Magyar Állami Eötvös Loránd Geofizikai Intézet
GEOFIZIKAI KÖZLEMÉNYEK
23

SEISMIC DATA PROCESSING USING A REDUCED NUMBER OF BITS

G. KORVIN*, I. PETROVICS*

Introduction

We present here what we believe to be the first systematic account on seismic data processing using a drastically reduced number of bits. After the formulation of the problem and an historical outline we use information theory to decide how many of the bits contained by a seismic trace is useful and actually needed. The problem of optimum quantizing and sampling rates is also touched upon. The paper is concluded by a description of a series of experiments that have recently been made to show that most of the routine steps of preliminary processing can be performed using very few bits—let alone the sign bits—of the seismic data.

Problem Discussion and Historical Outline

Scanning the seismic literature of the last few years there seems to be a difference of opinion among the authors about the dynamic range required for processing purposes. At the 8th World Petroleum Congress in Moscow, 1971, simultaneously with Savit and Mateker's "From Where? to What?", Polshkov mentions that seismic traces are almost uniquely determined by their extremal values. Two years after, it was a pleasant surprise of the 43rd SEG Meeting (Mexico City, 1973) and the 36th EAEG Meeting (Madrid, 1974) when—among the "bright spot" lectures—Savit announced that *the sign bit and four IFP bits* are sufficient to process Vibroseis data. We understand the 44th SEG Meeting in Dallas devoted special sections to direct CH detection and to sign bit techniques.

Digital seismic processing, as for its objectives and accuracy requirements, can be divided in two broad categories: *morphoseismics* and *lithoseismics*. The terminology originates from a paper of Leenhardt and Delserre (1974), a clear-cut distinction between the two trends has been made, of course, much before (Mateker, 1971, Savit and Mateker, 1971). Quoting Leenhardt and Delserre (op. cit.):

"...if by means of reflection, we can determine the shape of a reflector—it could be called *morphoseismics*".

"...as we have come to lithology by means of seismics, we propose to call this method *lithoseismics*"; i.e.: morphoseismics is concerned with the *geometry* of the reflecting boundaries, lithoseismics—beyond that—with the determination of *lithological characteristics.*

* Roland Eötvös Geophysical Institute
Manuscript received: 7, 2, 1976

The following are, among others, some of the basic tasks of lithoseismics:

a) Continuous, high-precision determination of interval velocities (as e.g. WESTERN's VELAN process, or the CONVEL of TELEDYNE Corp.).

b) Sand-shale ratio determination from interval velocities (TEGLAND, 1970)

c) Inference to overpressured zones from velocity anomalies (JANKOWSKY, 1971)

d) Estimation of density and elastic parameters from interval velocities (JANKOWSKY, 1971)

e) Classification of sedimentary rocks by a combined use of velocity and absorption coefficient (SAVIT and MATEKER, 1971; VOLAROVICH et al., 1969)

f) Finding stratigraphic traps on the basis of amplitude anomalies (SAVIT, 1960a, b; GAROTTA, 1971; BELYAYEVA et al., 1966; LYONS and DOBRIN, 1972; etc.)

g) Direct detection of hydrocarbons (DIEKMAN and WIERCZEYKO, 1970; CRAFT, 1973; SAVIT, 1973, 1974; LINDSEY, 1974; HILTERMAN, 1974; BACKUS and CHEN, 1974; STONE, 1974; QUARLES, 1973 etc.).

Most of the papers lay stress on the point that an increased dynamic range, both in the recording equipment and in the relative-amplitude-preserving processing system, is absolutely necessary to achieve the above goals. (CRAFT, 1973, e.g., speaks about floating point numbers ranging from 10^{-38} to 10^{38}!). We do not know of a single paper, however, which would have determined, at least approximately, how many bits are actually needed for lithoseismic tasks.

As for *morphoseismics,* it is generally known that the seismic data recorded by digital systems of large dynamic range are highly *redundant* for some of the classical processing purposes.

According to SAVIT's experiments (1973, 1974) for an explosive source 21-bit fixed point accuracy, for vibratory sources the sign bit and four IFP (or BGC) bits are sufficient for routine processing. Soviet authors claim (POLSHKOV et al., 1971, ZAHARCHENKO and KOROSHTISHEVSKY, 1973) that for all practical purposes the seismic traces can be represented by their local minima and maxima. There exist algorithms for reflection picking using only the times of peaks (PAULSON and MERDLER, 1968) and for automated velocity analysis starting out from zero-crossing times (BARR, 1971). It is a general practice that in many cases seismic processing can be successfully made with 4 or 8 msec sampling rates. Further data compression might be achieved by the RADEMACHER, WALSH, PALEY transforms (BOIS, 1974; WOOD, 1974).

Maybe the most striking way to reduce the redundancy of seismic records is to substitute each data by its polarity. The idea first occurred in a paper of MELTON and KARR (as early as 1957) who proposed polarity coincidence schemes for signal detection. (For theoretical background and further references see CARLYLE, 1968). Following up this idea we prepared a polarity-coincidence version of the Velocity Spectra program in 1972, a similar solution was published by COCHRAN (1973). An interesting point of COCHRAN's paper is his proposal about a possible hardware implementation of his algorithm.

It has been known since long in communication engineering that polarity coincidence correlation can be used for the estimation of power spectra and to reveal hidden periodicities of random processes (VAN VLECK, 1943, 1966; OSSENBERG, 1968). FARA and SCHEIDEGGER (1961) proposed similar techniques for the statistical description of porous media, CARRS and NEIDELL (1966) reports on an interesting geological cyclicity detected by means of polarity coincidence correlation. According to Bortfeld and Ristow (1969, personal communication) correlating field data with the sign bits of the Vibroseis sweep should be enough for a preliminary check or field-display.

Finally, it should be mentioned that while the dynamic range of seismic plotters used nowadays is some 7–10 bits, plotting as few as 1 bit could be surprisingly effective for certain purposes, as e.g. zero crossings in shallow investigations (Meidav, 1969) or the display of sign bits of a common-offset section.

Information Contained in a Seismic Trace

Let us suppose that on a magnetic tape or in the computer's memory each seismic data is represented by m bits, a seismic trace consists of $N+1$ data, its spectrum contains no frequencies above a certain limit f_{max}, the trace extends over a time interval T and it is sampled at a rate Δt. In case of a proper choice of the sampling rate

$$2 \cdot f_{max} \cdot \Delta t = 1\,; \qquad T = N \cdot \Delta t\,; \qquad N = T \cdot 2 f_{max}\,. \tag{1}$$

In the ideal case when all sampled data are independent of each other, the information content of the trace is maximal:

$$I = I_{max} = (N+1)\, m \text{ bits} \approx 2 \cdot f_{max} \cdot T \cdot m \text{ bits}. \tag{2}$$

The information of actual seismic traces is of course always less than I_{max}, for successive data are correlated and the spectrum is not white in $[0, f_{max}]$.

Suppose, e.g. that instead of a white spectrum we are given a trace which has a spectrum extending from 0 to $f_1 < f_{max}$ and that there is a resonance-like peak in the spectrum at some $f_0 < f_1$. The number of degrees of freedom of the trace is (cf. Brillouin, 1956, Eq. 8.51)

$$M = 2 f_1 T + 1$$

i.e. the amount of information is certainly less than

$$I = (2 f_1 T + 1)\, m \approx 2 f_1 T m < I_{max}\,. \tag{3}$$

Since the spectrum is not white in $[0, f_0]$, the actual information is even less. Let $f_0 \ll f_m$ (e.g. $f_0 = 30$ Hz, $f_m = 250$ Hz), let us suppose that the data are, together with their signs, m-bit fixed-point words, then, because of the resonance peak at f_0, the difference between successive values is less than

$$\max_t \left| 2^{m-1} \sin 2\pi f_0 t - 2^{m-1} \sin 2\pi f_0 (t + \Delta t) \right| \le$$

$$\le \max_t\, 2^{m-1} \Delta t \cdot 2\pi f_0 \left| \cos 2\pi f_0 t \right| = 2^{m-1} \Delta t \cdot 2\pi f_0 = \pi 2^{m-1} \frac{f_0}{f_{max}}\,. \tag{4}$$

Introducing the notation

$$\delta = \log_2 \left(\pi 2^m \frac{f_0}{2 f_{max}} \right); \tag{5}$$

Eq. (4) implies that the difference between two successive values cannot exceed δ in binary digits. Let us determine the number of all possible traces. The first value can be chosen in 2^m different ways, for the second we have $2 \cdot 2^\delta = 2^{\delta+1}$ possibilities, etc. Altogether, the number of possibilities is

$$N_{\text{possible}} = 2^m (2^{\delta+1})^N \tag{6}$$

i.e. the information in bits of a trace can be at most

$$I = \log_2 N_{\text{possible}} = m + (\delta + 1) N.$$

Combining Eqs. (2) and (5)

$$I = m + (\delta + 1) N = m + \left(m - \log_2 \frac{2f_{\max}}{\pi f_0} \right) N = I_{\max} - N \log_2 \frac{2f_{\max}}{\pi f_0} \tag{7}$$

that is appr. $\log_2 \frac{2f_{\max}}{\pi f_0}$ bits/datum is redundant. (In case of, say, $f_{\max} = 250$ Hz, $f_0 = 30$ Hz, the redundancy is about 2 bits/datum.)

