**Eligiusz PAWŁOWSKI**

Lublin University of Technology, Electrical Engineering and Computer Science Faculty, Department of Automatics and Metrology

doi:10.15199/48.2020.08.29

# Application of the Least Squares Method to the approximation of equally spaced samples in frequency measurement approach

***Abstract**. The paper discusses the application of the method of least squares to the linear approximation of the results of the measurements distributed evenly along the axis of abscissae or the axis of ordinates. This makes it possible to significantly simplify the formulae for calculating the values of the coefficients of the approximating straight line. The application of the proposed solution to frequency measurements is outlined.*

***Streszczenie.** W artykule rozpatruje się metodę najmniejszych kwadratów do liniowej aproksymacji wyników pomiarów rozmieszczonych równomiernie wzdłuż osi odciętych lub rzędnych. W takim przypadku otrzymuje się znaczne uproszczenie wzorów do obliczania wartości współczynników prostej aproksymującej. Przedstawiono zastosowanie proponowanego rozwiązania do pomiarów częstotliwości. (**Aplikacja metody najmniejszych kwadratów do aproksymacji równomiernie rozmieszczonych próbek w zastosowaniu do pomiarów częstotliwości**).*

**Keywords:**. Least Squares Method, LMS, equally spaced samples, frequency measurements.
**Słowa kluczowe:** Metoda Najmniejszych Kwadratów, MNK, równomiernie rozmieszczone próbki, pomiary częstotliwości.

## Introduction

The least squares method (LSM) is widely used for adjusting measurement results burdened with random errors [1]. The literature of the subject most frequently discusses the linear approximation of LSM as presented on Fig. 1. The straight line with equation $y=ax+b$ approximates the measurement points $(x_1, y_1)$, $(x_2, y_2)$, ..., $(x_i, y_i)$, ..., $(x_n, y_n)$ in such way as to minimise the sum of squares of accurately defined errors [2, 3]. Three types of LSM are distinguished: classical, inverse and orthogonal LSM [4, 5].

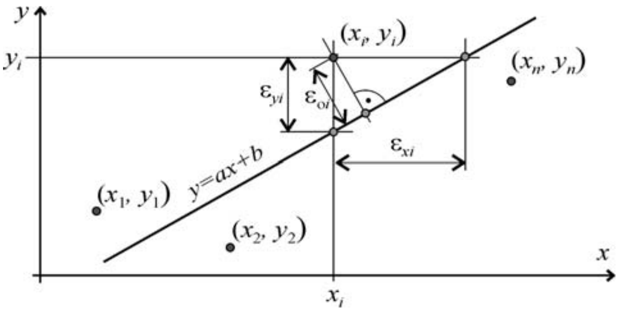


Fig.1. Illustration of classical, inverse and orthogonal LSM

In classical LSM, the sum of squares of errors of the ordinate values $\sum_{i=1}^{n}\varepsilon_{yi}^2$ is minimised [1], assuming that the values of abscissae $x_i$ are precisely known. In inverse LSM, the sum of squares of errors of the abscissa values $\sum_{i=1}^{n}\varepsilon_{xi}^2$ is minimised [4, 5], assuming that the values of ordinates $y_i$ are precisely known. Orthogonal LSM takes into consideration the errors of both abscissae and ordinates, and minimises the sum of squares of the distances of measurement points from the approximating straight line $\sum_{i=1}^{n}\varepsilon_{oi}^2$ [2, 3]. In practice, classical LSM is applied when the errors of abscissae $\varepsilon_{xi}$ are negligibly small as compared to the errors of ordinates $\varepsilon_{yi}$; in particular if it is possible that the values of abscissae $x_i$ are precisely known and differ by the same value, i.e. are equidistant. Likewise, inverse LSM is applied when the errors of ordinates $\varepsilon_{yi}$ are negligibly small as compared to the errors of abscissae $\varepsilon_{xi}$; in particular if it is possible that the values of ordinates $y_i$ are precisely known and differ by the same value, i.e. are equidistant. If the measurement errors of both coordinates are significant and similar in value, the orthogonal approximation should be applied [2, 3]. In such case, it is not possible to distribute the measurement points evenly.

The paper discusses the modification of LSM for a linear approximation of evenly distributed samples, and presents the examples of its application to frequency measurements.

