Inverse Problems 6 (1990) 55–76. Printed in the UK

Shannon's superresolution limit for signal recovery†

E L Kosarev
Institute for Physical Problems, USSR Academy of Sciences, 117334 Moscow, USSR

Received 3 May 1989, in final form 30 October 1989

Dedicated to Professor C E Shannon on the occasion of the 40th anniversary of his theorem

Abstract. It is shown there is an absolute limit for resolution enhancement in comparison with the Rayleigh classic diffraction limit. The maximum value of superresolution which can be obtained in principle is determined by noise and may be computed via the Shannon theorem concerning the maximum information transmission speed through the connecting channel having noise. A restoration algorithm based on the maximum likelihood method which has Shannon's supremum superresolution is described. Numerical tests of this algorithm are presented and results of its application to a nuclear magnetic resonance experiment are shown. The close connection of superresolution power and the uncertainty principle is discussed. Superresolution depends logarithmically on the signal-to-noise ratio.

1. Introduction

No measurement of any physical quantity ever gives the direct and precise result of the measured value owing to the ever-present measuring noise and also the finite resolution of each measuring system. Both these problems, correction for finite resolution of measuring system and decrease of noise effects, are closely connected and should be considered simultaneously. If there is no noise the resolution can be increased infinitely [1].

The problem of signal recovery (or inverse problem) in a simple case is reduced to the solution of the linear integral equation of the first kind in the form

$$\int_a^b K(x,y)f_0(y)\,dy = F_0(x) \qquad c \le x \le d \tag{1}$$

where $f_0(y)$ is the unknown function which should be determined from the measurements, $F_0(x)$ is the result of transformation of the function $f_0(y)$ by the measuring system which is characterised by the function $K(x,y)$, often called the apparatus function or point spread function (PSF). The result of measurements is the function $F(x)$ which equals the sum of $F_0(x)$ and random noise $N(x)$

$$F(x) = F_0(x) + N(x). \tag{2}$$

† A short version of this paper was presented as a report to the 8th International Maximum Entropy Workshop, Cambridge, UK, 1–5 August 1988.

0266-5611/90/010055+22$03.50

The statistical characteristics of the noise $N(x)$ (its distribution function) are considered to be known.

The goal of the inverse problem is to recover with maximum accuracy the unknown function $f_0(y)$ from the measurements of experimental data $F(x)$.

The Rayleigh classic resolution limit does not use at all the solution of the integral equation (1) and for the resolution it defines the effective width Δ of the point spread function $K(x, y)$, i.e. such value Δ which corresponds to the inequality

$$|K(x, y)/K(x, 0)| \ll 1 \qquad \text{at } |y| \gg \Delta. \tag{3}$$

This quality definition corresponds to the longstanding practice of distinguishing the unknown functions $f_0(y)$ according to the difference between the measurement functions $F(x)$—in optics the difference being the visually resolved one. Strictly speaking, such a definition was only possible at the time of Rayleigh because there were no effective algorithms for signal recovery.

After the development and documentation of such algorithms [2–14], some of which [12–14] have a resolution better than the Rayleigh limit, it is important (i) to choose from them the best method for signal recovery which has the best resolution providing a given signal-to-noise ratio at input, and (ii) to find out if there is any ultimate limit to the enhanced resolution of the distorted and noisy signals.

The existence of such a limit is very important from the theoretical point of view because in such a case it is possible not only to compare the different algorithms between themselves but also to compare all of them with the absolute limit of signal recovery enhancement, and to establish a new unit for measuring the efficiency of the different algorithms. In this paper it is shown that the existence of such a limit follows from the famous Shannon theorem [15] on the maximum information transmission speed through the noisy information channel and because of that it is reasonable to call this unit as 1 shannon and to measure the efficiency of any method for signal restoration in parts (or %) of this limiting value.

The problem of the resolution limit is considered in papers [16–27] and is even connected in papers [17, 20] with the Shannon theorem, but all of these papers deal only with signals which have a bounded spectrum support or with signals which can be simply parametrised, i.e. the analytical form is considered to be known and it is necessary to find a small number of unknown coefficients. In this paper these limitations are removed and the algorithm for general signal reconstruction with superresolution properties is presented, both in numerical tests and in real data reconstruction for the NMR experiment with a heavy fermion superconductor UBe_{13}.

2. Shannon's limit for superresolution enhancement

The basic equation (1) coincides with that in the information theory for signal transmission through a noisy information channel. The point spread function $K(x, y)$ simulates the characteristics of the information channel. If the properties of the channel do not depend on the time, the corresponding point spread function $K(x, y)$ depends only on difference $x - y$

$$K(x, y) = K(x - y) \tag{4}$$

and this function is time (or space) invariant.

The proof of the Shannon theorem given by the author himself in [15] (see also [28, 29]) is based on the expansion of the right-hand side $F(x)$ in the Kotelnikov series and therefore this proof is correct only for the functions having a bounded spectrum support. These functions are only a small part of the total function set which is really used in practice. For example, such widely used apparatus functions as a Gaussian profile

$$K_1 = \exp(-s^2) \qquad s = (x - y)/D \tag{5}$$

or a Lorentzian profile

$$K_2 = 1/(1 + s^2) \qquad s = (x - y)/D \tag{6}$$

are not the bounded spectrum support functions. The parameter D in (5) and (6) is a scaling factor of apparatus functions.

As a matter of fact, the analytical properties of the Kotelnikov (or *sinc*) basis functions

$$y_i(x) = \frac{\sin \pi(2Wx - i)}{\pi(2Wx - i)} \qquad i = 1, 2, 3, \ldots, n \tag{7}$$

are not used at all in Shannon's theorem proof. It is only important that there is a finite number of samples

$$n = 2WX \tag{8}$$

at finite observation interval

$$X = d - c. \tag{9}$$

W in formulae (7) and (8) is the frequency bandwidth of the function (7).

For unbounded spectrum support functions (such as (5) or (6) above) it is possible to use the other basis functions instead of Kotelnikov's functions. This method for solution of the first-kind integral equation (1) is documented in [6], where it is called the orthogonal expansion method (OEM), and in paper [30] it is shown that a reasonable number $m < n$ of basis functions should be taken in all expansion series in order to obtain the minimal mean-squared error of the equation solution.

In fact the assumption that the point spread function $K(x, y)$ has a finite spectrum support is not necessary for Shannon's theorem to be correct. According to Kolmogorov [31] it is sufficient for the dimension of signal space $F(x)$ to be finite. This is the case for almost all experimental data $F(x)$, because the function $F(x)$ is always measured in a finite number of points.

