СООБЩЕНИЯ
ОБЪЕДИНЕННОГО
ИНСТИТУТА
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ

Дубна

98-90

E11-98-90

B.Z.Belashev*, Yu.A.Panebratsev, E.I.Shahaliev, L.M.Soroko

MAXIMUM ENTROPY TECHNIQUE IN THE DOUBLET STRUCTURE ANALYSIS

*Karelian Institute for Geology, Petrozavodsk, Russia

1998

1. Introduction

The inverse problem in physics can be formulated as the processing of the initial experimental data with the aim to find the features of the physical process which cannot be measured directly. So by solving the inverse problem we get true causes of the signals detected by the physical device. To solve the inverse problem in the simple situations, the experimenter uses some semiintuitive approaches on the basis of the common sense and of some apriori information about the process to be investigated. Thus he substitutes the inverse problem by the array of several direct ones. In most cases he uses the regression methods with which the inverse problem can be reformulated as overestimated one with small number of free parameters.

In general case the inverse problem written mathematically is a noncorrected one. This means that even small errors in the experimental data and in the characteristic parameters of the physical device may introduce very great errors in our solution. One cause of this noncorrecticity is that in the accepted mathematical formulation the problem as a whole is a nondefinite one. To redefine it we must introduce some additional information and thus make more narrow the class of functions in which our solution is lying. This restricted class of functions must be compact.

It has been proved that in this formulation of the inverse problem, the latter can have many solutions, and so really we can get only approximated solution. We find the most probable solution. To construct such solution we use the Maximum Entropy Technique (MENT). The main point of this technique is that each possible approximated solution must have apriori probability of its realization. We accept that estimation of the solution which makes maximal the entropy functional of our problem. The mathematical structure of the entropy functional is defined by the kind of the quantum statistics of the physical support of the experimental information and by the value of the filling factor of the quantum degree of freedom. For electromagnetic nonlaser radiation and for observation directly of the adronic and γ -quantum spectra the entropy functional has the form

$$H = -\sum_{x=1}^{N} f(x) \cdot ln\, f(x), \tag{1}$$

where f(x) is the estimation of the solution and x is the independent variable.

In the contrary case, for example in the radioastronomy, the entropy functional has the form

$$H = -\sum_{x=1}^{N} ln\, f(x). \tag{2}$$

The first application of MENT made Burg J.P. [1] in 1975. The recurrence relation, constructed by Burg J.P. for corrected filter, has been used for estimation of the power spectra of the process to be investigated. This approach has been used in geophysics [2,3], in paleontology [4] and for data compression in the theory of information [5].

The MENT has been used for image reconstruction, in particular, of the surface of Ganimed [6], Jupiter satellite, in the thermonuclear synthesis with laser plasma [7], for control of active zone of nuclear reactors [8], in medical tomography [9], in economy and sociology [10].

The application of the MENT in the high energy physics was made in JINR, Dubna in 1981 [11]. In the frame of the additive noise model, the MENT was used for reconstruction of signals with slowly varying components and for sign alternating point spread functions.

The MENT has been used in solving of the inverse problem with poisson-like noises [12] for removing of the instrumental spreading at low events statistics. The results of the computer simulation by means of this variant of MENT were presented in [12] for input signals with unresolved doublets and also for linear input signals. It has been demonstrated that we can get good estimations at low events statistics, $\sim (10-20)$, for wide class of the input signals and for devices with very broad spreading functions of the space-invariant as well as of the space-variant structures.

So it has been proved that MENT can be used for solving of the inverse problems in high energy physics.

The analysis of the Lagrange multiples in the MENT was presented in [13] for poisson-like noise. To describe the joint action of the count statistics and of the instrumental spreading the concept of "temperature" of the experimental data has been introduced. The thermodynamical relations which arise naturally in the MENT for inverse problems with poisson-like noise were considered in [13]. The thermodynamical inequalities which have been found in the MENT formalism and in the Le Chatelier-Braun thermodynamical principles have been constructed. Namely, the variations of the Lagrange multiples are maximal when the variations of the observed signal cover only one sample. This relation between the Lagrange multiples and the observed signal is analogous to the thermodynamical relation between the temperature and the quantity of the heat exchange between two interacting subsystems of the isolated multicomponent system.

The interrelation between the MENT and the method of least squares has been analyzed in [14]. It has been shown that the maximum entropy condition induces the stability of the approximation with normal distribution of the probability density, decreases the spreading of the parameters to be searched for and specifies naturally the bounds of errors zone. The new form of the MENT for estimation of the solution in the poisson-like and gaussian white noises has been given in [14]. With this new form we can see directly the interrelation between the estimation of the solution of the inverse problem and the experimental data. The source of the immunity of the MENT against the noise has been considered. The stabilization operator which gives this immunity and defines the residual spreading in the MENT estimation has been found in [14]

In this paper we explain the MENT for solution of the inverse problems for the case of the additive noise. The effective computer program for solution of the nonlinear equation systems in the MENT has been developed and tested. The scopes of the MENT have been demonstrated in the doublet structure analysis of noisy experimental data. The comparison of the MENT results with the results of the Fourier algorithm technique without any regularization is presented. The relative tolerance noise level is equal to 30 % for MENT and only 0.1 % for pure Fourier algorithm.

2. Maximum entropy technique (MENT)

If the noise, incorporated in the measurements, does not correlate with the investigated spectrum and can be considered as additive noise, the relation between the true spectrum $f(\xi_n)$, the detected signal $s(x_m)$ and the point spread function of the device $h(x_m, \xi_n)$ can be presented as a linear convolution equation

$$S(x_m) = \sum_{n=1}^{N} h(x_m, \xi_n) \cdot f(\xi_n) + n(x_m), \qquad (3)$$
$$m = 1, ..., M; \quad n = 1, ..., N,$$

where M and N are the numbers of the samples at the input and output of the physical device. We consider only the case when the point spread function $h(x, \xi)$ is a space-invariant, that is,

$$h(x_m, \xi_n) = h(x_m - \xi_n). \qquad (4)$$

So we are searching for the estimation $\hat{f}(\xi_n)$ and $\hat{n}(x_m)$ of the spectrum $f(\xi_n)$ and of the noise $n(x_m)$. The functions $h(x_m, \xi_n)$ and $S(x_m)$ are known.