As an interesting application of Eq. (6) let us consider the usual seismic word format, when data are represented by means of mantissas of m_1 binary digits and four BGC bits, and suppose the gain cannot step more than 6 dB (i.e. 1 bit) during a sampling cycle. The number of bits used for a single data is $m = m_1 + 4$, that is the upper bound for the information in a trace is

$$I_{\max} = (N + 1) m = (N + 1) m_1 + (N + 1) 4.$$

Because of the restriction on the gain steps the possible number of traces differing in BGC values is

$$N_{\text{possible}} = 2^4 \cdot 2^{2N}$$

implying that about 2 bits/word should be enough to specify the BGC values—exactly as recommended by SEG for the 9 track tape format (Northwood et al., 1967).

If we accept the view of the Soviet school that seismic traces can be characterized by the extremal values (Zaharchenko and Koroshtishevsky, 1973) we get a further estimation of the redundancy of seismic data. Denoting the dominant frequency by f_0, and assuming a Poisson distribution of the times corresponding to peaks and troughs, the expected number of extrema will be $2N \cdot \Delta t \cdot f_0$. Representing extremal values by m bits each and their respective times by an array of altogether N bits, the information to be expected is

$$I = 2N \, \Delta t \, f_0 m + N = N \left(\frac{f_0}{f_{\max}} m + 1 \right) \ll I_{\max} \tag{8}$$

How many bits should a seismic trace contain?

Since computers cannot manufacture new information, a seismic trace, or an assembly of traces should contain at least that amount of information we hope to gain from their interpretation.

Let us address ourselves to the determination of interval velocities. According to some authors (KUNETZ, 1963; CLAERBOUT, 1968; EISNER, 1970; LINDSETH, 1972, etc.) there is a possibility to determine interval velocities (by a proper deconvolution and some kind of inversion of the computation of synthetic seismograms) from single least-offset traces. From the physical point of view, of course, it cannot be expected that layers thinner than a certain fraction of the dominant wave-length could be traced. Denoting the dominant frequency by f_0, the mean value of velocities by c, and adopting the optimistic view of CRAFT (1973) that the limit of resolution is 1/12th of the wavelength the thickness of the thinnest detectable layer is $c/12f_0$. If $T = N \cdot \Delta t$ is the total two-way travel-time, the corresponding depth range is

$$Z = \frac{cN \cdot \Delta t}{2}$$

and the number of degrees of freedom of the velocity distribution is

$$N^* \approx 2Z \cdot \frac{12f_0}{c} = cN\,\Delta t\,\frac{12f_0}{c} = 12N\,\Delta t\,f_0\,.$$

If N_v denotes the number of the different values of the velocities to be determined, the information we gain from the determination of velocities is

$$I_v = 12N \cdot \Delta t\, f_0 \log_2 N_v = 6N \left(\frac{f_0}{f_{\max}}\right) \log_2 N_v \tag{9}$$

in binary digits.

Estimating the amount of information contained in the trace by Eq. (8), it turns out as a necessary condition for the velocity determination that

$$I = N\left(\frac{f_0}{f_{\max}}\, m + 1\right) > N\,\frac{f_0}{f_{\max}}\, m \geq 6\,\frac{f_0}{f_{\max}}\, N \log_2 N_v$$

i.e.

$$m \geq 6 \cdot \log_2 N_v$$

should hold, where m is the number of bits used for representing seismic data.

The above inequality implies that a high-resolution, accurate velocity analysis requires a rather large dynamic range.

Contenting ourselves with detecting layers of thickness $\lambda/2$, the information of the velocity distribution is

$$I_v = 2N\,\Delta t\, f_0 \log_2 N_v = N\,\frac{f_0}{f_{\max}} \cdot \log_2 N_v \tag{10}$$

i.e. the number of bits should meet the less severe requirement:

$$m \geq \log_2 N_v\,. \tag{11}$$

If, e.g., we want to approximate the velocity distribution by 100 different values of interval velocities and by layers of thickness $\lambda/2$ the seismic data must be quantized by 6–7 bits. This means, that in case of a sixfold coverage there is theoretically nothing against to obtain a reasonable guess of the velocity function using as few as one bit for each data of a CDP gather.

As a matter of fact there are at least two successful experiments reported which prove that this can really be done, that of Cochran (1974) using sign bits, and of Barr (1971) who made use of zero-crossing times.

It should be noted that the amount of information of the velocity distribution in a sedimentary series is generally less than that given by Eqs. (9) and (10), for there is a correlation between the velocities of subsequent layers (Kats et al., 1969; Korvin, 1973).

It would be a difficult venture to guess along similar lines the number of bits needed to meet some other lithoseismic tasks. As for the absorption coefficient, it should be brought up that in case of a linear attenuation mechanism it could be computed from the fall-off of dominant frequencies (Bereznev and Malovichko, 1972; Huang Jen-Hu, 1961), and this latter can be estimated by simply counting zero-crossings or by Van Vleck's polarity correlation method (Carrs and Neidell, 1966). In certain cases, however, it has been found (Petrovics et al., 1975) that the main factor governing the attenuation of seismic waves is not linear in frequency and it is necessary to compute power spectra to separate absorption from scattering (Rapoport, 1969). In case of sufficiently long time-windows the power spectrum can be estimated from polarity coincidence correlation (Van Vleck, 1943, 1966) or from the auto-relay correlation (Ossenberg, 1968; Nuttal, 1958), in case of short time-windows, however, which are more frequent in CH detection problems, we must have a large dynamic range for a reasonable estimation of high frequencies.

Optimum Sampling and Quantizing: Number of Bits and Dynamic Range

In the previous section the number of bits submitted to seismic processing was considered from the point of view of geological information. We wish to turn now to the other facet of the problem: to the dynamic range of seismic signals and recording systems.

If we carry out seismic measurements in some time interval $[t_1, t_2]$, and denote the peak amplitude of the largest signal by η_1, that of the smallest one by η_2, the dynamic range of the seismic vibrations is defined as

$$D_A = 20 \lg \frac{\eta_1}{\eta_2} \text{ dB.} \tag{12}$$

Dynamic range can also be defined by means of the spectra of η_1 and η_2:

$$D(f) = 20 \lg \frac{|S_1(f)|}{|S_2(f)|} \text{ dB} \tag{13}$$

where S_1, S_2 are spectra of signals η_1 and η_2, respectively. The dynamic range defined by Eq. (13) is frequency-dependent, for practical purposes we can take the maximum of expression (13) over the useful range.

To find the amplitude and spectrum of the greatest signal we ought to know the source characteristics and the material properties of the near-surface layers. As for the smallest signals, their amplitudes and spectra can be estimated by taking into account spherical divergence, reflection losses and absorption (Born, 1941;

Posgay et al., 1971; Gurvich, 1973; Sheriff, 1973; Byakov and Ryazanova, 1974), for this task, however, we need velocity and absorption data. In the present state of the art maybe the best way to define D_A is to take the greatest amplitude which can be recorded for η_1, and to elect for η_2 the smallest signal which can be detected by available techniques from below the ambient ground noise level.

The usual definition of the dynamic range D_I of the seismic recording instrument is

$$D_I = 20 \lg \frac{A_{\max}}{\bar{A}_N} \text{ dB} \tag{14}$$

where $A_{\max}$ is peak amplitude of the greatest signal which can be recorded without distortion, and $\bar{A}_N$ is the RMS amplitude of instrument noise.

If seismic recording is made in digital format allotting m bits for each data, the dynamic range of registration is

$$D_d = 20 \log 2^{m-1} = 6(m-1) \text{ dB}. \tag{15}$$

It is advisable to choose the dynamic range of registration somewhat larger than that of the recording instrument in order that signals smaller than instrument noise should be detectable in course of computer processing (according to Gurvich, 1973, 6–7 bits should be devoted for such purposes).

In case of a properly chosen recording equipment, converter and digital representation

$$D_A \leq D_I \leq D_d \tag{16}$$

i.e. by Eqs. (12) and (15):

$$m - 1 > \frac{D_A}{6} \tag{17}$$

should hold for the number of bits required for digital registration.

What are the main causes of the paradox between the large number of bits justified by dynamic range and the redundancy of seismic data series experienced in data processing?

1. For a safe data-transmission and/or storage a certain degree of redundancy is by all means necessary. (It is a reasonable specification, e.g., of the SEG *A* format of nine track tapes that as a redundant check of the *U* and *G* bits the actual binary gain control value should appear at certain times.)