From this point we only consider the n-dimensional signal space

$$f_0(y) = \sum_{\alpha=1}^{n} c_\alpha \varphi_\alpha(y) \qquad (\varphi_\alpha, \varphi_\beta) = \delta_{\alpha\beta} \tag{10}$$

where the basis functions $\{\varphi_\alpha(y)\}$ have to be chosen according to the available physical information on the solution $f_0(y)$ of inverse problem (1) rather than the analytical

properties of integral operator (1). For the right-hand side $F_0(x)$ of equation (1) we have the expansion

$$F_0(x) = \sum_{\alpha=1}^{n} c_\alpha \psi_\alpha(x) \tag{11}$$

where

$$\psi_\alpha(x) = \int_a^b K(x, y)\varphi_\alpha(y)\,dy. \tag{12}$$

In order to find the coefficients c_α we introduce according to [6] the new orthogonal basis

$$\{e_\alpha\}: \quad (e_\alpha, e_\beta) = \delta_{\alpha\beta} \tag{13}$$

which is linearly connected with the basis $\{\psi_\alpha\}$

$$e_\alpha(x) = \sum_{\beta=1}^{n} u_{\alpha\beta}\psi_\beta(x). \tag{14}$$

Round brackets in formulae (10) and (13) denote the scalar products.

The matrix $\|u_{\alpha\beta}\|$ in (14) is the low triangle one

$$u_{\alpha\beta} = 0 \qquad \text{for } \beta > \alpha. \tag{15}$$

If we find the expansion

$$F_0(x) = \sum_{\alpha=1}^{n} s_\alpha e_\alpha(x) \tag{16}$$

then we find the unknown coefficients c_α

$$c_\alpha = \sum_{\beta=1}^{n} u_{\beta\alpha} s_\beta. \tag{17}$$

We find always the coefficients s_α by using the recurrent formula

$$s_\alpha = \left(F_0 - \sum_{\beta=1}^{\alpha-1} s_\beta e_\beta, e_\alpha\right) \tag{18}$$

rather than the usually used standard formula

$$s_\alpha = (F_0, e_\alpha) \tag{18'}$$

because recurrent formula (18) is more stable against truncation errors in comparison with the above standard formula.

While expansion (18) reduces formally to the standard formula (18'), the equivalence of both formulae is broken owing to the finite accuracy of computation. According

to the author's analysis both formulae (18) and (18′) are intrinstically unstable, which means that a small inaccuracy of any origin increases exponentially as a function of α, but the increments are different in both formulae.

For orthogonal polynomials the increments are 0.85 for (18) and 1.1 for (18′), and this difference is very remarkable: the accuracy of formula (18) is higher than that of (18′) by 4–5 decimal places for $\alpha \sim (15 \div 20)$ and real * 8 computation precision.

Since we have the signal and noise from the measurements only together (2) we have only the estimates $\hat{s}_\alpha$ of coefficients s_α

$$\hat{s}_\alpha = \left(F - \sum_{\beta=1}^{\alpha-1} \hat{s}_\beta e_\beta, e_\alpha\right) \tag{19}$$

and, of course, we have only the estimated solution

$$\hat{f}_0(y) = \sum_{\alpha=1}^{n} \hat{c}_\alpha \varphi_\alpha(y). \tag{20}$$

The very important point of OEM is to use the reasonable number $m < n$ of members in all series (10)–(20). This number m or in the more general case the optimal filter coefficients $\{k_\alpha\}$

$$\hat{f}_0(y) = \sum_{\alpha=1}^{n} \hat{c}_\alpha k_\alpha \varphi_\alpha(y) \tag{21}$$

we can find from the need to have the minimal mean-squared error $\overline{R}_m^2$, where

$$\overline{R}_m^2 = E\left[\int_a^b \left(F_0(x) - \sum_{\alpha=1}^{m} \hat{s}_\alpha e_\alpha(x)\right)^2 \mathrm{d}x\right] = \sum_{\alpha=m+1}^{n} s_\alpha^2 + \sum_{\alpha=1}^{m} \overline{(\hat{s}_\alpha - s_\alpha)^2} \tag{22}$$

or the minimal $\overline{R}^2$, where

$$\overline{R}^2 = E\left[\int_a^b \left(F_0(x) - \sum_{\alpha=1}^{n} \hat{s}_\alpha k_\alpha e_\alpha(x)\right)^2 \mathrm{d}x\right]. \tag{23}$$

The sign E in formulae (22) and (23) means averaging over different noise realisation $N(x)$. The minimisation of (22) gives the optimal value of m_{opt} in such a form

$$\text{for } \alpha > m_{\mathrm{opt}} \qquad s_\alpha^2 < \overline{(\hat{s}_\alpha - s_\alpha)^2} = D(\hat{s}_\alpha) \tag{24}$$

and the minimisation of (23) gives the optimal filter coefficients

$$k_\alpha = \frac{s_\alpha^2}{s_\alpha^2 + D(\hat{s}_\alpha)} \tag{25}$$

which coincides with the well known Wiener's filter coefficients.

While formula (24) is also a well known result, the series can only be expanded for eigenvalues that exceed the noise level, its generalisation, the Wiener filtering (25) proves to be very efficient for various applications [30].

We shall not deal here with more detailed specifications of OEM, referring readers for this information to the papers [6, 30] though we would like to mention only that the optimal value m_{opt} in (24) or effective frequency bandwidth of Wiener's filter (25) are both proportional to the resolution frequency and far from the resolution limit resulting from the Shannon theorem. Now we are going to prove this theorem for the signals having the unbounded spectrum support.

So we have n orthonormal basis functions $\{e_\alpha\}$ from (14) and for simplicity also just n measuring points x_i, $i = 1, 2, \ldots, n$. Let us expand the right-hand side (2) of the equation (1) upon these basis functions

$$F(x) = \sum_{\alpha=1}^{n} \hat{s}_\alpha e_\alpha(x) = \sum_{\alpha=1}^{n} (s_\alpha + n_\alpha) e_\alpha(x). \tag{26}$$

We only know from the measurements the sum $s_\alpha + n_\alpha$, but not each individual term separately. Let P_s denote the signal energy

$$P_s = \sum_{\alpha=1}^{n} s_\alpha^2 \tag{27}$$

and P_n the noise energy

$$P_n = \sum_{\alpha=1}^{n} \overline{n_\alpha^2} = \sum_{i=1}^{n} \overline{N(x_i)^2} = n\sigma^2 \tag{28}$$

where σ^2 is the standard deviation of noise in each measuring point

$$\overline{N(x_i)N(x_j)} = \sigma^2 \delta_{ij}. \tag{29}$$

We consider the noise to be stationary, ergodic and uncorrelated.

The energy of the received signal

$$\overline{\|F\|^2} = \sum_{\alpha=1}^{n} \overline{(s_\alpha + n_\alpha)^2} = P_s + P_n + 2\sum_{\alpha=1}^{n} s_\alpha \overline{n_\alpha} = P_s + P_n \tag{30}$$

is equal to the sum of P_s and P_n because of $\overline{n_\alpha} = 0$. The averaging in formulae (28)–(30) is also taken over a different noise realisation. The important formula (30) means that the noise vector having the coefficients $\{n_\alpha\}$ and length $\sqrt{P_n}$ is almost surely orthogonal to the transmitted signal vector having the coefficients $\{s_\alpha\}$ and length $\sqrt{P_s}$ providing the signal space dimension $n \gg 1$.