## Case of equidistant abscissae

Fig. 2 shows the case of application of classical LSM to linear approximation of measurement points distributed evenly along the axis of abscissae $x$ [1, 6, 7, 8].

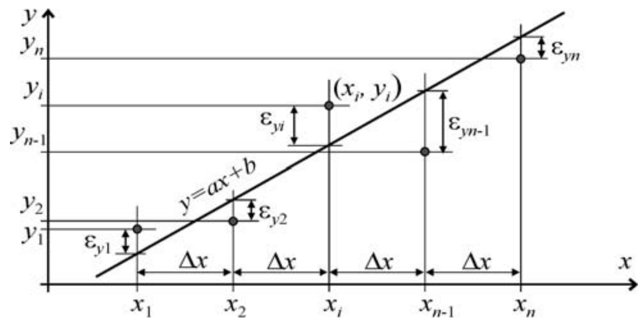


Fig.2. Least squares method for equidistant abscissae

For classical LSM, the coefficients of the approximating straight line $y=ax+b$ are determined with the following formula [1]:

$$a=\frac{n\sum_{i=1}^{n}x_i y_i-\sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i}{n\sum_{i=1}^{n}x_i^{\,2}-(\sum_{i=1}^{n}x_i)^2}, \tag{1}$$

$$b=\overline{y}-a\overline{x}. \tag{2}$$

The values of abscissae $x_1$, $x_2$, ... $x_i$, ... $x_n$ are precisely known with negligible error, and are equidistant, with the distance between them equivalent to the value of $\Delta x$:

$$x_i=i\,\Delta x. \tag{3}$$

By substituting (3) to (1) and taking into consideration the formulae for the sum of numerical series:

(4) $$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}\ ,\ \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}\ ,$$

after transformation we obtain the equation for the slope of the straight line:

(5) $$a = \frac{1}{\Delta x}\,\frac{6\sum_{i=1}^{n}\left(2\,i-n-1\right)y_i}{n\left(n^2-1\right)} = \frac{1}{\Delta x}\,\frac{\sum_{i=1}^{n} c_i y_i}{k_n}\ ,$$

where $k_n$, $c_i$ are coefficients calculated from the following equations:

(6) $$k_n = \frac{n(n^2-1)}{6}\ ,\ c_i = 2i-n-1\ .$$

It should be pointed out that $k_n$, $c_i$ are integers, clearly defined only by the number $n$ of the approximated measurement points. The examples of coefficient values $k_n$, $c_i$ are presented in Table 1. Considering the equation for mean values:

(7) $$\overline{x} = \frac{n-1}{2}\Delta x\ ,\ \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i\ \ ,$$

we obtain the equation for the intercept:

(8) $$b = \frac{1}{n}\sum_{i=1}^{n} y_i - a\frac{n-1}{2}\Delta x\ .$$

It should be pointed out that the obtained equations (5), (8) for coefficients $a$ and $b$ of the approximating straight line are much simpler than the general equations (1) and (2) which makes it possible to apply them to microcontroller systems with low computing power.

Table.1. The examples of coefficient values $k_n$, $c_i$

| $n$ | $k_n$ | $c_i$ | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4 | 10 | -3 | -1 | 1 | 3 | | | | | | | | | | | | |
| 7 | 56 | -6 | -4 | -2 | 0 | 2 | 4 | 6 | | | | | | | | | |
| 9 | 120 | -8 | -6 | -4 | -2 | 0 | 2 | 4 | 6 | 8 | | | | | | | |
| 10 | 165 | -9 | -7 | -5 | -3 | -1 | 1 | 3 | 5 | 7 | 9 | | | | | | |
| 15 | 560 | -14 | -12 | -10 | -8 | -6 | -4 | -2 | 0 | 2 | 4 | 6 | 8 | 10 | 12 | 14 | |
| 16 | 680 | -15 | -13 | -11 | -9 | -7 | -5 | -3 | -1 | 1 | 3 | 5 | 7 | 9 | 11 | 13 | 15 |

**Case of equidistant ordinates**

Fig. 3 shows the case of application of inverse LSM to linear approximation of measurement points distributed evenly along the axis of ordinates [12, 13, 14].

