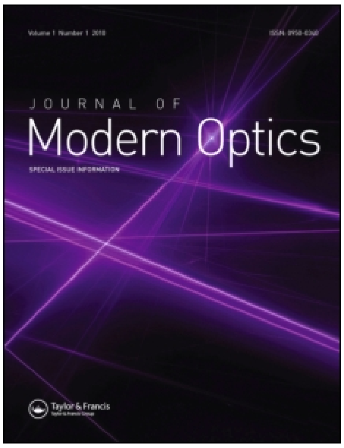

Journal of Modern Optics
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713191304

Resolution Beyond the Diffraction Limit for Regularized Object Restoration

M. Bertero[a]; G. A. Viano[a]; C. de Mol[b]
[a] Istituto di Scienze Fisiche dell'Università, Genova, Italy and Istituto Nazionale di Fisica Nucleare, Sezione di Genova, Italy. [b] Département de Mathématique, Université Libre de Bruxelles, Belgium.

To cite this Article Bertero, M. , Viano, G. A. and de Mol, C.(1980) 'Resolution Beyond the Diffraction Limit for Regularized Object Restoration', Journal of Modern Optics, 27: 3, 307 — 320
To link to this Article: DOI: 10.1080/713820228
URL: http://dx.doi.org/10.1080/713820228

OPTICA ACTA, 1980, VOL. 27, NO. 3, 307–320

Resolution beyond the diffraction limit for regularized object restoration

M. BERTERO and G. A. VIANO
Istituto di Scienze Fisiche dell'Università, Genova, Italy
and Istituto Nazionale di Fisica Nucleare, Sezione di Genova, Italy

and C. DE MOL
Département de Mathématique, Université Libre de Bruxelles, Belgium

(*Received* 2 *July* 1979)

Abstract. We propose a new formulation of Miller's regularization theory, which is particularly suitable for object restoration problems. By means of simple geometrical arguments, we obtain upper and lower bounds for the errors on regularized solutions. This leads to distinguish between ' Hölder continuity ' which is quite good for practical computations and ' logarithmic continuity ' which is very poor. However, in the latter case, one can reconstruct local weighted averages of the solution. This procedure allows for precise valuations of the resolution attainable in a given problem. Numerical computations, made for object restoration beyond the diffraction limit in Fourier optics, show that, when logarithmic continuity holds, the resolution is practically independent of the data noise level.

1. Introduction

Recently, many *linear* inverse problems have been considered in various fields. Let us mention a few examples : object reconstruction from radiographs (transaxial tomography) [1, 2] ; epicardial potential calculation from body surface maps [3] ; radar target shape estimation [4–6] ; near-field reconstruction from the scattered far-field [7, 8] and stepwise analytic continuation in order to identify an unknown scatterer [7, 9]. Further examples can be found in [10].

The mathematical formulation of the previous problems is the following : find a function f such that a known *linear* operator A transforms f into a given function g. Most often these problems are *improperly posed*, i.e., in a mathematical language, the inverse operator A^{-1} is not continuous. In practice an improperly posed problem, when discretized for numerical computation, presents instability. The remedy is to make a guess on some of the properties of the function f which has to be identified. This is the basis of regularization methods whose relevance for linear inverse problems has been already emphasized by many authors [11–14, 8]. In particular [11] is also a good tutorial paper, and we will try to derive our results without going beyond the mathematical tools used in that paper.

Various regularization methods can be found in the literature [15–21]. They are essentially equivalent in practice, since they always lead to a linear estimate of the unknown function f in terms of the data function g. Besides it is possible to prove [22] that probabilistic methods [19, 21] (Wiener filters) and functional analysis methods [15, 18, 20] lead to formally equivalent results.

An important question is to estimate bounds for error propagation in the inversion procedure. To this purpose we find particularly convenient to use the formulation of regularization theory due to Miller [18], and applied later by Miller and Viano [23] to problems of analytic continuation. Using this theory we have shown in previous papers [22, 24] the importance of the concept of *logarithmic continuity* for many linear inverse problems ; i.e. of the fact that the error on the restored solution is at best proportional to an inverse power of $|\ln \epsilon|$ where ϵ is the data error level.

The purpose of the present paper is twofold : firstly we present a simplified but efficient formulation of Miller regularization theory ; secondly we apply the theory to object restoration problems in order to discuss the improvement of resolution attainable by means of regularization methods. As in previous papers [22, 24], we focus on the problem of the restoration of coherent objects from their images through a diffraction limited optical system (perfect lowpass filter). For this problem accurate numerical calculations can be done. Besides, in this case, the improvement of resolution, due to the use of inversion techniques, has been already widely discussed [25, 26], so that a comparison of our results with computational practice is rather easy.

For one-dimensional objects, identically zero outside the interval $[-1, 1]$, the problem reduces to the inversion of the following linear integral operator :

$$(Af)(x) = \int_{-1}^{1} \frac{\sin [c(x-y)]}{\pi(x-y)} f(y)\, dy, \quad |x| \leqslant 1. \tag{1.1}$$

$R = \pi/c$ is the Rayleigh resolution distance. In order to specify the problem more precisely we must define the sets F and G to which the object f and the image g belong. Besides suitable norms have to be introduced in F and G (a simple and clear discussion of these points is contained in [11]). If we assume that both the object and the image have finite energy, then both sets F and G are spaces of square integrable functions on the interval $[-1, 1]$, i.e. $F = G = L^2(-1, 1)$. We denote by (f, h) the usual scalar product of two functions of L^2 :

$$(f, h) = \int_{-1}^{1} f(x) h^*(x)\, dx \tag{1.2}$$

and by $\|f\|$ the norm of a function f of L^2 :

$$\|f\|^2 = \int_{-1}^{1} |f(x)|^2\, dx. \tag{1.3}$$

In §§ 2 and 3 we will reformulate Miller's regularization theory deriving the main results by means of simple geometrical arguments. In particular we focus on the estimation of the restoration errors. In § 4 we discuss the problem of the inversion of the operator A, equation (1.1). Roughly speaking, we may summarize our results as follows : if a reasonable definition of resolution is introduced, then, by means of regularization methods, it is quite easy to get a resolution of about $R/2$. An improvement beyond this limit seems to be practically impossible, since too high signal-to-noise ratios would be required. In this fact we see essentially a consequence of the property of logarithmic continuity intrinsic to the problem.