2. The results of seismic data processing are usually displayed on plotters of much less dynamic range.

3. In a preliminary stage of processing (as e.g. straight stack) we are not always interested in the shape of the waves reflected from a horizon, only in their *arrival*, i.e. a *simultaneous occurrence* of signals on a number of CDP traces. In such cases signal detection can be performed by means of polarity coincidence methods (Carlyle, 1968).

4. When estimating the dynamic range by Eqs. (12) or (13) the role of ambient earth noise and instrument noise is sometimes overlooked i.e. the number of bits given by Eq. (17) is too high; or we do not have the necessary techniques to enhance signals below noise level.

5. In the definition of dynamic range by means of Eq. (12) it hasn't been taken into account that this range might be different in different time gates, or [when using Eq. (13)] in different frequency bands, i.e. digital recording with a constant number of bits all the time is certainly not optimal.

To clarify point 5. we review in some details *the problem of optimum sampling and quantizing* of analog signals (cf e.g. Brillouin, 1956; Goodman, 1966). The problem of sampling and quantizing can be briefly stated this way: The analog signal, $x(t)$ is passed through an ideal low-pass filter of cut-off frequency W, the output $y(t)$ is sampled at a rate $1/2W$ and each sample y_i is mapped into one of the set of n numbers $\hat{y}_1, \hat{y}_2, \ldots \hat{y}_n$. The numbers $\hat{y}_i$ are then encoded into $\log_2 n$ binary digits and are transmitted to a remote computer or to a magnetic tape over a channel of capacity of R bits/sec. On the receiver side the message is decoded and by means of digital or analog techniques we construct some approximation $\hat{x}(t)$ of the signal $x(t)$. Since the rate of information transfer (or storage) R bits/sec is fixed the task of optimum sampling and quantizing consists in finding such values of the sampling rate $1/2W$, of the number of steps n, and of $\hat{y}_1, y_2, \ldots \hat{y}_n$ that the expected error

$$E[(x(t) - \hat{x}(t))^2] \tag{18}$$

should be minimal.

Leaving out of consideration the optimization of $1/2W$, we discuss first the optimal choice of the quantizing steps $\hat{y}_1, \hat{y}_2, \ldots \hat{y}_n$. Suppose that x is a random variable uniformly distributed in [0, 1]. For an arbitrary subdivision:

$$0 = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = 1 \tag{19}$$

of interval [0, 1] we define a quantization

$$x = q_k \quad \text{if} \quad x_{k-1} \leq x < x_k \tag{20}$$

where $q_1, \ldots q_n$ are arbitrary numbers. The expected error of this quantization is

$$E = \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} (x - q_k)^2 \, dx = \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} \left(x - \frac{x_{k-1} + x_k}{2} \right)^2 dx + \sum_{k=1}^{n} (x_k - x_{k-1}) \left(\frac{x_{k-1} + x_k}{2} - q_k \right)^2.$$

The optimal choice of the q_k-s is evidently

$$q_k = \frac{x_{k-1} + x_k}{2}$$

the corresponding error being

$$E = \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} \left(x - \frac{x_{k-1} + x_k}{2} \right)^2 dx = \frac{1}{12} \sum_{k=1}^{n} (x_k - x_{k-1})^3. \tag{21}$$

Introducing auxiliary variables

$$\mu_k = x_k - x_{k-1} \qquad k = 1, 2, \ldots, n$$

we are faced with the conditional extremal value problem

$$\mu_1^3 + \mu_2^3 + \ldots + \mu_n^3 = \min$$

$$\mu_1 + \mu_2 + \ldots + \mu_n = 1.$$

Introducing a Lagrange parameter, the optimal solution is easily found:

$$\mu_1 = \mu_2 = \ldots = \mu_n = \frac{1}{n} \tag{22}$$

that is, the optimal quantization is given by the equidistant subdivision of [0, 1]:

$$q_k = \frac{2k-1}{2n} \quad \text{if} \quad \frac{k-1}{n} \leq x < \frac{k}{n}. \tag{23}$$

The expected error is, according to Eqs. (21) and (22)

$$E = \frac{1}{12} \sum_{k=1}^{n} (x_k - x_{k-1})^3 = \frac{1}{12} \frac{1}{n^2}. \tag{24}$$

We proceed now to the question of optimum quantization of a seismic trace digitized at the rate Δt. Suppose that the amplitudes have an exponentially decreasing envelope according to the law

$$e^{-at} = e^{-an\,\Delta t}. \tag{25}$$

For computational convenience we suppose the rate of information transfer is R nat/sec (1 nat = 1.442695 bit). Further, we suppose each of the first N samples is quantized by n_1 steps, the second N samples by n_2 steps each, and so on.

The errors in consecutive time intervals are, in turn:

$$E_1 = \frac{N}{12} \frac{1}{n_1^2}$$

$$E_2 = \frac{N}{12} \frac{1}{n_2^2} e^{-2aN\,\Delta t}$$

$$E_3 = \frac{N}{12} \frac{1}{n_3^2} e^{-4aN\,\Delta t}$$

$$\vdots$$

i.e. in altogether M successive time intervals the total error is

$$E = \frac{N}{12} e^{2aN\,\Delta t} \sum_{k=1}^{M} \frac{1}{n_k^2} e^{-2aNk\,\Delta t}.$$

Since transmission of a data quantized into n_i steps requires the transfer of $\ln n_i$ nats, in case of M time intervals we have to transfer

$$N \sum_{k=1}^{M} \ln n_k$$

nats during a time of $M \cdot N \cdot \Delta t$ sec, i.e.

$$\frac{N \sum_{k=1}^{M} \ln n_k}{NM\,\Delta t} = R \quad \text{nat/sec.}$$

In order to select optimally the number of quantization steps n_i we have to minimize the expression

$$\sum_{k=1}^{M} \frac{1}{n_k^2} e^{-2aNk\,\Delta t}$$

subject to the condition (26)

$$\sum_{k=1}^{M} \ln n_k = MR\,\Delta t = Q.$$

Introducing the notation $\gamma_k = \ln n_k$, Problem (26) simplifies to

$$\left.\begin{aligned} &\sum_{k=1}^{M} e^{-2\gamma_k - 2aNk\,\Delta t} = \min \\ &\sum_{k=1}^{M} \gamma_k = Q. \end{aligned}\right\} \tag{27}$$

Solving this conditional extremal value problem by standard methods of calculus, the optimum quantization is found to be:

$$\gamma_k = R \cdot \Delta t + \alpha N\,\Delta t \left(\frac{M+1}{2} - k\right) (k = 0,\ 1,\ \ldots,\ M) \tag{28}$$

Equation (28) has a clear-cut physical meaning: In case of a fixed sampling rate Δt and a (noiseless) transfer possibility of capacity R nat/sec the *optimum quantization of an analog message of exponentially decreasing envelope allocates exponentially decreasing numbers of binary digits to data belonging to successive time intervals*, i.e. the less the dynamic range the less the number of bits.

The problem discussed above has not been solved in the literature what we know of. Similar investigations are reported by Goblick (1965), he assumes however instead of the exponentially decreasing envelope a spectrum of that type.

Practical Examples for Seismic Processing Using a Reduced Number of Bits

All experimental processing reported have been performed on the MINSK-32 computer of the ELGI, with the seismic package developed in our Institute. The computer applied has 64 k words capacity, fixed and floating-point arithmetics, 37 bits word-length. Routine seismic processing is generally performed on fixed-point data normalized to 18 bits. The seismic materials used were explosive generated, multifold coverage sections recorded by the Hungarian SDT-1 and SDT-2 and the Hungarian–GDR SD-10 digital recording equipments. Displays were made on a plotter of 7 bit dynamic range.

One of our earliest experiments with sign bit techniques was in connection with the auto- and retrocorrelation sections (Anstey and Newman, 1966).

To save computer time, we use instead of these functions auto-relay correlation functions, and retro-relay correlation functions, which are computed by substituting the values of the lagged versions of the functions by their signs.

For instance, as compared with the usual definition of autocorrelation function

$$R_{xx}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T x(t)\, x(t-\tau)\, dt \tag{29}$$

the auto-relay correlation is defined as

$$R_{x\,\mathrm{sgn}\,x} = \lim_{T\to\infty} \frac{1}{T} \int_0^T x(t)\ \mathrm{sgn}\, x(t-\tau)\, dt, \tag{30}$$

where

$$\mathrm{sgn}\, x = \begin{cases} +1 & \text{if} \quad x > 0 \\ 0 & \text{if} \quad x = 0 \\ -1 & \text{if} \quad x < 0. \end{cases} \tag{31}$$

The evident advantage of Eq. (30) is that the multiplications occurring in the cross-products reduce to additions and subtractions.