For inverse LSM [4, 5], coefficient $a$ of the approximating straight line is determined from the following formula:

(9) $$a = \frac{n\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2}{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}\ ,$$

and the intercept is calculated from equation (2). The values of ordinates $y_1$, $y_2$, ... $y_i$, ... $y_n$ are precisely known with negligible error, and are equidistant, with the distance between them equivalent to the  from each other by the value of $\Delta y$:

(10) $$y_i = i\ \Delta y\ .$$

Substituting (10) to (9) and taking into consideration the formulae for the sums of numerical series (4), after similar transformations we obtain the equation for the slope:

(11) $$a = \Delta y\,\frac{n\,(n^2-1)}{6\sum_{i=1}^{n}(2i-n-1)x_i} = \Delta y\,\frac{k_n}{\sum_{i=1}^{n} c_i x_i}\ ,$$

where coefficients $k_n$, $c_i$ are calculated from formulae (6). Considering the equation for mean values:

(12) $$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i\ ,\ \overline{y} = \frac{n-1}{2}\Delta y\ ,$$

we obtain the equation for the intercept:

(13) $$b = \frac{n-1}{2}\Delta y - a\,\frac{1}{n}\sum_{i=1}^{n} x_i\ .$$

Also, in this case the obtained equations (11) and (13) for coefficients $a$ and $b$ of the approximating straight line are much simpler than the general equations (9) and (2).

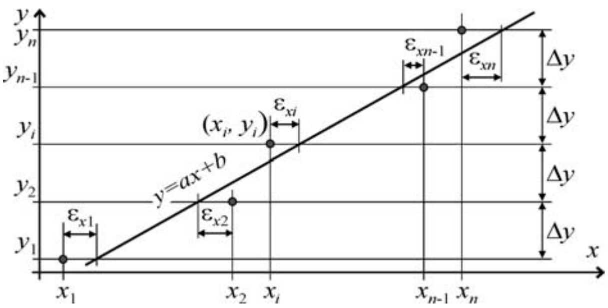


Fig.3. Least squares method for equidistant ordinates

**The fundamental principle of frequency measurements**

For a sinusoidal signal $u_s(t)$ with the frequency $f_s$ and amplitude $U_s$, defined by the equation:

(14) $$u_s(t) = U_s \sin(2\pi\ f_s t + \varphi_0)\ ,$$

phase $\varphi$ of the signal is a linear function of time:

(15) $$\varphi(t) = 2\pi\ f_s t + \varphi_0\ ,$$

where $\varphi_0$ is the initial phase. The frequency $f_s$ of the signal in question is determined by the derivative of phase $\varphi$ of the signal in relation to time [15]:

(16) $$f_s = \frac{1}{2\pi}\,\frac{d\varphi}{dt}\ .$$

This equation is the fundamental principle of the majority of frequency measurement methods, whereby the derivative is usually replaced with the difference quotient calculated for the set time interval $\Delta t$ or the set phase increment $\Delta\varphi$. The first method is carried out with a traditional digital frequency meter which counts the signal periods within a specified time [16], e.g. one second, and displays the measurement results in hertz. The second method is based on measuring the time in which the signal phase increments by angle $\Delta\varphi=2\pi$, i.e. by one full period, the inverse of which is the measured signal frequency [9-11, 13, 14, 16]. A number of modifications of both basic methods have been developed and reported in the literature of the subject. This paper considers frequency measurements with the use of the linear approximation of the dependence between angle $\varphi$ and time $t$, using LSM. The resulting slope of the approximating straight line equals the derivative of the signal phase in relation to time and makes it possible to calculate frequency according to equation (16).

**Measurement of the difference of standard frequencies**

The relative difference of frequencies of two standard frequency generators is very small, falling within the range of $10^{-6} \div 10^{-14}$, which is why it is measured using special methods. The generally applied approach is the so-called phase method [6, 7, 8]. The principle of measurement is provided on Fig.4, while the block diagram of the measurement system is presented on Fig.5. The measurement is conducted at reference frequency $f_{ref}$=1Hz, i.e. the periods of both signals have similar values of $T_{gen} \approx T_{ref}$=1s.