In the concluding remarks we discuss the extension of our results to other linear inverse problems.

2. Formulation of regularization theory

In this section we give a revisited version of Miller's regularization theory, having in mind the problem of inverting the operator A, equation (1.1). Indeed this operator has rather peculiar properties : it is self-adjoint, its eigenvalues are positive and decreasing to zero, its eigenfunctions (the linear prolate spheroidal wave functions) form an orthonormal basis in $L^2(-1, 1)$ [27].

We can now formulate the problem as follows : given a data function (noisy image) g, find a function f of F such that Af, equation (1.1), is approximately equal (in the sense of the norm of G) to g, i.e.

$$\|Af-g\| \leqslant \epsilon, \tag{2.1}$$

where ϵ is an estimate of the level of errors or noise. Since the eigenvalues of A tend to zero (i.e. the inverse operator A^{-1} is not continuous), it is easy to see that the set of the functions f satisfying condition (2.1) is *not* bounded. In other words, given an arbitrary positive number M, one can find two functions f_1, f_2, satisfying (2.1) and such that $\|f_1-f_2\|>M$. This fact is the main reason of the numerical instability arising when the problem is discretized for numerical calculations [16, 17].

In order to get numerical stability, one must restrict the class of admissible solutions by means of *a priori* bounds (as far as possible of physical origin) ; i.e. one has to guess some properties of the function f to be identified. According to Miller [18] we consider bounds of the following type (for a short discussion and comparison with other regularization methods see also [24]) :

$$\|Bf\| \leqslant 1, \tag{2.2}$$

where B is a linear operator whose inverse is bounded. Then the set K containing all the functions satisfying conditions (2.1), (2.2) is bounded. When its size is not too large, any function of K can be taken as a satisfactory estimate of the unknown object.

We have now the following problems : (1) how to exhibit at least one function of K ; (2) how to estimate the accuracy of the solution with respect to some given definition of the ' closeness ' of two functions.

To this purpose we introduce two sets K_0, K_1, sandwiching K. Indeed, if we consider the functional

$$\Phi(f)=\|Af-g\|^2+\epsilon^2\|Bf\|^2 \tag{2.3}$$

then it is obvious that the set K_0 of the functions f (if any) satisfying the condition $\Phi(f)\leqslant\epsilon^2$ is contained in K, while the set K_1 of the functions f such that $\Phi(f)\leqslant\epsilon^2$ contains K.

The sets K_0, K_1 have a simpler geometrical structure than the set K. Indeed, let us consider the operator

$$C=A^*A+\epsilon^2\,B^*B \tag{2.4}$$

where A^*, B^* are the adjoints of A, B (for the operator A of equation (1.1), $A^*=A$). The operator C has a bounded inverse, as a consequence of the assumption that B has a bounded inverse [18]. It follows that, for any given g, we can introduce the function :

$$\tilde{f}=C^{-1}\,A^*g. \tag{2.5}$$

Then the functional (2.3) can be written in the form

$$\Phi(f)=(C[f-\tilde{f}],[f-\tilde{f}])+\|g\|^2-(g,A\tilde{f}). \tag{2.6}$$

The operator C is self-adjoint and positive definite so that its eigenvalues are positive. If we assume that it has a complete orthonormal set of eigenfunctions (this assumption is not essential but simplifies the discussion of the problem), then the condition $\Phi(f)\leqslant a^2$ can be written in the following form :

$$\sum_{k=0}^{+\infty}\gamma_k(\epsilon)|f_k-\tilde{f}_k|^2\leqslant a^2+(g,A\tilde{f})-\|g\|^2 \tag{2.7}$$

where $\{\gamma_k(\epsilon)\}$ is the set of the eigenvalues of C and f_k, $\tilde{f}_k$ are the Fourier components of f, $\tilde{f}$ with respect to the basis $\{\phi_k\}$ of the eigenfunctions of C; i.e. $f_k=(f,\phi_k)$, $\tilde{f}_k=(\tilde{f},\phi_k)$.

At this point it is clear that the sets K_0, K_1 are infinite-dimensional ' ellipsoids ' having the same centre $\tilde{f}$ and the same principal axes, these being given by the eigenvectors of the operator C.

Now, does $\tilde{f}$ belong to the set K ? The answer is rather simple. Clearly $\tilde{f}$ is the function that minimizes Φ, and $\Phi(\tilde{f})=\|g\|^2-(g,A\tilde{f})\geqslant 0$. Then a sufficient condition for $\tilde{f}\in K$ is : $\Phi(\tilde{f})\leqslant\epsilon^2$. This condition is also a ' compatibility check ' for the conditions (2.1), (2.2), i.e. it ensures that there exists at least one function satisfying both conditions. It is more restrictive (for a factor of $\sqrt{2}$) than a compatibility check given by Miller [18]. Besides the set K_0 is non-void if and only if the condition $\Phi(\tilde{f})\leqslant\epsilon^2$ is satisfied. In such a case the situation is represented in figure 1. It is now clear that we may take $\tilde{f}$ as an estimate of the unknown solution (for a discussion of the relations between $\tilde{f}$ and Tikhonov regularized solution, Wiener filter etc., see [22], [24]). Moreover, it is clear that the sets K_0, K_1 can be used in order to find upper and lower bounds for the error on the restored solution. This point will be discussed in the next section. For simplicity we shall consider only the case where the spectrum of the self-adjoint operator C is discrete. However the extension to the case of a continuous spectrum is straightforward from the mathematical point of view.