The resolve the equations (3), we use the Maximum Entropy Technique (MENT). The whole problem is formulated as a variation one with conditional extremum and is resolved by means of the nonlinear approach. In line with this formulation we overcome the ambiguity problem and the problem of the noncorrectness.

To explain MENT let us consider the relation between the unknown probability density $\hat{p}(x)$ and the known M moments q_m of this density:

$$q_m = \int_{-\infty}^{\infty} dx \cdot x^m \cdot \hat{p}(x), \qquad m = 1, .., M \qquad (5)$$

E.T. Janes [5] had shown that the least displaced estimation $\hat{p}(x)$ makes the extremum of the entropy functional

$$H = -\int_{-\infty}^{\infty} dx \cdot \hat{p}(x) \cdot ln\, \hat{p}(x). \qquad (6)$$

Simultaneously, the condition of the maximum of the entropy functional gives the most smoothed estimation of $\hat{p}(x)$.

The MENT solution of this problem can be written as

$$\hat{p}(x) = exp\left(-\sum_{m=1}^{M} \lambda_m \cdot x^m\right) \qquad (7)$$

where the Lagrange multiples λ_m must be chosen to meet the momentum equation (5). From (7) we see that the MENT solution will be always a positive definite function.

In the formulation of the inverse problem accepted in this paper, equation (3) the additive noise $n(x_m)$ is also the unknown function to be estimated. The latter can be both positive and negative. For permission of the logarithm operation we introduce the shifted noise

$$\begin{cases} n_0(x_m) = n(x_m) + B, \quad B \geq 0 \quad \text{with} \\ B = \sup |n(x_m)|. \end{cases} \qquad (8)$$

The algebraical form of the inverse problem accepted in this paper can be written as:

$$\begin{cases} S(x_m) = \sum_{n=1}^{N} h(x_m, \xi_n) \cdot f(\xi_n) + n(x_m), \\ (m = 1, ..., M, \quad n = 1, ..., N) \\ \sum_{m=1}^{M} S(x_m) = A \end{cases} \qquad (9)$$

where A is the total number of the detected particles.

The Lagrange functional of our problem can be written as

$$\begin{cases} G\left(\hat{f}, \hat{n}, \lambda_m, \mu\right) = \\ -\sum_{n=1}^{N} \hat{f}(\xi_n) \cdot ln\, \hat{f}(\xi_n) - \rho \cdot \sum_{m=1}^{M} n_0(x_m) \cdot \hat{l}n\, n_0(x_m) - \\ -\sum_{m=1}^{M} \lambda_m \left[\sum_{n=1}^{N} \hat{f}(\xi_n) \cdot h(x_m - \xi_n) + \hat{n}_0(x_m) - B - S(x_m)\right] - \\ -\mu \cdot \left(\sum_{\xi=1}^{N} \hat{f}(\xi_n) - P_0\right) \rightarrow max \end{cases} \qquad (10)$$

where $\lambda(x)$ and μ are the Lagrange multiples to be searched for, and ρ is the parameter, which is introduced with the aim to smooth the estimation of the noise and which is really the ratio of the estimated signal dispersion to the estimated noise dispersion.

From the condition

$$\frac{\partial G}{\partial \hat{f}(\xi_n)} = 0 \quad (11)$$

we get the equation for the signal estimation

$$\hat{f}(\xi_n) = exp\left(-1 - \mu - \sum_{m=1}^{M} \lambda_m \cdot h(x_m, \xi_n)\right) \quad (12)$$

and from the condition

$$\frac{\partial G}{d\hat{n}_0(x_m)} = 0 \quad (13)$$

we get the equation for the noise estimation

$$\hat{n}_0(x_m) = exp\left(-1 - \lambda_m/\rho\right), \quad (14)$$

from the condition

$$\frac{\partial G}{\partial \lambda_m} = 0 \quad (15)$$

we get

$$S(x_m) = \sum_{n=1}^{N} \hat{f}(\xi_n) \cdot h(x_m, \xi_n) + \hat{n}_0(x_m) - B, \quad (16)$$

$$(m = 0, 1, .., M)$$

and from the condition

$$\frac{\partial G}{\partial \mu} = 0 \quad (17)$$

we get

$$P_0 = \sum_{n=1}^{N} \hat{f}(\xi_n). \quad (18)$$

To find (N+1) Lagrange multiples we get the following system of (N+1) - equations:

$$\begin{cases} S(x_m) = \sum_{n=1}^{N} h(x_m, \xi_n) \cdot exp[-1 - \mu - \sum_{m=1}^{M} \lambda_m \cdot h(x_m, \xi_n)] + \\ +exp(-1 - \lambda_m/\rho) - B \\ P_0 = \sum_{n=1}^{N} exp[-1 - \mu - \sum_{m=1}^{M} \lambda_m \cdot h(x_m, \xi_n)]. \end{cases} \quad (19)$$

The system of the nonlinear equations (19) is solving by the Newton-Rufson technique with relaxation [15].

3. Nonlinear equations system

To explain the Newton-Rufson technique let us consider the general form of the nonlinear equations

$$\begin{cases} f_1(x_1, ..., x_N) = 0 \\ \vdots \qquad\qquad \vdots \\ f_N(x_1, ..., x_N) = 0 \end{cases} \quad (20)$$

where $f_i(i = 1, ..., N)$ are the given functions of the unknown values $x_i(i = 1, ..., N)$.