It follows from the equality

$$\|F\| = \sum_{\alpha=1}^{n} (s_\alpha + n_\alpha)^2$$

that the value $\|F\|^2/\sigma^2$ has the non-central χ^2-distribution with n degrees of freedom and non-centrality parameter equal to P_s/σ^2. Using the formulae for the mean and

variance of a non-central χ^2-distribution (Abramowitz and Stegun [32] formula 26.4.37) we have the relative fluctuation

$$\frac{\delta\|F\|^2}{\|F\|^2} \sim \frac{1}{\sqrt{n}} \to 0 \qquad \text{for } n \to \infty$$

and hence the received signal vector having the components $\{s_\alpha + n_\alpha\}$ almost surely lies near the surface on an n-dimensional sphere having radius $\sqrt{P_s + P_n}$.

The total number M of distinct signals for $P_n \ll P_s$ is less than

$$M \leq \frac{S_n(\sqrt{P_s + P_n})}{S_n(\sqrt{P_n})} = \frac{V_{n-1}(\sqrt{P_s + P_n})}{V_{n-1}(\sqrt{P_n})} \tag{31}$$

where $S_n(r) = k_n r^{n-1}$ is the area of the n-dimensional sphere surface with radius r (k_n is the numerical coefficient) and $V_n(r) = k_n r^n/n$ is the volume of an n-dimensional ball with radius r.

The total number of information bits corresponding to this number M is

$$B = \log_2 M \leq \frac{n-1}{2} \log_2(1 + P_s/P_n) \approx \tfrac{1}{2} n \log_2(1 + P_s/P_n). \tag{32}$$

Formula (32) proves the one part of Shannon's theorem concerning the existence of an absolute limit of transmitted information for the signals having unbounded spectrum support.

This proof coincides completely with Shannon's original proof in [15]. There is only one exception: we use basis functions (14), which can be also determined for point spread functions $K(x, y)$ having an unbounded Fourier spectrum support. To reduce formula (32) to the standard form for PSFs having a bounded spectrum support width equal to $W(\text{cm}^{-1})$ let us determine space δ_W between measuring points. According to the Kotelnikov theorem this space δ_W is

$$\delta_W = 1/2W \tag{33}$$

and the total number of measuring points n, which is also the dimension of signal space, is given by

$$n = X/\delta_W = 2XW \qquad X = d - c. \tag{34}$$

From (32) and (34) we obtain the standard form of the Shannon formula

$$(B/X)(\text{bits cm}^{-1}) = W(\text{cm}^{-1}) \log_2(1 + P_s/P_n). \tag{35}$$

The physical sense of the Shannon formula (35) is that the ultimate frequency B/X (bits cm^{-1}), which determines the resolution, can be in principle greater than $W(\text{cm}^{-1})$ in the factor $\log_2(1 + P_s/P_n)$ for optimal restoration algorithms.

Since this resolution is not only greater than the Rayleigh classic limit (3) but also greater than the one given by OEM in (24) or (25), Shannon's resolution limit can really be called the superresolution limit. The Shannon theorem does not give us any definite algorithm for how to obtain this ultimate resolution and therefore this theorem is only an existence theorem.

In this paper we shall not prove the other part of the Shannon theorem concerning some theoretical possibility to reach the ultimate resolution by application of special encoding and decoding procedures to the transmitted and received signals.

We have no problem with the signal encoding procedure because it is done automatically by integral operator transform (1), but the decoding procedure should always be necessary and is called in our case the signal restoration algorithm. Instead of proving the other part of the Shannon theorem we shall present here in detail the restoration algorithm and results of its numerical tests which will show that this algorithm actually reaches Shannon's superresolution limit. According to the definition in the introduction of this paper, the efficiency of this algorithm will equal 1 shannon.

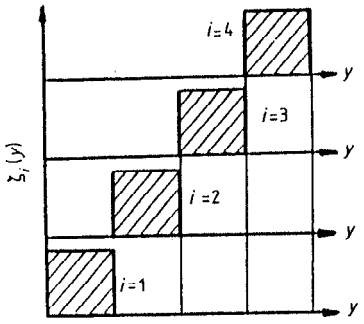

Figure 1. The basic functions $\zeta_i(y)$ for $B = 4$.

The formula (35) gives an answer to the questions about the limit of resolution for signal recovery. We cannot distinguish more than M signals providing the given signal-to-noise ratio (SNR) P_s/P_n and correspondingly cannot have more than B of information bits. Hence all the distinguished signals can be expanded over B different basis functions

$$f_0(y) = \sum_{\alpha=1}^{B} c_\alpha \zeta_\alpha(y) \tag{36}$$

which are presented in obvious graphic form in figure 1 for the case $B = 4$. The width of each basis function $\zeta_\alpha(y)$ equals

$$\varepsilon = X/B = 1/[W \log_2(1 + P_s/P_n)]. \tag{37}$$

It is seen from figure 2 that the resolution limit δ for signal f_0 in the form of equidistant lines, all having equal amplitudes, is equal to double the width ε of basis functions

$$\delta = 2\varepsilon = 2/[W \log_2(1 + P_s/P_n)] \tag{38}$$

or

$$1/W\delta = \tfrac{1}{2} \log_2(1 + P_s/P_n). \tag{39}$$

To proceed to the generalisation of Shannon's formula (35) for unbounded spectrum support functions let us write it in dimensionless form

$$\frac{B/X}{W} = \log_2(1 + P_s/P_n). \tag{40}$$

On the right-hand side of this formula there are only the energies of signal P_s and of noise P_n, which are both invariant to the choice of different basis functions. Geometrically this means only the rotation of the coordinate system in multidimensional signal space. The universal logarithmic law does not depend at all on the analytical form of the point spread function. Only one parameter of the PSF is significant—the width of its spectrum support W—and this appears only on the left-hand side of formula (40).

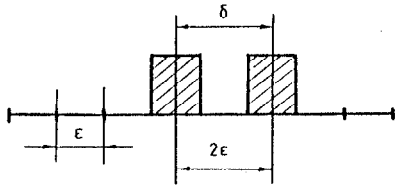

Figure 2. The relation between the width ε of each basic function $\zeta_\alpha(y)$ and the resolution limit δ.

From the dimensional analysis we can use instead of factor $1/W$ the width of the point spread function

$$\Delta = C/W \qquad C = \text{constant} \tag{41}$$

which is determined above in quality level. The numerical value of the constant C depends on the definition of the width of the point spread function $K(x, y)$. As seen from the Shannon theorem proof, the energetics characteristics of signals are most important, and therefore in this paper we shall determine the width Δ of the point spread function as a range

$$\Delta = \int_{-\infty}^{\infty} K^2(x)\,dx \tag{42}$$

providing the normalising condition

$$K(0) = 1. \tag{43}$$

This definition is suitable for the PSFs having both bounded and unbounded spectrum support width. Using this definition formula (39) for the resolution limit can be rewritten in a form

$$\Delta/\delta = \tfrac{1}{2}C \log_2(1 + P_s/P_n).$$

We find the numerical value of the constant C by calculating the frequency bound width W and space width Δ for the two following PSFs having the bounded spectrum support

$$K_3 = \left(\frac{\sin s}{s}\right)^2 \tag{44}$$

and

$$K_4 = \left(\frac{\sin s}{s}\right) \tag{45}$$

where $s = (x - y)/D$.

The first function is interesting because it is positive everywhere as well as the PSFs (5) and (6); the second one coincides with the Kotelnikov basis function (7).

