On the problems of object restoration and image extrapolation in optics

M. Bertero
Istituto di Scienze Fisiche dell'Università and Istituto Nazionale di Fisica Nucleare, Genova, Italy

C. De Mol
Département de Mathématique, Université Libre de Bruxelles, Belgium

G. A. Viano
Istituto di Scienze Fisiche dell'Università and Istituto Nazionale di Fisica Nucleare, Genova, Italy
(Received 22 November 1977)

In this paper we consider the problems of object restoration and image extrapolation, according to the regularization theory of improperly posed problems. In order to take into account the stochastic nature of the noise and to introduce the main concepts of information theory, great attention is devoted to the probabilistic methods of regularization. The kind of the restored continuity is investigated in detail; in particular we prove that, while the image extrapolation presents a Hölder type stability, the object restoration has only a logarithmic continuity.

1. INTRODUCTION

In Fourier optics a vast amount of literature has been devoted to the problem of object restoration. The interested reader is referred to the review papers by Frieden and Goodman.

One of the points, which has been largely debated, concerns restoration beyond the diffraction limit in the presence of noise. As already stated a significant improvement in resolution can be accomplished if the object is very poorly resolved by the optical system at the start; on the other hand, if the object is extremely complex at the start, improvement of resolution requires signal-to-noise ratios that are unrealistically high. One might argue that these limitations are due to the fact that the number of degrees of freedom of an image is finite. In other words the image is “ambiguous” in the sense that many different objects can have one and the same image (within a given accuracy) and therefore the observer must necessarily make use of some a priori knowledge.

This remark is in accordance with the fact that the object restoration belongs to a large class of linear inverse problems (relevant in many fields like identification of targets, probing of media, geophysical exploration, and in any other field of remote sensing) which are ill posed in the sense of Hadamard, since the solutions do not depend continuously on the data. Therefore, as it has been emphasized by many authors, these problems must be reconsidered according to the regularization theory of improperly posed problems. This is a means for introducing a priori knowledge under the form of precise a priori constraints (as far as possible of physical origin) which restrict the class of admitted solutions. Then regularization gives a continuous dependence of the solution on the data in the sense that a small variation of the data leads to a small change in the solution.

In this paper we consider an ideal, diffraction limited, space-invariant, imaging system. For one-dimensional, coherent objects, identically zero outside the interval $[-1,1]$, one has

$$y(t) = \int_{-1}^{1} \frac{\sin[c(t-s)]}{\pi(t-s)} x(s)ds + z(t), \quad |t| < 1, \quad (1.1)$$

where $y$ is the image (assumed to be known on the interval $[-1,1]$), $x$ is the object, and $z$ is the noise; the quantity $d = \pi/c$ is the Rayleigh resolution distance.

We shall analyze the following problems:

(1) object restoration, i.e., to estimate the object $x$ given $y$;

(2) image extrapolation, i.e., to estimate a band-limited function whose restriction to the interval $[-1,1]$ approximates $y$.

The main reason for our attention to the optical system described above is that, in this simple case, one knows the eigenfunctions of the integral operator of Eq. (1.1) [they are the linear prolate spheroidal wavefunctions $\psi_i(c,x)$, with the associated eigenvalues $\lambda_i(c)$] and their asymptotic behavior when $k \to + \infty$. From this fact it derives the possibility of obtaining very precise results on problems (1) and (2).

Now, as we said above, the main prescription in order to restore the continuity in the case of ill-posed problems, consists in restricting the admitted solutions by imposing a supplementary global bound; next one must investigate the kind of continuity which has been restored. In some problems it is possible to have a fairly satisfactory Hölder type stability; i.e., if we denote by $\epsilon$ the data accuracy, then the solution accuracy is dominated by a term proportional to $\epsilon^\alpha$ ($0 < \alpha < 1$). On the other hand there are problems where the solution accuracy is at best proportional to $|\ln \epsilon|^{-\tau}$ (logarithmic continuity). One of the main results of this paper consists in showing that the restored continuity is of the Hölder type in the extrapolation of a given image piece beyond its
borders [Problem (2)], while in the object restoration [Problem (1)] it is only logarithmic.

In the usual regularization theory of ill-posed problems, one supposes that the error and the solution lie in some bounded sets of the respective Hilbert spaces. These conditions could appear too rigid, since it should be preferable, from a physical point of view, to regard the noise as a stochastic process. Furthermore, after the advent of the Shannon information theory, many concepts which were originally elaborated in that context, were successively largely applied to the analysis of optical systems. These are the principal motivations for introducing probabilistic methods of regularization of ill-posed problems. We devote Sec. 3 to probabilistic methods in a Hilbert space setting, along the lines introduced by two of the authors; furthermore a comparison of the two procedures is discussed in detail.

The paper is organized as follows. In Sec. 2 we analyze the regularization method proposed by Tikhonov and Miller, with attention to the specific questions which are of interest in the optical problems here considered. In Sec. 3 we develop the probabilistic method. Section 4 is devoted to the problem of object restoration; here we prove that the restored continuity is only logarithmic. In Sec. 5 we analyze the problem of image extrapolation, showing that in this case the restored continuity is of the Hölder type. Finally in Sec. 6 we try some conclusions.

2. TIKHONOV–MILLER METHOD

In this Section we sketch a method first proposed by Tikhonov for Fredholm equations of the first kind and independently developed by Miller for improperly posed problems in a Hilbert space setting. Here we follow the paper of Miller for the presentation of the general method; we also give some results on stability estimates which, to our knowledge, are not available in the literature.

Let \( X, Y \) be separable Hilbert spaces and \( A : X \to Y \) a linear continuous operator such that the inverse operator \( A^{-1} \) exists; if \( A^{-1} : Y \to X \) is not continuous, the problem: find \( x \in X \) such that \( y = Ax \), where \( y \in Y \) is a given vector, is improperly posed in the sense of Hadamard. In practice this pathology is very serious, since the data are always affected by errors. More precisely, the data vector \( y \) can be viewed as the sum of two terms: \( y = Ax + z \), where \( x \in X \) is an unknown vector which has to be approximately determined and \( z \in Y \) is an unknown vector describing errors or noise. Since \( A^{-1} \) is not continuous, the knowledge of \( y \) and of a bound on the error \( z \) is not sufficient in order to find an approximation to \( x \). Further “a priori” knowledge on \( x \) is required.

A. The general method

The basic point is to assume for the error \( z = Ax - y \) and for the unknown vector \( x \) the following prescribed bounds: \( \|z\|_Y = \|Ax-y\|_Y \leq \varepsilon, \|Bx\|_Z \leq E \), where \( B : X \to Z \) (\( Z \) is a Hilbert space) is a linear operator (the so-called constraint operator), densely defined in \( X \), which has a bounded inverse; \( \varepsilon, E \) are given positive numbers (we do not consider the case where only one of the two numbers \( \varepsilon \) and \( E \) is known). Then, any vector \( x \in X \) satisfying the conditions:

\[
\begin{align*}
\|Ax-y\|_Y & \leq \varepsilon \\
\|Bx\|_Z & \leq E
\end{align*}
\]

(2.1)

(2.2)

can be called an approximation to the unknown vector \( x \).

Such a definition is reasonable if:

(i) there exists at least one vector \( \bar{x} \) satisfying (2.1), (2.2) (in such a case the pair \( \varepsilon, E \) is called permissible);

(ii) there exists a norm or seminorm \( \langle \cdot \rangle \) such that

\[
M(\varepsilon,E) = \sup\{\langle x \rangle | x \in X, \|Ax\|_Y \leq \varepsilon, \|Bx\|_Z \leq E \}
\]

(2.3)

tends to zero, as \( \varepsilon \to 0 \), for fixed \( E \).

When (ii) holds true, we say that Problem (2.1), (2.2), i.e., the problem of finding a vector \( x \) satisfying the conditions (2.1), (2.2) is stable with respect to the norm or seminorm \( \langle \cdot \rangle \). Any upper bound for \( M(\varepsilon,E) \) is called a stability estimate and \( M(\varepsilon,E) \) itself is called the best possible stability estimate, \( M(\varepsilon,E) \) gives the size, in the sense of the norm or seminorm \( \langle \cdot \rangle \), of the “packet” of the vectors \( \bar{x} \) satisfying (2.1), (2.2); indeed if \( x, \bar{x} \) satisfy (2.1), (2.2), then

\[
\|A(\bar{x} - x)\|_Y \leq 2\varepsilon, \|B(\bar{x} - x)\|_Z \leq 2E, \text{ so that}
\]

\[
\langle \bar{x} - x \rangle \leq 2M(\varepsilon,E).
\]

Now, if \( M(\varepsilon,E) \) is a permissible pair, one has to find a method to exhibit explicitly a vector which satisfies the constraints (2.1), (2.2).

Let \( K_{\varepsilon,E}^* \) be the set of all the vectors \( \bar{x} \) satisfying (2.1), (2.2). \( K_{\varepsilon,E}^* \) is convex and bounded; besides the unknown vector \( x \) belongs to \( K_{\varepsilon,E}^* \). If \( \bar{x} \in K_{\varepsilon,E}^* \), then \( \|\bar{x} - x\|_X \) is a measure, in the sense of the norm of \( X \), of the error which is done by taking \( \bar{x} \) as an approximation to \( x \). Since \( x \) is unknown, a pessimistic estimate of this error is given by

\[
\mathcal{E}(\bar{x}) = \sup\{\|\bar{x} - x\|_X | x \in K_{\varepsilon,E}^* \}.
\]

(2.4)

It is not difficult to prove that there always exists a vector \( \tilde{x} \) which minimizes the functional \( \mathcal{E}(\cdot) \) and belongs to \( K_{\varepsilon,E}^* \); then we can take such a vector as a “best-possible” approximation to \( x \). Besides, if there exists a center of symmetry of \( K_{\varepsilon,E}^* \), then it coincides with \( \tilde{x} \). Unfortunately the set \( K_{\varepsilon,E}^* \) usually does not have a center of symmetry; therefore it can be quite hard to find a vector which minimizes \( \mathcal{E}(\cdot) \). However we can combine the two constraints (2.1), (2.2) into a single one and introduce the convex, bounded set

\[
K_{\varepsilon,E}^* = \left\{ x | x \in X, \frac{1}{2E^2} \|Ax - y\|_Y^2 + \frac{1}{2E^2} \|Bx\|_Z^2 \leq 1 \right\}.
\]

(2.5)

Then \( K_{\varepsilon,E}^* \subset \bar{K}_{\varepsilon,E} \subset K_{\varepsilon,E}^* \). We do not make a large error if we consider \( \bar{K}_{\varepsilon,E} \) instead of \( K_{\varepsilon,E}^* \) and corresponding \( \mathcal{E}(\cdot) \) instead of \( \mathcal{E}(\cdot) \), \( \mathcal{E}(\cdot) \) being defined as in Eq. (2.4), where \( K_{\varepsilon,E}^* \) is replaced by \( \bar{K}_{\varepsilon,E} \). Since \( \bar{K}_{\varepsilon,E} \) has a center of symmetry which is given by

\[
\tilde{x} = A^*A + \left( \frac{\varepsilon}{E} \right)^2 B^*B \left[ -1 A^*y \right]
\]

(2.6)
\( \tilde{x} \) minimizes \( \tilde{f}(\cdot) \) and it can be taken as a "nearly-best-possible" approximation to \( x \). Indeed, recalling that
\[ \hat{K} \subseteq K \subseteq \hat{K} \subseteq \hat{K} \]
we can conclude that the approximation \( \tilde{x} \) is best-possible but for a factor of \( \sqrt{2} \). Finally, it is important to note that this method gives an approximation to \( x \), which is satisfactory independently of the choice of norm (or seminorm) for measuring the error. Indeed, one can prove that
\[ \langle \tilde{x} - x \rangle < \sqrt{2M(\epsilon,E)}, \]
where \( M(\epsilon,E) \) is defined by Eq. (2.3).