If for every arbitrary point $x(x_1, ..., x_N)$ there exist functions $f_i(x_1, ..., x_N)$ (i= 1,...,N), and their partial derivatives $\partial f_i(x_m)/\partial x_j$ (i=1,...,N), (j=1,...,N), and besides behavior of these derivatives can be considered as linear one, than the solution to be searching for can be estimated by the following technique. Namely, we expand each function $f_i(x_1, ..., x_N)$, (i=1,...,N) into the Newton series up to the second order obrivatives. Thus for increments $\Delta x_{j,m}$, which couple the solutions at (m+1) and m iterations

$$x_{j,m+1} = x_{j,m} + \Delta x_{j,m}, \quad (21)$$

we get the system of the linear equations:

$$\sum_{j=1}^{N} \frac{\partial f_i(x_m)}{\partial x_j}(x_{j,m+1} - x_{j,m}) = -f_i(x_{j,m}) \quad (22)$$

$$(i = 1, ..., N; m = 0, 1, 2...)$$

which we resolve by the iterative technique.

Then the equations (22) go to the equations:

$$\begin{cases} \frac{\partial S(\lambda,\mu)}{\partial\lambda}\Delta\lambda + \frac{\partial S(\lambda,\mu)}{\partial\mu}\Delta\mu = -S(\lambda,\mu) \\ \frac{\partial P_0(\lambda,\mu)}{\partial\lambda}\Delta\lambda + \frac{\partial P_0(\lambda,\mu)}{\partial\mu}\Delta\mu = -P_0(\lambda,\mu) \end{cases} \tag{23}$$

where the partial derivatives will be written as follows:

$$\begin{cases} \frac{\partial S(\lambda,\mu)}{\partial\lambda} = S_\lambda; \quad \frac{\partial S(\lambda,\mu)}{\partial\mu} = S_\mu \\ \frac{\partial P_0(\lambda,\mu)}{\partial\lambda} = P_{0\lambda}; \quad \frac{\partial P_0(\lambda,\mu)}{\partial\mu} = S_{0\mu}. \end{cases} \tag{24}$$

The corresponding increments $\Delta\lambda$ and $\Delta\mu$ are equal to:

$$\begin{cases} \Delta\lambda = \frac{P_0 \cdot S_\mu - S \cdot P_{0\mu}}{S_\lambda \cdot P_{0\mu} - S_\mu \cdot P_{0\lambda}}, \\ \Delta\mu = \frac{P_0 \cdot S_\lambda - S \cdot P_{0\lambda}}{P_{0\lambda} \cdot S_\mu - S_\lambda \cdot P_{0\mu}} \end{cases} \tag{25}$$

where,

$$S_\lambda = -\sum_{n=1}^{N} h(x_m,\xi_n)\cdot exp[-1-\mu-\sum_{m=1}^{M}\lambda_m \cdot h(x_m,\xi_n)]\cdot$$
$$\cdot\sum_{m=1}^{M} h(x_m,\xi_n) - \frac{1}{\rho}exp[-1-\lambda_m/\rho]$$
$$S_\mu = -\sum_{n=1}^{N} h(x_m,\xi_n)\cdot exp[-1-\mu-\sum_{m=1}^{M}\lambda_m \cdot h(x_m,\xi_n)] \tag{26}$$
$$P_{0\lambda} = -\sum_{n=1}^{N} exp[-1-\mu-\sum_{m=1}^{M}\lambda_m \cdot h(x_m,\xi_n)]\cdot\sum_{m=1}^{M} h(x_m,\xi_n)$$
$$P_{0\mu} = -\sum_{n=1}^{N} exp[-1-\mu-\sum_{m=1}^{M}\lambda_m \cdot h(x_m,\xi_n)]$$

4. Computer program

We have used the computationally efficient and numerically stable algorithm for MENT [15], by means of which we have processed the nonlinear equations system (22).

Our initial input signal consisted of two delta-functions, blurred by the given point spread function

$$h(x,\xi) = exp[(x-\xi)^2/2\sigma^2]. \tag{27}$$

The sum was normalized at unity. The additive noise is superimposed then. We have used one such realization as the "measured signal " S(x).

At the first stage we find the Lagrange multiplies, which define the solution in the nonexplicit fashion. The iterative technique gives the solution after 5-8 iterations at noise level 30 %. We have changing the doublet splitting, the relative intensities of the doublet components and the auxiliary parameter ρ. The number of samples was equal to m = l = 120. In some cases we used the additional smoothing of the data.

The auxiliary parameter ρ has been used in this paper for base smoothing of our estimations, $\hat{f}(x)$ and $\hat{n}(x)$. It is evident that at high ρ we can see doublet with harrow splitting with broad point spread function. We have introduced a new parameter

$$\eta = \frac{max|n_r|}{max|A|}, \tag{28}$$

where n_r -maximum of the reconstructed noise and A - maximum of the reconstructed spectrum. The influence of those two parameters, ρ and η on the results of the analysis will be presented in §6.

5. Results of MENT reconstruction

We have performed the computer simulations for MENT doublet structure analysis of the noisy experimental data.