Comparison of an exact auto- and retrocorrelation section with the corresponding auto-relay, retro-relay sections is shown in Fig. 1, for a detailed survey of auto-relay correlation we refer to the Ph.D. thesis of Ossenberg (1968).

Since auto-relay and retro-relay sections obviously meet the requirements of routine processing, we have tried to find out which other processes might be successful with a reduced number of bits, let alone sign bits.

Data compression was performed this way: We substituted the original 14–18 bit fixed point data by a three-valued step function, according to either of the transformations

$$x \longrightarrow c \cdot \mathrm{sgn}\, x = \begin{cases} +c & \text{if} \quad x > 0 \\ 0 & \text{if} \quad x = 0 \\ -c & \text{if} \quad x < 0 \end{cases} \tag{32}$$

or

$$x \longrightarrow c[u(x-A) - u(-x-A)] = \begin{cases} +c & \text{if} \quad x \geq A \\ 0 & \text{if} \quad -A < x < A \\ -c & \text{if} \quad x \leq -A \end{cases} \tag{33}$$

where in Eq. (33) $u(x)$ is the Heaviside unit step function, c was chosen as 2^{11}, while A can be made an arbitrary power of 2. The nonlinear transformations (32) and (33) will be referred to in what follows as *sign-transform* (signing) and *combined signing and clipping at level A* (Fig. 2).

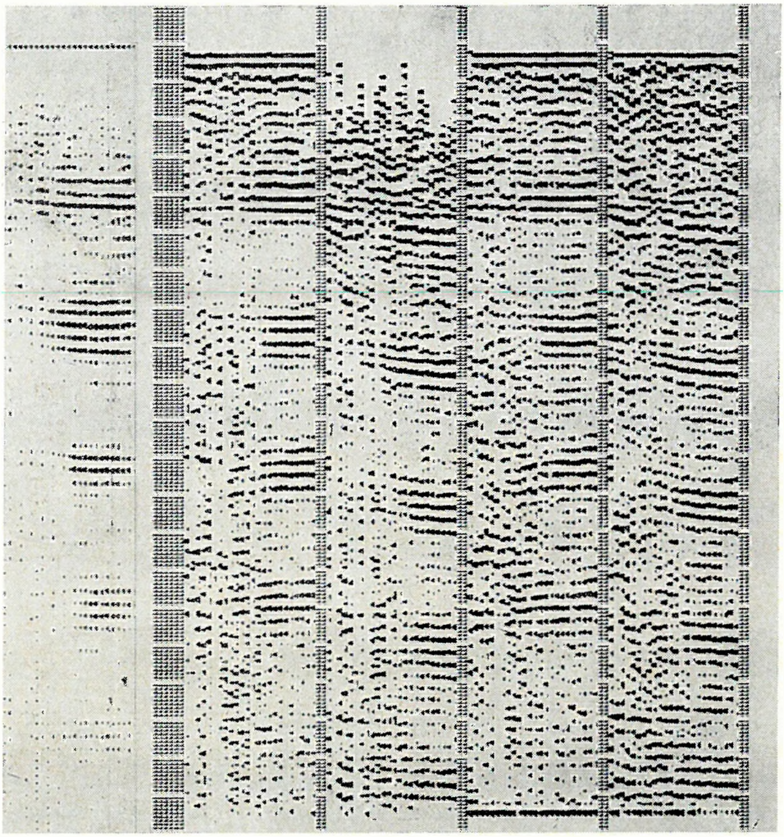

Fig. 1 Comparison of auto- and retrocorrelation sections with auto-relay resp. retro-relay sections. (From left to right: original, autocorrelation, retrocorrelation, auto-relay, retro-relay)

1. ábra. Autokorrelációs, ill. retrokorrelációs szelvény összehasonlítása az auto-relais, ill. retro-relais eredményekkel (eredeti, autokorreláció, retrokorreláció, auto-relais, retro-relais)

Рис. 1. Сравнение автокорреляционного и ретрокорреляционного разреза с результатами авто-реле и ретро-реле

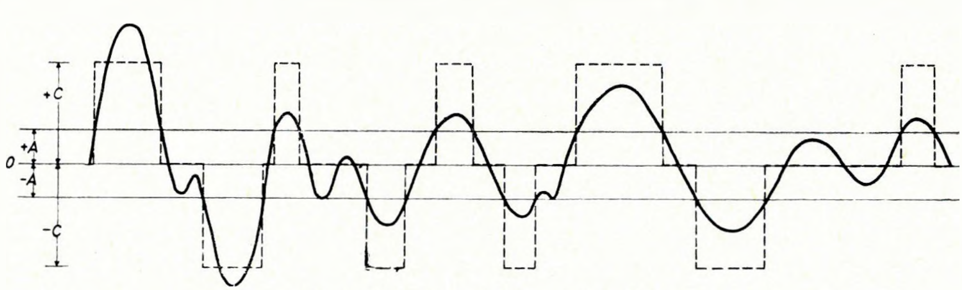

Fig. 2 Principle of data compression

2. ábra. Az adatkompresszió szemléltetése

Рис. 2. Пояснение принципа компрессии данных

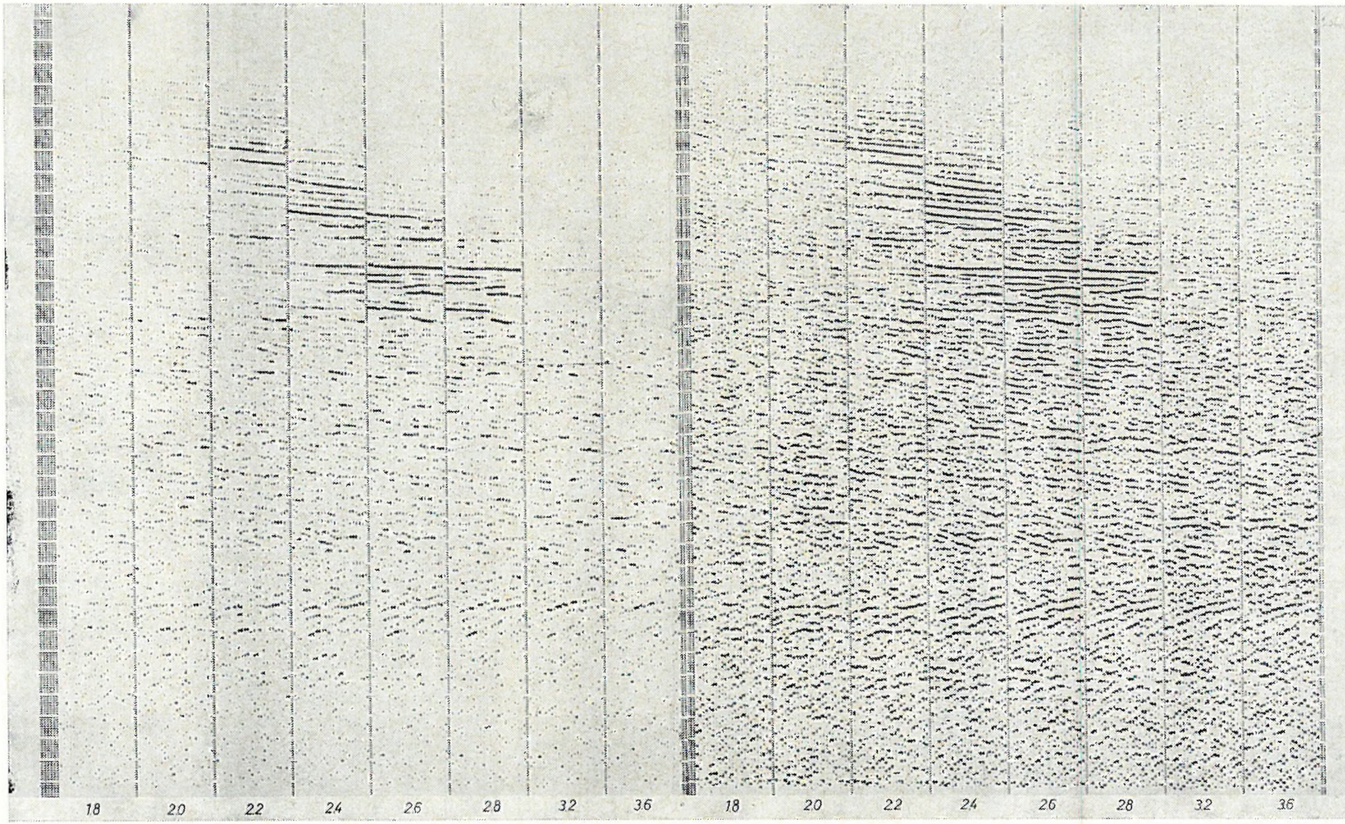

Fig. 3 Comparison of Velocity Scans performed on full dynamic range data and on signed data

3. ábra. Teljes bitszámú anyagon és „előjelezett" anyagon végrehajtott velocity scan eredmények összehasonlítása

Рис. 3. Сравнение результатов сканирования скоростей, полученных по материалу с полным числом разрядов и по «массиву знаков»

In Figure 3 we compare Constant Velocity Scans performed on full-range data and on data reduced to signs. On the signed version the deeper reflections appear more clearly, their interpretation is easier. This is possibly due to the fact that signs are invariant to amplitude changes along an individual trace and between CDP traces.