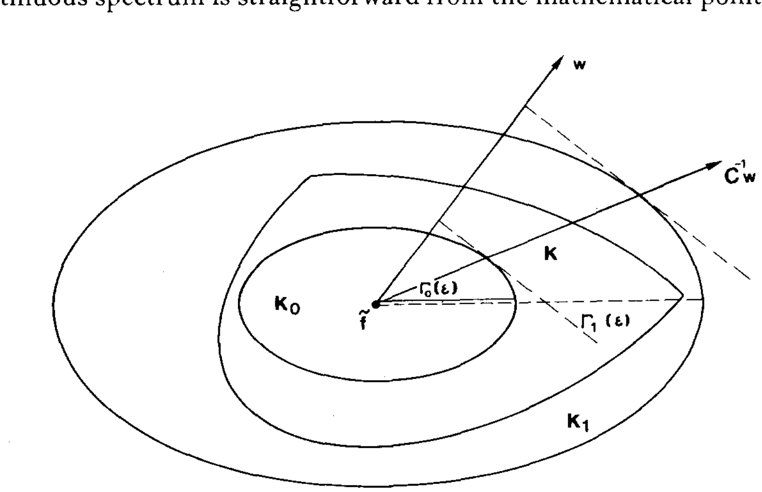

Figure 1. Schematic representation of the relation between the sets K_0, K, K_1. The sets K_0, K_1 are represented as two homothetic ellipses with centre $\tilde{f}$.

3. Estimation of restoration errors

3.1. *Absolute mean square errors*

We define the absolute mean square error as the maximum value taken by $\|f-\tilde{f}\|$ in K, the norm being defined by equation (1.3). We write:

$$\mathscr{E}_K(\epsilon)=\sup_{f\in K}\|f-\tilde{f}\|. \tag{3.1}$$

In figure 1, $\mathscr{E}_K(\epsilon)$ is the maximum distance between any point of K and $\tilde{f}$; it is clear that this maximum value is attained at the boundary of K.

Let us now denote by $\Gamma_0(\epsilon)$ and $\Gamma_1(\epsilon)$ the maximum length of the axes of K_0 and K_1 respectively. Then by looking at figure 1, we see immediately that

$$\Gamma_0(\epsilon)\leqslant\mathscr{E}_K(\epsilon)\leqslant\Gamma_1(\epsilon). \tag{3.2}$$

The quantities $\Gamma_0(\epsilon)$, $\Gamma_1(\epsilon)$ may be easily related to the data g and to the eigenvalues of C. Indeed, let us denote by $\gamma(\epsilon)$ the *smallest* eigenvalue of C, i.e. $\gamma(\epsilon)=\inf_k\{\gamma_k(\epsilon)\}$. Then, from equation (2.7) we get

$$\Gamma_i(\epsilon)=\left(\frac{1}{\gamma(\epsilon)}\left[a_i^2+(g,A\tilde{f})-\|g\|^2\right]\right)^{1/2},\quad i=0,1 \tag{3.3}$$

where $a_0^2=\epsilon^2$ and $a_1^2=2\epsilon^2$. Therefore $\Gamma_0(\epsilon)$ and $\Gamma_1(\epsilon)$ can be computed in practical cases and by means of equation (3.2) we get respectively a lower and an upper bound on the absolute mean square error $\mathscr{E}_K(\epsilon)$.

If we recall that the condition $\Phi(\tilde{f})\geqslant 0$ is equivalent to: $(g,A\tilde{f})-\|g\|^2\leqslant 0$, then equations (3.2) and (3.3) show that there exists a bound for $\mathscr{E}_K(\epsilon)$ independent of the data function g. More precisely, we have $\mathscr{E}_K(\epsilon)\leqslant\mathscr{E}(\epsilon)$ where

$$\mathscr{E}(\epsilon)=\sqrt{\left(\frac{2\epsilon^2}{\gamma(\epsilon)}\right)}. \tag{3.4}$$

The quantity $\mathscr{E}(\epsilon)$ is called by Miller [18] *stability estimate*. The reason for choosing this name is clear: if $\mathscr{E}(\epsilon)\to 0$ for $\epsilon\to 0$, then the error on the solution of the inverse problem tends also to zero (this corresponds to the collapse of the ellipses K_0, K_1 in figure 1 into a point). In other words, if the noisy image g tends to a noiseless image $g_0=Af_0$, then the estimated solution $\tilde{f}$ tends to the true solution f_0. In such a case one also says that ‘ continuous dependence of the solution on the data has been restored ’.

However, in order to know whether it is possible to get accurate results in a particular inverse linear problem, it is not enough to show that $\mathscr{E}(\epsilon)$ tends to zero; one must know *how fast* $\mathscr{E}(\epsilon)$ tends to zero. As we have already remarked elsewhere [22, 24], in various relevant inverse problems we have to distinguish between two cases: the first when $\mathscr{E}(\epsilon)\propto\epsilon^\alpha$ $(0<\alpha\leqslant 1$; Hölder continuity) and in this case continuity is rather good; the second when $\mathscr{E}(\epsilon)\propto|\ln\epsilon|^{-\alpha}$ $(\alpha>0$; *logarithmic continuity*) and in this case continuity is very poor. As regards the problem of the inversion of a linear integral operator, one can say that the situation is roughly the following: if the kernel is not too smooth (for instance, it has continuous derivatives of order less or equal to a fixed integer n) then the restored continuity is usually of the Hölder type; if the kernel is very smooth (for instance an entire function of finite order) then the restored continuity is generally of the logarithmic type [24]. The previous remark contains the essential points of the question but it is not very precise since, of course, the type of continuity is rather a property relative to the couple of operators A, B than a property relative to the operator A alone.

An elementary discussion of the last point can be done if we assume that the operator A^*A and B^*B commute. Then, let $\{\phi_k\}$ be the set of the eigenfunctions of A^*A and B^*B (and therefore also of C, equation (2.4)) and let $\{\lambda_k^2\}$ and $\{\beta_k^2\}$ be respectively the sets of their eigenvalues. Then the eigenvalues of the operator C are given by: $\gamma_k(\epsilon)=\lambda_k^2+\epsilon^2\beta_k^2$. From equation (3.4) we get

$$\mathscr{E}(\epsilon)=\sqrt{(2)}\epsilon \sup_k (\lambda_k^2+\epsilon^2\beta_k^2)^{-1/2}. \tag{3.5}$$

Now, if the β_k form an increasing sequence (at least for $k>k_0$) and $\lim \beta_k=+\infty$ for $k\to\infty$, then a simple argument shows that $\mathscr{E}(\epsilon)\to 0$ for $\epsilon\to 0$. Indeed, since the λ_k form a decreasing sequence, the denominator of equation (3.5) has a minimum for a value of the index $N=N(\epsilon)$, which tends to infinity when $\epsilon\to 0$. Then, from the inequality $\mathscr{E}(\epsilon)\leqslant\sqrt{(2)}\beta_N^{-1}$, it follows $\mathscr{E}(\epsilon)\to 0$. More precise results can be obtained if stronger assumptions are done on the β_k. For instance, if we assume that $\beta_k\propto\lambda_k^{-\mu}$ $(\mu>0)$, then we find that the denominator of equation (3.5), which is a function of the squared eigenvalues like $f(t)=t+\epsilon^2 t^{-\mu}$ $(t=\lambda_k^2)$, has a minimum for $t=(\mu\epsilon^2)^a$, $a=(\mu+1)^{-1}$. By means of elementary calculations one finds that $\mathscr{E}(\epsilon)\propto\epsilon^\alpha$, $\alpha=\mu/(\mu+1)$. Therefore we have Hölder continuity. Logarithmic continuity is obtained when the eigenvalues λ_k decrease exponentially fast, while the β_k grow like a power of k: $\beta_k\propto k^\mu$ $(\mu>0)$ [22, 24].

