Tomáš LOKAJÍČEK[1], Karel KLÍMA[2]

STANOVENÍ ČASU PŘÍCHODU SEISMICKÝCH/ULTRAZVUKOVÝCH SIGNÁLŮ POMOCÍ STATISTICKÝCH MOMENTŮ VYŠŠÍCH ŘÁDŮ

DETERMINATION OF FIRST ARRIVAL TIME OF SEISMIC/ULTRASONIC SIGNALS BY MEANS OF HIGH ORDER STATISTIC APPROACH

Abstract

One of the fundamental problems in non-destructive testing and geophysics is the precise determination of the first arrival of acoustic emission (AE) signals (or seismic signals) recorded by multi-channel systems. The knowledge of this time is very important, mainly in the case of automatic localization of individual AE events. Several approaches are routinely used in practice as crossing of the threshold level, analysis of the STA/LTA (short-time average/long-time average), etc. In our contribution, an approach based on high-order statistics (HOS) that is able to carry out precise arrival time determination without human intervention is presented. The approach was tested using real AE data recorded by 8 channel recording system. This simple, accurate and quite fast method is predetermined to be used in automatic processing of transient waveform data from acoustic emission, seismic signals, ultrasonic sounding, etc.

Keywords: high-order statistics, acoustic emission, first arrival identification

Introduction

Acoustic emission, micro seismic tremors or earthquakes are technical or geophysical phenomena that are studied based on elastic waves. These waves are radiated during the energy release process. Radiated waves are recorded by different types of pick-ups, such as piezoceramic transducers or geophones, etc. One of the fundamental problems is determination of the seismo-acoustic source origin. Such a problem is generally solved by the determination of the time of signal arrival, or signal-phase arrival – compressional (P-wave), transversal (S-wave), surface wave (Rayleigh, Love waves), etc. that are recorded by various recording sensors located at different directions and distances from the source of radiation. Most important is time of the first arrival of the signal to the sensor – arrival of compressional wave (P-wave), as this arrival is predominantly used for acoustic/seismic source location or different types of signal frequency analysis, mainly in the case of rock mechanics, or in the study of seismic/acoustic energy release, as such a study is carried out on 3D rock samples, dimension of which is from centimeters up to several tens of centimeters. Under these studies body waves are recorded. In case of recording of AE first arrival of Lamb mode HOS approach can be used successfully too, but only when compressional wave will not be pronounced in the recorded signal. Accurate time of this first arrival can be very useful for a number of different applications, such as seismic/acoustic source location, mechanism determination, structure description, seismicity designation, hazard assessment, etc. All of the above mentioned phenomena produce huge amount of data, which call for automatic recorded data processing to determine individual phase arrival without human intervention by sophisticated approaches that enable a precise automatic phase determination. In past years, different approaches were used for phase arrival determination such as exceeding of threshold level, ratio of short-time average (STA) to long-time average (LTA) published by Baer and Kradolfer (1987); seismic wave polarity assumption published by Jurkevic (1988); neural networks published by Zhao and Takano (1999); wavelet transform published by Anant and Dowla (1997). Saragiotis et al. (2002) published the high-order statistics based approach how to automatically determine P-phase arrival of radiated seismic signals by skewness and kurtosis values of the recorded seismic waves. These parameters were applied in P-wave arrival determination of seismic signals recorded in Greece.

[1] Ing., CSc., Institute of Geology, Academy of Sciences of the Czech Republic, v.v.i., lokajicek@gli.cas.cz
[2] † RNDr., CSc., Institute of Geology, Academy of Sciences of the Czech Republic, v.v.i.

The aim of our contribution was to analyze high-order statistic parameters and to apply high-order statistic approach to automate first arrival determination of the acoustic emission (AE) signals radiated from stressed rock samples. In our contribution we also designed an empirical formula (parameter) that follows from the definition of skewness and kurtosis – S_6 parameter. This parameter is also based on high order statistics approach.