We have for the function K_3

$$\Delta_3 = \frac{2\pi}{3}D \qquad W_3 = \frac{1}{\pi D} \qquad \Delta_3 = \frac{2}{3W_3} \tag{46}$$

and for the function K_4

$$\Delta_4 = \pi D \qquad W_4 = \frac{1}{2\pi D} \qquad \Delta_4 = \frac{1}{2W_4}. \tag{47}$$

From these relations we receive the sought-for formulae for the limit of resolution in terms of the ratio of the PSF width Δ to the minimal space δ between signals as a function of signal-to-noise ratio

$$\Delta/\delta = \tfrac{1}{3}\log_2(1 + P_s/P_n) \tag{48}$$

for the PSF K_3 and

$$\Delta/\delta = \tfrac{1}{4}\log_2(1 + P_s/P_n) \tag{49}$$

for the PSF K_4. For reference purposes we cite the numerical values of width Δ for the PSFs (5) and (6) (Gaussian and Lorentzian profiles)

$$\Delta_1 = \sqrt{\frac{\pi}{2}}D \qquad \Delta_2 = \frac{\pi}{2}D. \tag{50}$$

Before concluding let us emphasise that Shannon's superresolution limit in a form (35) or (48) and (49) is an ultimate superresolution limit for any restoration method. This is a rigorous mathematical theorem and it says nothing about efficiency and resolution of some specific methods for signal restoration. In next parts of the paper we are going to describe the restoration algorithm based on the maximum likelihood method, to test it and to show its efficiency for superresolution.

3. Maximum likelihood method for signal restoration

For restoration we use the maximum likelihood (ML) method [33] which is the generalisation of Tarasko's iteration algorithm [34]. For the case when the right-hand side $F(x_i)$ has binomial (or Poisson) distribution in each separate point x_i and jointly polynomial distribution for the whole set of $\{F(x_i)\}$ values, for $i = 1, 2, \ldots, n$, the iteration formula can be written in the form

$$g_k^{(s+1)} = g_k^{(s)} + h g_k^{(s)} \sum_{i=1}^{n} p_{ik}\left(\frac{f_i}{\sum_{j=1}^{m} p_{ij} g_j^{(s)}} - 1\right). \tag{51}$$

In this formula $s = 1, 2, \ldots$ is the iteration number, the vector g_k, $k = 1, 2, \ldots, m$, equals the values of the unknown function $f_0(y_j)$ in points $y_1, y_2, \ldots, y_m$; the vector

f_i, $i = 1,2,\ldots,n$, is proportional to the right-hand side function $F(x_i)$ in points $x_1, x_2, \ldots, x_n$; and matrix p_{ij} is equal to the values of the PSF

$$p_{ij} = K(x_i, y_j). \tag{52}$$

h is the length step in the space of unknown vector $\{g_k\}$ in the direction which is close to the gradient direction. At $h = 1$ the iteration procedure (51) coincides with Tarasko's procedure. The actual computer program implementation will be described in a forthcoming publication.

In the procedure the values of f_i and p_{ij} have to be normalised according to relations

$$\sum_{i=1}^{n} f_i = 1 \qquad \sum_{i=1}^{n} p_{ij} = 1 \qquad \text{for } j = 1,2,\ldots,m. \tag{53}$$

The normalising condition and positiveness of the vector g_k are automatically kept at any step length h according to relations (51), although in our computer program it does so after each iteration at $h > 1$ for more stability against truncation errors. All discrete convolutions for the PSF in the form (4) are computed by the fast Fourier transform (FFT) program.

The value of step length h is chosen according to steepest ascent of logarithmic likelihood function

$$L = \text{constant} + N \sum_{i=1}^{n} f_i \ln p_i \tag{54}$$

where

$$p_i = \sum_{k=1}^{m} p_{ik} g_k \qquad N = \sum_{i=1}^{n} N_i \qquad N_i = F(x_i)$$

by computing of derivatives $\mathrm{d}L/\mathrm{d}h|_{h=0}$ and $\mathrm{d}^2L/\mathrm{d}h^2|_{h=0}$ based on the exact analytical formulae one time per each iteration. The optimal value of h is sometimes equal to the value $h \sim 10^3 \div 10^4$ and this is the explanation why our procedure (51) works faster than the original procedure of Tarasko.

For the case when the right-hand side $F(x)$ has the Gaussian distribution the iteration formula (51) should be modified, but according to our numerical tests the use of (51) for this case (instead of the more correct procedure specially written for the Gaussian case) decreases the efficiency of restoration only slightly.

Iterations (51) should be continued while the value

$$chi\,2 = \sum_{i=1}^{n} \frac{(\delta N_i)^2}{N_i} \tag{55}$$

where $\delta N_i = N(p_i - f_i)$, is larger than the level of $\chi^2_{n-1}(P_1)$, where P_1 is the significance level of the χ^2 criterion, which is usually assumed equal to 5%, and the iterations should be stopped as soon as this value drops below the level

$$chi\,2 \leq \chi^2_{n-1}(P_2) \tag{56}$$

where $P_2 = 80$–95%.

All computation results have been obtained on an HP-1000 minicomputer at the Institute for Physical Problems with the accuracy of 39 bits for mantissa ($\sim 1.8 \times 10^{-12}$). For number of data points $n = 512$ every 50 iterations took about 5 min of CPU time. The total number of iterations depends first of all on how close is the value δ to the resolution limit and is sometimes as many as 5–10 thousand.

Because of such long CPU times, even using the FFT programs, a question arises: could it be possible to obtain similar results on superresolution by using some linear restoration methods, i.e. based on the Fourier transform for convolution equation (1) together with the optimal filtering? Regretfully the answer is negative because in the cases where superresolution is interesting and important the width of the Fourier transform spectrum of function $F(x)$ on the right-hand side of equation (1) is always much less than the one of the restored signal $f_0(y)$ which we look for. And any kind of linear restoration methods can only modify the amplitudes of the Fourier harmonics but cannot generate the new ones which are absent in the input data or were lost in the input noise.

The nonlinear restoration method used in this work does actually extrapolate the Fourier spectrum having input data of much higher frequencies compared with our input data. This is an explanation in essence of the efficiency of the ML method for signal restoration.

4. The numerical tests and their analysis

In numerical tests we examined the dependence of resolution limit for two- and three-line signals as a function of signal-to-noise ratio for the different PSFs. For each of the functions (5), (6), (44) and (45) the solution in the form of two (or three) narrow Gaussian lines has been taken

$$f_0(y) = \sum_{i=1}^{m} \exp(-u_i^2) \qquad m = 2 \text{ or } 3 \tag{57}$$

where $u_i = (y - y_i)/D_1$, $D_1 = D/40$, and a convolution integral

$$F_0(x) = \int_{-\infty}^{\infty} K(x - y) f_0(y)\, dy \tag{58}$$

is computed. The parameter D in (5), (6), (44) and (45) is chosen to be equal to such a value that the PSFs decrease up to zero inside the interval (a, b) and in this case the integral equation (1) can be considered to have infinite limits for integration. All functions in calculations are determined on a discrete grid of points

$$x_i = i \qquad i = 1, 2, \ldots, n \tag{59}$$

and the integral is computed as a finite sum.

