Analysis of the MGGL seismic data by a noise-filtered Fourier transform method

Péter Kicsiny
Wigner Research Centre for Physics, H-1525 Budapest, Konkoly Thege Miklós út 29-33, Hungary; Physics Department, Technische Universität München, Garching 85748, Germany; pkicsiny@gmail.com

László Ábel Somlai
Wigner Research Centre for Physics, H-1525 Budapest, Konkoly Thege Miklós út 29-33, Hungary; Institute of Physics, University of Pécs , H-7624 Pécs, Ifjúság út 6, Hungary; somlai.laszlo@wigner.mta.hu

Zoltán Zimborás
Wigner Research Centre for Physics, H-1525 Budapest, Konkoly Thege Miklós út 29-33, Hungary; zimboras.zoltan@wigner.mta.hu

Kulcsszavak: Szeizmológia, Fourier-transzformáció, MGGL

keywords: Seismology, Fourier transform, MGGL

ABSTRACT:

We study a noise-filtered Fourier transform method that has until now only been tested on synthetic data. After a brief theoretical summary and a test of the method on synthetic data, we apply it also to real data obtained from the Mátra Gravitational and Geophysical Laboratory. We describe the difficulties in implementing the method and discuss the outlook for improved methods.

1 INTRODUCTION

Reducing the influence of noise in seismic data is an essential task in geophysics. Since different variants of the Fourier transform play a particularly important role in seismic data analysis, various noise filtering techniques have been developed that can be used before and after the computation of a Fourier transform or its inverse. In the present paper, we will investigate one such filtering technique that was developed in Refs. (Dobróka et al. 2012,2017). This method was tested only on synthetic data until now. Here we present a report on the application of this method to real-world data obtained from the Mátra Gravitational and Geophysical Laboratory (MGGL) (Barnaföldi et al. 2017, 2016). The Laboratory is located near Gyöngyösoroszi, 88 m below the surface in the cavern system of an unused ore mine. The aim of the MGGL is to measure and analyze the advantages of the underground installation of third generation gravitational wave detectors, in particular the Einstein Telescope (ET Science Team, 2011). As part of the Einstein Telescope design phase, ground motion studies are accomplished at various sites all around the world, see, e.g. (Beker et al. 2015), and such seismic studies are also the among the key goals of the MGGL. It should also be mentioned that beside seismic investigations, other specialized instruments have been also installed at the MGGL to measure infrasound, electromagnetic noise, and the variation of the cosmic muon flux.

The paper is structured in the following way. In Section 2, we briefly summarize the noise-filtered Fourier transform of Ref. (Dobróka et al. 2017). Section 3 is devoted to the analysis of the MGGL seismic data with this method, and this is followed by a brief summary and outlook in Section 4.

2 DESCRIPTION OF THE NOISE-FILTERED FOURIER TRANSFORM METHOD

In order to make the presentation self-contained, we present briefly the theoretical background of Fourier transform method introduced in Refs. (Dobróka et al. 2012, 2017). The starting step is a discretization of the frequency spectrum,

$$\mathrm{U}(\omega) = \sum_{\mathrm{n}=1}^{\mathrm{M}} \mathrm{B_n}\, \psi_{\mathrm{n}}(\omega) \tag{1}$$

where we expanded $\mathrm{U}(\omega)$ in terms of the basis functions ψ_{n}-s with $\mathrm{B_n}$ complex coefficients, taking into account only M number of unknown series expansion coefficients. From this we can get the time series by applying inverse Fourier transformation and denoting the kth discrete time sample with t_k :

$$\mathrm{u}(\mathrm{t_k}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Big(\sum_{\mathrm{n}=1}^{\mathrm{M}} \mathrm{B_n}\, \psi_{\mathrm{n}}(\omega)\Big) \mathrm{e}^{\mathrm{i}\omega \mathrm{t_k}}\, \mathrm{d}\omega = \sum_{\mathrm{n}=1}^{\mathrm{M}} \mathrm{B_n} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi_{\mathrm{n}}(\omega) \mathrm{e}^{\mathrm{i}\omega \mathrm{t_k}} \mathrm{d}\omega = \sum_{\mathrm{n}=1}^{\mathrm{M}} \mathrm{B_n}\, \mathrm{G_{k,n}} \tag{2}$$

with

$$\mathrm{G_{k,n}} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi_{\mathrm{n}}(\omega) \mathrm{e}^{\mathrm{i}\omega \mathrm{t_k}} \mathrm{d}\omega \tag{3}$$