Finally we remark that we cannot restore the continuity by choosing bounded β_k. Indeed, if we take for instance $\beta_k=1$ (i.e. $B=1$) in equation (3.5), then we have, for any ϵ, $\mathscr{E}(\epsilon)=\sqrt{2}$ and therefore $\mathscr{E}(\epsilon)$ does not tend to zero.

3.2. *Absolute errors on smeared objects*

The previous analysis does not say anything about the resolution attainable by means of regularized inversion methods. Only in the case of logarithmic continuity it seems to suggest that, if some resolution has been attained for some reasonable error level, say $\epsilon=10^{-3}$, then even a lowering of the noise of many orders of magnitude does not produce a great improvement of resolution.

We can analyse theoretically this point as follows. Let us try to identify not the unknown f but a smeared object given by (neglecting edge effects)

$$F(x)=\int_{-1}^{1} w(x-y)f(y)\,dy \tag{3.6}$$

where the smearing function w has the usual properties: it is positive, even, peaked upon the point $x=0$ and its integral is equal to one. We can also define a resolving length associated to w as

$$d=\left(\int_{-1}^{1} x^2\,w(x)\,dx\right)^{1/2}. \tag{3.7}$$

In other words we are considering the problem of restoring local averages of f over some resolving length d. We recall that this is just what is usually done in the reconstruction of objects from radiographs [2] or in some geophysical inverse problems [28].

When the problem is approximately space invariant, then, in order to estimate the error on $F(x)$ it is enough to estimate the error on $F(0)=(f, w)$, the scalar product being defined by equations (1.2). Since the estimate of $F(0)$, according to the theory of § 2, is $\tilde{F}(0)=(\tilde{f}, w)$, the error is

$$\mathscr{E}_K(\epsilon\,;\ w)=\sup_{f\in K}|(f-\tilde{f}, w)| \tag{3.8}$$

and therefore it is the maximum value of the component of $f-\tilde{f}$ along the direction of the vector w. If we look at figure 1, we clearly understand that $\mathscr{E}_0(\epsilon\,;\ w)\leqslant\mathscr{E}(\epsilon\,;\ w)\leqslant\mathscr{E}_1(\epsilon\,;\ w)$, where $\mathscr{E}_0(\epsilon\,;\ w)$ and $\mathscr{E}_1(\epsilon\,;\ w)$ are quantities analogous to (3.8), the supremum being taken respectively over the sets K_0 and K_1. $\mathscr{E}_0(\epsilon\,;\ w)$ and $\mathscr{E}_1(\epsilon\,;\ w)$ can be easily computed.

Indeed, if we use again figure 1 as a schematic representation of our infinite dimensional problem, we see that the component along w of a vector $h=f-\tilde{f}$ of K_0 is maximal when h coincides with that point h_0 of the boundary of K_0 such that the tangent to the ellipse at h_0 is orthogonal to w. Now, if we write the equation of the ellipse as $(Ch, h)=b^2$, then the equation of the tangent in h_0 is given by $(Ch_0, h)=b^2$ and therefore the tangent is orthogonal to the vector Ch_0. If we require this vector to be parallel to w, we get $h_0=\mu C^{-1}w$. Finally h_0 belongs to the boundary of K_0 if $\mu=b(C^{-1}w, w)^{-1/2}$, so that: $\mathscr{E}_0(\epsilon\,;\ w)=|(h_0, w)|=b(C^{-1}w, w)^{1/2}$. Using a similar argument for K_1 and recalling equation (2.7), we have

$$\mathscr{E}_i(\epsilon\,;\ w)=(a_i^2+(g, A\tilde{f})-\|g\|^2)^{1/2}(C^{-1}w, w)^{1/2}, \quad i=0, 1, \tag{3.9}$$

where $a_0^2=\epsilon^2$ and $a_1^2=2\epsilon^2$. For infinite dimensional ellipsoids the previous argument can be made completely rigorous using the Schwarz inequality.

Again we can find an upper bound on the error, which is independent of g, i.e. $\mathscr{E}_K(\epsilon\,;\ w)\leqslant\mathscr{E}(\epsilon\,;\ w)$, where

$$\mathscr{E}(\epsilon\,;\ w)=\sqrt{(2)}\epsilon(C^{-1}w, w)^{1/2}. \tag{3.10}$$

This quantity coincides with the stability estimate computed by Miller ([18], Lemma 5). It is possible to prove [22] that $\mathscr{E}(\epsilon\,;\ w)$ tends to zero when $\epsilon\to 0$, provided that the constraint operator B should have a bounded inverse. Hence, in this case, we are allowed to take $B=1$. This type of stability can be called *weak stability* (or weak continuity).

3.3. *Relative errors on smeared objects*

The definition of relative errors is rather natural when stochastic regularization (Wiener filters) is considered [19, 21]. Using the formal analogy between stochastic regularization and Miller's regularization theory [22], we might introduce a quantity which could be called an estimate of the relative errors. However, let us justify this by the following argument.