The previous method suggests the introduction of the following stability estimate,
\[ \tilde{M}(\epsilon,E) = \mbox{sup} \left\{ \langle x, x \rangle \mid x \in X, \frac{1}{2\epsilon} \|Ax - y\|^2 \right\} + \frac{1}{2\epsilon^2} \|Bx\|^2 < 1 \right\}. \]

(2.8)

It is clear that \( \tilde{M}(\epsilon,E) \) is an upper bound for \( M(\epsilon,E) \); more precisely one has
\[ M(\epsilon,E) < \tilde{M}(\epsilon,E) < \sqrt{2M(\epsilon,E)} \]
and, instead of inequality (2.7), one gets
\[ \langle \tilde{x} - x \rangle < \tilde{M}(\epsilon,E). \]

(2.9)

### B. Stability estimates

Here we give results on the stability of Problem (2.1), (2.2) with respect to those norms and seminorms which are more appropriate for the problems considered in this paper; the stability can be achieved by properly choosing the constraint operator \( B \).

(i) Let us assume that the operator \( B \) has a bounded inverse, without any further property; then Problem (2.1), (2.2) is stable with respect to the family of seminorms:
\[ \langle x, u \rangle = \langle x, u \rangle, \|u\|_X = 1 \] (where \( u \) is a fixed but arbitrary vector on the unit sphere).

The following result is proved in the paper of Miller:

**Theorem 2.1:** If \( C = A^*A + (\epsilon/E)^2B^*B \), then
\[ \tilde{M}(\epsilon,E) = \sqrt{2\epsilon(C^{-1}u,u)}, \]
\[ \tilde{M}(\epsilon,E) \] being defined by Eq. (2.8) with \( \langle \cdot, \cdot \rangle \) replaced by \( \langle \cdot, u \rangle \).

From Theorem 2.1 we can derive the following result:

**Theorem 2.2:** \( \tilde{M}(\epsilon,E) \) tends to zero, as \( \epsilon \to 0 \), for fixed \( E \); besides \( \tilde{M}(\epsilon,E) = O(\epsilon) \) if \( u \in \epsilon \) range (A *).

**Proof:** If we write \( |B| = (B^*B)^{1/2} \) and introduce the operator \( S = A^*B^{-1} \) (S^{-1} exists since \( A^{-1} \) exists), then Eq. (2.11) can be rewritten as follows,
\[ \tilde{M}(\epsilon,E) = \sqrt{2\epsilon \left( \|S \| \|B \|^{-1} \right)^{-1}} \left( |S^{-1}u| B^{-1} |u| \right)^{1/2} \]
\[ \tilde{M}(\epsilon,E) = \sqrt{2\epsilon \left( \|S \| \|B \|^{-1} \right)^{-1}} \left( |S^{-1}u| B^{-1} |u| \right)^{1/2}. \]

(2.12)

For the sake of simplicity we suppose that the operator \( A \) is compact (however the Theorem holds true even without this assumption). Then the operator \( S \) is compact and we denote by \( \{s_k\}_{k=0}^\infty \) (s_k > 0, s_0 > s_0 > s_2 > ...) the set of the eigenvalues of the operator \( S^*S \) and by \( \{u_k\}_{k=0}^\infty \) the set of the corresponding eigenvectors. \( \{u_k\}_{k=0}^\infty \) is an orthonormal basis in \( X \) so that we can write
\[ \tilde{M}(\epsilon,E) = \sqrt{2\epsilon \left( \sum_{k=0}^\infty \frac{s_k^2}{s_k^2 + (\epsilon/E)^2} \right)^{1/2}}, \]

where \( M(\epsilon,E) \) is defined by Eq. (2.3).

(2.13)

From this equation it is clear that \( \tilde{M}(\epsilon,E) \) tends to zero, as \( \epsilon \to 0 \), for fixed \( E \). Besides, let us assume that \( u \in \epsilon \) range (A *); then there exists a vector \( v \in Y \) such that \( u = A^*v \). If we introduce the vectors \( v_k = s_k^{-1}S_k(u_k), \{v_k\}_{k=0}^\infty \) is an orthonormal basis in range \( (S^*S) \), then Eq. (2.13) becomes
\[ \tilde{M}(\epsilon,E) = \sqrt{2\epsilon \left( \sum_{k=0}^\infty \frac{s_k^2}{s_k^2 + (\epsilon/E)^2} \right)^{1/2}}, \]

(2.14)

so that \( \tilde{M}(\epsilon,E) \to 0 \) as \( \epsilon \to 0 \), \( \tilde{M}(\epsilon,E) = O(\epsilon) \).

Now, let us assume \( \tilde{M}(\epsilon,E) = O(\epsilon) \), then there exists a constant \( \delta_u \) such that
\[ \sum_{k=0}^\infty \frac{s_k^2}{s_k^2 + (\epsilon/E)^2} \leq \delta_u, \]

(2.15)

If \( K \) is the maximum integer such that \( s_k > \epsilon/E \), then \( K \to + \infty \) when \( \epsilon \to 0 \); therefore
\[ \sum_{k=0}^K s_k^{-2} \left( |B|^{-1} u_k, u_k \right)^2 \leq 2 \sum_{k=0}^\infty \frac{s_k^2}{s_k^2 + (\epsilon/E)^2} \leq 2\delta_u, \]

(2.16)

for any \( K \) we get
\[ \sum_{k=0}^\infty s_k^{-2} \left( |B|^{-1} u_k, u_k \right)^2 < 2\delta_u, \]

(2.17)

or, in other words, \( |B|^{-1} u \in \epsilon \) range (S *). Then there exists a vector \( u \in Y \) such that \( |B|^{-1} u \in S^\epsilon \) and it follows that \( u = A^*v \). The Theorem is proved.

(ii) Let us assume now that the operator \( B \) has a compact inverse. Then the set \( K^E = \{x \in X, \|Bx\| < \epsilon \} \) is a compact subset of \( X \); analogously the set \( J^E = AK^E \) is a compact subset of \( Y \) and the operator \( A \) defines a continuum, one to one mapping of \( K^E \) onto \( J^E \). From a well known topological lemma it follows that the inverse mapping from \( J^E \) onto \( K^E \) is continuous; such a result implies that \( \tilde{M}(\epsilon,E) = \sup \{ \|x\|_X \mid x \in X, \|Ax\| < \epsilon \} \]
\[ = \sup \{ \|A^{-1}y\|_Y \mid y \in Y, \|y\| < \epsilon \} \]
\[ = \sup \{ \|B^{-1}y\|_X \mid y \in Y, \|y\| < \epsilon \} \]
\[ \mbox{tends to zero, as } \epsilon \to 0, \mbox{ for fixed } E. \] Therefore, one can conclude that:

**When the operator \( B \) has a compact inverse, Problem (2.1), (2.2) is stable with respect to the norm of \( X \).**

The previous result is essentially qualitative; we can have a more precise result (see Theorem 2.3) in the following case. Let us assume that \( A : X \to Y \) is a compact operator (such an assumption is satisfied for the problems considered in Secs. 4 and 5); then, let \( \{\alpha_k\}_{k=0}^\infty \) (\( \alpha_k > 0 \); \( \alpha_0 > \alpha_1 > \alpha_2 > \ldots \)

Bertero, DeMol, and Viano
be the set of the eigenvalues of the operator $A^*A$ and \( \{u_k\}_{k=0}^{\infty} \) be the set of the corresponding eigenvectors,

\[
A^*Ax = \sum_{k=0}^{\infty} \alpha_k^2 x_k u_k, \quad x_k = \langle x, u_k \rangle_x.
\]  

(2.18)

In such a case a considerable simplification is introduced if one considers a constraint operator $B$ such that $B^*B$ and $A^*A$ commute, i.e.,

\[
B^*Bx = \sum_{k=0}^{\infty} \beta_k^2 x_k u_k, \quad x_k = \langle x, u_k \rangle_x.
\]  

(2.19)

Then the operator $B$ has a compact inverse iff

\[
\lim_{k \to \infty} \beta_k = + \infty. \quad \text{It is self evident that the condition is necessary. In order to show that it is also sufficient, one must prove that any set } K_E = \{ x \in X \mid \| Bx \|_x < E \} \text{ is conditionally compact. Such a result follows easily from the proposition}^{15}: \text{A subset } K \text{ of } p \text{ space, } p > 1, \text{ is conditionally compact iff it is bounded and } \lim_{n \to \infty} \sum_{k=n}^{\infty} |x_k|^p = 0 \text{ uniformly for } \{ x \}_{k=0}^{\infty} \in K.
\]

Now, let $I_{\epsilon/E}$ be the set defined as follows,

\[
I_{\epsilon/E} = \{ k \mid | \alpha_k | E \beta_k \}.
\]  

(2.20)

Since $A^*A$ is compact, the eigenvalues $\alpha_k^2$ accumulate to zero; on the other hand, the eigenvalues $\beta_k^2$ of $B^*B$ tend to infinity. Therefore, the set $I_{\epsilon/E}$ contains only a finite number of points.