Another version of velocity analysis (Velocity Spectra) based on sign information is found in COCHRAN (1973).

The examples presented on Figs. 4 and 5 show that a surprisingly small amount

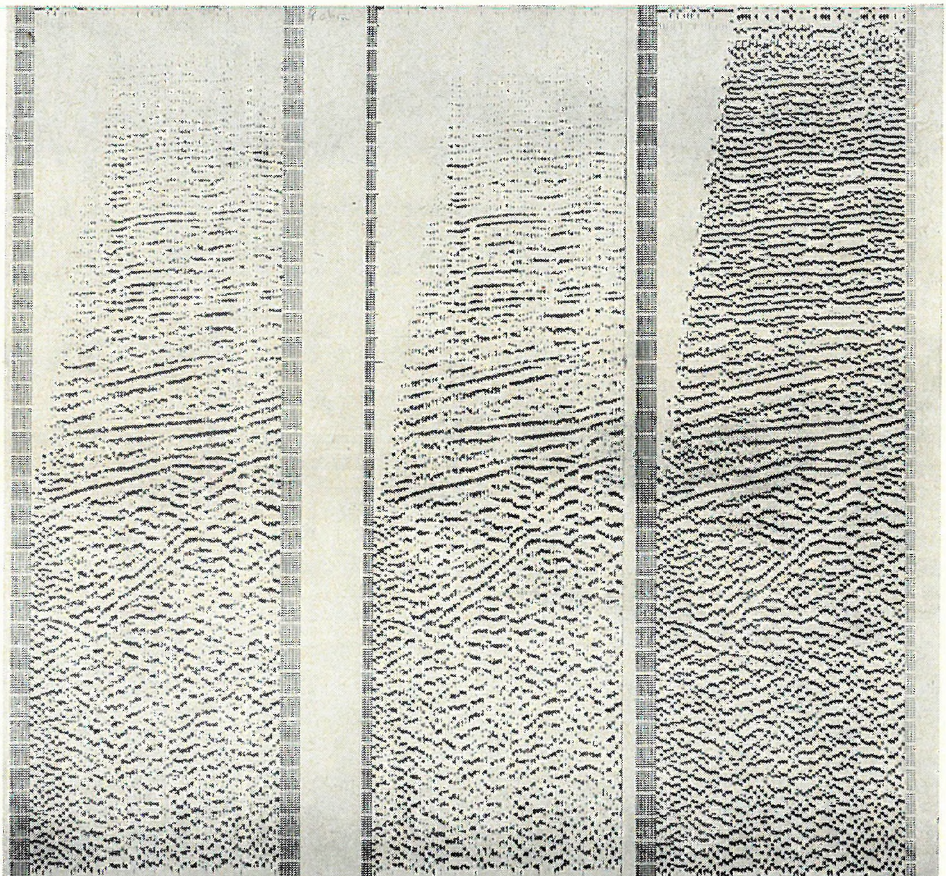

Fig. 4 From left to right: Complete processing of compressed data; Data compression before stack; Data compression after stack, before playback

4. ábra. Balról jobbra: Teljes feldolgozás az adatkompresszió után; Adatkompresszió az összegezést megelőző fázisban; Adatkompresszió a végső eredményen, kiírás előtt

Рис. 4. Полная обработка после компрессии данных. Компрессия данных перед суммированием. Компрессия данных окончательного результата перед их графическим изображением (if any) of the information is lost if we use only sign bits throughout a routine processing.

As Fig. 5 shows even linear migration works successfully on signed data. It sholud be noted the sections presented had been displayed in wide band (0–74 Hz) without any digital or analog smoothing or low-pass filtering.

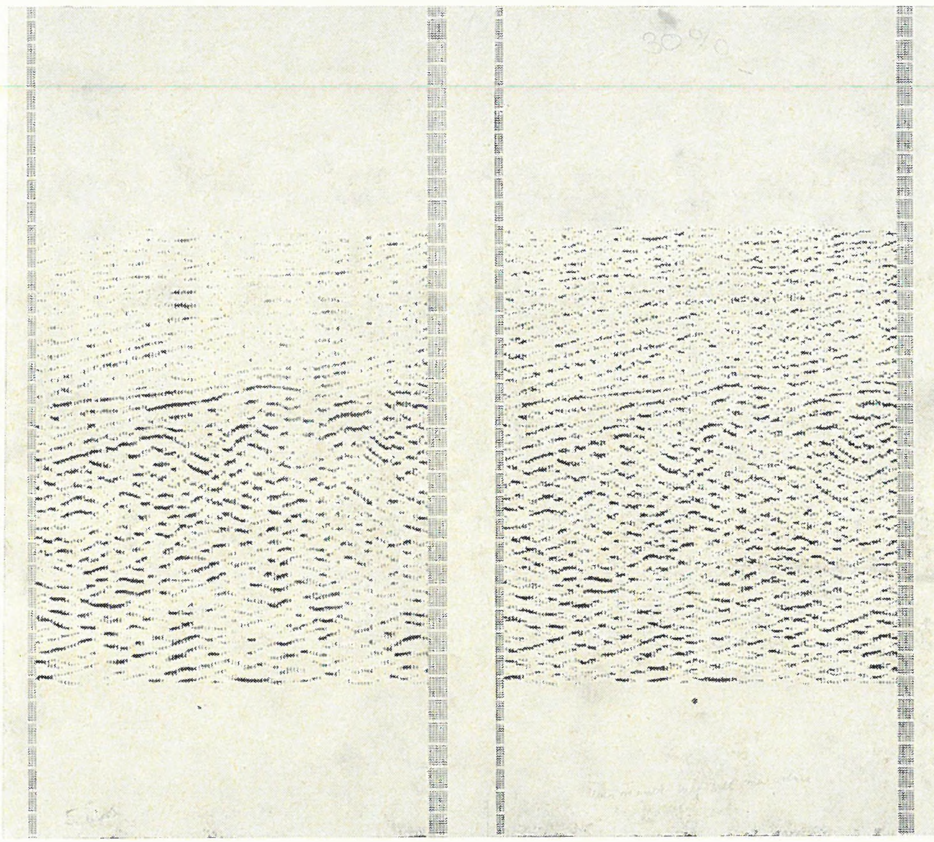

Fig. 5 Migration of full-range data, and of compressed data

5. ábra. Migráció: teljes bitszámú anyagon és adatkompresszió utáni anyagon

Рис. 5. Миграция: на массиве с полным количеством разрядов и на массиве после компрессии

Figure 6 demonstrates the possibility of deconvolution of signed materials. Starting out from sign-transformed traces (*1*) we computed first a retro-correlation section (*2*). Then we executed spike deconvolution on the signed data (*3*), passed the output through a digital band-pass filter (*5*) and at the same time computed

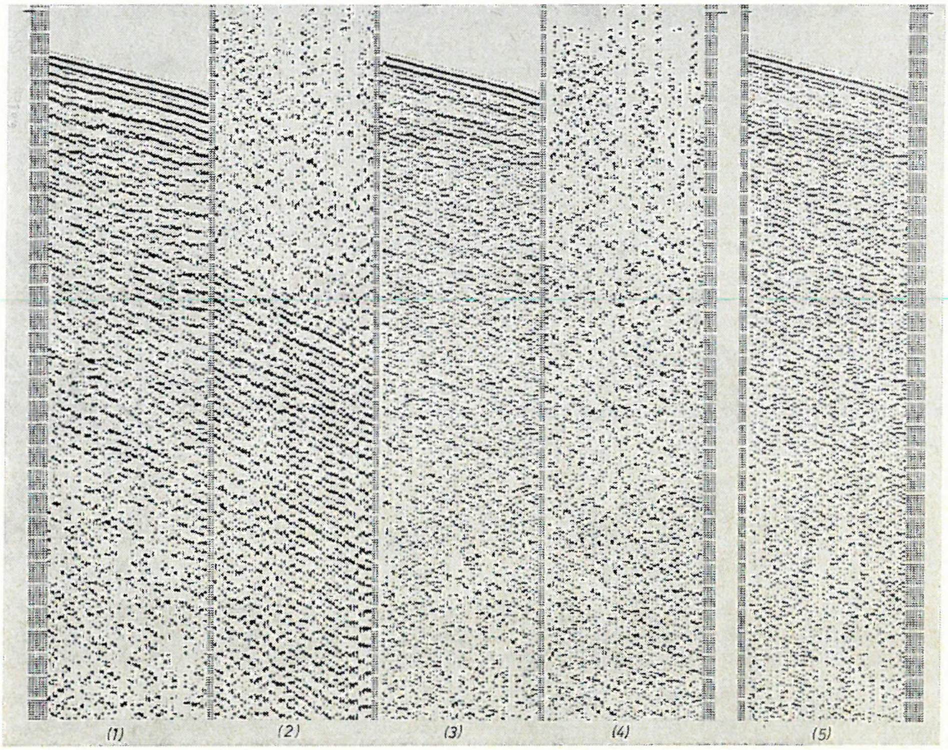

Fig. 6 Retrocorrelation and deconvolution experiments on the signs of data

6. ábra. Retrokorrelációs és dekonvolúciós vizsgálatok „előjelezett" anyagon

Рис. 6. Ретрокорреляционные и деконволюционные исследования на «массив знаков»

the retrocorrelation of the deconvolved data (*4*) to judge the effectiveness of deconvolution.