being the Jacobian matrix, i.e., the inverse Fourier transform of the basis functions. During the computation we estimate the $\mathrm{B_n}$ coefficients by solving equation (2). Choosing the ψ_{n} functions to be the Hermite polynomials times an exponential term $e^{\frac{-\omega^2}{2}}$ turns out to be useful as they are the eigenfunctions of the Fourier transformation operator with eigenvalues i^n. Thus, our basis will be given by

$$H_n^{(0)}(\omega) = \frac{e^{\frac{-\omega^2}{2}} h_n^{(0)}(\omega)}{\sqrt{\sqrt{\pi} n! 2^n}} \tag{4}$$

where $h_n^{(0)}(\omega)$ is the nth Hermite polynomial and the denominator comes from the normalization induced by the orthogonality condition of the $H_n^{(0)}$ Hermite functions. (Note that the "(0)" upper index notation only denotes that these are the unscaled basic forms of the functions.)

In order to be able to apply this to real problems, the functions need to be scaled by a factor α:

$$H_n(\omega, \alpha) = \frac{e^{\frac{-\alpha\omega^2}{2}} h_n(\omega, \alpha)}{\sqrt{\sqrt{\frac{\pi}{\alpha}} n! (2\alpha)^n}} \tag{5}$$

Then these are again normalized and orthogonal eigenfunctions of the Fourier transformation. Furthermore, the well-known recursive formula for calculating the $h_n(\omega, \alpha)$ Hermite polynomials (now scaled with α) can be generalized as

$$h_{n+1}(\omega, \alpha) = 2\omega\alpha h_n(\omega, \alpha) - 2n\alpha h_{n-1}(\omega, \alpha) \tag{6}$$

With this and the transformations $\omega' = \sqrt{\alpha}\omega$ and thus $t' = \frac{t}{\sqrt{\alpha}}$, the Jacobian $G_{k,n}$ can be written as

$$G_{k,n} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} H_n(\omega, \alpha) e^{i\omega t_k} d\omega = \frac{1}{\sqrt[4]{\alpha}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} H_n^{(0)}(\omega', \alpha) e^{i\omega' t_k'} d\omega' = \frac{1}{\sqrt[4]{\alpha}} \mathcal{F}^{-1}\big[H_n^{(0)}(t_k')\big] \tag{7}$$

Since these functions have eigenvalues i^n , the corresponding equation reads

$$\frac{1}{\sqrt[4]{\alpha}} \mathcal{F}^{-1}\big[H_n^{(0)}(t_k')\big] = \frac{i^n}{\sqrt[4]{\alpha}} H_n^{(0)}\left(\frac{t_k}{\alpha}\right) \tag{8}$$

This is useful for us as we do not need to solve any integration to get the Jacobian.

In our work, we regard the Fourier transform problem as a least squares inverse problem, which is the first method described in the source article (Dobróka et al. 2017). Once we calculated $\mathbf{G}$ we can write up the following equation:

$$\mathbf{G}^T \mathbf{G} \vec{\mathbf{B}} = \mathbf{G}^T \vec{\mathbf{u}}^{(measured)} \tag{9}$$

This derives from taking the L_2norm of the so-called deviation vector and search for its minimum:

$$E_2 = \sum_{k=1}^{N} e_k^2 = \sum_{k=1}^{N} \big(u_k^{(\mathrm{measured})} - u_k^{(\mathrm{calc})}\big)^2 = \sum_{k=1}^{N} \big(u_k^{(\mathrm{measured})} - \sum_{n=0}^{M} B_n\, G_{k,n}\big)^2 = \min. \tag{10}$$

Here we have conducted an inverse Fourier transform on Eq. (1) and inserted Eq. (3) into it since$G_{k,n} = \mathcal{F}^{-1}[\psi_n(\omega')]$. With this we obtained the calculated time series $u_k^{(calc)} = u_k^{(calc)}(t_k)$ which we then plugged into the last equation.
Then we estimate the complex coefficient B n values by multiplying both sides with $(\mathbf{G}^T\mathbf{G})^{-1}$ from the left:

$$\vec{\mathbf{B}} = (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\vec{\mathbf{u}}^{(measured)} \tag{11}$$