The number of points is usually equal to $n = 512$ but in some tests this number is also equal to $n = 2048$.

To each of the computed values $F_0(x_i)$ has been added the noise $N(x_i)$ having the Gaussian distribution and non-correlated values in different points according to (29). Parameter σ^2 has been determined from the signal-to-noise ratio P_s/P_n where

$$P_s = \int_{-\infty}^{\infty} F_0^2(x)\, dx \tag{60}$$

and

$$P_n = n\sigma^2 \tag{61}$$

according to the definition of decibel units $dB = 10\lg(P_s/P_n)$

$$\sigma^2 = \frac{1}{n}\frac{P_s}{10^{dB/10}}. \tag{62}$$

The formulae (27) and (60) for the energy of the signal are equivalent according to the Parseval relation.

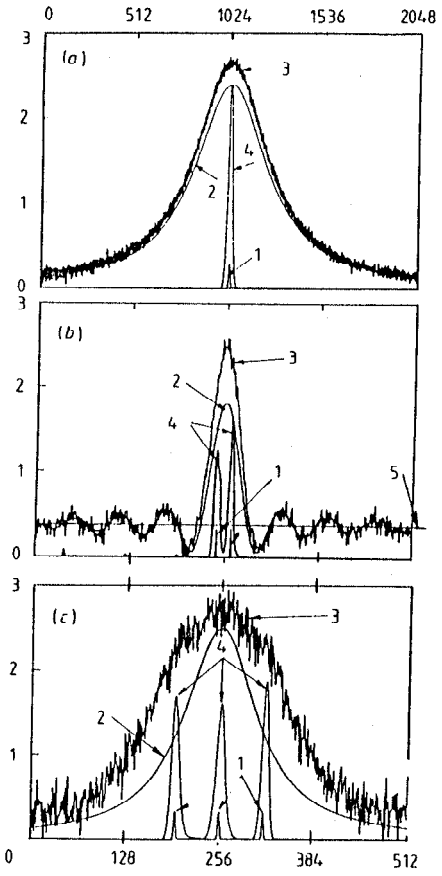

Figure 3. Examples of restoration for (*a*) one, (*b*) two and (*c*) three lines, illustrating the original lines (curves 1), the point spread functions (curves 2), the initial data (curves 3) for the restoration, and the results (curves 4) of restoration, and in (b) the zero level (curve 5) for the PSF. (*a*) The Lorentzian PSF with $D = 240$, number of points $n = 2048$, signal-to-noise ratio 30 dB, 500 iterations, $\chi^2/n = 1.0055$. (*b*) PSF $= \sin x/x$ with $D = 10$, $n = 512$, SNR $= 17.5$ dB, separation of two original lines $\delta_2 = 20$, 1200 iterations, $\chi^2/n = 1.0269$. In this example the new version of restoration procedure was used, which was specially written for the data having a Gaussian noise distribution. (*c*) the Lorentzian PSF with $D = 60$, $n = 512$, SNR $= 20$ dB, separation of original lines $\delta_2 = 60$, 1000 iterations, $\chi^2/n = 0.99037$.

For each given signal-to-noise ratio (SNR) the distance between two (or three) lines

$$\delta_2 = y_2 - y_1 \tag{63}$$

is decreased up to such a value for which these lines can be well resolved by the restoration algorithm.

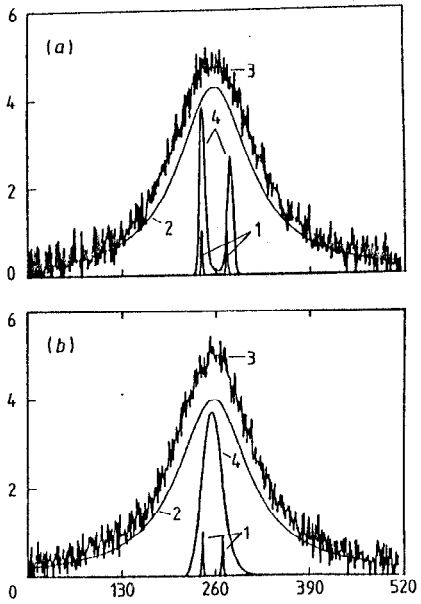

Figure 4. Restoration of two lines convolved with the Lorentzian shape point spread function and signal-to-noise ratio 20 dB for (*a*) separation of lines $\delta = 35$ greater than resolution limit (67) and (*b*) separation $\delta = 28$ less than resolution limit (67), illustrating the original lines (curves 1), the point spread functions (curves 2), the initial data (curves 3) for the restoration, and the results (curves 4) of restoration.

The examples of restoration for one, two and three lines are shown in figure 3. An example of two-line restoration is presented in figure 4 for the Lorentzian PSF (6) with parameter $D = 60$, SNR = 20 dB and $\delta_2 = 35$ (the lines are well resolved) and $\delta_2 = 28$ (the lines are not resolved). The superresolution coefficient, SR, in this example is

$$\text{SR} = \Delta/\delta_2 = \tfrac{1}{2}\pi D/\delta_2 = 2.7. \tag{64}$$

It is worth noting that the width of transition gap between domains where there is resolution and there is not is quite small—of about 20%.

The next example for the Gausian PSF with parameters $D = 120$, SNR = 45 dB and $\delta_2 = 30$ is shown in figure 5(*a*). The superresolution coefficient for this example is

$$\text{SR} = \Delta/\delta_2 = \sqrt{\tfrac{1}{2}\pi}D/\delta_2 = 5.01. \tag{65}$$

Figure 5(*b*) shows the Fourier spectra of the noisy input data $F(x)$, of the precise solution $f_0(y)$ and of the restoration result $\hat{f}_0(y)$ as computed by iteration formula (51).

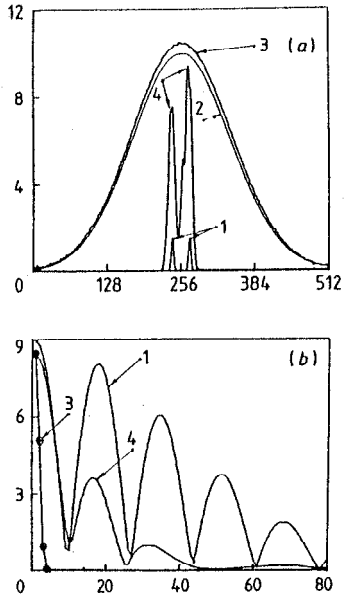

Figure 5. (*a*) Restoration of two lines convolved with the Gaussian PSF and SNR = 45 dB, separation of lines $\delta = 30$ greater than resolution limit (67), illustrating the original lines (curves 1), the point spread functions (curves 2), the initial data (curves 3) for the restoration, and the results (curves 4) of restoration. (*b*) Corresponding Fourier spectra in a linear scale, for curves 1, 3 and 4 as in (*a*).

It is seen from this figure that the Fourier spectrum of the restored signal has some first minima at the correct places because the restoration result has a correct form of the doublet lines. It is very important that the width of the Fourier spectrum for the restored signal is much broader than the one for input noisy data. So the restoration algorithm based on the nonlinear ML method really does extrapolate an input Fourier spectrum to much higher frequency.