The following figures should serve to call attention to the noise-reducing effect of signing combined with clipping. On the deep part of the 600%-coverage, conventionally processed section of Fig. 7 (consisting of 18 bit data) short-period multiples and poor signal-to-noise ratio made impossible to trace reflection segments or even diffractions.

To improve signal-to-noise ratio the section was first spike-deconvolved and normalized to 14 bits (Fig. 8) then we performed a "clipping analysis" to find the most appropriate level A (Fig. 9). Cutting off all values in absolute value less than $A = 2^{10}$, substituting the rest by their polarity, and performing digital filtering, the resulting section (Fig. 10) was much more apt to interpretation.

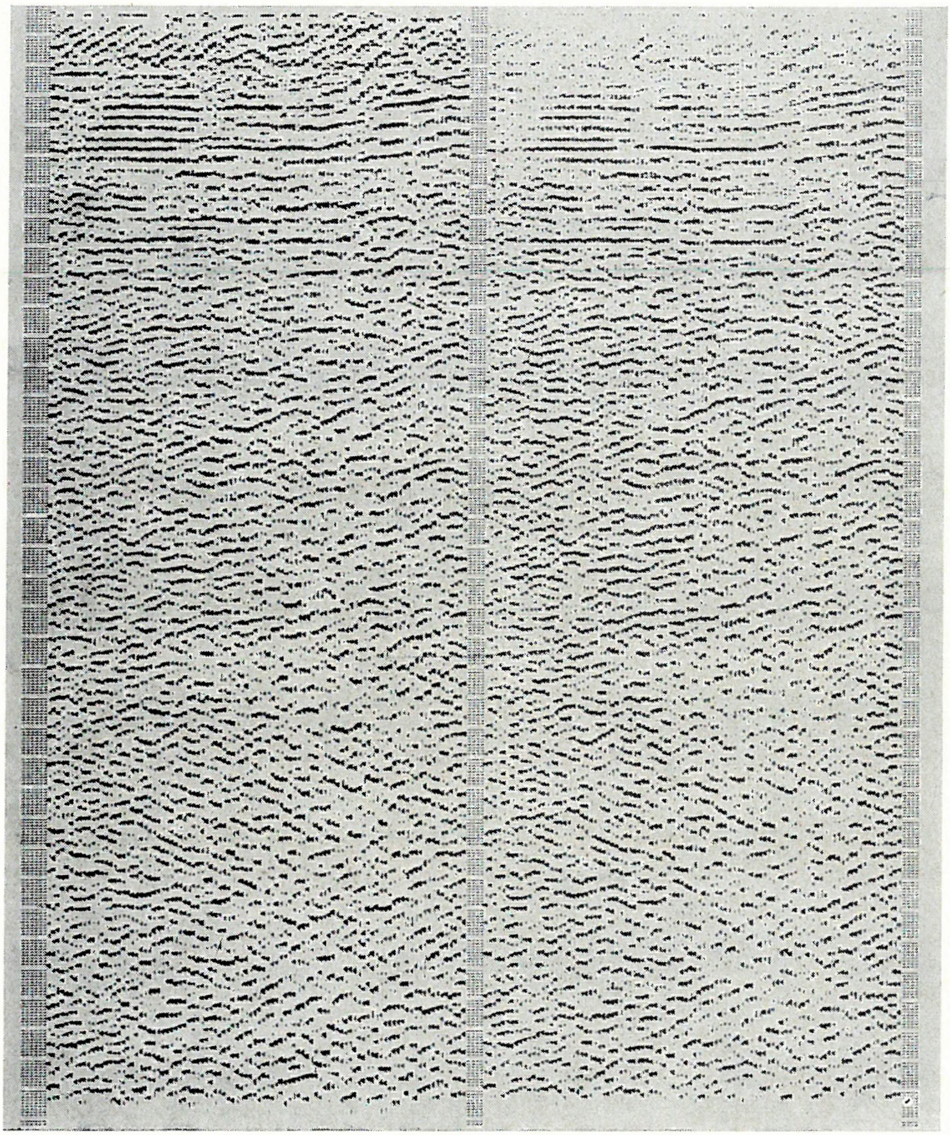

Fig. 7 600%-coverage conventional section played back in wide band (0/74 Hz) with and without AGC

7. ábra. Stacking szelvény szélessávú (0/74 Hz) AGC-s és AGC nélküli kiírása

Рис. 7. Широкополосное (0—74 гц) графическое изображение профиля с шестикратным перекрытием данных с АРУ и без АРУ

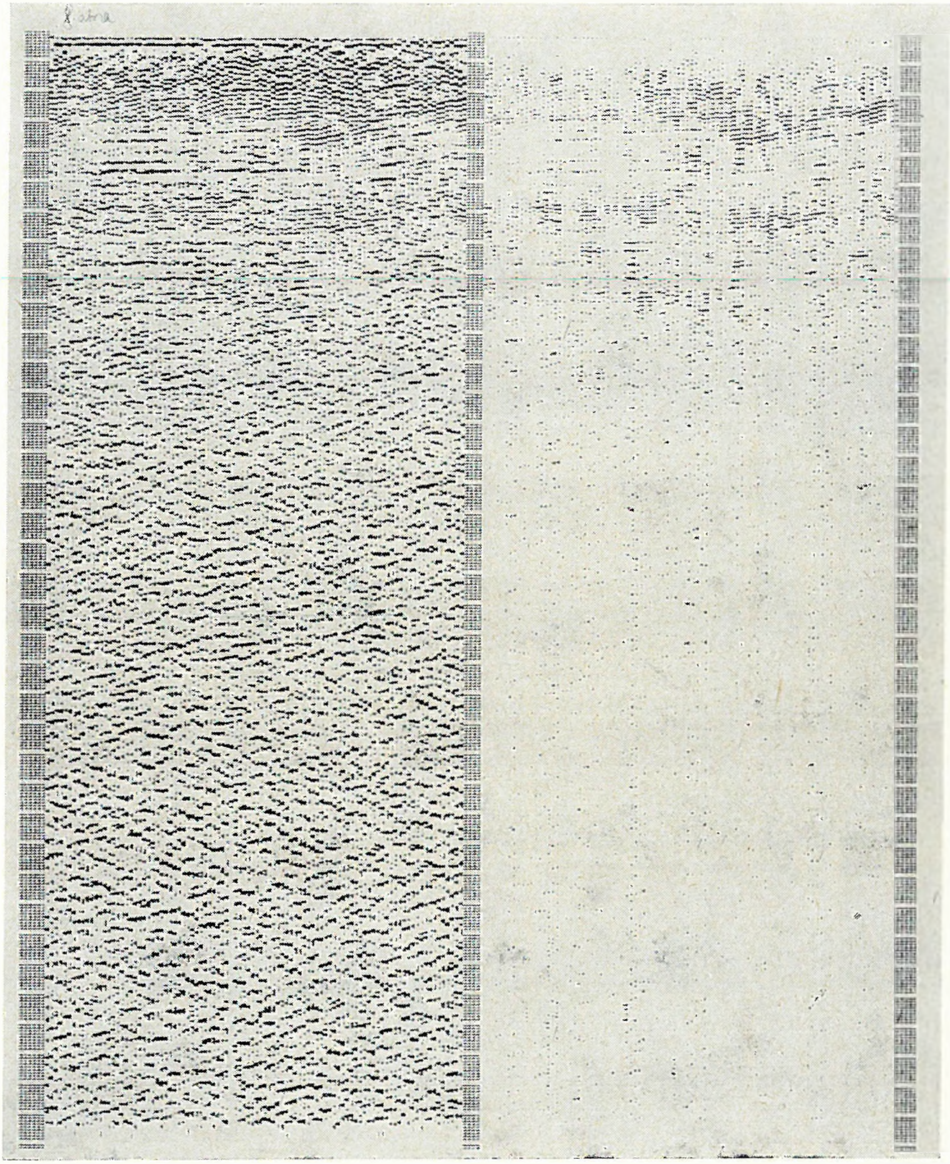

Fig. 8. Section shown on Fig. 7., after spike-deconvolution. With and without AGC