Besides, if we write

\[
\alpha \left( \frac{\epsilon}{E} \right) = \min \{ | \alpha_k | k \in I_{\epsilon/E} \},
\]

\[
\beta \left( \frac{\epsilon}{E} \right) = \min \{ | \beta_k | k \notin I_{\epsilon/E} \},
\]

(2.21)

we have the following result:

**Theorem 2.3:** The inequality holds,

\[
\frac{1}{2} \left( \frac{\epsilon}{\alpha(\epsilon/E)} + \frac{E}{\beta(\epsilon/E)} \right) \leq M(\epsilon,E) \leq \frac{\epsilon}{\alpha(\epsilon/E)} + \frac{E}{\beta(\epsilon/E)},
\]

where $M(\epsilon,E)$ is given by Eq. (2.3) with $\langle \cdot \rangle = \| \cdot \|_x$.

**Proof:** Inequality (2.22) is a consequence of a lemma (three norm lemma) proved by Miller\textsuperscript{15} which states that

\[
\frac{1}{2} [L(\epsilon,E) + H(\epsilon,E)] \leq M(\epsilon,E) \leq L(\epsilon,E) + H(\epsilon,E),
\]

(2.23)

where:

\[
L(\epsilon,E) = \sup \left\{ \left( \sum_{k \in I_{\epsilon/E}} |x_k|^2 \right)^{1/2} | x \in X, \right. \}
\]

\[
\left. \left( \sum_{k \notin I_{\epsilon/E}} \alpha_k^2 |x_k|^2 \right)^{1/2} \leq \epsilon \right\}.
\]

(2.24)

\[
H(\epsilon,E) = \sup \left\{ \left( \sum_{k \in I_{\epsilon/E}} |x_k|^2 \right)^{1/2} | x \in X, \right. \}
\]

\[
\left. \left( \sum_{k \notin I_{\epsilon/E}} \beta_k^2 |x_k|^2 \right)^{1/2} \leq \epsilon \right\}.
\]

\[
\left( \sum_{k \notin I_{\epsilon/E}} \beta_k^2 |x_k|^2 \right)^{1/2} \leq \epsilon \right\}.
\]

(2.25)

Since the functionals involved in Eqs. (2.24) and (2.25), are linear in the parameters $|x_k|^p$, it is easy to show that

\[
L(\epsilon,E) = \frac{\epsilon}{\alpha(\epsilon/E)}, \quad H(\epsilon,E) = \frac{E}{\beta(\epsilon/E)},
\]

(2.26)

and the theorem is proved.

In the particular case $B = (A^*)^{-1}$ we can obtain a more precise result:

**Theorem 2.4:** If $B = (A^*)^{-1}$, then

\[
M(\epsilon,E) \leq \sqrt{\epsilon E},
\]

where $M(\epsilon,E)$ is defined by Eq. (2.3) with $\langle \cdot \rangle = \| \cdot \|_x$. Besides the equality holds when the ratio $\epsilon/E$ is equal to one of the eigenvalues of the operator $A^*A$.

**Proof:** From the Schwarz inequality we have, for any $x \in$ range $(A^*)^{-1}$

\[
|x|_x^2 = (A^*Ax, x) = (Ax, (A^*)^{-1}x) \leq \|Ax\|_y \| (A^*)^{-1}x \|_y,
\]

(2.27)

and the inequality (2.27) follows from the constraints $\|Ax\|_y \leq E, \| (A^*)^{-1}x \|_y \geq E$. Besides, if the pair $| \epsilon, E |$ is such that $\epsilon/E = \alpha_k^2$, then $x = Ea_k u_k$ is an eigenvector of $A^*A$ and if we take $x = Ea_k u_k$ $u_k$ is the eigenvector corresponding to $\alpha_k^2$, then $\|Ax\|_y = Ea_k^2 = \epsilon, \| (A^*)^{-1}x \|_y = E$; on the other hand, $\|x\|_y = E \alpha_k = \sqrt{\epsilon E}$ and we can conclude that, in such a case, the equality holds in (2.27).

(iii) Let us assume now that the inverse of $B^*B$ is an operator of the trace class. In other words, there exists an orthonormal basis $\{u_k\}_{k=0}^{\infty}$ in $X$ such that the operator $B^*B$ has the spectral representation (2.19) and

\[
\sum_{k=0}^{\infty} \beta_k^{-2} < + \infty.
\]

(2.29)

In the domain of $B$ we can introduce the following norm (which shall be used in Sec. 5),

\[
\| x \| = \sum_{k=0}^{\infty} |x_k|, \quad x = \langle x, u_k \rangle_x;
\]

\[
\| x \| = \sum_{k=0}^{\infty} |x_k|, \quad x = \langle x, u_k \rangle_x;
\]

(2.30)

indeed, by the Schwarz inequality

\[
\| x \| \leq \left( \sum_{k=0}^{\infty} \beta_k^{-2} \right)^{1/2} \left( \sum_{k=0}^{\infty} \beta_k^{-2} |x_k|^2 \right)^{1/2}
\]

\[
= \left( \sum_{k=0}^{\infty} \beta_k^{-2} \right)^{1/2} \| Bx \|_x.
\]

(2.31)

Now, the set $K_E = \{ x \in X \mid \| Bx \|_x < E \}$ is compact with respect to the norm (2.30); indeed, by the Schwartz inequality and condition (2.29),

\[
\sum_{k=n}^{\infty} |x_k| < E \left( \sum_{k=n}^{\infty} \beta_k^{-2} \right)^{1/2}, \quad n \to + \infty
\]

(2.32)

uniformly for $x \in K_E$. The compactness criterion in $l_p, p > 1$, already recalled\textsuperscript{19} implies that $K_E$ is a conditionally compact
subset of \( l_i \). Besides it is easy to see that the set \( K_E \) is closed with respect to the norm (2.30).

If we put \( \beta = \min \{ |\beta_k| \} > 0 \) and we remark that from the constraint \( \| Bx \|_2 < E \) it follows that \( |x_k| < E/\beta \), we get, for any \( x \in K_E \), \( \| x \|_x < (E/\beta)^{1/2} \| x \|_E \), so that

\[
\| Ax \|_Y \leq \| A \|_Y \| x \|_x < \sqrt{E \over \beta} \| A \| \| x \|_E.
\]  

(2.33)

Therefore, if we set \( J_E = AK_E \), the operator \( A \) defines a continuous mapping of \( K_E \) onto \( J_E \), when \( K_E \) is normed with the norm (2.30) and \( J_E \) is normed with the norm of \( Y \). The topological lemma already recalled, \(^{13}\) implies that the inverse mapping from \( J_E \) onto \( K_E \) is continuous. Therefore, \( M(e, E) \), defined by Eq. (2.3) with \( \langle \cdot, \cdot \rangle = \| \cdot \|_Y \| \cdot \|_E \|, \) tends to zero, as \( e \to 0 \), for fixed \( E \); we can conclude that: When condition (2.29) is satisfied, Problem (2.1), (2.2) is stable with respect to the norm (2.30).

When \( A \ast A \) and \( B \ast B \) are given by Eqs. (2.18), (2.19), we can get the following more precise result.

**Theorem 2.5:** If \( M(e, E) \) is defined by Eq. (2.3) with \( \langle \cdot, \cdot \rangle = \| \cdot \|_Y \| \cdot \|_E \|, \), then

\[
{e \over 2} \left( \sum_{k \in I_{r, E}} \alpha_k^{-2} \right)^{1/2} + \frac{E}{2} \left( \sum_{k \in I_{r, r}} \beta_k^{-2} \right)^{1/2} < M(e, E) < e
\]

\[\times \left( \sum_{k \in I_{r, E}} \alpha_k^{-2} \right)^{1/2} + \frac{E}{2} \left( \sum_{k \in I_{r, r}} \beta_k^{-2} \right)^{1/2},
\]

(2.34)

where \( I_{r, E} \) is the set defined in Eq. (2.20).

**Proof:** By the three norm lemma\(^1\) we have for \( M(e, E) \) an inequality like (2.23) where now

\[
L(e, E) = \sup \left\{ \sum_{k \in I_{r, r}} \| x_k \| \left( \sum_{k \in I_{r, r}} \alpha_k^2 \| x_k \|^2 \right)^{1/2} < e \right\},
\]

(2.35)

\[
H(e, E) = \sup \left\{ \sum_{k \in I_{r, r}} \| x_k \| \left( \sum_{k \in I_{r, r}} \beta_k^2 \| x_k \|^2 \right)^{1/2} < e \right\},
\]

(2.36)

Then, by the Schwarz inequality (which is precise) we have

\[
L(e, E) = e \left( \sum_{k \in I_{r, r}} \alpha_k^{-2} \right)^{1/2}, \quad H(e, E) = E \left( \sum_{k \in I_{r, r}} \beta_k^{-2} \right)^{1/2},
\]

(2.37)

and the Theorem is proved.

We conclude with a result which is a modified version of Theorem 2.5.

**Theorem 2.6:** If \( \bar{M}(e, E) \) is defined by Eq. (2.8) with \( \langle \cdot, \cdot \rangle = \| \cdot \|_Y \| \cdot \|_E \|, \) then

\[
\bar{M}(e, E) = \sqrt{2} e \left( \sum_{k = 0}^{+\infty} \frac{1}{\alpha_k^2 + (e/E) \beta_k^2} \right)^{1/2}.
\]

(2.38)

**Proof:** Indeed, by the Schwarz inequality, we have

\[
\sum_{k = 0}^{+\infty} |x_k|^2 = \sum_{k = 0}^{+\infty} \left( \frac{\alpha_k^2 + (e/E) \beta_k^2}{\alpha_k^2 + (e/E) \beta_k^2} \right)^{1/2} |x_k|^2
\]

\[
\leq \left( \| Ax \|_Y^2 + \frac{e^2}{E^2} \| Bx \|_Z^2 \right)^{1/2}
\]

Since the Schwarz inequality is precise, we get Eq. (2.38).

**C. Methods of eigenfunction expansions**

In the case where the operator \( A : X \to Y \) is compact and the operators \( A \ast A \) and \( B \ast B \) commute [see Eqs. (2.18), (2.19)], Eq. (2.6) takes the form

\[
\bar{x} = \sum_{k \in I_{r, r}} \alpha_k^{-1} (e/E) \beta_k^{-1} y_k u_k, \quad y_k = \langle y, u_k \rangle_Y,
\]

(2.40)

where \( u_k = \alpha_k^{-1} A u_k, \langle u_k, v_k \rangle_Y \) is an orthonormal basis in \( \text{range}(A)^\perp \). An approximation, even more simple than (2.40), is given by

\[
\bar{x} = \sum_{k \in I_{r, E}} \frac{y_k}{\alpha_k} u_k,
\]

(2.41)

where the set \( I_{r, E} \) is defined by Eq. (2.20) and \( y_k, u_k \) are as in Eq. (2.40).