Using this vector, we can get the time series according to Eq. (2), which is simply the discretized inverse Fourier problem:

$$u^{(\mathrm{calc})}(t_k) = \sum_{n=1}^{M} B_n\, G_{k,n} \tag{12}$$

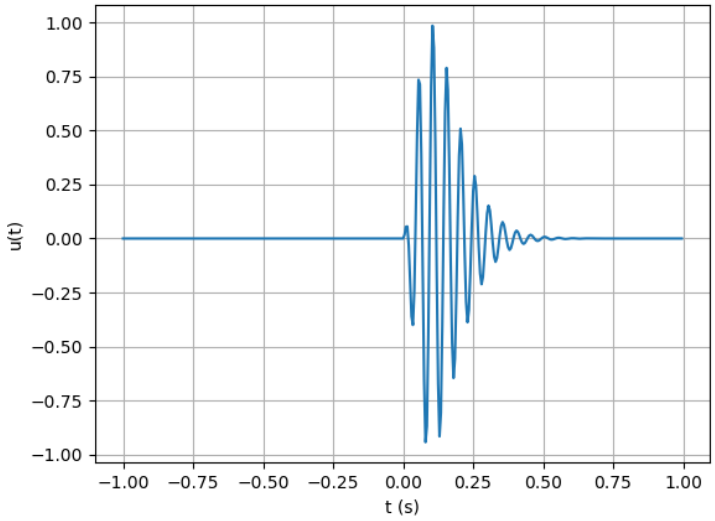

FIG. 1: Input signal data without noise according to Eq. (13).

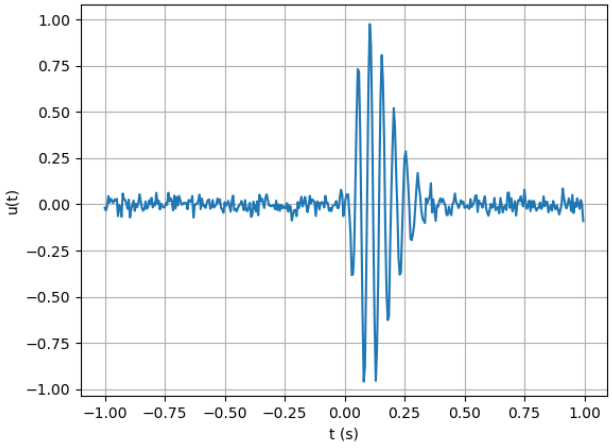

FIG. 2:(a) Noisy data $\vec{u}^{(measured)}$.

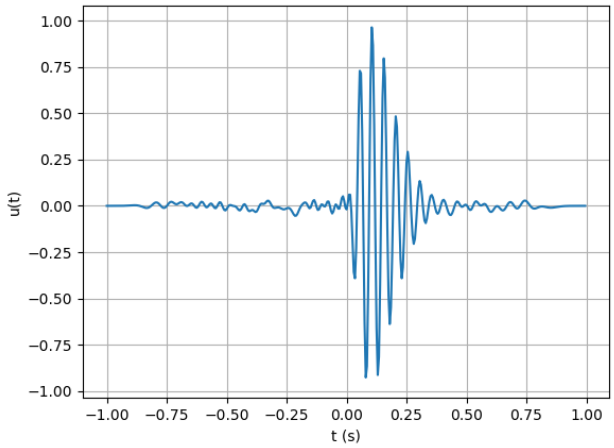

FIG. 2:(b) Filtered data $\vec{u}^{(calc)}$.

FIG. 2: Synthetic test data with Gaussian noise and the filtered data after applying the algorithm.

We have tested this algorithm at first on the following special artificial data series, which was also used in the source article (Dobróka et al. 2017), i.e., for the signal

$$u(t) = 738.91 \cdot t^2\, e^{-20t} sin\left(40\pi + \frac{\pi}{4}\right) \tag{13}$$

randomized with Gaussian noise with a standard deviation $\sigma = 0.03$. In Fig. 1 the noiseless signal is shown, while Fig. 2(a) and 2(b) depicts the noisy signal and the filtered signal, respectively. In this test simulation we have used $M = 100$ dimensional basis and a scaling factor of $\alpha = 0.004$. The algorithm seems to reduce high frequency noise by smoothing and it becomes stronger towards the sides of the window. The distance in L_2-norm between the filtered signal and the noiseless signal has decreased compared to that between the noisy and the noiseless signal, showing the effectiveness of the method. In the following numerical calculations, we will keep this α value and change the dimension of our basis of Hermite functions.