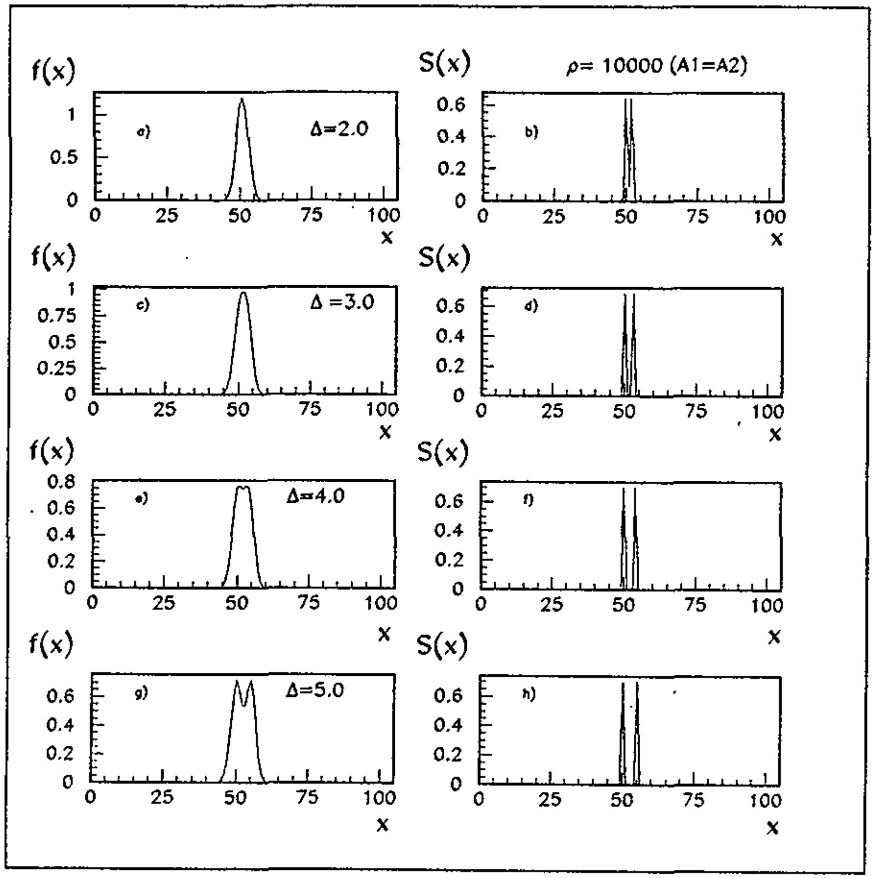

Fig.1. Results of the computer simulations for equal intensities of the doublet components with gaussian point spread function for various doublet splitting without noise.

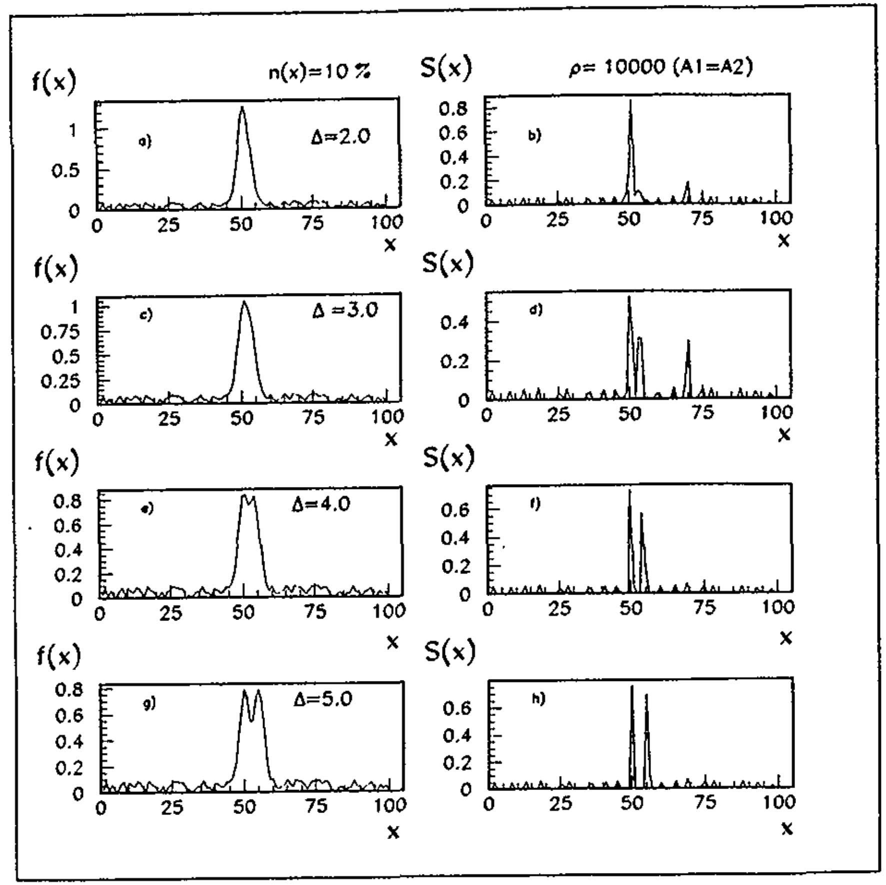

Fig.2. Results of the computer simulations for equal intensities of the doublet components with noise level $n_0(x) = 10\%$.

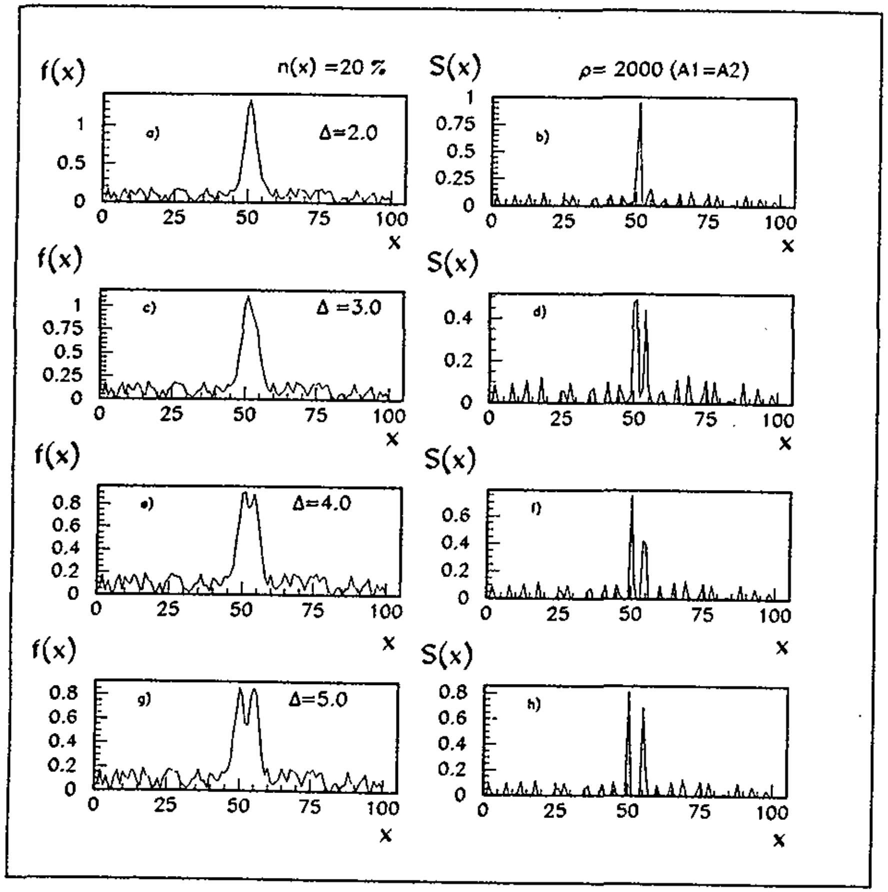

Fig.3. The same as in Fig.2, but with noise level $n_0(x) = 20\%$.