Theory

Four basic statistic moments are used in mathematical statistics: first statistic moment - mean value – S1, second statistic moment - power of standard deviation – S2, and third and fourth statistic moments are used in empirical formulas as skewness – S3, and kurtosis – S4 (see NIST/SEMATECH), but generally statistic moment of any order can be used in statistic computations.

For univariate data Y1, Y2, ..., YN variance (first statistic moment) - S1 is defined as

$$S_1 = \bar{Y} = \frac{\sum_{i=1}^{N} Y_i}{N}, \qquad (1)$$

where $\bar{Y}$ is the mean value and N is the number of data points.

Standard deviation (second statistic moment) – S2 is defined as

$$S_2 = \sigma^2 = \frac{\sum_{i=1}^{N} (Y_i - \bar{Y})^2}{N-1}. \qquad (2)$$

Skewness - S3 is a measure of symmetry, or more precisely, the lack of symmetry. The distribution, i.e. data set, is symmetric, if it looks the same to the left and right to the peak point.

The formula for skewness is:

$$S_3 = skewness = \frac{\sum_{i=1}^{N} (Y_i - \bar{Y})^3}{(N-1)\sigma^3} \qquad (3)$$

The skewness for a normal distribution is zero, and any symmetric data should have skewness near zero. Negative values for the skewness indicate data that is skewed left and positive values for the skewness indicate data that they is skewed right. It follows from formula (3) that skewness is defined as the 3rd statistic moment $\frac{\sum_{i=1}^{N} (Y_i - \bar{Y})^3}{(N-1)}$ to be divided by the relevant power of standard deviation.

On the contrary, kurtosis is a measure of whether the data are peaked or flat in relation to normal distribution. That is, data sets with high kurtosis tend to have a distinct peak near the mean value, decline rather rapidly, and have heavy tails. Data sets with low kurtosis tend to have a flat top near the mean rather than a sharp peak. A uniform distribution would be the extreme case. The formula for kurtosis is:

$$S_4 = kurtosis = \frac{\sum_{i=1}^{N}(Y_i - \bar{Y})^4}{(N-1)\sigma^4} - 3 \quad (4)$$

The formula $\frac{\sum_{i=1}^{N}(Y_i - \bar{Y})^4}{(N-1)\sigma^4}$ has a value 3 for a standard normal distribution. For this reason, the kurtosis is defined as in formula (4). Positive kurtosis indicates a "peaked" distribution and negative kurtosis indicates a "flat" distribution.

Similarly to definition of skewness S3 and kurtosis S4, where relevant statistic moment of 3rd and 4th orders is divided by the appropriate power of standard deviation (second statistic moment), also empirical formula S6 was designed as,

$$S_6 = \frac{\sum_{i=1}^{N}(Y_i - \bar{Y})^6}{(N-1)\sigma^6} - 15 \quad (5)$$

The formula $\frac{\sum_{i=1}^{N}(Y_i - \bar{Y})^6}{(N-1)\sigma^6}$ has a value 15 for a standard normal distribution. For this reason, the S6 is defined as in formula (5).

In the following analysis only even parameters S2, S4 and empirical S6 were used for the study of determination of time arrival of the signal.

A methodical study based on high-order statistics was carried out on synthetic signals (data sets) that were designed to have a normal distribution. On such a data sets, parameter S1 and even parameters S2, S4 and S6 were calculated (see Table 1). Odd parameter S3 was not calculated due to the fact that it changed the sign and also its dependence was not sufficiently clear pronounced. Three different data sets were prepared with one hundred, two hundred and one thousand data points. All three above mentioned data sets had practically normal distribution, their mean value was zero and standard deviation σ equaled 1. Such data sets were deformed in our analysis by adding one point with a value equal to 3σ, or by adding two points with a value equal to 3σ and 4σ, respectively. Graphical presentation of three different data sets is depicted in Fig. 1. Data set with normal distribution is denoted by white color. Gray color shows data set distribution, which is deformed by adding 3□, black color display distribution deformed by adding 3□ and 4□ data points. One point and also a pair of two points were used with a positive value. Using this approach, the collection of three data sets of individual size was obtained for mutual comparison, see:

normal distribution values

3σ, + normal distribution values

4σ, 3σ, + normal distribution values

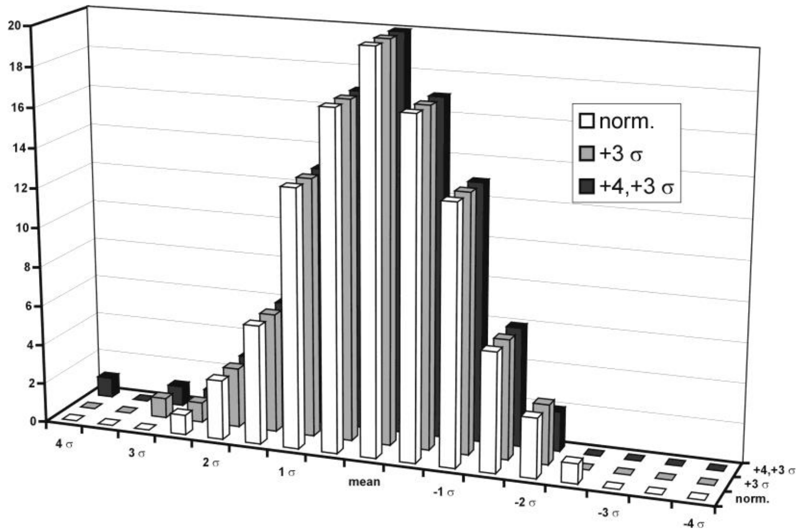

Fig.1 Frequency distribution of three synthetic data sets. Data set with normal distribution was distorted by adding one or two points having value 3σ or 3σ, 4σ, respectively.

Negative values added to the data with normal distribution simulated the signal arrival with a negative signal onset. On the contrary, positive values simulated an arrival of the signal with a positive onset.

Tab.1 Results of methodical analysis of data sets with normal distribution that were distorted by adding data 3σ, 4σ with positive and negative values. For resultant data sets parameters S_1, S_2, S_4, and S_6 were calculated.

points of signal	signal deformation	S parameter			
		S_1	S_2	S_4	S_6
100	**0**	0	0.9937	-0.1622	-2.9211
	3σ	0.0558	1.0055	-0.0112	-0.5575
	3σ, 4σ	0.1175	1.0601	0.9655	18.0453
200	**0**	0	0.9967	-0.0992	-1.9681
	3σ	0.029	0.9996	-0.0604	-1.3108
	3σ, 4σ	0.0612	1.0245	0.5681	11.6829
1000	**0**	0	0.9993	-0.0275	-0.6759
	3σ	0.0063	0.9984	-0.0543	-1.1638
	3σ, 4σ	0.0133	1.002	0.075	1.7954

Experiment

Laboratory measurements were carried out on a migmatite rock sample. Cylindrical rock sample with diameter 50 mm and 100 mm length was used in the study. Acoustic emission was recorded by eight wide band transducers - WD (Physical Acoustic Corporation - PAC). WD transducer has a flat amplitude transfer characteristics in the frequency range from 50 kHz to about 800 kHz. The transducers were located mainly on the rock sample surface. Two transducers were fixed in loading plates. Plane of these two transducers was perpendicular to the sample axis. Remaining 6 transducers were evenly distributed on the sample surface. 3 transducers were fixed on the circle positioned 30 mm from the bottom. The transducers had 120 degrees angular distribution. Remaining 3 transducers were also distributed on the circle, but 70 mm above the surface of the lower compressional plate. The sample was loaded by a constant stress rate 0.5 kN/min, achieved by a controlled hydraulic press. Total strength of the migmatite rock sample was 125 MPa.

Acoustic emission signals were recorded by eight-channel recording system. The sampling frequency was 10 MHz and each waveform had a time length of 204.8 μs (2048 points), resolution of A/D converter was 16 bits. The special object oriented software was designed to determine the arrival time by means of HOS-high order statistics method.