The stability estimate (3.10) is the maximum value of $|(f, w)|$ under the constraint $(Cf, f)\leqslant 2\epsilon^2$. This *a posteriori* constraint is compatible with the *a priori* constraint $(B^*Bf, f)\leqslant 2$. The maximum value of $|(f, w)|$ under the *a priori* constraint is $\sqrt{2}([B^*B]^{-1}w, w)^{1/2}$ (one can use the same argument as in § 3.2). Therefore we can define as an estimate of the relative error the ratio between the *a posteriori* and the *a priori* maximum value of $|(f, w)|$, i.e.

$$\mathscr{E}_{\text{rel}}(\epsilon\,;\ w)=\epsilon\frac{(C^{-1}w,\ w)^{1/2}}{([B^*B]^{-1}w,\ w)^{1/2}}. \tag{3.11}$$

Definition (3.11) is not contradictory since $\mathscr{E}_{\text{rel}}(\epsilon\,;\ w)\leqslant 1$. Indeed, this property results from the fact that the operator C is ' greater ' than the operator $\epsilon^2[B^*B]$. In other words, for any function v in the domain of B^*B we have : $(Cv, v)\geqslant \epsilon^2(B^*Bv, v)$.

Let us analyse in more details the case $B=1$. In this case, if we denote by w_k the Fourier components of w with respect to the basis $\{\phi_k\}$ of the eigenvectors of A^*A, i.e. $w_k=(w, \phi_k)$, we have from equation (3.11)

$$\mathscr{E}_{\text{rel}}(\epsilon\,;\ w)=\frac{1}{\|w\|}\left(\sum_{k=0}^{+\infty}\frac{\epsilon^2}{\lambda_k^2+\epsilon^2}|w_k|^2\right)^{1/2}. \tag{3.12}$$

We consider now a family of smearing functions w_η such that, when $\eta\to 0$, w_η tends to the Dirac delta measure. We denote by $\mathscr{E}_{\text{rel}}(\epsilon, \eta)$ the corresponding relative error and we write equation (3.12) as follows :

$$\mathscr{E}_{\text{rel}}(\epsilon, \eta)=\left(1-\frac{1}{\|w_\eta\|^2}\sum_{k=0}^{+\infty}\frac{\lambda_k^2}{\lambda_k^2+\epsilon^2}|w_{\eta,k}|^2\right)^{1/2}. \tag{3.13}$$

Now it is not difficult to see that $\mathscr{E}_{\text{rel}}(\epsilon, \eta)\to 1$ when $\eta\to 0$ (for fixed ϵ). Indeed, the series at the right-hand side of equation (3.13) has a finite limit when $w_\eta\to\delta$ since it is bounded by the convergent series $\sum_k \lambda_k^2|\phi_k(0)|^2$; on the other hand $\|w_\eta\|\to+\infty$, as it is easy to verify. We conclude that, when $B=1$, we have 100 per cent error for a pointwise reconstruction of f, whatever be the noise level ϵ. This result is related to the remark done in § 3.1, that the constraint operator $B=1$ does not ensure continuity in the sense of mean square errors.

4. Numerical results

In this section we apply the previous general method to the problem of restoring an object when we know its image given by the perfect lowpass filter (1.1). Our purpose is to illustrate how the estimates of restoration errors provide a deeper understanding of the features of the inversion procedure.

At first we recall some well-known properties of the integral operator (1.1). It is a compact, self-adjoint, non-negative operator in $L^2(-1, 1)$. Its eigenvalues λ_k have a step behaviour : they are approximately equal to one for values of the index less than $2c/\pi$ and then fall off to zero exponentially like $\exp[-2k\ln(k/ce)]$ [29]. The eigenfunctions associated to the eigenvalues λ_k are the so-called linear prolate spheroidal wave-functions $\psi_k(x)$. They satisfy the differential equation

$$-[(1-x^2)\psi'_k(x)]'+c^2x^2\psi_k(x)=\chi_k\psi_k(x) \tag{4.1}$$

where

$$\chi_k=k(k+1)+\tfrac{1}{2}c^2+0\left(\frac{1}{k^2}\right),\quad k\to+\infty. \tag{4.2}$$

Usually the ψ_k are normalized in such a way that their norm in $L^2(-1, 1)$ is equal to $\lambda_k^{1/2}$. Then the eigenfunctions $\phi_k=\lambda_k^{-1/2}\psi_k$ form an orthonormal basis in $L^2(-1, 1)$ [27, 30].

4.1. *Mean square errors*

We consider only the case where the constraint operator B commutes with the operator A. Then condition (2.2) becomes

$$\|Bf\|^2 = \sum_{k=0}^{+\infty} \beta_k^2 |f_k|^2 \leqslant 1 \tag{4.3}$$

where the β_k are the eigenvalues of B and $f_k = (f, \phi_k) = \lambda_k^{-1/2}(f, \psi_k)$. As we have already remarked in § 3.1, if $\beta_k \to +\infty$ when $k \to +\infty$, then the stability estimate (3.5) tends to zero when $\epsilon \to 0$.

There are two choices of the β_k which are quite naturally related to the structure of the operator A. The first one is to take $\beta_k = \lambda_k^{-1}$. In this case, as it has been already remarked in a previous paper [24], we have $\mathscr{E}(\epsilon) \leqslant \sqrt{\epsilon}$ and therefore Hölder continuity holds true. However the constraint (4.3) with $\beta_k = \lambda_k^{-1}$ is very restrictive. Indeed it implies that the object f has significant Fourier components only for $k < 2c/\pi$ [24].

The second choice is to take $\beta_k = \chi_k^{1/2}$ (see equations (4.1) and (4.2)). In this case, using equation (4.1) and (4.3), we find through an integration by parts that

$$\|Bf\|^2 = \int_{-1}^{1} (1-x^2)|f'(x)|^2 \, dx + c^2 \int_{-1}^{1} x^2 |f(x)|^2 \, dx \leqslant 1. \tag{4.4}$$

Therefore we get essentially a bound on the unknown object f and on its first derivative.

Figure 2 shows $1/\mathscr{E}(\epsilon)$ as a function of $\log_{10}(1/\epsilon)$ for $c = 10$ and $\beta_k^2 = k(k+1)+1$ (see equation (4.2)). It clearly appears that $1/\mathscr{E}(\epsilon)$ grows more slowly than (constant) $\times$ $|\log_{10} \epsilon|$ so that we have logarithmic continuity. More generally one can show that logarithmic continuity arises whenever bounds are imposed only on a finite number of derivatives of the unknown object f [22].