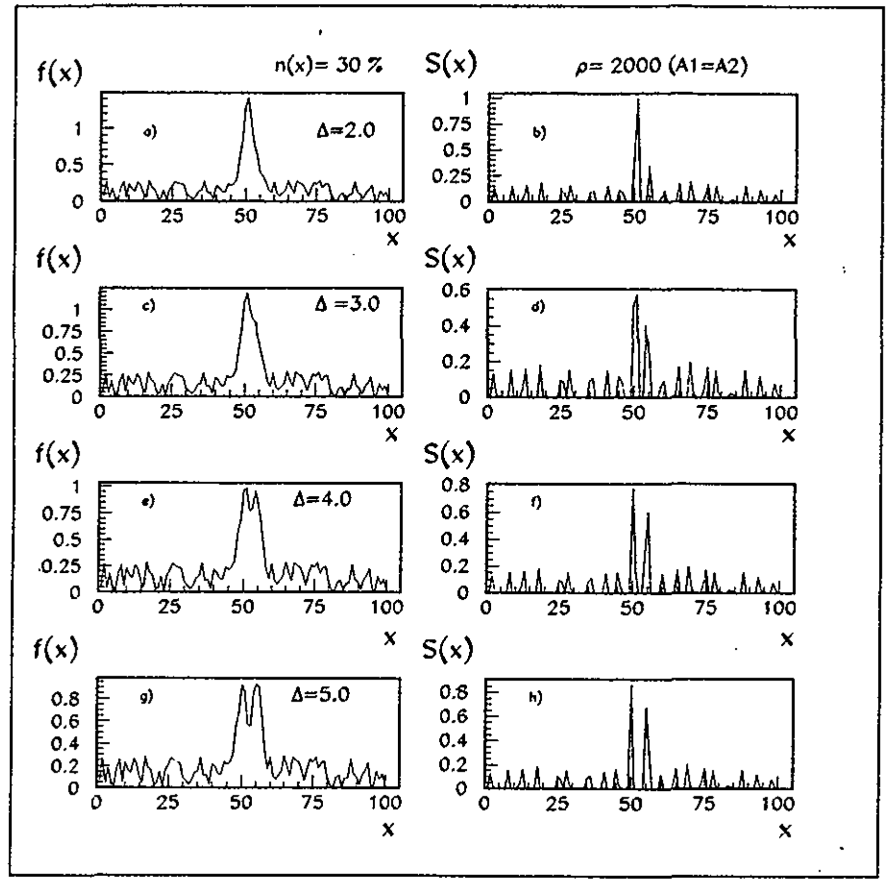

Fig.4. The same as in Fig.3, but with noise level $n_0(x) = 30\%$.

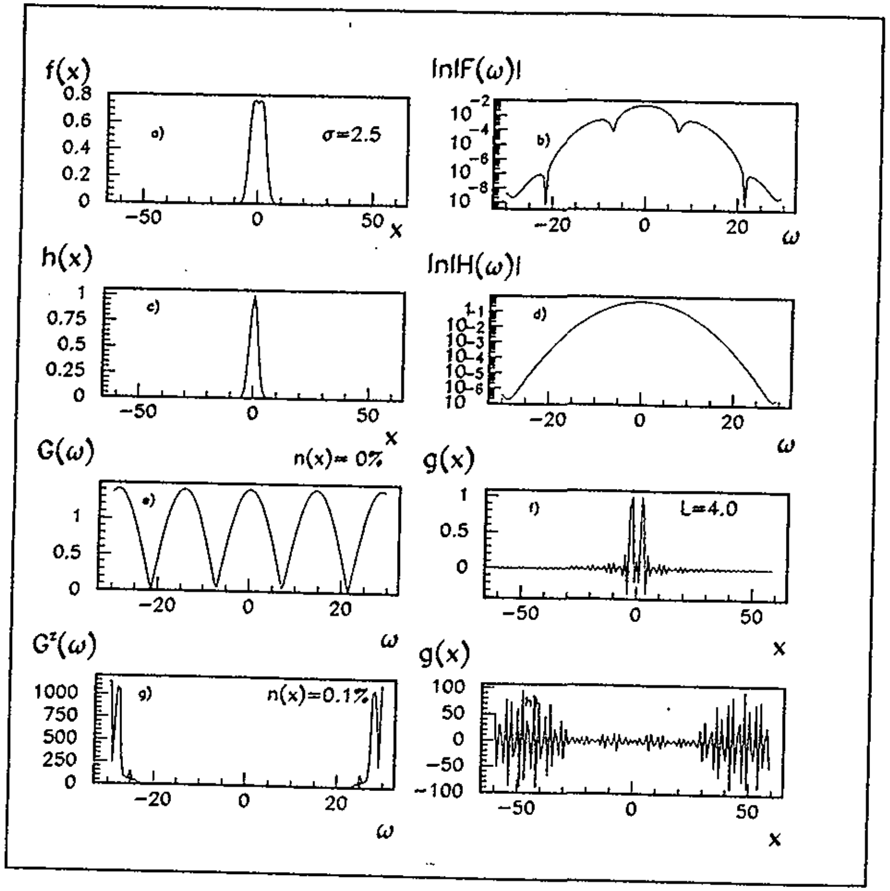

Fig.5. Results of the computer simulations of the Fourier Transform reconstruction for doublet splitting $\Delta = 4$ and for gaussian point spread function with $\sigma = 2.5$ and for two noise levels: $n = 0\%$ and $n = 0.1\%$.