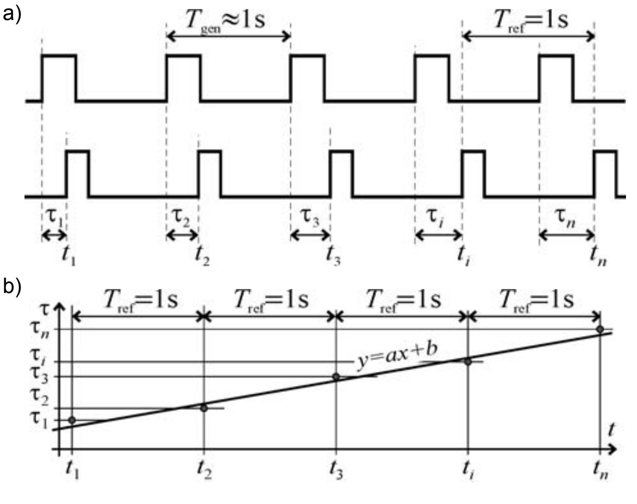


Fig.4. Measurement of the difference of standard frequencies: a) time intervals of the signals, b) approximation of the phase time at equidistant points in time

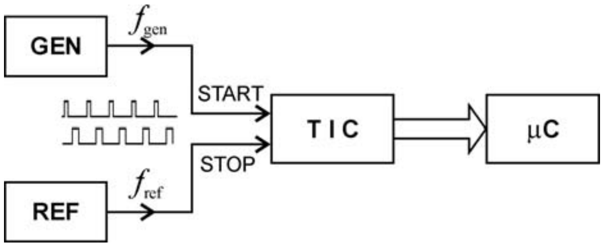


Fig.5. Block diagram of the system for measuring the difference of standard frequencies

The phase $\varphi_{ref}$ of signal of the reference generator REF (Fig.5) is a linear function of time $t$:

$$\varphi_{ref}(t) = 2\pi \frac{t}{T_{ref}} , \tag{17}$$

where $T_{ref}$ is the period of the reference signal (Fig. 4a). The signal of the analysed generator GEN (Fig.5) is shifted in time with respect to the reference generator by the so-called phase time $\tau$ [6], i.e. the phase $\varphi_{gen}$ of the signal of the analysed generator GEN can be written down as:

$$\varphi_{gen}(t) = \varphi_{ref}(t) + 2\pi \frac{\tau}{T_{ref}} = \frac{2\pi}{T_{ref}}(t+\tau) . \tag{18}$$

Considering (16), frequency $f_{gen}$ of the signal of the analysed generator GEN is:

$$f_{gen} = \frac{1}{2\pi}\frac{d\varphi_{gen}}{dt} = \frac{1}{2\pi}\frac{d}{dt}\left(\varphi_{ref}(t) + 2\pi\frac{\tau}{T_{ref}}\right). \tag{19}$$

If the phase time has a constant value of $\tau$=*const.*, then both frequencies are equal $f_{gen}=f_{ref}$; whereas if $\tau$ is changing, the relative difference of frequency equals:

$$\delta f_{gen} = \frac{f_{gen} - f_{ref}}{f_{ref}} = \frac{d\tau}{dt} = a , \tag{20}$$

where $a$ is the slope of the approximating straight line illustrated on Fig. 4b. The measurements of subsequent values of phase time $\tau_i$ are carried out by Time Interval Counter TIC (Fig.5) at the points in time $t_i$ equidistant from each other by the reference signal period, so $T_{ref}$=1s. Therefore, classical LSM for equidistant abscissae ($\Delta x = T_{ref}$) can be applied (Fig.2), whereby:

$$x_i = t_i = i\,T_{ref} \; , \; y_i = \tau_i . \tag{21}$$

Considering equation (5), the relative difference of frequency of both generators equals:

$$\delta f_{gen} = a = \frac{1}{T_{ref}} \frac{6\sum_{i=1}^{n}(2i-n-1)\tau_i}{n(n^2-1)} = \frac{1}{T_{ref}}\frac{\sum_{i=1}^{n} c_i \tau_i}{k_n} , \tag{22}$$

where coefficients $c_i$ and $k_n$ are calculated from formulae (6) and presented in Table 1. Finally the relative difference of frequency is calculated by microcontroller μC (Fig. 5). The method presented herein can also be applied to discipline the local generator VCXO with respect to one pulse-per-second (1PPS) signal of the Global Positioning System (GPS). The obtained measurement results confirm the possibility of synchronizing a local OCXO generator using the presented method to a 1 PPS GPS signal with the error $\delta f_{gen}$ not worse than $5 \cdot 10^{-11}$ Hz/Hz [7, 8].