If \( \bar{x} \) is any vector satisfying the conditions (2.1), (2.2), then

\[
\| A(\bar{x} - x) \|_Y \leq \sqrt{2} e, \quad \| B(\bar{x} - x) \|_Z \leq \sqrt{2} e;
\]

(2.42)

indeed, if \( \bar{x}_k = \langle \bar{x}, u_k \rangle \),

\[
\| A(\bar{x} - x) \|_Y^2 = \sum_{k \in I_{r, r}} |y_k - \alpha_k \bar{x}_k|^2 + \sum_{k \in I_{r, r}} |\alpha_k \bar{x}_k|^2
\]

\[
\leq \sum_{k \in I_{r, r}} |y_k - \alpha_k \bar{x}_k|^2
\]

\[
+ \left( \frac{e}{E} \right)^2 \sum_{k \in I_{r, r}} \beta_k^2 |x_k|^2
\]

\[
< 2 \| A \bar{x} - y \|_Y^2 + \left( \frac{e^2}{E^2} \right) \| B \bar{x} \|_Z^2 \leq 2 e.
\]

(2.43)

In a similar way one can prove the second inequality (2.42). Therefore, we get

\[
\langle \bar{x} - x \rangle \leq \sqrt{2} \bar{M}(e, E),
\]

(2.44)

\( \bar{M}(e, E) \) being defined by Eq. (2.8). We can conclude that, if Problem (1), (2) is stable with respect to the norm or seminorm \( \langle \cdot, \cdot \rangle \), both \( x \) and \( \bar{x} \) converge (in the sense of \( \langle \cdot, \cdot \rangle \)) to the "true" solution of the problem, as \( e \to 0 \), for fixed \( E \).

3. PROBABILISTIC METHOD

Let us assume, as in Sec. 2, that the data vector \( y \) is the sum of two terms, \( y = Ax + z \), where \( x \) is the vector which has to be approximately determined and \( z \) is an unknown vector describing errors or noise; then, the basic point in the probabilistic method is to consider the vectors \( x, y, z \) as the values of random variables \( \xi, \eta, \zeta \).\(^{11,16}\)
More precisely, let \((\Omega, \mathcal{F}, P)\) be a probability space, i.e., \(\Omega\) is an abstract point set, \(\mathcal{F}\) is a \(\sigma\)-algebra of subsets of \(\Omega\), and \(P\) is a measure on \(\mathcal{F}\) with \(P(\Omega) = 1\); besides, let \(X, Y\) be separable Hilbert spaces and let \(\xi: \Omega \rightarrow X, \zeta: \Omega \rightarrow Y\) be weak random variables\(^9\) (hereafter shortened to w.r.v., i.e., the mappings \(\xi, \zeta\) induce cylinder measures \(\mu_{\xi}, \mu_{\zeta}\) on \(X, Y\), respectively; then we consider the w.r.v. \(\eta: \Omega \rightarrow Y\) given by

\[
\eta = A\xi + \zeta,
\]

where \(A: X \rightarrow Y\) is a linear continuous operator whose inverse \(A^{-1}\) exists but is not continuous.

If we assume that the joint measure of the pair of w.r.v. \(\xi, \zeta\) is known, then the problem is the following: Given a value \(y\) of the w.r.v. \(\eta\), find the best possible mean square estimate of the w.r.v. \(\xi\).

Now, we assume (as it is reasonable in many practical applications) that the w.r.v. \(\xi, \zeta\) are Gaussian and independent and that they have zero mean; in such a case their joint measure is uniquely characterized by the covariance operators \(R_\xi, R_\zeta\). Then the covariance operator \(R_\eta\) is expressed in terms of \(R_\xi, R_\zeta\) by\(^11\)

\[
R_\eta = AR_\xi A^* + R_\zeta. \tag{3.2}
\]

Besides, the w.r.v. \(\xi, \eta\) are not independent and their cross-covariance operator is given by

\[
R_{\xi\eta} = R_\eta A^*. \tag{3.3}
\]

In the following we shall also assume that the inverse of the operator \(R_\zeta\) exists. Indeed, if this condition is not satisfied, there exists a vector \(v_0\) such that \(R_\zeta v_0 = 0\); then, the random variable \((\xi, v_0)\) takes the value zero with probability one or, in other words, the component of the data vector in the direction of \(v_0\) is not affected by error or noise. Since such a situation is not realistic, we assume that the inverse of \(R_\zeta\) exists. From Eq. (3.2) it follows that the inverse of the operator \(R_\eta\) also exists; besides, range \((R_\zeta)\) and range \((R_\eta)\) are both dense in \(Y\).

The covariance operators \(R_\xi, R_\zeta\) have a role, in the probabilistic method, similar to that of the bounds \(E, \tilde{E}\) and of the constraint operator \(B\) in the Tikhonov–Miller method. For this reason we assume that: \(R_\zeta = \epsilon^2 N\), where \(\epsilon\) is a “small” positive number and \(N: Y \rightarrow Y\) is a linear bounded operator independent of \(\epsilon\).

A. The general method

If \(L: Y \rightarrow X\) is any linear continuous operator, we call the w.r.v. \(\xi_L = L\eta\) a linear estimate of the w.r.v. \(\xi\). Then, for any \(u \in X\), the reliability of the estimate for the random variable \((\xi, u)_Y\) is measured by the mean square error (here \(E\) denotes the expectation)

\[
Q(u; L\epsilon) = E\left[ (\xi - L\eta, u)_X \right]^2
\]

\[
= [(R_\xi - R_\xi A^* L^* + L R_\xi A^* L^*) u, u]_X. \tag{3.4}
\]

If there exists a unique linear continuous operator \(L_0: Y \rightarrow X\) which minimizes \(Q(u; L\epsilon)\) for any \(u \in X\), then the w.r.v.

\(\hat{\xi} = L_0\eta\) is said to be the best linear estimate of the w.r.v. \(\xi\).

We have the following result\(^11\):

**Theorem 3.1:** There exists a unique best linear estimate \(\hat{\xi}\) of the w.r.v. \(\xi\) iff the operator \(R_\xi R^{-1}_\xi\) is bounded on the range \((R_\xi)\); in such a case

\[
\tilde{L}_0 = R_\xi R^{-1}_\xi = R_\xi A^* (A R_\xi A^* + \epsilon^2 N)^{-1}. \tag{3.5}
\]

**Remark 3.1:** If \(R_\zeta\) has a bounded inverse, then the operator \(R_\xi\) also has a bounded inverse. In such a case the product \(R_\zeta R^{-1}_\xi\) is a bounded operator and therefore there is no question in the interpretation of Eq. (3.5). On the other hand, if the inverse of \(R_\zeta\) is not continuous, then \(R_\zeta R^{-1}_\xi\) is the product of two operators where \(R_\xi\) is bounded while \(R_\zeta^{-1}\) is unbounded and densely defined. [see Eq. (3.2)–we recall that \(A^{-1}\) is not continuous.] If \(R_\zeta R^{-1}_\xi\) is bounded on range \((R_\xi)\), then Theorem 3.1 states that \(L_0\) is the usual extension (by continuity) of \(R_\xi R^{-1}_\xi\) to the closure of range \((R_\xi)\), i.e., to the whole Hilbert space \(Y\). For simplicity we still denote by \(R_\xi R^{-1}_\xi\) [as in Eq. (3.5)] such an extension.

**Remark 3.2:** If we consider the problem formulated at the beginning of this section, then the solution is: Given a value \(y\) of the w.r.v. \(\eta\), the best possible mean-square estimate of the w.r.v. \(\xi\) is

\[
\hat{\xi} = R_\xi A^* (A R_\xi A^* + \epsilon^2 N)^{-1} y. \tag{3.6}
\]

It is interesting to compare formally Eq. (3.6) with Eq. (2.6). If we put \(N = I\) and \(R_\zeta = E(B^* B)^{-1}\) in Eq. (3.6) and we use the identity

\[
(A^* R_\xi^{-1} A + R_\zeta^{-1}) R_\xi A^* = A^* R_\xi^{-1} (A R_\xi A^* + R_\zeta), \tag{3.7}
\]

we see that, in such a case, Eqs. (3.6) and (2.6) coincide.

**Remark 3.3:** In the case \(N = I\) (i.e., in the case where \(\zeta\) is the so-called “white noise”) there is an interesting interpretation of the least-square estimate (3.6). We introduce the operator \(T = R^{1/2} A\) (recall that a covariance operator is a bounded, nonnegative, self-adjoint operator), so that we can write \(L_0 = R^{1/2} T (T^* T + \epsilon^2 I)^{-1}\); besides we denote by \([\tau_k]_{k=0}^\infty\) the set of the nonzero singular values of \(T\), i.e., \(T^* u_k = \tau_k u_k, T u_k = \tau_k u_k\), where \([u_k]_{k=0}^\infty\) is an orthonormal basis in range \((T) \subset X\) and \([v_k]_{k=0}^\infty\) is an orthonormal basis in range \((T^*) \subset Y\); then, from Eq. (3.6), by means of the relation \(R_\xi^{-1/2} u_k = \tau_k^{-1/2} v_k\), we get

\[
\hat{\xi} = \sum_{k=0}^\infty \frac{\tau_k}{\tau_k^2 + \epsilon^2} y_k R_\xi^{-1/2} u_k
\]

\[
= \sum_{k=0}^\infty \frac{\tau_k^2}{\tau_k^2 + \epsilon^2} y_k A^* v_k, \tag{3.8}
\]

where \(y_k = (v_k, u_k)_Y\). Now, the random variables \(\xi_k = (\xi, R_\xi^{-1/2} u_k)_Y, \eta_k = (\eta, u_k)_Y\) (\(k\) fixed but arbitrary) are Gaussian and: \(E \left[ |\xi_k|^2 \right] = 1, E \left[ |\eta_k|^2 \right] = \tau_k^2 + \epsilon^2, E \left[ \xi_k \eta_k \right] = \tau_k\); therefore their correlation coefficient is
\[ r_k = r_k (r_k^2 + 8^2)^{-1} \] and their average mutual information is
\[ J_k = J (\xi_k, \eta_k) = -\frac{1}{2} \ln (1 - r_k^2) = \frac{1}{2} \ln (1 + r_k^2 / 8^2). \]  \( (3.9) \)

From Eq. (3.9) we get
\[ r_k^2 (r_k^2 + 8^2)^{-1} = 1 - \exp (-2J_k), \] so that we can write Eq. (3.8) as follows
\[ \hat{x} = \sum_{k=0}^{\infty} (1 - e^{-2J_k}) y_k A^{-1} v_k. \]  \( (3.10) \)

From this expression we see that the best linear estimate \( \hat{x} \) can be obtained, from the formal solution \( A^{-1} y = \sum_{k=0}^{\infty} y_k A^{-1} v_k \) by means of a penalty on the coefficients in terms of the amount of information on the random variable \( \xi_k \) contained in the random variable \( y_k \).