I. Equal intensities of the doublet components ($A_1 = A_2$).

In Fig.1(a,c,e,g) we presented the input signals f(x) in the form of the doublet with gaussian point spread function at various doublet splitting $\Delta = 2, 3, 4, 5$ in units of x axis. In Fig.1(b,d,f,h) we see the corresponding signals reconstructed by means of MENT for different doublet splitting Δ.

In Fig.2(a,c,e,g) we presented the corresponding signals as in Fig.1(a,c,e,g) with additive noise $n_0(x) = 10\%$. In Fig.2(b,d,f,h) we see the corresponding signals reconstructed by means of MENT for different doublet splitting Δ and for auxiliary parameter $\rho = 10^4$. The analogous results are presented for noise levels n=20 % and 30 % in Fig.3 (for $\rho = 2000$) and in Fig.4 (for $\rho = 2000$).

We see that even for noise level n=30 % the doublet structure could be reconstructed for doublet splitting $\Delta \geq 3$.

Now let us compare the results for MENT shown above (Fig.1-,2,3,4) with results reconstructed by the pure linear algorithms, such as Fourier algorithm technique without regularization.

In Fig.5 the Fourier Transform reconstruction is shown for doublet splitting $\Delta = 4$ and for gaussian point spread function with $\sigma = 2.5$. In Fig.5a the initial signal f(x) is shown. In Fig.5b the absolute value of the Fourier Transform $|F(\omega)|$ in the logarithmic scale is given. The point spread function h(x) is shown in fig.5c, and its absolute value of the Fourier Transform $|H(\omega)|$ in the logarithmic scale is presented in Fig.5d. In Fig.5e we see ratio $G(\omega) = F(\omega)/H(\omega)$ as a periodic function. In Fig.5f we see the reconstructed function g(x), as inverse Fourier Transform of the function $G(\omega)$ in the absence of the noise for $\Delta = 4$. The function $G(\omega)$ for noise level $n(x) = 0.1\%$ is shown in Fig.5g. and the "reconstructed " signal in Fig.5h.

II. The skew doublet with $A_1 = 0.2 \cdot A_2$.

In Fig.6(a,c,e,g) we see the initial signals with doublet splittings $\Delta = 2, 3, 4, 5$. The MENT reconstructed signals are shown in Fig.6(b,d,f,h) in the absence of the noise ($n_0(x) = 0\%$). The reconstruction is indeed well for splittings $\Delta \geq 3.0$. Here we observe that the most complex example of signal reconstruction for $A_1 = 0.2 \cdot A_2$ is demonstrated by means of the MENT algorithm.

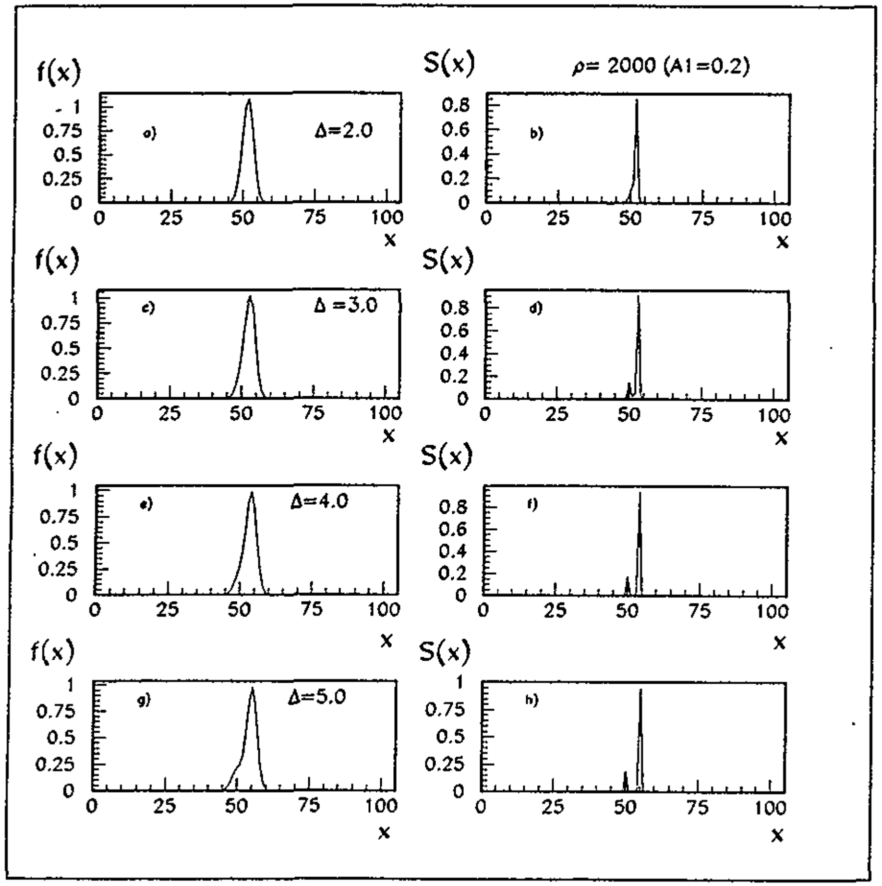

Fig.6. Results of the computer simulations for shew doublet with $A_1 = 0.2 \cdot A_2$ for doublet splitting $\Delta = 2, 3, 4, 5$ and for noise level $n = 0\%$.