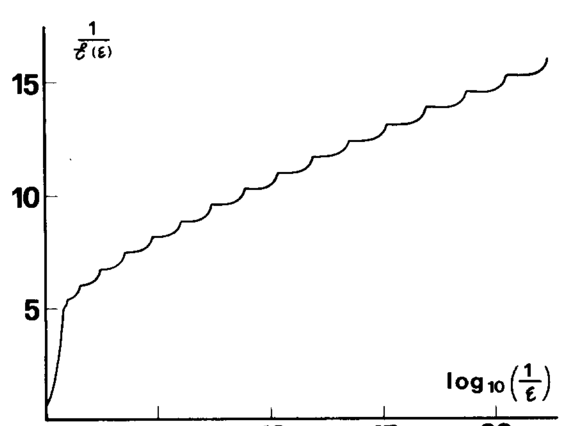

Figure 2. Stability estimate for absolute mean square errors plotted as a function of the noise level ϵ in the case $c = 10$.

4.2. Errors on smeared objects

We apply here the analysis made in §§ 3.2 and 3.3. We focus on the estimate of relative errors as defined in equation (3.11) and we consider only the case $B=1$, which is the most popular in many analysis of regularization methods [8, 11, 13, 17]. For our problem, condition (2.2) with $B=1$ has the following physical meaning : we shall restore objects whose energy does not exceed one.

When c is not too large and ϵ not too small, formula (3.13) can be used for numerical computations. We have considered the case $c=10$, corresponding to a Rayleigh resolution distance $R=0{\cdot}314$. Now, if the smearing function w is even, the odd terms are zero in the series of equation (3.13). Besides, if we choose ϵ in the range 10^{-1}–10^{-5}, we have $(\lambda_{14}/\epsilon)^2<10^{-9}$. Therefore, in order to get a sufficient accuracy, it is enough to take eight terms in the series (3.13), corresponding to $k=0, 2, \ldots, 14$.

A computer program was written using single-precision arithmetic (twelve digits). The linear prolate spheroidal wave-functions for $c=10$ have been computed by means of their expansion as a series of Legendre polynomials [31], the series being truncated in order to have eight significant digits. For the computation of the Fourier coefficients the Gauss–Legendre quadrature method was used.

First we analyse the dependence of the relative error on the smearing function w_η. We have considered the following cases :

$$w_\eta^{(1)}(x)=N^{(1)}\theta\left(\frac{x}{D}\right)\operatorname{sinc}\left(\frac{x}{D}\right),\qquad w_\eta^{(2)}(x)=N^{(2)}\theta\left(\frac{x}{D}\right)\operatorname{sinc}^2\left(\frac{x}{D}\right)$$

$$w_\eta^{(3)}(x)=N^{(3)}\exp\left(-x^2/2D^2\right),\qquad w_\eta^{(4)}(x)=N^{(4)}\theta\left(\frac{x}{D}\right)\left(1-\frac{|x|}{D}\right)$$

$$w_\eta^{(5)}(x)=N^{(5)}\theta\left(\frac{x}{D}\right)$$

where $\theta(t)$ denotes the function which is 1 for $|t|<1$ and 0 for $|t|>1$, $\operatorname{sinc}(t)=\sin(\pi t)/\pi t$ and the $N^{(i)}$ $(i=1, \ldots, 5)$ are normalization constants such that the integral of $w_\eta^{(i)}$ over the interval $[-1, 1]$ is 1. For each function $w_\eta^{(i)}$, the resolution parameter η has been defined as $\eta=d/R$, where d is the resolving distance given by equation (3.7).

In figure 3 we give the relative error as a function of η, in the case $\epsilon=10^{-2}$ and for the previously defined smearing functions. We recall that, as follows from the remark done in § 3.3, the value of $\mathscr{E}_{\rm rel}(\epsilon, \eta)$ for $\eta=0$ is always 1. Now, as we see, the curves corresponding to different smearing functions have a common feature : they are rapidly decreasing up to a value of η of about 0·5 and then become rather flat. This typical behaviour seems to suggest that, for $\epsilon=10^{-2}$ and independently of the smearing function, a resolution of about $R/2$ can be obtained. Of course, along these lines, a more precise definition of resolution would require a specification firstly of the smearing function and then of the acceptable restoration error. If we compare the effect of two smearing functions like, for instance, $w_\eta^{(2)}$ and $w_\eta^{(5)}$, then it is clear that $w_\eta^{(5)}$ is less smoothing than $w_\eta^{(2)}$ and therefore it is more discriminating than $w_\eta^{(2)}$. In other words a greater restoration error has to be acceptable in the case of $w_\eta^{(5)}$, and this agrees with the relative position of the curves for $w_\eta^{(2)}$ and $w_\eta^{(5)}$ in figure 3.

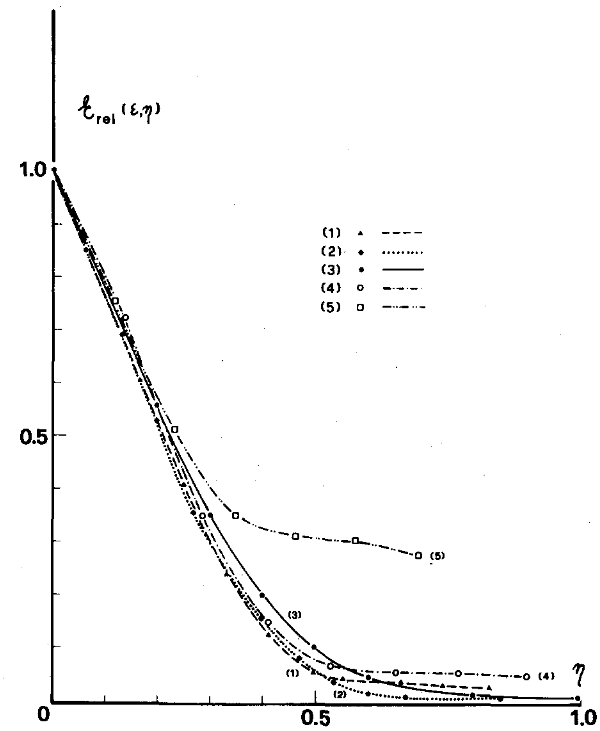

Figure 3. Relative error versus the resolution parameter $\eta = d/R$ ($c = 10$) for $\epsilon = 10^{-2}$ and for various smearing functions. The label of a curve coincides with the label of the corresponding smearing function, as defined in the text.