Results and Discussion

Acoustic emission signals radiated from the stressed rock samples have a very wide range of amplitudes, energy, frequency or time duration. Hundreds of thousands of recorded acoustic emission signals were processed by the procedure described bellow. For each recorded signal even parameters S_2, S_4 and S_6 and also the derivative S_6 were calculated. For all processed signals the moving window having the length of 100 points was used.

If a moving window is applied to the studied signal, then if this window contains only the points of the recorded data before the signal arrival, its recorded noise is close to normal distribution and thus parameters S_4 or S_6 are practically equal zero. As soon as the first point of the recorded signal occurs in the window, then the normal distribution is distorted by this point, which results in a change of all even S parameters, as S_2, S_4 and S_6.

"AE_1", see Figure 2, is an example of the weakest signals, where the real signal of the acoustic emission is comparable to noise of the complete recording system. The left column of *"AE_1"* displays the whole length of the recorded signal, as well as dependence of the computed individual S parameters based on high-order statistics. The time length of all signals is 80 μs. Right column of *"AE_1"* shows the time interval of 10 μs of the signal and relevant S parameters.

This time interval is indicated by a couple of dashed lines. As can be seen from left column of Fig.2, all parameters show their amplitude change, which coincide with visually observed arrival time of the signal. Higher slope of this dependence coincides with higher value of parameter S index. Parameter S_2 (standard deviation) also shows higher amplitude fluctuations before the visually observed signal arrival. The dependence of S_4 and S_6 parameters before the signal arrival has a nearly constant value with only low time fluctuations. This is due to the normal distribution of noise of the recorded signal. The recorded acoustic emission signal distorted this normal distribution, which resulted in a rapid change of S_4 and S_6 dependence. There were several attempts made to determine the time of their first arrival. Crossing over the threshold level was found to be very unreliable due to the low frequency fluctuations of the signal.

Because rapid changes of S_4 and S_6 parameters determine the signal time arrival, we decided to use the first derivative of both of them in our study. In the data presentation only S_6 derivative will be shown. Before taking the derivative, a moving average window of 3 points was used to smooth the S_6 parameter dependence. This smoothing eliminates noise fluctuations of the signal. After this, the STA/LTA method was used to determine the rapid change of S_6 derivative. The determined point coincides with a high accuracy with the time of arrival of the acoustic emission signals. More than several thousands of AE events recorded by multichannel acoustic emission system were tested.

Recording of the signals started when the root-mean-square (RMS) of the signal exceeded a multiple of 3 times the RMS of the background noise level.