The next example for the Gaussian PSF and distance between two lines $\delta_2 = 25$ is presented in figure 6(*a*). In contrast with figure 5, this distance is less than the resolution limit and the restoration algorithm cannot indeed resolve these two lines. The Fourier spectra corresponding to this case are presented in figure 6(*b*). While the Fourier spectrum of the restored signal is broader than the ones of the input data, the result of extrapolation is incorrect: instead of oscillation we have a smoothed Gaussian-like spectrum.

The most obvious example to show the superresolution efficiency is presented in figure 7 for the PSF (44). The Fourier transform of this PSF is

$$\mathscr{K}(\omega) = \int_{-\infty}^{\infty} K_3(x)\mathrm{e}^{\mathrm{i}\omega x}\,\mathrm{d}x = \begin{cases} \pi D(1 - |\omega|D/2) & \text{for } |\omega| < 2/D \\ 0 & \text{for } |\omega| \geq 2/D. \end{cases} \tag{66}$$

Thus there is no information in the input data concerning the unknown signal $f_0(y)$ out of the frequency range $|\omega| \geq 2/D$. Nevertheless if the distance between two lines is larger than the resolution limit, the restoration algorithm generates absent harmonics and resolves these two lines.

A summary of such tests is presented in figure 8 for the three PSFs: (5), (6) and

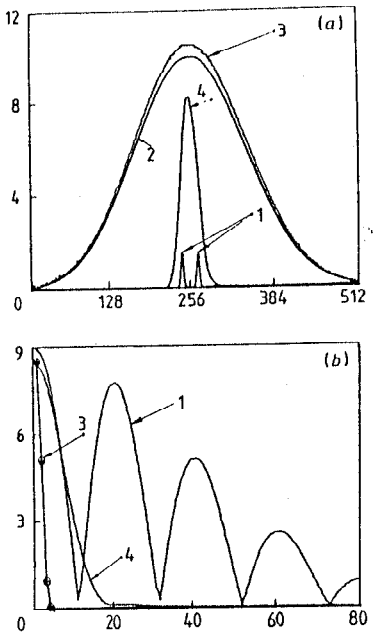

Figure 6. (*a*) Restoration of two lines convolved with the Gaussian shape PSF and SNR = 45 dB, for separation of lines $\delta = 25$ less than resolution limit (67), illustrating the original lines (curves 1), the point spread functions (curves 2), the initial data (curves 3) for the restoration, and the results (curves 4) of restoration. (*b*) Corresponding Fourier spectra in a linear scale, for curves 1, 3 and 4 as in (*a*).

(44) with SNR between 10 and 50 dB. This figure shows the Shannon resolution limit

$$\text{SR} = \tfrac{1}{3}\log_2(1 + P_s/P_n) \tag{67}$$

the regression line for the numerical experiment data and the 95% confidence interval for this line. There is a good agreement between them.

It follows from figure 8 that the superresolution depends roughly linearly on signal-to-noise ratio in dB units, and this law does not depend in practice on the analytical form of the point spread functions. This is the first conclusion from our numerical tests.

This result is quite a new one and it differs from the power relationship of superresolution as a function of SNR in dB units stated in previous papers [14, 20, 27]. This is because we do not use any parametrisation for restoration. If we could have some information concerning the unknown signal $f_0(y)$ to be restored then the superresolution limit (67) could be exceeded.

Parametric methods can, in principle, have better resolution than the Shannon limit (67) but these methods are only adequate for the restoration of a small set of signals. It is probably more reasonable to use such a two-step (or adaptive) restoration procedure: at first one should use the nonparametric algorithm and only after obtaining some *a posteriori* information about unknown signals one could use more accurately the parametric algorithms.

We can also see from figure 8 that the resolution which is achieved in practice is approximately equal to the Shannon resolution limit and this fact is the demonstration of the other part of Shannon's theorem which we have mentioned above but have not

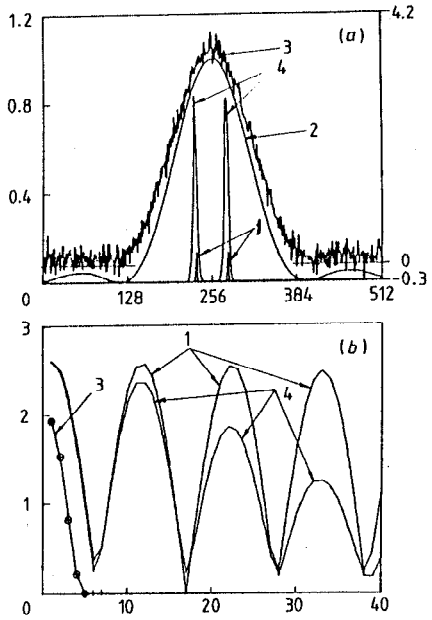

Figure 7. (*a*) Restoration of two lines, convolved with $(\sin x/x)^2$ point spread function with $D = 45$ and SNR = 22.5 dB, for separation of lines $\delta = 48$ greater than resolution limit (67), illustrating the original lines (curves 1), the point spread functions (curves 2), the initial data (curves 3) for the restoration, and the results (curves 4) of restoration. (*b*) Corresponding Fourier spectra in a linear scale for curves 1, 3 and 4 as in (*a*) for 5000 iterations, $\chi^2/n = 0.96057$. The numbers on the right-hand vertical axis are the ordinate values for initial data 3.

proved: the principal possibility to reach the limit of resolution by choosing the special encoding and decoding procedures.

Due to this the efficiency of the ML restoration algorithm presented in this paper is approximately equal to 1 shannon. From our point of view the practical demonstration of a restoration algorithm having the Shannon supremum efficiency has the same importance as the theoretical proof of this part of the Shannon theorem.

From the Shannon theorem and the numerical tests presented here we come to the second conclusion of this paper: no restoration algorithm can have better resolution than the ML algorithm documented here. Of course there are other algorithms which may be faster and demand less computer memory but none of them can have better resolution.

5. Application to a nuclear magnetic resonance (NMR) experiment

The result of applying the ML algorithm to an NMR experiment [35] with the heavy-fermion superconductor UBe_{13} is presented in figure 9. It is clear from the figure that some lines overlap to a considerable degree. To resolve these lines it is necessary above all to find the PSF. Using the spectral lines farthest to the left and right (the 'boundary' lines), which have insignificant overlap with remaining components, we can reconstruct the shape of an isolated line of this spectrum.

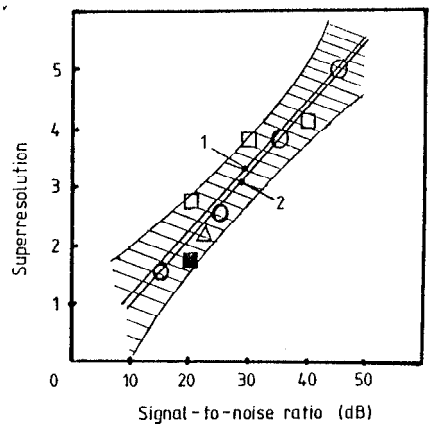

Figure 8. Summary of numerical tests for two- and three-line restoration. Open circles indicate results for the two-line Gaussian PSF; open squares indicate the two-line Lorentzian PSF; the full square indicates the three-line Lorentzian PSF; the open triangle indicates the two-line PSF in the form of $(\sin x/x)^2$. Curve 1 is the Shannon superresolution limit (67); curve 2 is the regression line for numerical tests, which coincides with the Shannon limit within the corridor of errors. The shaded strip shows the 95% confidence interval between domains above and below the superresolution limit.