**Frequency measurement of noisy signal**

Frequency measurement according to equation (16) can also be performed by approximating the signal phase at the points where the signal passes through zero level with the same sign of derivative [12]. The principle of measurement is provided on Fig. 6, while the block diagram of the relevant measurement system is presented on Fig. 7.

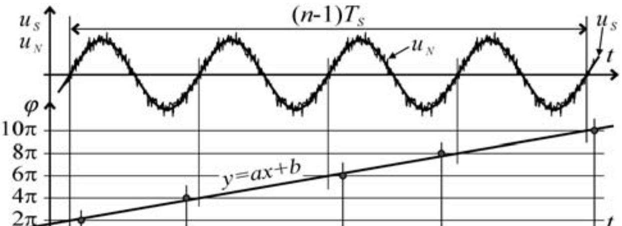


Fig.6. Frequency measurement of a noisy sinusoidal signal

We are considering the measurement of frequency $f_s$ of signal $u_N(t)$ which is the sum of the sinusoidal signal $u_s$ (14) and noise $N(t)$ with normal distribution and zero expected value:

$$u_N(t) = u_s(t) + N(t) = \hat{U}_s \sin(2\pi\, f_s t + \varphi_0) + N(t) . \tag{23}$$

For the sinusoidal signal $u_s$, its phase $\varphi$ is a linear function of time, the values of which are equal to the multiple of round angle $2\pi$ at subsequent points in time $t_1$, $t_2$, ..., $t_i$, ..., $t_n$ when signal $u_s$ is passing through zero level (Fig. 6). Frequency $f_s$ of signal $u_s$ can be determined from the following derivative [12, 15]:

(24) $$f_s = \frac{1}{2\pi}\frac{d\varphi}{dt} = \frac{a}{2\pi},$$

where $a$ is the slope of the approximating straight line illustrated on Fig. 6.

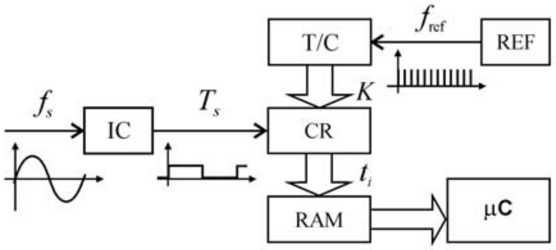


Fig.7. Block diagram of the system for measuring the frequency of a noisy signal

The measurements of subsequent points in time $t_i$ when the signal passes through zero level are carried out by Timer/Counter T/C (Fig. 7); the current state $K$ of T/C is read "on-the-fly" by Capture Register CR and stored in RAM. Input Circuit IC detects subsequent points in time $t_i$ when the signal passes through zero level, which correspond to identical increments of the signal phase precisely by the angle equal to $2\pi$. However, the points in time $t_i$ are not precisely measured due to the presence of noise $N(t)$ in the measured signal $u_N(t)$. Therefore, inverse LSM for equidistant ordinates ($\Delta y=2\pi$) can be applied (Fig.3), whereby:

(25) $$x_i = t_i\ ,\ y_i = 2\pi i\ .$$

Considering equation (11), the signal frequency $f_s$ is equals:

(26) $$f_s = \frac{a}{2\pi} = \frac{n(n^2-1)}{6\sum_{i=1}^{n}(2i-n-1)t_i} = \frac{k_n}{\sum_{i=1}^{n}c_i t_i}$$

where coefficients $c_i$ and $k_n$ are calculated from formulae (6) and presented in Table 1. Finally the signal frequency $f_s$ is calculated by microcontroller μC (Fig.7). The method presented herein makes it possible to effectively measure the frequency of a power grid in the presence of noise. The trigger error introduced by Input Circuit IC is an important error component of digital measurements of slow-edge signal frequencies in the presence of noise. This error can be effectively reduced by estimating the frequency using the method proposed (26), calculating the slope of the linear regression using the LSM [12]. The same method can be used for the frequency pulse signal, for which we can determine only increases of the phase equal to a full period $\Delta\varphi=2\pi$ [13].