We have also analysed the ϵ-dependence of the relative error for a fixed smearing function. We have chosen a gaussian smearing (i.e. function $w_\eta^{(3)}$). Computations have been done for ϵ ranging in the interval 10^{-1}–10^{-5}. Some results are reported in figure 4.

If we consider as acceptable an error on the smeared object of about 10 per cent (this error corresponds to $\eta = 0{\cdot}5$ in the case $\epsilon = 10^{-2}$), then from figure 4 we see that for $\epsilon = 10^{-5}$ we obtain a value of the resolution parameter which is approximately 0·43. Therefore, a lowering of the noise of three orders of magnitude gives an improvement in resolution of about 14 per cent. This result is in a qualitative agreement with previous analysis [25, 32, 33]. However, for greater values of c, we can presume that the resolution does depend still more weakly on the noise. Indeed it has been shown, by means of a very simple and nice argument [32], that in the case $2c/\pi = 10^4$ a lowering of the noise by three orders of magnitude gives an increase of only 0·15 per cent in the amount of available information.

We have tested this assumption by computing relative errors in the case $c = 20$, $\epsilon = 10^{-2}$. We found values lying above the curve obtained for $c = 10$, $\epsilon = 10^{-2}$ (see figure 5). For this computation we used equation (3.11) (with $B = 1$) and the simple trapezoidal representation to discretize functions (181 points). This method is not very accurate since the function $C^{-1} w$ is strongly oscillating. However we have tested the method in the case $c = 10$, $\epsilon = 10^{-2}$,

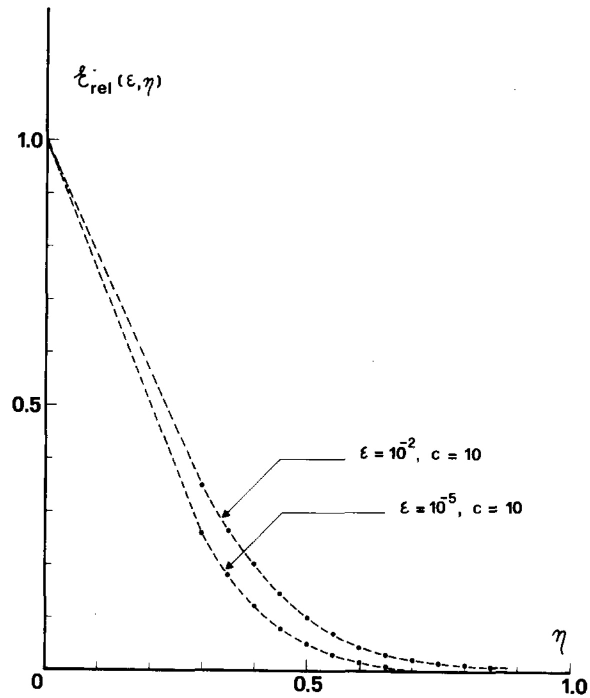

Figure 4. Relative error versus the resolution parameter $\eta = d/R$ for $\epsilon = 10^{-2}$ and $\epsilon = 10^{-5}$ ($c = 10$ and gaussian smearing function).

10^{-3}, and we found agreement with the previous results within about 2 per cent. By means of this method we have also computed the relative errors for restoration of incoherent objects, i.e. for the inversion of the integral operator

$$(Af)(x) = \frac{\pi}{c} \int_{-1}^{1} \left[\frac{\sin [c(x-y)]}{\pi(x-y)} \right]^2 f(y)\, dy, \quad |x| \leqslant 1. \tag{4.5}$$

Since the kernel of equation (4.10) is an entire function of finite order, then the corresponding eigenvalues have an exponential tail [34] and therefore we expect a resolution weakly dependent on the noise level ϵ. Computations have been done for $c = 10$, 20, $\epsilon = 10^{-2}$ and gaussian smearing functions. Results are reported in figure 5.

5. Conclusions

The example considered in this paper is quite simple so that the interpretation of the results is rather easy. They show that regularization methods for improperly posed problems are powerful but not miraculous. Indeed, these methods never could remedy to a fundamental lack of information, but should allow for an optimal use of available *a priori* knowledge. In particular, whenever logarithmic continuity holds true, it should be possible to estimate a resolution, attainable in the inversion procedure, which is rather weakly dependent on the data accuracy. This is the case for restoration of coherent and incoherent objects as well as, for instance, for the problem of near-field reconstruction [8].

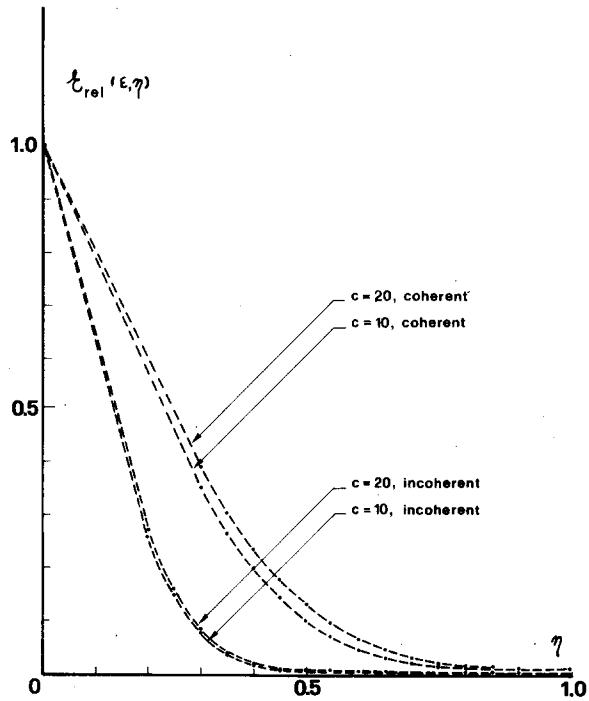

Figure 5. Relative errors versus the resolution parameter η for coherent and incoherent illumination ($\epsilon = 10^{-2}$).

This result does not exclude that a significant improvement in resolution might be obtained when other supplementary constraints, like for instance positivity, can be imposed on the solution of the problem. In the case of restoration of incoherent objects many methods have been proposed taking into account positivity [26], however, as far as we know, no rigorous mathematical analysis of restoration errors has been done.