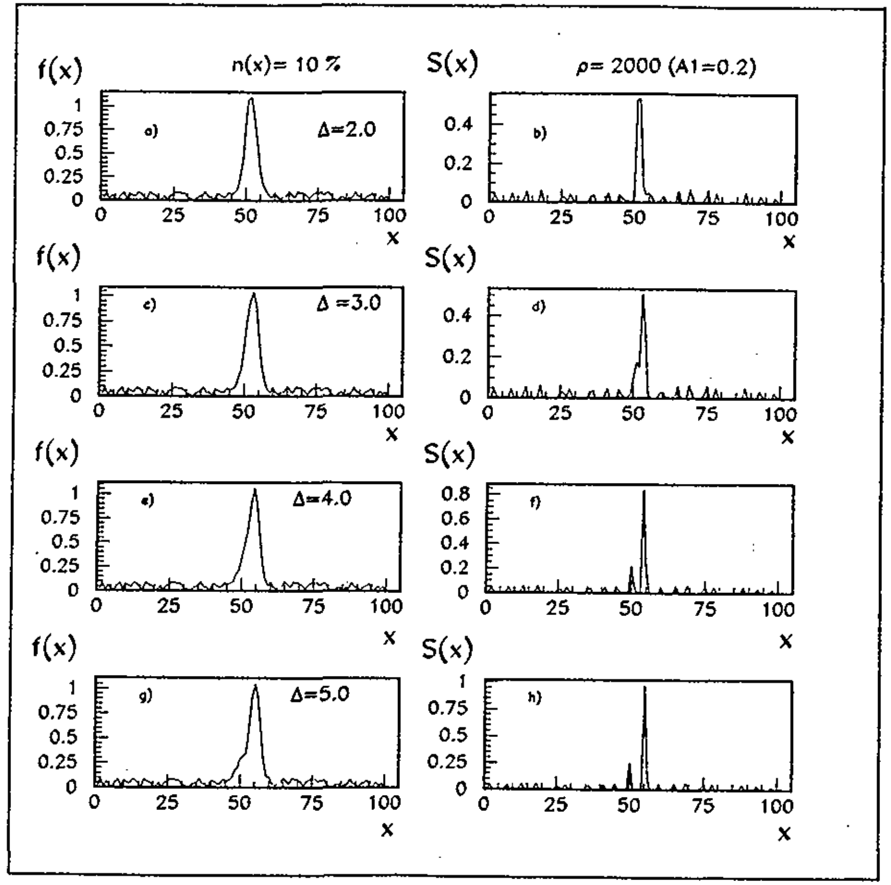

Fig.7. The same as in Fig.6, but for noise level n=10 %.

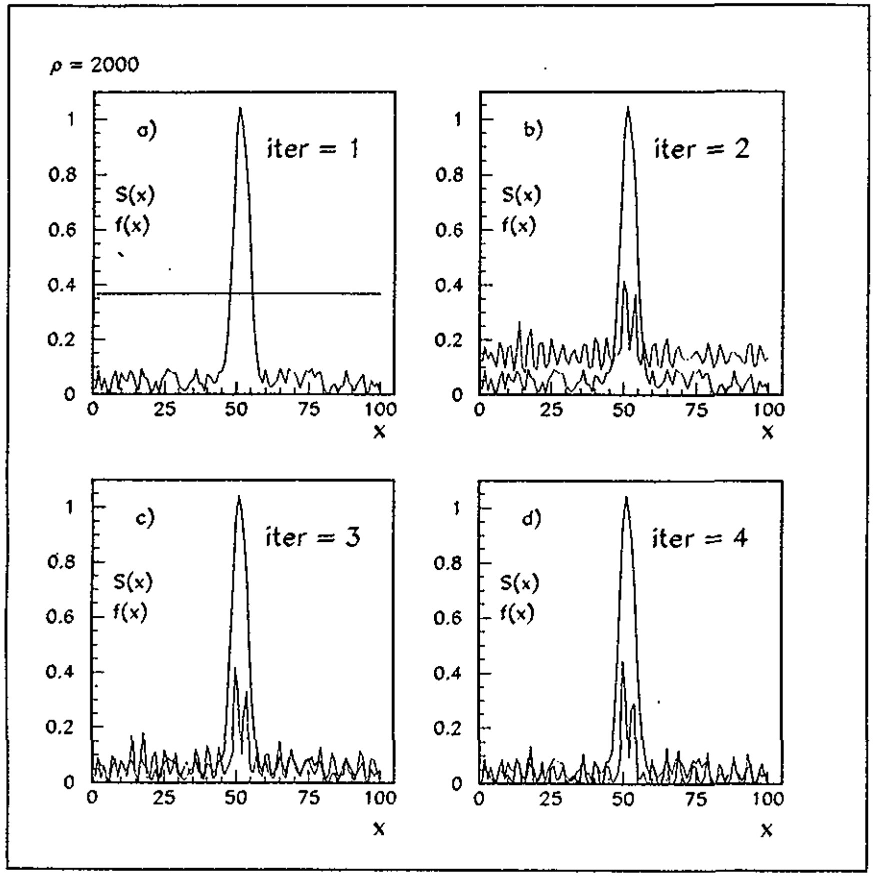

Fig.8. The dynamics of MENT iterations (see text).