**Summary**

The paper presents the application of LSM to the linear approximation of the results of the measurements distributed evenly along the axis of abscissae or the axis of ordinates. The relevant mathematical equations and examples of application of the method to frequency measurement were discussed. The advantage of the proposed solution is the significant simplification of the necessary mathematical calculations, thus making it possible to use it in microcontroller systems with low computing power.

***Author:*** *dr inż. Eligiusz Pawłowski, Lublin University of Technology, Electrical Engineering and Computer Science Faculty, 38A Nadbystrzycka str., 20-618 Lublin, Poland, E-mail: e.pawlowski@pollub.pl*

REFERENCES

[1] Taylor, J.R.: An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements. 2nd edn. *University Science Books*, Sausalito, California (1997)
[2] Dorozhovets M., Taking into consideration uncertainties of the both results in linear regreesion, *Pomiary Automatyka Kontrola*, 54 (2008), nr 9, 612--615
[3] Dorozhovets M., Uncertainty of linear orthogonal regression, *Pomiary Automatyka Kontrola*, 53 (2007), nr 9 bis, 31-34
[4] Jason J.Z. Liao, An insight into linear calibration: univariate case, *Statistics & Probability Letters*, 56 (2002), No. 3, 271-281
[5] Shalabh, Toutenburg H., Consequences of departure from normality on the properties of calibration estimators, *Journal of Statistical Planning and Inference*, 136 (2006), No. 12, 4385-4396
[6] Lombardi M., An Introduction to Frequency Calibration, *NIST Frequency Measurement and Analysis System: Operator's Manual - Appendix A*, (2009), 55-86
[7] Pawłowski E., Comparison of reference frequencies in the LabVIEW environment, *Zeszyty Naukowe Wydziału Elektrotechniki i Automatyki PG*, 34 (2013), 53-56
[8] Pawlowski E., Warda P., Method and system for disciplining a local reference oscillator by GPS 1PPS signal, *Przegląd Elektrotechniczny*, 94 (2018), No 7, 38-41
[9] Patil T., Ghorai S., Robust zero-crossing detection of distorted line voltage using line fitting, *2016 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT),* Mysuru, (2016), 92-96
[10] Mog G. E., Ribeiro E. P., Zero crossing determination by linear interpolation of sampled sinusoidal signals, *2004 IEEE/PES Transmision and Distribution Conference and Exposition: Latin America, Sao Paulo,* (2004), 799-802
[11] Pawlowski E., Power Grid Frequency Estimation Based on Zero Crossing Technique Using Least Squares Method to Approximate Sampled Voltage Signal Around Zero Level, *in: Hanus R., Mazur D., Kreischer C. (eds) Methods and Techniques of Signal Processing in Physical Measurements. MSM 2018. Lecture Notes in Electrical Engineering*, Springer, Cham, 548 (2019), 248-268
[12] Pawlowski E., Warda P., Power grid frequency measurement in LabVIEW environment using the least mean squares method to signal phase approximation in the presence of noise, *2017 International Conference on Electromagnetic Devices and Processes in Environment Protection with Seminar Applications of Superconductors (ELMECO & AoS)*, Lublin, (2017), 1-4
[13] Świsulski D., Pawłowski E., Dorozhovets M. Digital Processing of Frequency–Pulse Signal in Measurement System, *in: Mazur D., Gołębiowski M., Korkosz M. (eds) Analysis and Simulation of Electrical and Computer Systems. Lecture Notes in Electrical Engineering, Springer, Cham*, 452 (2018), 319-332
[14] Świsulski D., Hanus R., Zych M., Petryka L., Methods of measurement signal acquisition from the rotational flow meter for frequency analysis, *The European Physical Journal Web of Conferences,* 143, 02124 (2017), 1-6
[15] Boashash B.: Estimating and Interpreting The Instantaneous Frequency of a Signal-Part 1: Fundamentals, *Proceedings of The IEEE*, 80 (1992), No. 4, 520-538
[16] Fundamentals of the electronic counters, Application Note 200, *Hewlett Packard Co.*, (1997)