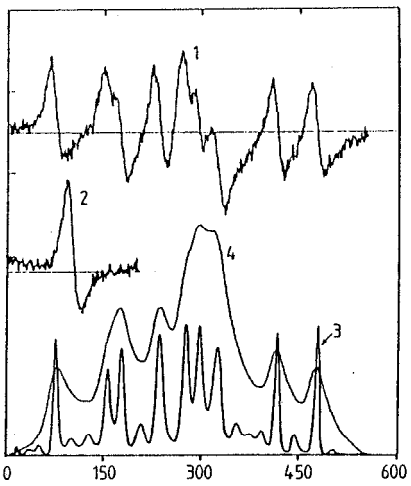

Figure 9. Application of ML method algorithm (51) to NMR data processing [35]. Curve 1 is the original experimental data for NMR absorbtion. Curve 2 is the PSF, which was determined by gluing together the left and the right tails of curve 1. Curve 3 is the restoration result for NMR absorbtion. Curve 4 is the numerical integral of curve 1.

Here we should note (together with the authors of [35]) that in this case the width and shape of the lines are apparently identical for all the spectral components, since they are primarily determined by the distribution of magnetisation in the sample (due to the large value of the susceptibility and the sample's non-ellipsoidal shape) and by a dispersive admixture to the signal due to the fact that plate thicknesses in the sample are of the order of the skin depth.

Thus we have both functions: the experimental data (curve 1) and the PSF (curve 2).

Curve 3 is the result of restoration and curve 4 is the numerical integral of curve 1.

The superresolution coefficient is equal to the ratio of the width of separate lines in curve 4 to the width in curve 3: SR ≈ 6. This value was expected from the formula (67) because in this experiment the SNR is about equal to 55 dB. All restored lines are quite distinct in this figure and fully interpreted in [35]. This example demonstrates not only the efficiency of the restoration algorithm but also reveals information in the raw experimental data which would be completely unrecognised and would be lost if modern restoration methods were not used.

6. Superresolution and uncertainty principle

From the mathematical point of view the Heisenberg uncertainty principle for coordinate and momentum in quantum mechanics, the uncertainty principle for coherence time and spectral frequency width in optical coherence theory and the uncertainty principle for any two conjugate variables in general spectral analysis are only consequences arising from the Fourier integral theory.

For readers' convenience, we outline here the basic assumptions of this principle following mainly [36, 37]. Let $f(x)$ be some considered signal and $\mathscr{F}(k)$ be its Fourier transform

$$\mathscr{F}(k) = \int_{-\infty}^{\infty} f(x)\mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}x. \tag{68}$$

We may determine the effective mean-squared range of the function $f(x)$ as

$$(\Delta x)^2 = \frac{\int_{-\infty}^{\infty}(x-x_0)^2 f^2(x)\,\mathrm{d}x}{\int_{-\infty}^{\infty} f^2(x)\,\mathrm{d}x} \tag{69}$$

where

$$x_0 = \int_{-\infty}^{\infty} x f^2(x)\,\mathrm{d}x \tag{70}$$

is the mean value† of x, and similarly for the Fourier transform $\mathscr{F}(k)$

$$(\Delta k)^2 = \frac{\int_{-\infty}^{\infty}(k-k_0)^2|\mathscr{F}(k)|^2\,\mathrm{d}k}{\int_{-\infty}^{\infty}|\mathscr{F}(k)|^2\,\mathrm{d}k} \tag{71}$$

where

$$k_0 = \int_{-\infty}^{\infty} k|\mathscr{F}(k)|^2\,\mathrm{d}k \tag{72}$$

is the mean value of k. Then, providing $k_0 = 0$, it is true by the general uncertainty principle that

$$(\Delta x)(\Delta k) \geq \mu = 1/4\pi = 0.079577\ldots \tag{73}$$

† Using the change of variables $x' = x - x_0$ we may consider $x_0 = 0$.

corresponding to the Heisenberg quantum mechanics uncertainty principle

$$(\Delta x)(\Delta p) \geq h/4\pi$$

where $h = 6.626 \times 10^{-27}$ erg s is the Planck constant.

If $k_0 \neq 0$ it was shown in [37] that

$$\frac{1}{12\pi} < \mu < \frac{\pi - 2}{4\pi^2} \qquad 0.0265\ldots < \mu < 0.0289\ldots \tag{74}$$

and hence μ is a discontinuous function of k_0 (finite discontinuity from 0.02... to 0.079... at $k_0 = 0$).

It is worth noting that our earlier definition (42) for the range of a function is suitable for a larger set of functions in comparison with the new definition (69), but using (42) instead of (69) merely changes the numerical value of constant μ.

Thus for any k_0 there is a lower limit for the product of the ranges of $f(x)$ and its Fourier transform $\mathscr{F}(k)$. This is a rigorous mathematical theorem and, of course, we do not intend to refute it.

However, there is another way which gives us a possibility not to get over but to go round the uncertainty principle.

Let $f(x)$ be a signal having a *large* range Δx and therefore a *small* uncertainty Δk of its Fourier spectrum. If we cannot directly observe the original function $f(x)$ but only its cut-down part

$$\hat{f}(x) = f(x) \cdot A(x) \tag{75}$$

where the cutting function

$$A(x) = 0 \qquad \text{for } |x| \gg \Delta A \tag{76}$$

then for the Fourier spectrum of the cut function $\hat{f}(x)$

$$\hat{\mathscr{F}}(k) = \int_{-\infty}^{\infty} f(x)A(x)e^{ikx}\,dx \tag{77}$$

we have a greater uncertainty than for the original function $f(x)$

$$\Delta\hat{\mathscr{F}} > \Delta k. \tag{78}$$

This is true by the uncertainty principle.

But, from the convolution theorem in Fourier integral theory, it follows that

$$\hat{\mathscr{F}}(k) = \int_{-\infty}^{\infty} \mathscr{F}(k')\mathscr{A}(k - k')\,dk'. \tag{79}$$

Here $\mathscr{A}(k)$ is the Fourier transform of $A(x)$

$$\mathscr{A}(k) = \int_{-\infty}^{\infty} A(x)e^{ikx}\,dx \tag{80}$$

and $\mathscr{F}(k)$ is that for $f(x)$.

The basic relation (79) between $\hat{\mathscr{F}}(k)$ and $\mathscr{F}(k)$ is the convolution integral equation of the first kind. This paper demonstrates the ultimate resolution to find the function $\mathscr{F}(k)$ from the equation (79) determined only by the noise level in the function $\hat{\mathscr{F}}(k)$. If this level is very small we can find $\mathscr{F}(k)$ from (79) with less uncertainty than $\Delta\hat{\mathscr{F}}$.