Let us assume now that the w.r.v. \( \xi \) has a finite variance, i.e., \( E \langle \| \xi \|_2^2 \rangle < +\infty \) (recall that \( \xi \) has a zero mean); as is known, this condition is satisfied iff \( R_\xi \) is an operator of the trace class.\(^{11}\) In such a case, we can try to define a "global" mean square error in the estimate of \( \xi \) (by means of \( \xi_\ell = L \eta \)) as the variance of the w.r.v. \( \xi - L \eta \), i.e.,
\[ Q (L, e) = E \langle \| \xi - L \eta \|_2^2 \rangle \]
\[ = \text{tr} (P_{\xi} - R_{\xi} A^* L^* - L A R_{\xi} + L R_{\xi} L^*). \]  \( (3.11) \)

\( Q (L, e) \) is finite iff the w.r.v. \( L \eta \) has also a finite variance. Next we can define a best linear estimate \( \hat{\xi} = L \eta \) in the sense of the mean square error (3.11), if there exists a bounded operator \( L_0 \) which minimizes \( Q (L, e) \).

We must distinguish two cases. The first is when the noise \( \xi \) has a finite variance; then, from Eq. (3.2) it follows that \( R_{\xi} \) is of the trace class and therefore the w.r.v. \( L \eta \) has also a finite variance for any bounded linear operator \( L : Y \rightarrow X \). In such a case, if there exists a bounded operator \( L_0 \) which minimizes \( Q (u; L, e) \) then it also minimizes \( Q (L, e) \) and vice versa. The second case is when \( \xi \) (and therefore also \( \eta \)) does not have a finite variance. Now, since the covariance operator of \( L \eta \) is \( L R_{\xi} L^* \), it follows that \( L \eta \) has a finite variance iff the operator \( L \) is of the Schmidt class. Therefore, there exists a best linear estimate in the sense of the mean-square error (3.11) iff the operator \( R_{\xi \eta} R_{\eta \eta}^{-1} \) has a bounded extension of the Schmidt class.

### B. Least mean-square errors

We can define the least mean-square error in the estimate of \( \langle \xi, u \rangle \) as
\[ \delta(u, e) = \inf_L \langle Q(u; L, e) \rangle^{1/2}, \]  \( (3.12) \)

where \( Q(u; L, e) \) is defined by Eq. (3.4). If the best linear estimate exists, then (3.12) gives a measure of the reliability of the estimate; indeed, we have
\[ \delta(u, e) = \langle Q(u; L_0, e) \rangle^{1/2}, L_0 \] being given by Eq. (3.5). After some simple calculations we get
\[ \delta(u, e) = \langle [R_{\xi \eta} - L_0 R_{\xi} R_{\eta} L_0^*] u, u \rangle^{1/2}. \]  \( (3.13) \)

In the general case (i.e., even if the best linear estimate does not exist) we have the following result,\(^{11}\)
\[ \delta(u, e) = \langle (1 - V^* V) R_{\xi \eta}^{1/2} u, R_{\xi \eta}^{1/2} u \rangle^{1/2}, \]  \( (3.14) \)

where \( V : X \rightarrow Y \) is the unique linear continuous operator (with \( \| V \| < 1 \)) such that \( R_{\eta \eta} = R_{\eta \eta}^{1/2} V R_{\xi \eta}^{1/2} \).

**Remark 3.4:** \( \delta(u, e) \) has a role, in the probabilistic method, similar to that of the best stability estimate (2.3), with \( \langle \cdot , \cdot \rangle = \langle \cdot , \cdot \rangle \) in the Tikhonov–Miller method. Indeed, if we put \( N = I \) and \( L_{\xi} = E \langle (B + B^*) \rangle^{-1} \) in Eq. (3.5), from Eqs. (3.7) and (3.13) we get
\[ \delta(u, e) = \epsilon \left[ A \ast A + \left( \frac{E}{e} \right)^2 B \ast B \right]^{-1/2} u, u \rangle^{1/2}; \]  \( (3.15) \)

this expression coincides, except for a factor of \( \sqrt{2} \), with Eq. (2.11).

The following result\(^{11}\) gives conditions which guarantee that \( \delta(u, e) \) tends to zero when \( e \to 0 \):

**Theorem 3.2:** If the operator \( R_{\xi} = e^2 N \) has a bounded inverse, then, for any \( u \in \mathcal{X} \)
\[ \lim_{e \to 0} \delta(u, e) = 0 \]  \( (3.16) \)

iff the problem
\[ \begin{align*}
\text{range}(R_{\xi}^{1/2}), & \quad A x = 0
\end{align*} \]  \( (3.17) \)
has only the trivial solution.

**Remark 3.5:** In the problems considered in this paper, the inverse of the operator \( A \) exists; therefore problem (3.17) has only the trivial solution.

The following result is the analog of Theorem 2.2:

**Theorem 3.3:** If the operator \( R_{\xi} = e^2 N \) has a bounded inverse and if the inverse of the operator \( R_{\xi} \) exists, then
\[ \delta(u, e) = O(e), \quad e \to 0 \]  \( (3.18) \)
iff \( u \in \text{range}(A \ast N^{-1/2}) \).

**Proof:** From the identity (3.7) and Eq. (3.13) we get
\[ \delta(u, e) = e \left( [A \ast N^{-1/2} + e^2 R_{\xi}]^{-1} - I \right) u, u \rangle^{1/2} \]  \( (3.19) \)
and, if we introduce the operator \( S = N^{-1/2} A R_{\xi}^{1/2} \), we can write
\[ \delta(u, e) = e \left( [S \ast S + e I]^{-1} - R_{\xi}^{1/2} u, R_{\xi}^{1/2} u \rangle^{1/2}. \]  \( (3.20) \)
At this point the proof proceeds like the proof of Theorem 2.2.

When the inverse of the operator \( R_{\xi} = e^2 N \) is not bounded, it seems that no general result on the behavior of \( \delta(u, e) \), when \( e \to 0 \), can be derived; assumptions on the ranges of the operators \( (AR_{\xi} A)^{-1/2} \) and \( R_{\xi}^{1/2} = e^2 N^{1/2} \) have to be done. However there exists an important case where assumptions of this type can be proposed in a quite natural way: We mean when both the operators \( R_{\xi}, R_{\eta} \) are of the trace class. In such a case the covariance operator \( R_{\xi} \) belongs to the same class and one can introduce the average mutual information \( J(\xi, \eta) \) of the w.r.v. \( \xi, \eta \). Our prescription is to require that \( J(\xi, \eta) \) is finite [in the problem of Sec. 4, \( J(\xi, \eta) \) is the amount of information about the object contained in the optical image and the finite variances of \( \xi \) and \( \eta \) are the average energies of the object and of the image, respectively].
According to the results of Baker\(^n\) one can give necessary and sufficient conditions on the covariance operators \(R_\xi R_\xi\) in order that \(J(\xi, \eta) < + \infty\); if we take into account these results, then the previous assumptions can be formulated as follows:

(i) \(\text{tr}(R_\xi) < + \infty\), \(\text{tr}(N) < + \infty\);
(ii) \(AR_\xi A^* = N^{1/2} SN^{1/2}\) where \(S: Y \to Y\) is a linear operator of the trace class.

The following result holds\(^p\): if the assumptions (i), (ii) are satisfied, then Eq. (3.16) holds for any \(u \in X\), iff the problem (3.17) has only the trivial solution.

When the w.r.v. \(\xi\) has a finite variance, we can consider the mean-square error (3.11) and define a least mean-square error as follows,

\[
\delta(\varepsilon) = \inf_{L} Q(L, \varepsilon). \tag{3.21}
\]

If the assumptions (i), (ii) are satisfied, then it is possible to prove\(^q\) that \(\delta(\varepsilon) \to 0\) when \(\varepsilon \to 0\).

C. Methods of eigenfunction expansions

We consider now the case where the operator \(A: X \to Y\) is compact. We assume that the operators \(A^* A\) and \(R_\xi\) commute; the same assumption is done on the operators \(A^* A\) and \(R_\xi\) = \(e^N\).

If \(A\) and \(A^*\) both have an inverse operator (such a condition is satisfied in the problems considered in this paper), then all the singular values \(\alpha_k\) of \(A\) are non-zero; besides, \(\{u_k\}_{k=0}^{\infty}\) (the set of the eigenvectors of \(A^* A\)) and \(\{v_k\}_{k=0}^{\infty}\) (the set of the eigenvectors of \(A A^*\); \(v_k = \alpha_k^{-1} A u_k\)) are an orthonormal basis in \(X, Y\), respectively.

Thanks to the previous assumption we can write

\[
R_\xi x = \sum_{k=0}^{\infty} \rho_k x_k u_k, \quad R_\xi y = e^\varepsilon \sum_{k=0}^{\infty} y_k v_k, \tag{22.22}
\]

where \(x_k = (x, u_k)_X, \ y_k = (y, v_k)_Y\).

Now, the operator \(L_\varepsilon\), Eq. (3.5), is bounded iff \(\sup(\alpha_k^2 \rho_k v_k^{-1}) < + \infty\);\(^i\) in such case Eq. (3.6) takes the form

\[
\tilde{x} = \sum_{k=0}^{\infty} \alpha_k^2 \rho_k /\alpha_k^2 + \varepsilon^2 v_k, \tag{23.23}
\]
or a form similar to (3.10),

\[
\tilde{x} = \sum_{k=0}^{\infty} \left(1 - e^{-2\varepsilon} \right) \frac{y_k}{\alpha_k} u_k, \tag{24.24}
\]

where now \(J_k = \frac{1}{2} \ln \left[1 + (\alpha_k^2 \rho_k) / (\varepsilon^2 v_k)\right]\) is the average mutual information of the random variables \(\xi_k = (\xi, u_k)_X\) and \(\eta_k = (\eta, v_k)_Y\).

As regards the least-mean-square error (3.12) we have

\[
\delta(u, \varepsilon) = \epsilon \left(\sum_{k=0}^{\infty} \frac{\rho_k v_k}{\alpha_k^2 + \varepsilon^2 v_k} \left| (u, u_k)_X \right|^2 \right)^{1/2} \tag{25.25}
\]

and we see that \(\delta(u, \varepsilon) \to 0\), when \(\varepsilon \to 0\), for any \(u \in X\), without any further condition on the covariance operators \(R_\xi R_\xi\). Such a peculiarity is a consequence of the fact that, now, the operators \(AR_\xi A^*\) and \(R_\xi\) commute.