After the synthetic test examples, we move on to apply the method to real data. The data to be used was obtained at underground facility of the MGGL by a Guralp CMG-3T seismometer, where three independent sensor mass position outputs are recorded with the counting-number data rescaled to millivolts (mV). An example of a time series obtained in this way (on 2017 June 26) is shown on Fig. 3.

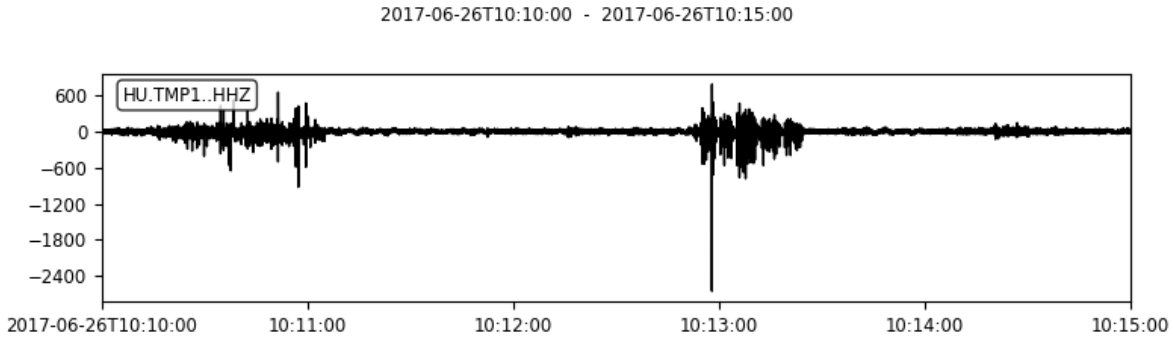

FIG. 3: Seismic data series used to test the filter algorithm.

This time series data on Fig. 3 has been detrended and then highpassed in order to get rid of the low frequency noise components. We have cut out shorter chunks in the neighborhood of large peaks and performed the filtering. Figs. 4-11 show the input and output data for different number of basis functions (M). Our program code applies the described filtering method, using equations (11) and (12) for determining $\vec{u}^{(calc)}$.

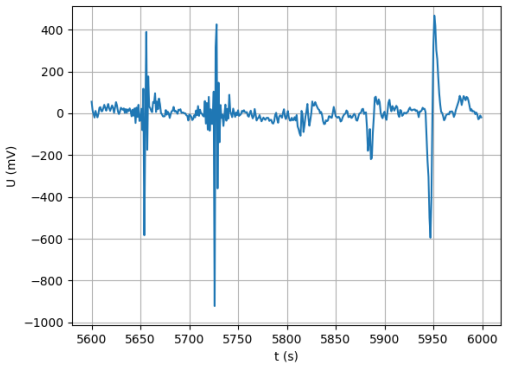

FIG. 4: Noisy data $\vec{u}^{(measured)}$.

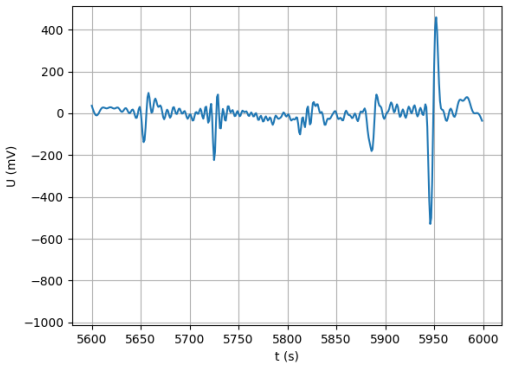

FIG. 5: Filtered data $\vec{u}^{(calc)}$, M = 100.

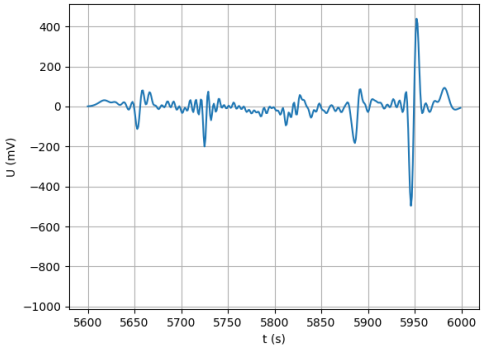

FIG. 6: Filtered data $\vec{u}^{(calc)}$, M = 120.