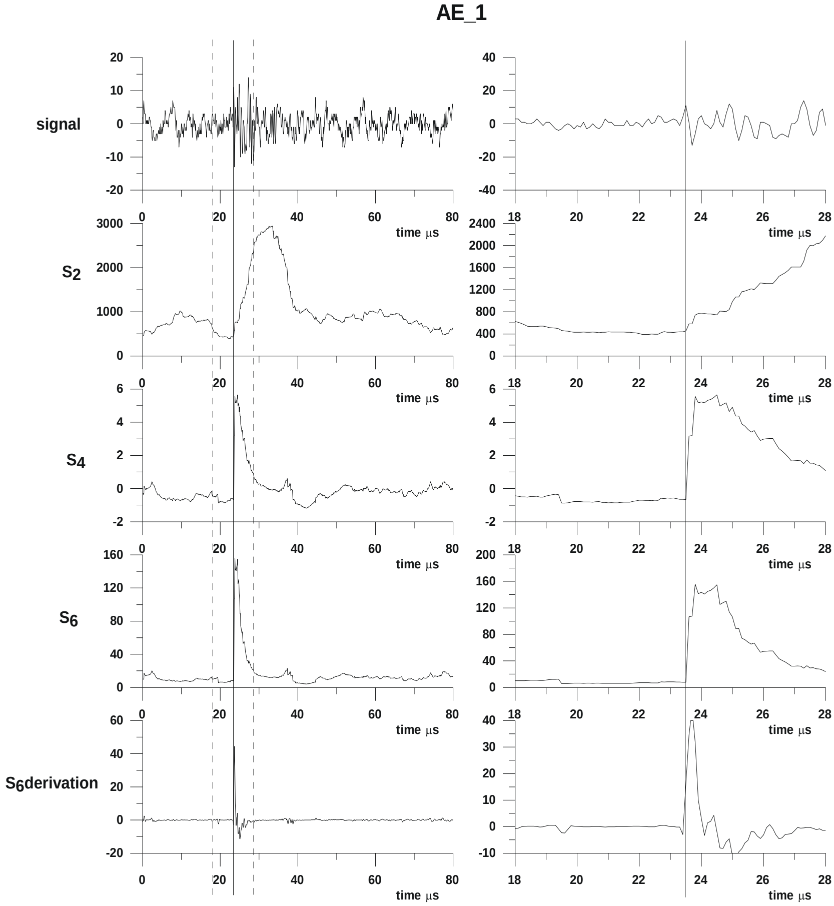

Fig.2 Signal "AE_1" is an example of weak signal. Left column shows the basic part of the signal and analysis results, their time duration is 80 µs. The dashed lines determine the time interval that is shown in the right column. The position of the solid line indicates the determined time arrival based on the S_6 derivative analysis.

"AE_2", see Figure 3, is an example of HOS based analysis, applied on common signals recorded by the system, where signal/noise ratio is more then 10. It means that the maximum amplitude of the signal is more then 10 times higher than the mean value of the recorded signal noise. In this case, there is a very well pronounced waveform of the recorded signal of acoustic emission.

This waveform clearly differs from the signal noise recorded before the AE signal arrival. Again, the S_2 parameter displays amplitude change, which coincides with the signal arrival time, but its slope is too small. Parameters S_4 and S_6 have a far more pronounced amplitude change, which agrees with the signal arrival. In this case, the ratio between the maximum values of S_6 and S_4 parameters is about 50. Again, the determined arrival time coincides very well with the reading seen by the naked eye. The staircase dependence of both S_4 and S_6 parameters could be explained by different arrivals of the signal due to reflections from the rock sample boundaries, or by the more complicated mechanism of the acoustic emission release. Also, in this case, the automatically determined arrival time coincides with time arrival manually determined by the operator.