In the case where the w.r.v. \(\xi\) has a finite variance \(\Sigma_\xi^{-\frac{1}{2}} \rho_k < + \infty\), we have, for the least mean-square error (3.21), the expression

\[
\delta(\varepsilon) = \epsilon \left(\sum_{k=0}^{\infty} \frac{\rho_k v_k}{\alpha_k^2 + (\varepsilon / E^2) \beta_k} \right)^{1/2}, \tag{26.26}
\]

and we find that \(\delta(\varepsilon) \to 0\) when \(\varepsilon \to 0\). If we take \(N = 1\) (i.e., \(v_k = 1\)) and \(R_\xi = E(B^2 B^*)^{-1}\) (i.e., \(\rho_k = E^2 \beta_k^{-2}\)) we get

\[
\delta(\varepsilon) = \epsilon \left(\sum_{k=0}^{\infty} \frac{1}{\alpha_k^2 + (\varepsilon / E) \beta_k} \right)^{1/2}, \tag{27.27}
\]
i.e., we obtain formally Eq. (2.38) except for a factor of \(\sqrt{2}\).

An estimation, even more simple that (3.23) and similar to (2.41) can be obtained as follows: We consider the class of the linear bounded operators \(L: Y \to X\) defined by

\[
L_\varepsilon = \sum_{k=0}^{\infty} \alpha_k^{-1} y_k u_k, \tag{28.28}
\]

where \(I\) is an arbitrary finite set and \(y_k = (y, v_k)_Y\); then we look for the operator \(L_\varepsilon\) in such a class, which minimizes the mean square error (3.4). Now let us consider the set

\[
I_\varepsilon = \{k | \rho_k \geq \varepsilon^{-1} (v_k / \rho_k)\}; \tag{29.29}
\]

\(I_\varepsilon\) contains only a finite number of points for any \(\varepsilon > 0\) if \(\alpha_k^2 \rho_k v_k^{-1}\) tends to zero when \(\varepsilon \to + \infty\). When this condition is satisfied, it is easy to see that \(L_\varepsilon\) exists and it is the operator corresponding to \(I_\varepsilon\). Therefore, given the value \(y\) of \(\eta\), we have the following estimate for \(\xi\),

\[
\tilde{x} = \sum_{k \in I_\varepsilon} y_k u_k. \tag{30.30}
\]

Equation (3.29) can be obtained from Eq. (3.24) replacing the factor \(1 - \exp(-2J_k)\) by \(1\) when \(J_k > \frac{1}{2} \ln 2\) and by \(0\) when \(J_k < \frac{1}{2} \ln 2\).

The minimum of the mean-square error (3.4) over the class of the operators \(L_\varepsilon\) is

\[
\bar{\delta}(u, \varepsilon) = \inf_{L_\varepsilon} Q(u, L_\varepsilon, \varepsilon) = \sum_{k \in I_\varepsilon} \frac{v_k^2}{\alpha_k^2} \left| (u, u_k)_X \right|^2 + \sum_{k \in I_\varepsilon} \rho_k \left| (u, u_k)_X \right|^2; \tag{31.31}
\]

besides, when \(\xi\) has a finite variance, the minimum of the mean square error (3.11) is

\[
\bar{\delta}(\varepsilon) = \inf_{L_\varepsilon} Q(L_\varepsilon, \varepsilon) = \sum_{k \in I_\varepsilon} \frac{v_k^2}{\alpha_k^2} + \sum_{k \in I_\varepsilon} \rho_k; \tag{32.32}
\]

Since the condition \(\alpha_k^2 \rho_k v_k^{-1} \to 0, k \to + \infty\) is assumed, the
expressions (3.30) and (3.31) are finite and we have the following results: (a) $\delta(u,e) \rightarrow 0$, when $e \rightarrow 0$, for any $u \in X$; (b) if the w.r.v. $\xi$ has a finite variance, then $\delta(e) \rightarrow 0$, when $e \rightarrow 0$.

In order to prove (a), we remark that Eq. (3.30) can be written as follows:

$$\delta_i(u,e) = \sum_{k=0}^{+\infty} \delta_k(e) |(u,u_k)_X|^2,$$

(3.32)

where $\delta_k(e) = \rho_k k e_k l$ and $\delta_k(e) = e^2 y_k' \alpha_k^{-2}$, $keL$. Then statement (a) follows easily observing that $\delta_k(e) \leq \rho_k$ for any $e > 0$ and $\delta_k(e) \rightarrow 0$, when $e \rightarrow 0$. Analogously statement (b) can be proved recalling that, by assumption, $\sum_{k=0}^{+\infty} \rho_k < +\infty$.

4. OBJECT RESTORATION

In this section we consider the problem of object restoration for the optical system described in the introduction—see Eq. (1.1); we recall that the problem is to estimate the object $x$, given the image $y$ on the interval $[-1,1]$. If we assume that the object and the image have a finite energy, then both $x$ and $y$ belong to $L^{2}(-1,1)$. Besides, the integral operator $A$, defined as follows

$$A(x) = \int_{-1}^{1} \frac{\sin(c(t-s))}{\pi(t-s)} x(s) ds, \quad |t| < 1$$

(4.1)
is a linear, compact, self-adjoint, nonnegative operator in $L^2(-1,1)$; its inverse $A^{-1}$ exists but is not continuous, i.e., the problem of object restoration is an improperly posed problem. Therefore, we can apply the general methods of Secs. 2 and 3 by setting $X = Y = Z = L^2(-1,1)$.

We give here a list of the main properties of the operator $A$ which are useful for our analysis:

(i) $A$ is an operator of the trace class; indeed it is the first iterate of the finite Fourier transform operator, which is of the Schmidt class.

(ii) The eigenvectors of the operator $A$ are given by

$$u_k(t) = \lambda_k^{-1/2} \psi_k(c,t), \quad k = 0,1,2,..., \quad |t| < 1,$$

(4.2)

where $\psi_k(c,t)$ are the linear prolate spheroidal wavefunctions and $\lambda_k = \lambda_k(c)$ are the corresponding eigenvalues; $\{u_k\}_{k=0}^{+\infty}$ is an orthonormal basis in $X = L^2(-1,1)$.

(iii) The eigenvalues $\lambda_k = \lambda_k(c)$ form a decreasing sequence: $\lambda_0 > \lambda_1 > \lambda_2 > ...$, bounded away from 1 and approaching 0 as $c \rightarrow +\infty$. More precisely we have the following asymptotic behavior

$$\lambda_k = O\left(\frac{1}{K} \left(\frac{e c}{k}\right)^2\right), \quad k \rightarrow +\infty,$$

(4.3)

which can be derived from the power series expansion of $\psi_k(c,t)$ as a function of $c$. Besides we have

$$\sum_{k=0}^{+\infty} \lambda_k \leq \text{tr}(A) = 2c/\pi.$$
First let us observe that, for any vector \( x \in X \) we can always find a constraint operator \( B \) like (2.19), with \( \lim_{x \to \infty} \beta_k = + \infty \), and a positive number \( E \) such that the vector \( x \) belongs to the compact subset of \( X: K_E = \{ x \in X : \| x \| < E \} \). Besides, if we recall the properties of the linear prolate spheroidal wavefunctions, we can interpret the constraint (2.19) as a condition on the concentration of the finite Fourier transform of \( x \). Then one can easily understand that the goodness of the reconstruction of \( x \) is as much satisfactory, as far as the finite Fourier transform of \( x \) is concentrated in the interval \( [-c/2\pi,c/2\pi] \), or equivalently as rapidly as \( \beta_k \) tends to infinity.

Now we take \( B = A^{-1} \), i.e., we assume that the object \( x \) belongs to range(\( A \))—see Eq. (4.10); therefore, condition (2.19) holds with \( \beta_k = \lambda_k^{-1} \) and Theorem 2.4 can be applied. We see that in this case we have very satisfactory behavior of the stability estimate when \( \epsilon \to 0 \) (Hölder type stability). We can understand this result if we note that we are considering a class of objects whose finite Fourier transform is rather strongly concentrated in the interval \( [-c/2\pi,c/2\pi] \). However, usually, it is too restrictive to assume that the object belongs to this class. It is more interesting to suppose that the eigenvalues \( \beta_k \) grow like a power of \( k \) for \( k \to + \infty \). Let us spend a few words in order to justify this assertion.

From property (iv)—Eq. (4.5)—of the linear prolate spheroidal wavefunctions it follows that

\[
x_k = (x,u_k)_x = O(x_k), \quad k \to + \infty,
\]

where

\[
x_k = (k + \frac{1}{2})^{1/2} \int_{-1}^{1} x(t)P_k(t)dt.
\]

In other words the asymptotic behavior of the coefficients of the expansion of \( x \) as a series of linear prolate spheroidal wavefunctions is given by the asymptotic behavior of the coefficients of the expansion of \( x \) as a series of Legendre polynomials.

Now suppose that \( x \) has continuous derivatives up to the order \( m \) and vanishes outside the interval \( [-1 + \delta,1 - \delta] \) (\( \delta > 0 \), fixed but arbitrary), then the coefficients \( x_k \) decrease, for \( k \to + \infty \), at least as \( k^{-m} \). Indeed, by means of a well-known asymptotic formula of Legendre polynomials

\[\int_{-1}^{1} x(t)P_k(t)dt = O((k)^{m+1/2}),\]

so that, thanks to Eqs. (4.11) and (4.12), \( x_k = O(k^{-m}) \). As a consequence of this remark, we can conclude that, if we take \( \beta_k = k^\mu \) in Eq. (2.19), then the set \( K_E = \{ x \in X : \| x \| < E \} \) contains functions which vanish outside an interval \( [-1 + \delta,1 - \delta] \) and whose derivatives are continuous up to the order \( m > \mu + \frac{1}{2} \).

We have now the following result:

**Theorem 4.1:** If, for \( k > 2c/\pi \), \( \beta_k = \gamma k^\mu (\gamma \mu > 0) \), then there exists functions \( F_i, F_f \) such that

\[
EF_i(\frac{\epsilon}{E}) < M(\epsilon,E) < EF_f(\frac{\epsilon}{E})
\]

where \( M(\epsilon,E) \) is defined by Eq. (2.3) with \( \langle \rangle = \| \cdot \|_\epsilon \) and, for \( \epsilon \to 0 \)

\[
F_i(\epsilon) = O(\| \ln(1 - \gamma) \|_\epsilon), \quad F_f(\epsilon) = O(\| \ln(1 - \mu(1 - \eta)) \|_\epsilon)
\]

where \( \eta \) may be taken as small as desired.