8. ábra. A 7. ábrán bemutatott stacking szelvény spike-dekonvolúció után, AGC-s és AGC nélküli kiírással

Рис. 8. Графическое изображение разреза, показанного на фиг. 7, после обратной фильтрации

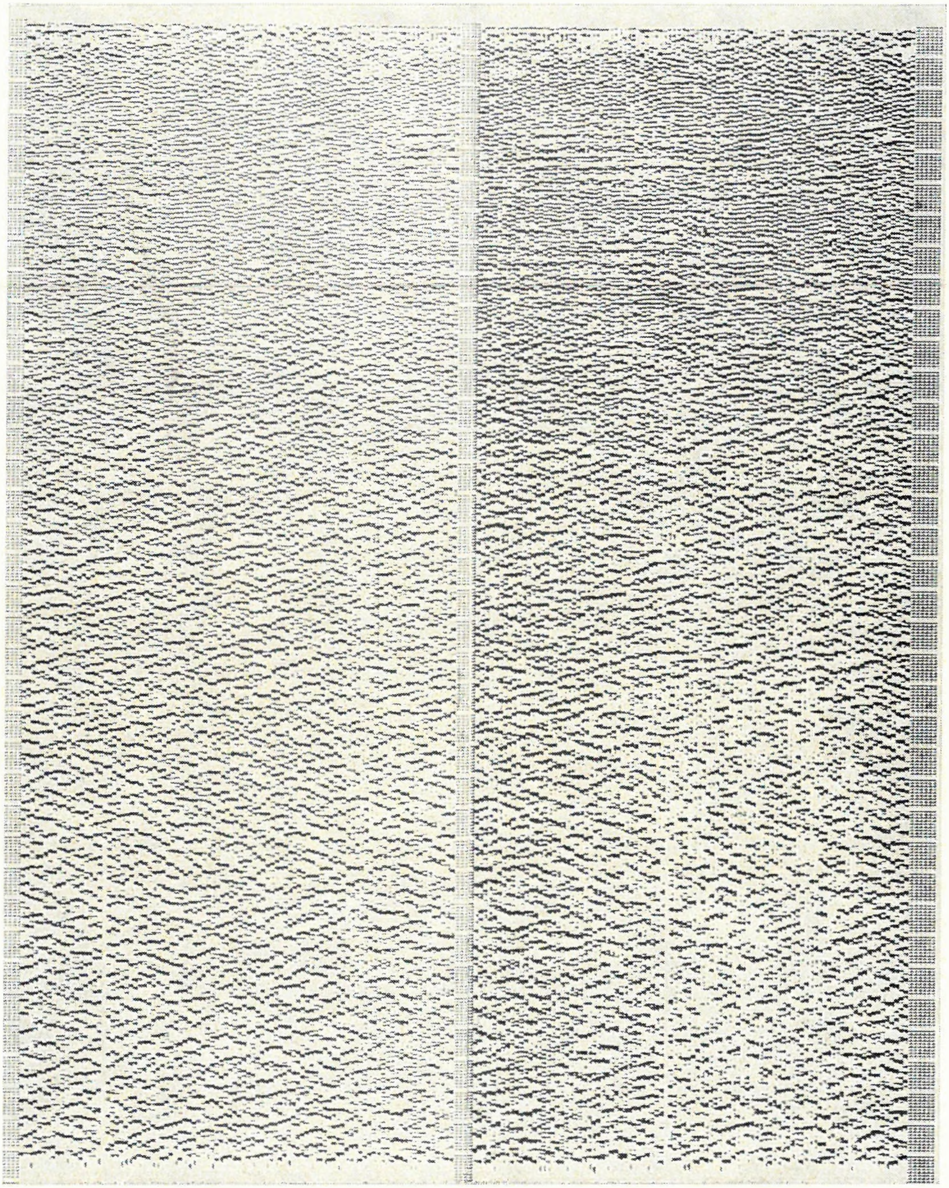

Fig. 9 "Signing combined with clipping" at levels A 2^5 and 2^{10}

9. ábra. A dekonvolvált szelvény (8. ábra) „levágással kombinált előjelzése", A értéke 2^5 és 2^{10}

Рис. 9. Замена деконвольвированного профиля (фиг. 8) с массивом знаков после «отрезания» значений, меньших A; $A = 2^5, 2^{10}$

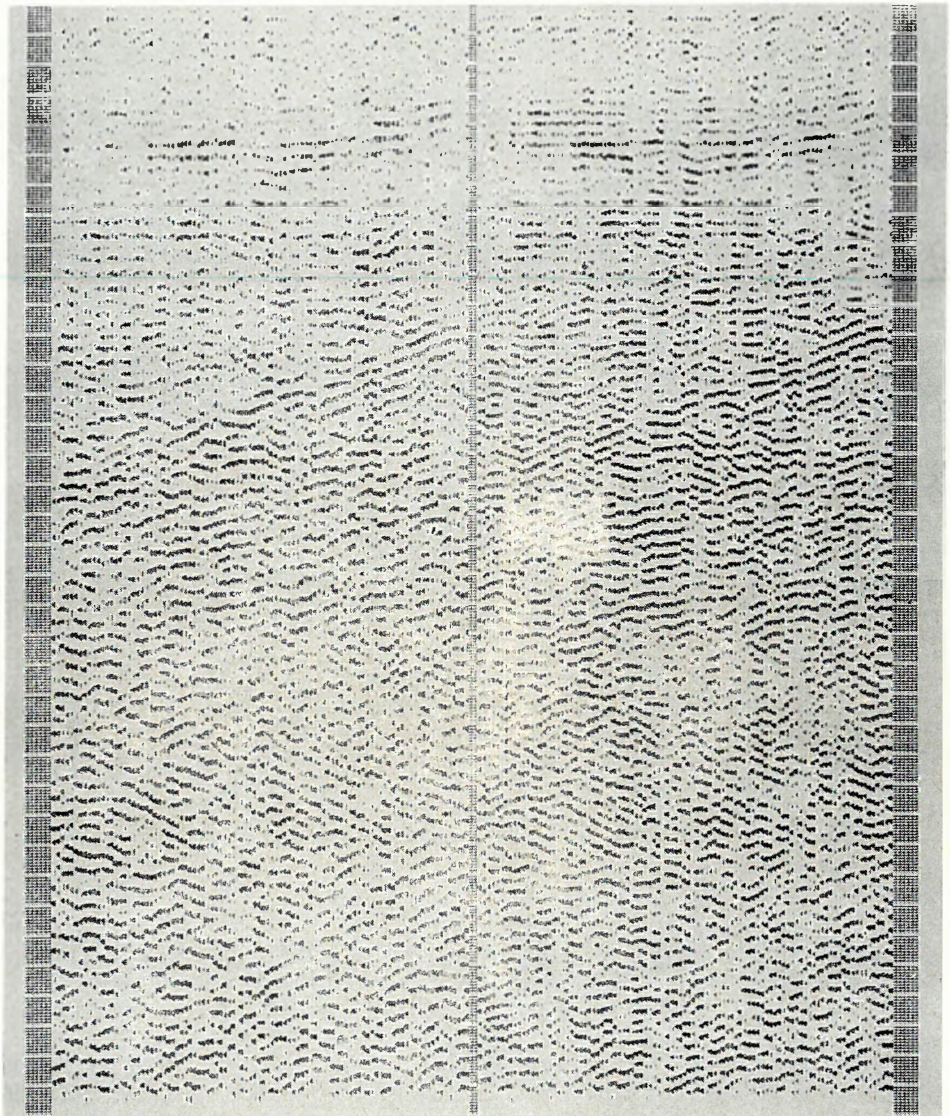

Fig. 10 Digital filter analysis on compressed data, using (15–25) Hz and (20–30) Hz bandpass filters

10. ábra. Digitális szűrőanalízis az adatkompresszió utáni anyagon (15—25) Hz-es és (20—30) Hz-es sávszűrőkkel

Рис. 10. Цифровой фильтрационный анализ после компрессии данных с использованием полосовых фильтров (15—25 гц) и (20—30 гц)

Conclusions

The survey of literature given in the first part of the paper, theoretical considerations on the redundancy of seismic traces and most of all the results of processing sign bit data prove beyond doubt that the dynamic range of present day digital recording is much more than required by routine morphoseismic tasks. Keeping in mind the goals of lithoseismics, it is of course out of question to advocate any reduction of dynamic range. It can also be foreseen that lithoseismic tasks will have to be solved on very fast, floating point computers—but it's a very uneconomic way to employ the same machines to low-accuracy preliminary processing.

As an optimal solution we look forward in the near future to the use of fast, cheap pre-processors having only a few bits and special arithmetics, for most tasks of preliminary processing. We hope the realization of the amount of information contained in sign bits would give a new impetus to the development of field computers.