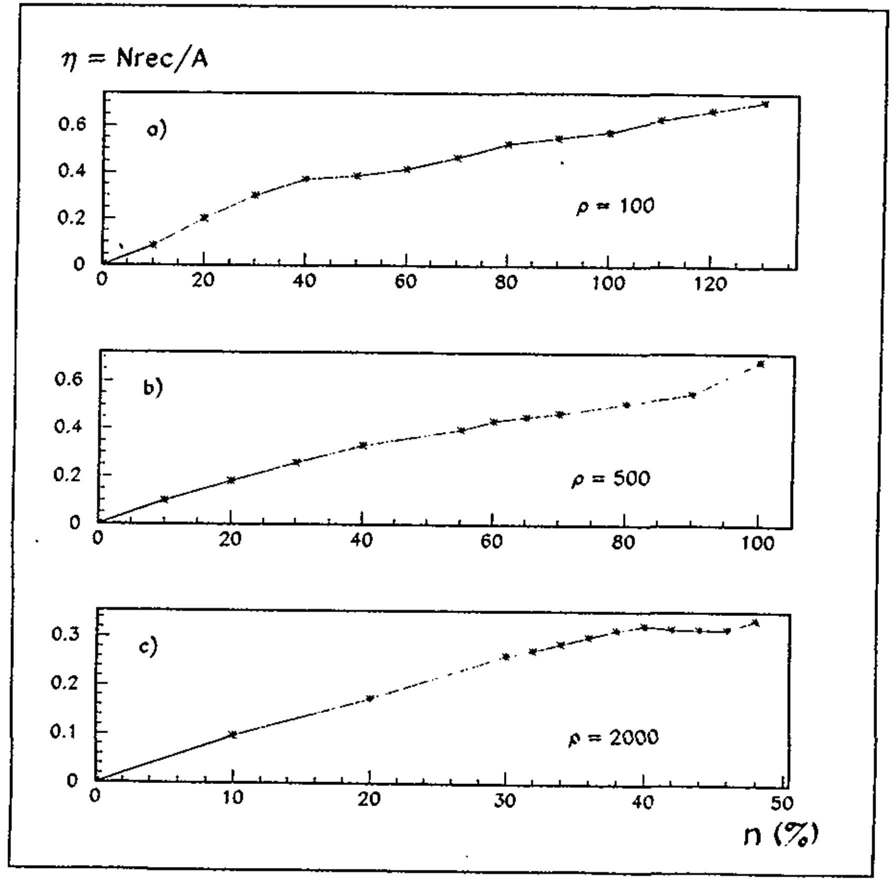

Fig.9. The dependencies of the parameter η versus noise level n(%) for different ρ parameter.

The case of the skew doublet with noise level $n = 10\%$ is shown in Fig.7. The reconstructed doublet can be seen at doublet splittings $\Delta \geq 4$.

The efficiency of the MENT program is define by the number of iterations necessary for good doublet resolution. Let us consider the dynamism of the iterative process for the case of noise level n=10 %. In Fig.8a we see the noise input signal S(x) and the first iteration for f(x) (constant value at ≈ 0.37). The results offer the 3-d and 4-th iterations are given in Fig.8c and Fig.8d, respectively. The convergence is attended at the 4-th iteration, that is the convergence of the MENT program is extremely high.

6. Influence of the auxiliary parameter ρ

To demonstrate the stability of the results of the MENT reconstruction against the auxiliary parameter ρ [11], introduced in §2, we have analyzed the dependencies of the parameter η versus noise level n(%) for different ρ parameter. In Fig.9 we see three such functions $\eta(n)$ for $\rho = 100$, 500 and 2000. The indicator of this stability is the initial slope of the curve $\eta(n)$ for different parameter ρ, which is equal to $0.93 \cdot 10^{-2}$ for $\rho = 100, 0.88 \cdot 10^{-2}$ for $\rho = 500$ and $0.92 \cdot 10^{-2}$ for $\rho = 2000$. In the range of ρ parameter 20:1 the various of the initial slope is only $\pm 2.8\%$. So we have demonstrated that the MENT program is indeed very stable against large variations of the auxiliary parameter ρ.

7. Conclusions

1. The Maximum Entropy Technique (MENT) for solution of the inverse problems is explained.
2. The effective MENT computer program for solution of the system of the nonlinear equations is developed and tested.
3. The doublet structure analysis of the noisy experimental data has been performed by means of MENT.
4. The comparison of the MENT results with the results of the Fourier algorithm technique without any regularization has been made. The relative tolerance noise level is equal to 30 % for MENT and only 0.1 % for pure Fourier algorithm.
5. It is shown that MENT reconstruction algorithm demonstrates high stability to the variations of the ρ parameter.

The authors (Yu.A. Panebratsev and E.I. Shahaliev) would like to thank the Russian Foundation for Fundamental Research for financial support of their investigations (Grant No. 97-27-15 L with project 96-15-96423)

References

1. J.P. Burg, Ph.D.Thesis, Department of Geophysics, Stanford University, Stanford, Calif., 1975.
2. T.J. Ulrych, -J. Geophys.Res., 1972, v. 77, p. 1396.
3. T.J. Ulrych, et.al -J.Geophys.Res., 1973, v. 78, N0.29, p. 10.
4. T. Ulrych, -Nature, 1972, v. 235, p. 218.
5. A.K. Джайн -ТИИЭР, 1981,т.69, No.3, с.71.
6. B.R. Friden, Swindell W.-Science, 1976, v. 191, p. 1237.
7. J.E. Brolley, et.al.-In: 1976 SPSE Conference Proceedings (ed. R. Shaw).SPSE, Washington D. C., 1977, p. 244.
8. J. Sanderson, -IEEE Trans., 1979, v. NS-26, p. 2685.
9. Л.М.Сороко, -ЭЧАЯ, 1981,т.12,вып.3,с.754
10. А.Дж.Вильсон, Энтропийные методы моделирования сложных систем. -М.: Наука, 1978.
11. Б.З.Белашев, В.И.Данилов, Л.М.Сороко, Исследование спектров методом максимума энтропии, Сообщения ОИЯИ, Р10-82-101, 1982, Дубна.
12. Б.З.Белашев, Л.М.Сороко, Метод максимума энтропии и пуассоновский шум, Сообщения ОИЯИ, Р10-85-884, 1985,Дубна.
13. Б.З.Белашев, Л.М.Сороко, Метод максимума энтропии и "температура" Экспериментальных данных, Сообщения ОИЯИ, Р10-85-885, 1985, Дубна.

14. Б.З.Белашев, Л.М.Сороко,
Решения обратной задачи методом максимума энтропии и проблема устойчивости оценок, Сообщения ОИЯИ, Р10-86-100,1986, Дубна.

15. I. Barrodale, R.E. Erickson, Algorithms for least-squares linear prediction and maximum entropy spectral analysis-Part I: Theory, Geophysics, 1980, v.45, N0.3, p.420.

Received by Publishing Department
on April 15, 1998.