**Proof:** If we recall the step function behavior of the eigenvalues \( \lambda_k \) and the increasing behavior of the constraints \( \beta_k \) for \( k > 2c/\pi \), we can assert that the set \( I(\epsilon) \) in Eq. (2.20), contains all the values of \( k \) until a certain \( k > 2c/\pi \), provided that the ratio \( \epsilon/E \) is sufficiently small. Therefore, in Eq. (2.22), Theorem 2.3, we have \( \alpha(\epsilon/E) = \lambda_k, \quad k(\epsilon/E) = \gamma k_0, \) and, from the inequalities \( \lambda_{k_0} > \frac{\epsilon}{E} \gamma k_0, \) \( (k_0 + 1)^\mu > k_0^\mu, \) we get

\[
\frac{E}{2\gamma} (k_0 + 1)^{1 - \eta} < M(\epsilon,E) < \frac{2E}{\gamma} k_0^{-\mu}.
\]

where \( \gamma \) is a function of the ratio \( \epsilon/E \). In order to prove the theorem we have to show that there exists an upper bound for \( k_0 \) which is \( O(\| \ln(\epsilon/E) \|) \) and a lower bound which is \( O(\| \ln(\epsilon/E) \|^1 - \eta) \), where \( \eta \) may be made as small as desired.

First we find an upper bound for \( k_0 \). Property (iii), and more precisely, Eq. (4.3), implies that there exists, for fixed \( c \), a constant \( \gamma \) such that \( \lambda_k > \gamma \exp[-(k)^{1-\mu}], \) \( k > 2c/\pi \), then, if we set \( k_0 = \sup\{ k \gamma \exp[-(k)^{1-\mu}] \} \), we have \( k_0 > k_0 \). On the other hand, \( k_0 = [\gamma] \), where \( [\cdot] \) denotes the greatest integer such that \( [\cdot] < \epsilon \) and \( \epsilon \) is the solution of the equation:

\[
\epsilon^{-\mu} \exp[-(\epsilon)^{1 - \mu}] = (\gamma\epsilon) / (\gamma\epsilon).
\]

It is not difficult to see that, when \( \epsilon \to 0, \epsilon = O(\| \ln(1 - \gamma) \|) \) so that, from the first half of inequality (4.16) we get the first half of inequality (4.14).

Then we find a lower bound for \( k_0 \). Equation (4.3) implies also that there exists (for fixed \( c \)) constants \( \gamma, \delta \) such that \( \lambda_k > \gamma \exp[-(k)^{1-\mu}], \) \( k > 2c/\pi \), where \( \delta \) may be made as small as desired; then, if we set

\[
k_0 = \sup\{ k \gamma \exp[-(k)^{1-\mu}] > (\gamma\epsilon) / (\gamma\epsilon) \}
\]

we have \( k_0 > k_0 \). On the other hand, \( k_0 > [\gamma] \), where \( [\cdot] \) is the solution of the equation:

\[
\gamma^{-\mu} \exp[-(\gamma)^{1 - \mu}] = (\gamma\epsilon) / (\gamma\epsilon).
\]

It is not difficult to see that, when \( \epsilon \to 0, \chi_1 = O(\| \ln(\epsilon/E) \|^1 - \eta) \), where \( \eta \) may be made as small as desired; from the second half of inequality (4.16) we get the second half of inequality (4.14).

Theorem 4.1 can be interpreted by saying that, when we do not make too restrictive assumptions on the unknown vector \( x \), then the restored continuity in the problem of object restoration is, at most, logarithmic.

Now we come to the probabilistic method. As regards the least mean-square error \( \delta(u,e) \), it is enough to observe that it has properties very similar to the best stability estimate \( M_\infty(\epsilon,E) \) (see, for instance Theorem 2.2 and Theorem 3.3).

It is more interesting to analyze the least mean-square error \( \delta(e) \). If we consider the case where all the operators \( A \),
\(R_3, R_4\) are diagonal, then \(\delta(e)\) is given by Eq. (3.26) with \(\alpha_k = \lambda_k\), now \(\delta(e)\) is finite if \(R_3\) is an operator of the trace class, i.e., if \(\xi\) has a finite variance. This assumption means that we are considering a class of optical objects whose average energy is finite. The same assumption is reasonable for \(\xi\); if we normalize the operator \(N\) in such a way that \(\text{tr}(N) = 1\), then \(e^2\) is just the average energy of the noise. We can easily find upper and lower bounds for \(\delta(e)\) if we further assume that \(\lambda_k^2 \rho_k v_k^{-1} \rightarrow 0, k \rightarrow +\infty\), i.e., we suppose that the amount of information on the component \(\xi_k = (\xi_k, u_k)\), of the object, contained in the corresponding component \(\eta_k = (\eta_k, u_k)\), of the image, tends to zero for \(k \rightarrow +\infty\)—see Eq. (3.24). Now we can say that the set \(I_\rho\), Eq. (3.28), contains only a finite number of points and we have the following result,

\[
\frac{1}{2} \delta(e) < \delta(e) < \delta(e),
\]

where

\[
\delta(e) = \left( \sum_{k \in I_\rho} \frac{v_k}{\lambda_k^2} \right)^{1/2} + \left( \sum_{k \in I_\rho} \rho_k \right)^{1/2}.
\]

The inequalities (4.17) follow from Eq. (3.26) when we observe that:

\[
\begin{align*}
\frac{1}{k} \sum_{\lambda_k} \frac{v_k}{\lambda_k^2} & \leq \frac{\rho_k v_k}{\lambda_k^2} + e^2 v_k \left( \sum_{k \in I_\rho} \frac{v_k}{\lambda_k^2} \right)^{-1} \\
\frac{1}{\rho_k} \left( \sum_{\lambda_k} \rho_k v_k \right)^{1/2} & \leq \frac{\rho_k v_k}{\lambda_k^2} + e^2 v_k \left( \sum_{k \in I_\rho} \frac{v_k}{\lambda_k^2} \right)^{-1}.
\end{align*}
\]

Now, the eigenvalues \(\lambda_k\) decrease for increasing \(k\); moreover, if the signal to noise ratios \(\rho_k / v_k\) also decrease (at least for \(k > S\)), then, for \(e\) sufficiently small, we have

\[
\delta(e) \leq \left( \sum_{k \in I_\rho} \frac{v_k}{\lambda_k^2} \right)^{1/2} + \left( \sum_{k \in I_\rho} \frac{v_k}{\lambda_k^2} \right)^{1/2} \leq \frac{e}{\lambda_k} + \frac{1}{\beta_k + 1},
\]

where \(\beta_k = \sqrt{\rho_k v_k^{-1}}\). As a consequence we see that, if the signal to noise ratios decrease like an inverse power of \(k\), \(\rho_k v_k^{-1} = (\text{constant}) \cdot k^{-\omega}\), then \(\delta(e)\) is bounded by a function which is \(O(\max\{\nu(1 - \nu)\})\), for \(e \rightarrow 0\) \((T\) may be taken as small as desired; the proof is just the same as in Theorem 4.1).

### 5. IMAGE EXTRAPOLATION

The problem of extrapolation of optical image data has been already considered; in that paper a stabilization condition was imposed by requiring the boundedness of the total energy of the image. Here we prefer to reinforce the stabilization constraint, requiring explicitly that the functions, which we want to extrapolate, are optical images corresponding to objects whose energy is finite and bounded by \(E\). Moreover, observing that the magnification of the optical system described by Eq. (1.1) is one, we assume that these optical images are measured in the interval \([-1, +1]\). In this case we can show that the restored continuity, in the problem of extrapolation of an image piece beyond its borders, is of the Hölder type. Let us remark that this result holds true only if the images, which we want to extrapolate, are approximately known in an interval which is equal to the support of the object. Otherwise, if the images are measured in a smaller interval, it is reasonable to conjecture that the restored continuity is not of the Hölder type, but weaker.

The problem of image extrapolation can be formulated as follows: Let \(x = x(t), -\infty < t < +\infty,\) be a square-integrable band-limited function,

\[
x \in L^2(\mathbb{R}),
\]

\[
x(t) = \int_{-\infty}^{\infty} e^{2\pi i t \xi} \hat{x}(\xi) d\xi.
\]

(5.1)

Let \(A \in \mathbb{R}\) be the restriction of \(x\) to the interval \([-1, 1]\), and \(y = \hat{x} + z\) the image measured in the interval \([-1, 1]\); \(z\) describes the experimental errors. Then the problem of image extrapolation is to estimate \(x\) given \(y\).

Now, in order to apply the methods of Secs. 2 and 3, we must specify the spaces \(X, Y,\) and the main properties of the operator \(A\).

In \(L^2(\mathbb{R}, + \infty)\) let \(X\) be the subspace of the bandlimited functions (5.1); since \(X\) is closed, then it is a separable Hilbert space with respect to the inner product of \(L^2(\mathbb{R})\). Besides, let \(Y \subseteq L^2(\mathbb{R}, + \infty)\) be the (closed) subspace of those functions whose support is contained in the interval \([-1, 1]\); \(Y\) is also a separable Hilbert space and is isomorphic to \(L^2(\mathbb{R})\). Then \(A : X \to Y\) is the operator defined by

\[
(Ax)(t) = x(t), \quad |t| \leq 1; \quad (Ax)(t) = 0, \quad |t| > 1.
\]

(5.2)

\(A\) is a compact operator of the Schmidt class; in other words the operator \(\mathcal{A}^*\) is of the trace class. At this purpose, let us remark that the adjoint \(\mathcal{A}^*: Y \to X\) is given by

\[
(A^* y)(t) = \int_{-1}^{1} \frac{\sin[(t - s)]}{\pi(t - s)} y(s) ds, \quad y \in Y.
\]

(5.3)

as easily follows from the identity for band-limited functions

\[
x(t) = \int_{-\infty}^{\infty} \frac{\sin[(t - s)]}{\pi(t - s)} x(s) ds, \quad x \in X.
\]

(5.4)

Indeed, from Eqs. (5.3) and (5.4), we get, for any \(x \in X, y \in Y, (Ax, y) = (x, A^* y)\). Then, from Eqs. (5.3) and (5.2) we obtain that the operator \(\mathcal{A}^*\) coincides with the operator (4.1) which is of the trace class, as it was shown in Sec. 4.