REFERENCES

Anstey, N. A.–Newman, P., 1966: The sectional autocorrelogram and the sectional retro. correlogram. Geophysical Prospecting 14. No. 4, pp. 389–426

Backus, M. L.–Chen, R. L., 1974: Flat spots in HC exploration. 36th EAEG Meeting, Madrid

Barr, F. J. Jr., 1971: An automatic X^2—T^2 velocity analysis. 41st SEG Meeting, Houston

Belyayeva, A. I.–Rogoza, O. I.–Semenova, S. A., 1966: Certain results of dynamic analysis of reflected waves in Kuleskovka oil field. Internat. Geol. Rev. 10. No. 9, pp. 1067–1072

Bereznev, V. A.–Malovichko, A. A., 1972: Determination of the absorption coefficient from the change of dominant frequencies. Vopr. Obrabotki i Interpretatsii Geof. Nabl. No. 292, pp. 122–124 (in Russian)

Bois, P., 1974: Compatibility of seismic data processing with compression of seismic information. 36th EAEG Meeting, Madrid

Born, W. T., 1941: The attenuation constant of earth materials. Geophysics 6 No. 1, pp. 132–148

Brillouin, L., 1956: Science and Information Theory. Academic Press, Inc. New York

Byakov, Yu. A.–Ryazanova, T. A., 1974: Estimation of the dynamic range of reflected waves at normal incidence. Vopr. Obrabotki i Interpretatsii Geof. Nabl. No. 11, pp. 141–144 (in Russian)

Carlyle, J. W., 1968: Nonparametric methods in detection theory. In: Communication Theory, Ed. by A. V. Balakrishnan. McGraw-Hill, Inc. New York

Carrs, B. W.–Neidell, N. S., 1966: A geological cyclicity detected by means of polarity coincidence correlation. Nature 212 Oct. 8, pp. 136–137

Claerbout, J. F., 1968: Synthesis of a layered medium from its acoustic transmission response. Geophysics 33 No., 2, pp. 264–269

Cochran, M. D., 1973: Seismic signal detection using sign bits. Geophysics 38 No. 6, pp. 1042–1052

Craft, C., 1973: Detecting hydrocarbons for years, the goal of Exploration Geophysics. Oil and Gas J. Feb. 19.

Diekmann, E.–Wierczeyko, E., 1970: A possible method for determining the extent of spread of the gas in an aquifer storage by seismic techniques. 11th International Gas Conference, Moscow

Eisner, E., 1970: Inversion of seismograms. Geophysics 35 No. 3, pp. 404–406

Fara, H. D.–Scheidegger, A. E., 1961: Statistical geometry of porous media. Journal of Geoph. Res. 66. No. 10, pp. 3279–3284

Garotta, R., 1971: Selection of seismic picking based upon the dip, move-out and amplitude of each event. Geophysical Prospecting 19, No. 3, pp. 357–370

Goblick, T. J. Jr., 1965: Theoretical limitations on the transmission of data from analog sources. IEEE Trans. Inf. Theory IT–11, pp. 558–567

Goodman, L. M., 1966: Optimum sampling and quantizing rates. Proc. IEEE 54, No. 1, pp. 90–92

Gurvich, I. I., 1973: Dynamic range of registrations in the seismic reflection method. Razv. Geof. 60, pp. 3–8 (in Russian)

Hilterman, F., 1974: Direct detection of hydrocarbons using reflection data. University of Houston, Geology Dept. Manuscript

Huang Jen-hu, 1961: Change in the spectral properties of seismic waves during generation, propagation and recording. C. Sc. Thesis, Budapest (in Hungarian)

Jankowsky, W., 1971: New aspects of quantitative interpretation of velocity data and their impact on the geological evaluation of exploration prospects. AAPG Convention, Houston

Kats, S. A.–Kondratovich, Yu. V.–Isaev, V. S.–Vilkova, E. S., 1969: Effect of a random structure of a bed member upon dynamic characteristics of the reflected waves. Prikl. Geof. 57 pp. 70–80 (in Russian)

Korvin, G., 1973: Certain problems of seismic and ultrasonic wave propagation in a medium with inhomogeneities of random distribution. Geof. Közl. 21, Nos. 1–4 pp. 5–57

Kunetz, G., 1963: Quelques examples d'analyse d'enregistrements sismiques. Geoph. Prospecting 11. No. 4, pp. 409–422

Leenhardt, O.–Delserre, Marie-Anne, 1974: The value of lithoseismics. The Phone, Winter 1974, pp. 36–38

Lindseth, R. O., 1972: Lithology information from seismic traces. 42nd SEG Meeting, Anaheim, Calif.

Lindsey, J. P., 1974: Bright spot: A progress report and look ahead. World Oil, Apr. 1974

Lyons, P.–Dobrin, M. B., 1972: Seismic exploration for stratigraphic traps. In: Stratigraphic Oil and Gas Fields—Classification, Exploration Methods, and Case Histories. AAPG Memoir 16 and SEG Spec. Publ. No. 10, pp. 225–243

Mateker, E. J. Jr., 1971: Lithologic predictions from seismic reflections. Oil and Gas J. Nov. 8. 1971

Meidav, T., 1969: Hammer reflection seismics in engineering geophysics. Geophysics 34, No. 3, pp. 383–395

Melton, B. S.–Karr, P. R., 1957: Polarity coincidence scheme for revealing signal coherence. Geophysics 22, No. 3, pp. 553–564

Northwood, E. J.–Weisinger, R. C.–Bradley, J. J., 1967: Recommended standards for digital tape formats. Geophysics 32, No. 6, pp. 1073–1084

Nuttal, A. H., 1958: Theory and application of the separable class of random processes. MIT Research Lab. Electronics, Tech. Rept. No. 343

Ossenberg, K., 1968: Vergleich der Eigenschaften des Relaiskorrelators mit denen des klassischen Korrelators. Ph. D. Dissertation, Hannover

Paulson, K. V.–Merdler, S. C., 1968: Automatic seismic reflection picking. Geophysics 33, No. 3, pp. 431–440

Petrovics, I.–Jánvári, J.–Korvin, G.–Sipos, J., 1975: Investigation of the correlation of reflecting horizons via digital filtering and energy analysis. Magyar Geofizika, in press (in Hungarian)

Polshkov, M. K. et al.: Interpretation of seismic exploration data using digital computers. 8th World Petr. Congress, Moscow

Posgay, K.–Korvin, G.–Vincze, J., 1971: Concepts of seismic digital instrumental and methodological development in the ELGI. Geof. Közl. 20, Nos. 1–2, pp. 9–16

Quarles, M. Jr.–Mayne, W. H., 1973: Correlation of Extracted Seismic Properties (ESP) with known oil and gas deposits. Geoph. Soc. Houston Seminar, 1973. Oct.

Rapoport, M. B., 1969: Certain seismic applications of correlation theory. Prikl. Geof. 56, pp. 21–30 (in Russian)

Savit, C. H., 1960a: Preliminary report: a stratigraphic seismogram. Geophysics 25, No. 1, pp. 312–321

Savit, C. H., 1960b: Use seismic data to find stratigraphic traps. Oil and Gas J. April 11, 1960

Savit, C. H., 1974a: Accuracy requirements of seismograph recording systems. 36th EAEG Meeting, Madrid

Savit, C. H., 1974b: Bright spot in the energy picture. Ocean Industry, Febr. 1974

Savit, C. H.–Goldberg, S. S., 1973: How many bits does a seismograph need? 43rd SEG Meeting, Mexico City

Savit, C. H.–Mateker, E. J. Jr., 1971: From "where?" to "what?". 8th World Petr. Congress, Moscow

Sheriff, R. E., 1973: Factors affecting amplitude—a review of physical principles. Geoph. Soc. Houston Seminar, Oct. 8–9, 1973

Stone, C. B., 1974: Geophysical hydrocarbon indicators. 36th EAEG Meeting, Madrid

TEGLAND, E. R., 1970: Sand—shale ratio determination from seismic interval velocity. 23rd Ann. Midwestern Meeting of SEG and AAPG, Dallas

VAN VLECK, J. H., 1943: The spectrum of clipped noise. Radio Res. Lab. Rept. 51. Harvard. Reprinted in Proc. IEEE 54, No. 1, pp. 2–19 (1966)

VOLAROVICH, M. P.–GZOVSKY, M. V.–LEVYKIN, A. I.–OSOKINA, D. N., 1969: The correlation between attenuation decrements and the elastic moduli of rocks. Izv. Earth. Phys. No. 12, pp. 18–26

WOOD, L. C., 1974: Seismic data compression methods. Geophysics 39, No. 4, pp. 499–525

ZAHARCHENKO, V. V.–KOROSTISHEVSY, M. B., 1973: Representation of reflection seismic records by means of their extremal values. Razv. Geof. 58, pp. 35–39 (in Russian)