This approach, going round the uncertainty principle, opens new ways for the rigorous substantiation and practical development of effective algorithms for signal recovery in optics, superresolution antennas, spectral analysis and, apparently, in quantum mechanics.

In the last case additional problems arise in experimental measurements of complex wavefunctions and in the physical sense of noise in such measurements. The author hopes these problems can be resolved in future.

7. Conclusion

In this paper the classical Shannon theorem is shown to be closely connected to the ultimate superresolution limit of the general non-parametric signal restoration methods. It is demonstrated that the ML restoration algorithm has a supremum efficiency equal to 1 shannon.

Of course, it does not follow from this conclusion that the ML algorithm is better than any other. Much faster algorithms or those demanding less computer memory may exist but none of them could have better resolution than the Shannon resolution limit

$$\mathrm{SR} = \tfrac{1}{3}\log_2(1 + P_s/P_n).$$

This formula is the main result of this paper.

It is not shown here why the ML algorithm has a supremum resolution, neither is the mechanism of its efficiency explained. This is the subject for future research.

Acknowledgments

The author would like to thank Professors L A Vainstein and K Sh Zigangirov for their interest in this work and their critical comments and his colleagues E R Podolyak and V I Gelfgat for helping in program implementation. The author extends his thanks to the reviewers for their helpful and constructive comments.

References

[1] Gorelik G S 1952 Application of modulation method in optical interferometry *Dokl. Akad. Nauk* **83** 549–52 (in Russian)
Bernstein I L and Gorelik G S 1952 On theory of the Michelson star interferometer *Dokl. Akad. Nauk* **86** 47–50 (in Russian)
[2] Rautian S G 1958 Real spectral apparatus *Sov. Phys.–Usp.* **66** (1) 245–73
[3] Turchin V F, Kozlov V P and Malkevich M S 1971 The use of mathematical statistics methods in the solution of incorrectly posed problems *Sov. Phys.–Usp.* **13** 681–840
[4] Vainstein L A 1972 Noise filtering for numerical solution of the first kind integral equations *Dokl. Akad. Nauk* **204** 1067–70 (in Russian)
Vainstein L A 1972 On numerical solution of the first kind integral equations with using of *a priori* information on sought-for function *Dokl. Akad. Nauk* **204** 1331–34 (in Russian)
[5] Frieden B R 1979 Image enhancement and restoration *Picture Processing and Digital Filtering* ed T S Huang (Berlin: Springer)
[6] Kosarev E L 1980 Applications of the first kind integral equation in experimental physics *Comput. Phys. Commun.* **20** 69–75
[7] Tichonov A N and Arsenin V Ya 1977 *Solution of Ill-posed Problems* (New York: Wiley)
[8] Tichonov A N, Gontcharskii A V, Stepanov V V and Jagola A G 1983 *Regularization Algorithms and a priori Information* (Moscow: Fizmatgiz) (in Russian)
[9] Mendel J M 1983 *Optimal Seismic Deconvolution* (New York: Academic)
[10] Jansson P A 1984 *Deconvolution with Application in Spectroscopy* (New York: Academic)
[11] Vasilenko G I and Taratorkin A M 1986 *Restoration of Images* (Moscow: Radio i swjaz) (in Russian)
[12] Pollon G E and Lank G E 1968 Angular tracking of two closely spaced radar targets *IEEE Trans. Aerospace Electron. System* **AES-4** 541–50
[13] Gerchberg R W 1974 Superresolution through error energy reduction *Opt. Acta* **21** 709–20
[14] Gabriel W F 1980 Spectral Analysis and Adaptive Array Superresolution Techniques *Proc. IEEE* **68** 654–66

[15] Shannon C E 1949 Communication in the presence of noise *Proc. IRE* **37** 10–21
[16] Toraldo di Francia G 1955 Resolving power and information *J. Opt. Soc. Am.* **45** 497–501
[17] Fellgett P B and Linfoot E H 1955 On the assesment of optical images *Phil. Trans. R. Soc.* Ser. A **247** 369–407
[18] Helstrom C W 1967 Image restoration by the method of least squares *J. Opt. Soc. Am.* **57** 297–303
[19] Rushfort C K and Harris R W 1968 Restoration, resolution and noise *J. Opt. Soc. Am.* **58** 539–45
[20] Bershad N J 1969 Resolution, optical-channel capacity and information theory *J. Opt. Soc. Am.* **59** 157–63
[21] Khalfin L A 1969 On resolving power of optical devices *Opt. Spektrosk.* **26** 1065–7 (in Russian)
[22] Khurgin Ja I and Jakovlev V P 1971 *Finite Functions in Physics and Techiques* (Moscow: Fizmatgiz) (in Russian)
[23] Stremel M A 1972 The limits of diffractometric analysis possibilities in fine structure measurements *Dokl. Akad. Nauk* **203** 570–3 (in Russian)
[24] Vainstein L A and Vakman D E 1983 *Resolution of Frequences in Theory of Oscillations and Waves* (Moscow: Fizmatgiz) (in Russian)
[25] Cathey W T, Frieden B R, Rhodes W T and Rushforth C K 1984 Image gathering and processing for enhanced resolution *J. Opt. Soc. Am.* A **1** 241–50
[26] Bakut P A, Derjugin A I and Kuraschov V N 1985 Image restoration of partly coherently quasi-homogeneous sources *Radiotechnika Electronika* **30** 1119–25 (in Russian)
[27] Karavaev V V and Molodtzov V S 1987 Accuracy characteristics of superresolution antenna *Radiotechnika Electronika* **32** 22–6 (in Russian)
[28] Kharkevitch A A 1955 *Outlines of General Communication Theory* (Moscow: Gostechizdat) (in Russian)
[29] Wozenkraft J M and Jacobs I M *Principles of Communication Engineering (New York: Wiley) ch 5*
[30] Kosarev E L and Pantos E 1983 Optimal smoothing of 'noisy' data by fast Fourier transform *J. Phys. E: Sci. Instrum.* **16** 537–43
[31] Kolmogorov A N 1956 Theory of information transmission *Theory of Information and Theory of Algorithms* Collected papers (Moscow: Nauka 1987) (in Russian)
[32] Zelen M and Severo N C 1972 Probability functions *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* ed M Abramowitz and I A Stegun (New York: Wiley)
[33] Kosarev E L, Peskov V D and Podolyak E R 1983 High resolution soft x-ray spectrum reconstruction by MWPC attenuation measurements *Nucl. Instrum. Methods* **208** 637–45
[34] Tarasko M Z 1969 On the one method for solution of the linear system with stochastics matrixes *Preprint* Physics and Energetics Institute, Obninsk, PEI-156 (in Russian)
[35] Alekseevskii N E and Nikolaev E G 1986 Nuclear magnetic resonance in the heavy-fermion superconductor UBe_{13} *Sov.Phys.-JETP* **64** (5) 1078–84
[36] Born M and Wolf E 1968 *Principles of Optics* (Oxford: Pergamon) section 10.7.3
[37] Mayer A G and Leontovitch E A 1934 On some inequality relating to Fourier's integral *Sov. Math. Dokl.* **4** 353–60