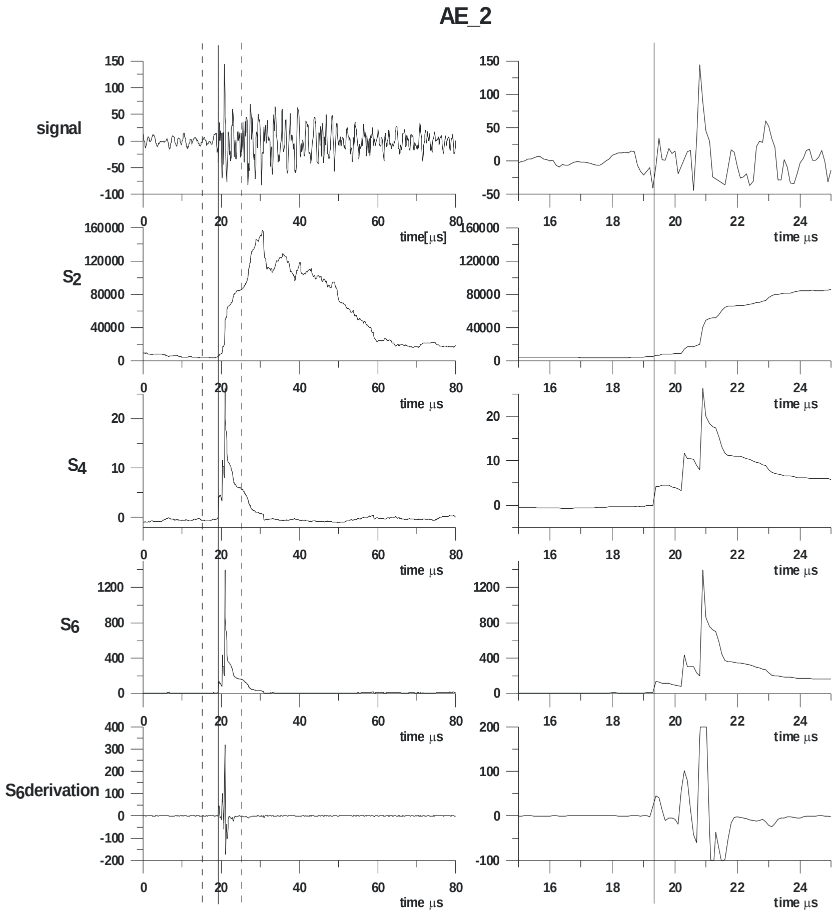

Fig.3 Signal "AE_2" is an example of common signal. The left column shows the basic part of the signal and analysis results, their time duration 80 μs. The dashed lines determine the time interval that is shown in the right column. The position of the solid line indicates the determined time arrival based on the S_6 derivative analysis.

Conclusions

It was proved that high order statistics (HOS) could be used for reliable and accurate determination of P-wave arrivals of the acoustic emission signals radiated from the stressed rock samples. HOS can be applied for the seismic/acoustic/ultrasonic signals in a wide energy range. This procedure seems to be efficient also for weak signals, where the amplitude of the signal is just above the signal noise. The theoretical analysis and experimental results showed higher sensitivity of the S_6 parameter with respect to parameter S_4.

HOS was successfully tested by Saragiotis et al. (2002) on real seismic data. Our results show that also HOS approach can be successfully used in determination of the first arrival time of AE data. This approach is applicable, when the recorded signal converts from random distribution to non-random one. But such an approach is not suitable for the determination of arrival time of multipath signals, since only the first arrival time can be determined, and times of following arrivals would be very probably hidden in the tail of previous signal.

Arrival times of several thousands of AE signals were analysed manually and automatically by high order statistic approach – by computation of S_6 parameter derivative. It was found that more than 95% events that were analyzed are within accuracy of ± 200 ns, which means within accuracy of ±2 sampling points. Such accuracy, for the signal sampled by 10 MHz, is quite sufficient. Due to the fact that in the processing of AE signals there is a need to process tens of thousands of individual events, this accuracy is sufficient. Due to the mathematical calculation procedure applied, S_6 computation is suitable for real-time implementation in acoustic/seismic determination of arrival time of the signal.

Influence of the high-order statistic approach for location accuracy of acoustic emission events will be subjected to subsequent analysis. Also the applicability of high order statistic approach for determination of the arrival times of the signals recorded by a resonant type transducer should be subjected to further analysis.

Acknowledgment

This study was partly supported by projects GACR 205/04/0088 and GACR 205/06/0906. Basic part of this contribution was done in the Institute of Rock Structure and Mechanics, Academy of Sciences of the Czech Republic.

References

[1] Anant K.S. & Dowla F.U. (1997): Wavelet Transform Methods for Phase Identification in Three-Component Seismograms. BSSA, 87, 1598-1612.

[2] Baer M. & Kradolfer U. (1987): An Automatic Phase Picker for Local and Teleseismic Events, BSSA, 77, 1437-1445.

[3] Jurkevic A. (1988): Polarization Analysis of Three Component Array Data, BSSA, 78, 1725-1743.

[4] NIST/SEMATECH e-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook

[5] Saragiotis Ch.D., Hadjileontiadis L.J. & Panas S.M. (2002) PAI-S/K: A Robust Automatic Seismic P Phase Arrival Identification Scheme, IEEE Transactions on Geoscience and Remote Sensing, 40, No. 6.

[6] Zhao Y. & Takano K. (1999): An Artificial Neural Network-Based Seismic Detector. BSSA, 89, 670-680.