Next we want to show that all the singular values of both the operators \(A\) and \(A^*\) are nonzero. In other words we want to prove that: The equations \(Ax = 0\) and \(A^* y = 0\) have only the trivial solutions \(x = 0\) and \(y = 0\), respectively. In fact, \(Ax = 0\) means that \(x\) is a.e. zero in the interval \([-1, 1]\); since \(x\) is an entire function, we conclude that it is zero on the whole \((-\infty, +\infty)\). Analogously \(A^* y = 0\) means that the Fourier transform of \(y\), i.e., \(\hat{y}\), is a.e. zero in the interval \([-c/2\pi, c/2\pi]\); since \(\hat{y}\) is an entire function, they \(y\) is zero on the whole \((-\infty, +\infty)\), i.e., \(y = 0\).
Now, let us denote by \(| u_k \rangle_{k=0}^{+\infty}\) the set of the eigenvectors of the operator \(A^*A\) and by \(| v_k \rangle_{k=0}^{+\infty}\) the set of the eigenvectors of the operator \(AA^*\); \(| u_k \rangle_{k=0}^{+\infty}\) is an orthonormal basis in \(X\) and \(| v_k \rangle_{k=0}^{+\infty}\) is an orthonormal basis in \(Y\). If we recall that \(AA^*\) coincides with the operator (4.1) we have

\[
\psi_k(r, t) = \psi_k(c, t), \quad |r| < +\infty;
\]

\[
v_k(r) = \lambda_k^{-1/2} \psi_j(c, r), \quad |r| < 1; \quad k = 0, 1, 2, \ldots
\]

(5.5)

\(\psi_k(c, t)\) being normalized to one with respect to the norm of \(L^2(-\infty, +\infty)\)—see also property (5) of Sec. 4. The vectors \(u_k, v_k\) satisfy the relations

\[
Au_k = \sqrt{\lambda_k} v_k, \quad A^* v_k = \sqrt{\lambda_k} u_k, \quad k = 0, 1, 2, \ldots
\]

(5.6)

At this point we can specify the stabilization constraint. We assume that the unknown vector \(x\) belongs to a class of optical images corresponding to optical objects whose energy is finite and bounded by \(E\),

\[
x(t) = \frac{\sin[c(t - s)]}{\pi(t - s)} \psi(s)ds, \quad |r| < +\infty,
\]

(5.7)

\[\int_{-\infty}^{+\infty} |\psi(s)|^2 ds \leq E,\]

i.e., \(x = A^* v, \|v\|_E \leq E\). This condition can be written in the form (2.2) if we take \(B = \{A^*\}^1\). Since the operator \((B^* B)^{-1} = A^* A\) is of the trace class, the problem is stable with respect to the norms or seminorms considered in Sec. 2. Besides, from Theorem 2.4, it follows that

\[
M(\varepsilon, E) \leq \sqrt{E},
\]

(5.8)

where \(M(\varepsilon, E)\) is given by Eq. (2.3) with

\[
\langle x \rangle = \|x\|_\infty = \left( \int_{-\infty}^{+\infty} |x(t)|^2 dt \right)^{1/2}.
\]

(5.9)

The approximation for the unknown vector \(x\) is then given by

\[
\hat{x}(t) = \sum_{k=0}^{+\infty} \lambda_k^{-3/2} \frac{1}{\lambda_k^2 + (\varepsilon/E)^2} y_k \psi_k(c, t),
\]

(5.10)

where

\[
y_k(t) = \frac{1}{\sqrt{\lambda_k}} \int_{-\infty}^{+\infty} y(t) \psi_k(c, t) dt.
\]

(5.11)

Thanks to Eq. (2.7), \(\hat{x}\) gives an approximation to the unknown image \(x\) [in the sense of the norm of \(L^2(-\infty, +\infty)\)] with an error which is of the order of \(\sqrt{\varepsilon}\) (Hölder continuity).

Now, if we consider the best stability estimate \(M(\varepsilon, E)\), given by Eq. (2.3) with

\[
\langle x \rangle = \|x, u\|_X = \left( \int_{-\infty}^{+\infty} |x(t)u(t)|^2 dt \right),
\]

\[
\int_{-\infty}^{+\infty} |u(t)|^2 dt = 1,
\]

(5.12)

then, from inequality \(|\langle x, u \rangle| \leq \|x\|_X\) and Eq. (5.8) we obtain

\[
M(\varepsilon, E) \leq M(\varepsilon, E) \leq \sqrt{E}.
\]

(5.13)

Inequality (5.13) means that, if one pretends to reconstruct only some weighted averages of \(x\), then the error is, at most, of the order of \(\sqrt{\varepsilon}\). The situation can be more favorable if one takes \(\langle x, u \rangle\)—see Theorem 2.2.

Finally, we come to the norm \(\|\cdot\|_E\) defined in Eq. (2.30). Such a norm is interesting in the problem of image extrapolation, since it is stronger than the uniform norm. Indeed, from inequality (4.8) we have

\[
\|x\|_E = \sup_{t} |x(t)| = \sum_{k=0}^{+\infty} |x_k| \sup_{t} |\psi_k(c, t)|
\]

\[
< \sqrt{c/\pi} \|x\|_E.
\]

(5.14)

As a consequence, any stability estimate for \(M(\varepsilon, E)\) with \(\langle x \rangle = \|\cdot\|_E\) is a stability estimate for the uniform convergence. Now, from Theorem 2.5, we have

\[
\tilde{M}(\varepsilon, E) = \sqrt{2c} \left( \sum_{k=0}^{+\infty} \frac{\lambda_k^{5/2}}{\lambda_k^2 + (\varepsilon/E)^2} \right)^{1/2}.
\]

(5.15)

Then, from the inequality

\[
\tilde{M}(\varepsilon, E) \leq \sqrt{2c} C_{\alpha} \left( \sum_{k=0}^{+\infty} \frac{\lambda_k^{5/2}}{\lambda_k^2 + (\varepsilon/E)^2} \right)^{1/2} \left( \frac{\alpha}{E} \right)^{\alpha - 1}
\]

\[
= \sqrt{2} \Gamma_{\alpha} E \left( \frac{\varepsilon}{E} \right)^{\alpha},
\]

(5.16)

where \(\Gamma_{\alpha} < +\infty, \alpha > \frac{1}{2}\)—see Eq. (4.3). If we denote by \(M_{\alpha}(\varepsilon, E)\) the best stability estimate (2.3) with \(\langle x \rangle = \|\cdot\|_\infty\), then we have the result

\[
M_{\alpha}(\varepsilon, E) \leq \left( \text{constant} \right) E \left( \frac{\varepsilon}{E} \right)^{\alpha},
\]

(5.17)

where \(\alpha\) is any number between 0 and \(\frac{1}{2}\).

Now we come to the probabilistic method. The stabilization constraint (5.7) can be translated in a probabilistic language by saying that there exists a \(Y\)-valued w.r.t. \(w\) such that \(\xi = A^* u\); therefore, one has: \(R_\xi = A^* R_u A\) and the w.r.t. \(\xi\) has a finite variance.

We have the simplest case assuming for \(R_\xi\) and \(R_u\) the following expressions,

\[
R_u = E^2 I, \quad R_\xi = \varepsilon^2 I,
\]

(5.18)

where \(I\) is the identity operator in \(Y\). Then it is easy to verify that, in this case, Eq. (3.6) gives Eq. (5.10). Moreover, since

\[
\text{tr}(R_\xi) = E^2 \text{tr}(A^* A) = E^2 \sum_{k=0}^{+\infty} \lambda_k = (2c/\pi)E^2 < \infty—see
\]

Eq. (4.4)—both the least mean-squares errors (3.25) and (3.26) can be analyzed; more precisely we obtain
\[\delta(u, \epsilon) = \epsilon \left( \sum_{k=0}^{+\infty} \frac{\lambda_k}{\lambda_k^2 + (\epsilon/E)^2} \right)^{1/2} \left( |u, u_k, x\rangle \right)^2 \tag{5.19}\]

and

\[\delta(\epsilon) = \epsilon \left( \sum_{k=0}^{+\infty} \frac{\lambda_k}{\lambda_k^2 + (\epsilon/E)^2} \right)^{1/2}. \tag{5.20}\]

Then we have \(\delta(u, \epsilon) = 2^{-1/2} \overline{M}(\epsilon, \epsilon)\) and \(\delta(\epsilon) = 2^{-1/2} \overline{M}(\epsilon, \epsilon)\)—see Eq. (5.15)—and therefore the results previously obtained for the best stability estimates hold true also for the least mean-square errors.

Next we assume in Eq. (3.26), \(\rho_k = E \lambda_k w_k\), with the condition \(\sum_{k=0}^{+\infty} w_k = 1\); furthermore we assume, for the noise, the condition \(\sum_{k=0}^{+\infty} v_k = 1\). Then \(\delta(\epsilon)\) reads as follows,

\[\delta(\epsilon) = \epsilon \left( \sum_{k=0}^{+\infty} \frac{w_k \lambda_k}{\lambda_k^2 + (\epsilon/E)^2} \right)^{1/2}. \tag{5.21}\]

From the inequality used in Eq. (5.16) we get

\[\delta(\epsilon) < C_{\alpha} \epsilon^{1/2} E^{-1} \left( \sum_{k=0}^{+\infty} \frac{w_k \lambda_k}{\lambda_k^2 + (\epsilon/E)^2} \right)^{1/2} \tag{5.22}\]

with \(0<\alpha<1\). The series at the rhs of Eq. (5.22) is certainly convergent for \(0<\alpha<\frac{1}{2}\) (in the specific case \(\alpha = \frac{1}{2}\) it becomes \(\sum_{k=0}^{+\infty} \sqrt{w_k v_k} \leq 1\)). The series can also converge for \(\alpha > \frac{1}{2}\) if the sequence \(|v_k|/w_k\) is bounded; in the latter case we have a stronger Hölder continuity.

6. CONCLUSIONS

The main result of our paper is the proof that in the object restoration problem, the continuity, in most realistic cases, is at best logarithmic. This result derives essentially from the rapid exponential fall of the eigenvalues \(\lambda_k\) of the operator \(A\). Now we can try to extend this result. In fact an asymptotic behavior of the same kind [more precisely \(\lambda_k = O(\exp(-D_k \ln k))\) where \(D_k\) is a constant] holds true for every integral operator whose kernel is an entire analytic function of finite order. So we may argue that we get at best logarithmic continuity whenever, for inverting such an integral operator, we impose \textit{a priori} bounds on a finite number of derivatives of the solutions. Analytic kernels are involved in some inverse problems such as the near-field reconstruction, or the Bojarski–Lewy inverse scattering method when one has information only over a finite frequency band (bandpass kernels).\(^5\)

ACKNOWLEDGMENTS

The authors are indebted to Professor K. Miller and Professor G. Talenti for many useful discussions on the mathematical aspects of the problem, and to Professor A. Consortini, Professor L. Ronchi, and Professor G. Toraldo di Francia for discussing the questions of optics of this paper. They want also to thank Prof. W. M. Boerner for many suggestions and for a list of references.

\(^1\)B.R. Frieden, Prog. Opt. 9, 311 (1971).


\(^13\)A. Tikhonov and V. Arsenine, Méthodes de Résolution de Problèmes mal posés (Moscou, 1976), p. 29.


