous inversion-based Fourier transformation algorithm, the Legendre-Polynomials Least-Squares Fourier Transformation (L-LSQ-FT) and Legendre-Polynomials Iteratively Reweighted Least-Squares Fourier Transformation (L-IRLS-FT) methods. The new algorithms use Legendre polynomials as basis functions for discretizing the resultant Fourier spectrum employing series expansion. In doing so, the

expansion coefficients were estimated by resolving an over-determined inverse problem. The iteratively reweighted least squares method was introduced as a mean of robustifying the entire procedure. With this method, each data contributes to the solution according to its error margin, as individual data were weighted iteratively using the Cauchy-Steiner weights. The final algorithm was very tough and resistant to data outliers, especially sparsely distributed ones. This advantaged position was reached by comparing it with the traditional DFT algorithm in reduction to the pole of magnetic data. The L-IRLS-FT algorithm was found to be a very efficient, robust and resistant procedure with a higher noise rejection capability. Also, a test on synthetic data showed similar conclusions. Consequently, the newly developed L-IRLS-FT can be considered as a better alternative to the traditional DFT.

Acknowledgements The research is supported by the European Union, co-financed by the European Social Fund and the GINOP-2.3.2-15-2016-00010 “Development of enhanced engineering methods with the aim at utilization of subterranean energy resources” project in the framework of the Széchenyi 2020 Plan, funded by the European Union, co-financed by the European Structural and Investment Funds.

Author contributions Conceptualization: MD; Methodology: MD, PV, MKB; Formal analysis and investigation: DOBN, MD; Software: DOBN, MD, PV; Writing—original draft preparation: DOBN; Visualization: DOBN; Writing—review and editing: MD, PV, MKB.

Funding Open access funding provided by University of Miskolc. The research was carried out in the GINOP-2.3.2-15-2016-00010 “Development of enhanced engineering methods with the aim at utilization of subterranean energy resources” project in the framework of the Széchenyi 2020 Plan, funded by the European Union, co-financed by the European Structural and Investment Funds.

Availability of data and material Because of the data confidentiality, the experimental data is not published.

Code availability Because of the data confidentiality, the code is not published.

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Al-Dossary S, Marfurt KJ (2007) Lineament-preserving filtering. *Geophysics* 72:P1–P8
- Amundsen L (1991) Comparison of the least-squares criterion and the Cauchy criterion in frequency-wave-number inversion. *Geophysics* 56:2027–2038
- Askari R, Siahkoobi HR (2008) Ground-roll attenuation using the s- and x-f-k-transforms. *Geophys Prospect* 56:105–111
- Deighan AJ, Watts DR (1997) Ground-roll suppression using the wavelet transform. *Geophysics* 62:1896–1903

- Dobróka M, Völgyesi L (2010) Series expansion based inversion IV. Inversion reconstruction of the gravity potential. *Magyar Geofizika* 51:143–149
- Dobróka M, Gyulai A, Ormos T, Csókás J, Dresen L (1991) Joint inversion of seismic and geoelectric data recorded in an under-ground coal mine. *Geophys Prospect* 39:643–665
- Dobróka M, Szegedi H, Vass P, Turai E (2012) Fourier transformation as inverse problem: an improved algorithm. *Acta Geodaetica Et Geophysica Hungarica* 47:185–196. <https://doi.org/10.1556/AGeod.47.2012.2.7>
- Dobróka M, Szegedi H, Somogyi Molnár J, Szűcs P (2015) On the reduced noise sensitivity of a new Fourier transformation algorithm. *Math Geosci* 47:679–697. <https://doi.org/10.1007/S11004-014-9570-X>
- Dobróka M, Szabó NP, Tóth J, Vass P (2016) Interval inversion approach for an improved interpretation of well logs. *Geophysics* 81:D155–D167
- Dobroka M, Szegedi H, Vass P (2017) Inversion-based Fourier transform as a new tool for noise rejection. *INTECH*. <https://doi.org/10.5772/66338>
- Kiss K (1981) *Magnetic methods of applied geophysics*. ISBN 978-963-284-057-4
- Kunaratnam K (1981) Simplified expressions for the magnetic anomalies due to vertical rectangular prisms. *Geophys Prosp* 29:883–890
- Liu C, Liu Y, Yang B, Wang D, Sun J (2006) A 2D multistage median filter to reduce random seismic noise. *Geophysics* 71:V105–V110
- Liu Y, Liu C, Wang D (2009) A 1D time-varying median filter for seismic random, spike-like noise elimination. *Geophysics* 74:V17–V24
- Nuamah DOB, Dobroka M (2019) Inversion-based Fourier transformation used in processing non-equidistantly measured magnetic data. *Acta Geod Geoph* 54:411–424. <https://doi.org/10.1007/s40328-019-00266-4>
- Scales JA, Gersztenkorn A, Treitel S (1988) Fast Lp solution of large, sparse, linear systems: application to seismic travel time tomography. *J Comput Phys* 75:314–333
- Steiner F (1988) *Most Frequent Value procedures (a short monograph)*. *Geophys Trans* 34:139–260
- Steiner F (1997) *Optimum methods in statistics*. Academic Press, Budapest
- Szabó NP (2004) Global Inversion of Well Log Data. *Geophys Trans* 44:313–329
- Szegedi H, Dobróka M (2014) On the use of Steiner's weights in inversion-based Fourier transformation: robustification of a previously published algorithm. *Acta Geophys* 49:95–104. <https://doi.org/10.1007/s40328-014-0041-0>
- Szűcs P, Civan F, Virág M (2006) Applicability of the most frequent value method in groundwater modeling. *Hydrogeol J* 14:31–43. <https://doi.org/10.1007/s10040-004-0426-1>
- Turai E, Dobroka M, Herczeg A (2010) Series expansion based inversion III. Procedure for inversion processing of induced polarization (IP) data (in Hungarian). *Magyar Geofizika* 51:88