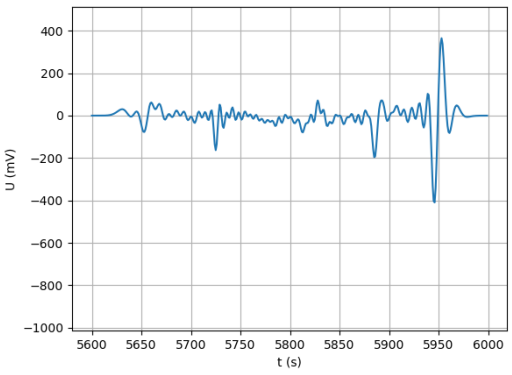

FIG. 7: Filtered data $\vec{u}^{(calc)}$, M = 140.

We can learn a number of lessons from these tests. If within the selected time window a sharp peak is contained near the ends, where the difference between two neighboring data points is large, the filter can ruin it by significantly reducing its size as it tries to make it smoother. It can be seen that there are three large peaks, two of which are sharp and one less sharp. After the filtering the size of the sharper peaks reduces critically and only the third peak survives. The algorithm also seems to filter out high frequency oscillations and dampen the sharpness of the peaks. The damping is the strongest if we use

relatively small M. The higher the dimension, the better result we get in terms of preserving the large peaks.

In order to avoid the disappearing of large peaks, we have cut out each peak from the chunk whose amplitude was larger than a given value and positioned the peaks into the middle of a window. Then we applied the filter on these shorter windows always with a peak in the middle. In each of the following four figures such a case is presented with one peak in the middle and two small peaks on each side. If two chunks happen to lie close to each other, we can reduce the window size so that there is only one large peak in each window. Of course, sometimes there will be cases when we cannot avoid having at least one large side peak really close to the middle peak. For those we can try to set the middle point in between the peaks.

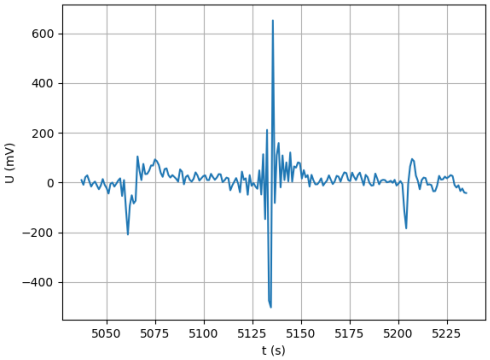

FIG. 8: Noisy data $\vec{u}^{(measured)}$.

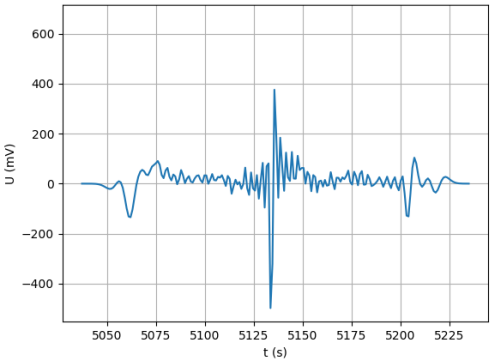

FIG. 9: Filtered data $\vec{u}^{(calc)}$, M = 100.

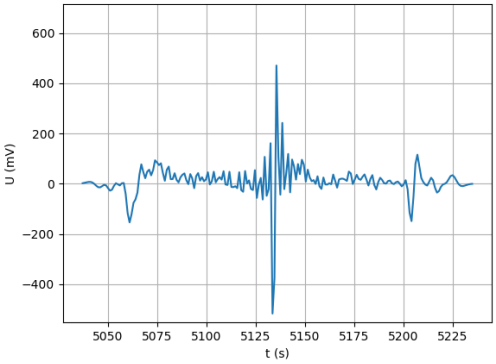

FIG. 10: Filtered data $\vec{u}^{(calc)}$, M = 120.

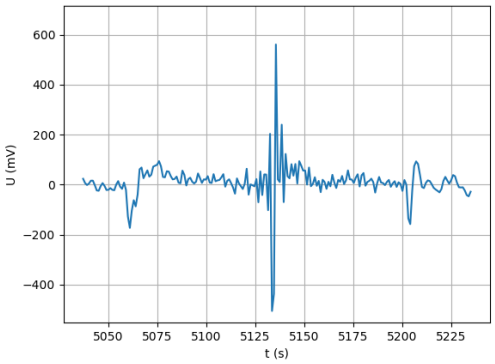

FIG. 11: Filtered data $\vec{u}^{(calc)}$, M = 140.