Finally we want to remark that the previously discussed limitations on resolution improvement do not apply to those inverse problems for which Hölder continuity holds true. This happens when the kernel of the integral operator is not strongly smoothing, like, for instance, in the case of object reconstruction from projections [1, 2, 6].

Acknowledgments

The authors wish to thank Professor W. M. Boerner for many suggestions and references.

It is also a pleasure to thank Professors A. Consortini, L. Ronchi, G. Toraldo di Francia and G. Talenti for very helpful discussions.

One of us (M.B.) is indebted to Professor J. Reignier for the kind hospitality at the Vrye Universiteit Brussel, where computations have been done. He is also indebted to the Solvay Institute for Physics and Chemistry for a grant.

Christine De Mol is an Aspirant du Fonds National Belge de la Recherche Scientifique.

Wir schlagen eine neue Formulierung der Millerschen Regularisierungstheorie vor, welche sich insbesondere für Objektrestaurationsprobleme eignet. Mit Hilfe einfacher geometrischer Argumente erhalten wir untere und obere Grenzen für die Fehler der regularisierten Lösungen. Dadurch kann zwischen der ' Hölder Kontinuität ', die sich für praktische Rechnungen recht gut eignet und der ' logarithmischen Kontinuität ' welche ziemlich schlecht ist, unterschieden werden. Allerdings kann man im letzteren Fall lokal gewichtete Mittel der Lösungen rekonstruieren. Dieses Verfahren ermöglicht präzise Abschätzungen der bei einem gegebenen Problem erzielbaren Auflösung. Numerische Rechnungen zur Objektrestauration jenseits der Beugungsgrenze der Fourieroptik zeigen, daß im Falle logarithmischer Kontinuität die Auflösung praktisch unabhängig vom Rauschpegel der Daten ist.

References

[1] SMITH, K. T., SOLOMON, D. C., and WAGNER, S. L., 1977, *Bull. Am. math. Soc.*, **83,** 1227.

[2] BARRETT, H. H., and SWINDELL, W., 1977, *Proc. Inst. elect. electron. Engrs*, **65,** 89.

[3] COLLI FRANZONE, P., TACCARDI, B., and VIGANOTTI, C., 1977, *Adv. Cardiol.*, **21,** 167.

[4] KENNAUGH, E. M., and MOFFAT, D. L., 1965, *Proc. Inst. elect. electron. Engrs*, **53,** 893.

[5] LEWIS, R. M., 1969, *I.E.E.E. Trans. Antennas Propag.*, **17,** 308.

[6] DAS, Y., and BOERNER, W. M., 1978, *I.E.E.E. Trans. Antennas Propag.*, **26,** 274.

[7] IMBRIALE, W. A., and MITTRA, R., 1970, *I.E.E.E. Trans. Antennas Propag.*, **18,** 633.

[8] CABAYAN, H. S., MURPHY, R. C., and PAVLASEK, T. J. F., 1973, *I.E.E.E. Trans. Antennas Propag.*, **21,** 346.

[9] AHLUWALIA, H. P. S., and BOERNER, W. M., 1974, *I.E.E.E. Trans. Antennas Propag.*, **22,** 673.

[10] COLIN, L. (editor), 1971, *Mathematics of Profile Inversion*, NASA Technical Memorandum NASA TM–X–62, p. 150.

[11] DESCHAMPS, G. A., and CABAYAN, H. S., 1972, *I.E.E.E. Trans. Antennas Propag.*, **20,** 268.

[12] BARAKAT, R., and BLACKMAN, E., 1973, *Optics Commun.*, **9,** 252.

[13] PERRY, W. L., 1974, *I.E.E.E. Trans. Antennas Propag.*, **22,** 826.

[14] VIANO, G. A., 1976, *J. Math. Phys.*, **17,** 1160.

[15] TIKHONOV, A., 1963, *Soviet Math.*, **4,** 1035.

[16] PHILLIPS, B. L., 1962, *J. Ass. comput. Mach.*, **9,** 84.

[17] TWOMEY, S., 1965, *J. Franklin Inst.*, **279,** 95.

[18] MILLER, K., 1970, *SIAM J. Math. Analysis*, **1,** 52.

[19] FRANKLIN, J. N., 1970, *J. math. Analysis Applic.*, **31,** 682.

[20] NASHED, M. Z., and WAHBA, G., 1974, *SIAM J. Math. Analysis*, **5,** 974.

[21] BERTERO, M., and VIANO, G. A., 1978, *Boll. Un. mat. ital.* B, **15,** 483.

[22] BERTERO, M., DE MOL, C., and VIANO, G. A., 1979, *J. Math. Phys.*, **20,** 509.

[23] MILLER, K., and VIANO, G. A., 1973, *J. Math. Phys.*, **14,** 1037.

[24] BERTERO, M., DE MOL, C., and VIANO, G. A., 1978, *Optics Lett.*, **3,** 51.

[25] GOODMAN, J. W., 1970, *Progress in Optics*, Vol. VIII, edited by E. Wolf (Amsterdam : North-Holland), p. 1.

[26] FRIEDEN, B. R., 1975, *Picture Processing and Digital Filtering*, edited by T. S. Huang (Berlin, Heidelberg, New York : Springer-Verlag), p. 177.

[27] SLEPIAN, D., and POLLAK, H. O., 1961, *Bell. Syst. tech. J.*, **40,** 43.

[28] BACKUS, G., and GILBERT, F., 1968, *Geophys. J. R. astr. Soc.*, **16,** 169.

[29] LANDAU, H. J., 1965, *Trans. Am. math. Soc.*, **115,** 242.

[30] FRIEDEN, B. R., 1971, *Progress in Optics*, Vol. IX, edited by E. Wolf (Amsterdam : North-Holland), p. 311.

[31] FLAMMER, C., 1957, *Spheroidal Wave Functions* (Stanford University Press).

[32] TORALDO DI FRANCIA, G., 1969, *Riv. nuovo Cim.*, **1,** 460.

[33] RUSHFORTH, C. K., and HARRIS, R. W., 1968, *J. opt. Soc. Am.*, **58,** 539.

[34] HILLE, E., and TAMARKIN, J. D., 1931, *Acta math., Stockh.*, **57,** 1.