From the plots above we can see that the algorithm smooths the most in case of the sides of the window. However as we increase the dimension of our basis this effect reduces. We can also observe that near the big middle peak the high frequency oscillation gets slightly bigger instead of getting smoother. This effect increases with increasing dimension. Several other results confirm these conclusions. Since the results show that the smoothing occurs mainly at the sides of the window they will become dependent on the cutting. After the filtering we have reinserted the peaks into our original unfiltered data. An example is shown in Fig. 12.

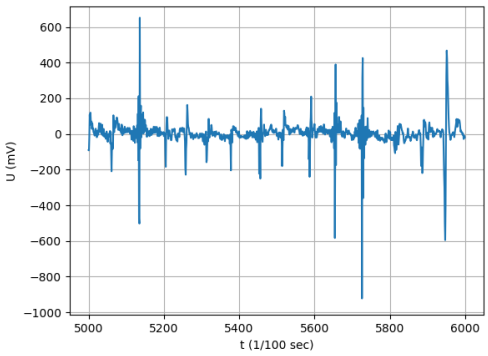

FIG. 12:(a) Input data.

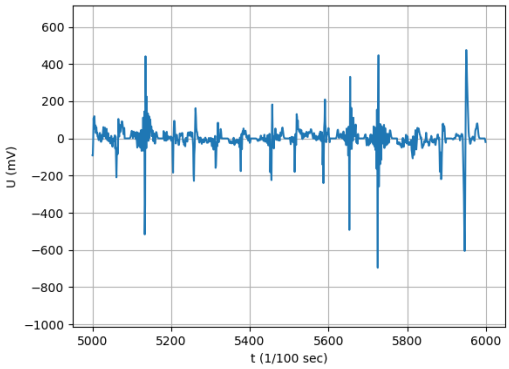

FIG. 12:(b) Filtered data.

FIG. 12: Reassembly of data series after filtering the windows separately. M = 100.

The result of the filtering is the most visible in between the peaks where the signal has become smoother thanks to that these were the sides of the windows that we had cut out.

For the reason that in this way the result will largely depend on the cuttings and the type of the data the testing of the second method described by the authors in (Dobróka et al. 2017) is required where more promising results are expected.

4 SUMMARY AND OUTLOOK

We reviewed a novel noise-filtered Fourier transform method that has until now only been tested on synthetic data. By developing a code that implements this transform, we analyzed real-world data, namely the MGGL seismic data. From this analysis it became clear that much care needs to be taken when implementing such an approach. By using this approach, we could filter out noise around a few peaks, however, overall the data did not change much. Nevertheless, one could hope that using even more advanced methods (such as the the second method of Ref. (Dobróka et al. 2017)) filtered Fourier transforms could become useful tool also in practice, which would, for instance, even allow for a more precise determination of the power spectral density plots.

5 ACKNOWLEDGMENT

We would like to thank the MGGL collaboration, in particular Péter Ván, Mátyás Vasúth, Előd Csongor Debreceni and Péter Lévai, for their help and encouragement, and would like to acknowledge the useful correspondence with Mihály Dobróka and Péter Vass concerning the implementation of the noise-filtered Fourier transform. We are also grateful for the continuous support given by Ádám Váradi and Vilmos Rofrits from Nitrokémia Zrt and Róbert Weisz from GEO-FABER Zrt to the MGGL collaboration. This work was supported by the Hungarian National Research, Development and Innovation Office (NKFIH Grants No. 124366 and 124508), and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

References:

M. Dobróka, H. Szegedi, P. Vass, E. Turai (2012): Fourier transformation as inverse problem? an improved algorithm; Acta Geo- daetica et Geophysica Hungarica, 47, 185-196.

M. Dobróka, H. Szegedi, J. S. Molnár, P. Szűcs (2017): On the reduced noise sensitivity of a new Fourier transformation algorithm; Math. Geosci. 47, 679-697.

Barnaföldi G.G. és tsi. (2017), First report of long term measurements of the MGGL laboratory in the Mátra mountain range. Classical and Quantum Gravity, 34:114001(22), (arXiv: 1610.07630).

Barnaföldi G.G. és tsi. (2016): A Mátrai Gravitációs és Geofizikai Laboratórium első mérései és mérési programja. Magyar Geofizika, 57(4):152.169.

ET Science Team (2011): Einstein gravitational wave telescope, conceptual design study Technical Report ET-0106C-10 (www.et-gw.eu/etdsdocument)

M. G. Beker, J. F. J. van den Brand és D. S. Rabeling (2015): Subterranean ground motion studies for the Einstein Telescope. Class. Quantum Grav. 32, 025002.