HAL
open science

Generalized sentinels defined via least squares

Guy Chavent

To cite this version:

Guy Chavent. Generalized sentinels defined via least squares. [Research Report] RR-1932, INRIA. 1993. inria-00074742

HAL Id: inria-00074742

https://inria.hal.science/inria-00074742

Submitted on 24 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Generalized Sentinels Defined Via Least Squares

Guy CHAVENT

N° 1932
Juin 1993

PROGRAMME 6

Calcul scientifique,
modélisation et
logiciels numériques

Rapport de recherche

1993

Generalized Sentinels Defined Via Least Squares

Des Sentinelles Généralisées Définies via les Moindres Carrés

Guy Chavent *†

May 17, 1993

Abstract

We address the problem of monitoring a linear functional $(c, x)_E$ of an unknown vector x of an Hilbert space E, the available data being the observation z, in an Hilbert space F, of a vector Ax depending linearly of x through some known operator $A \in \mathcal{L}(E\ ;F)$. When $E = E_1 \times E_2$, $c = (c_1, 0)$ and A is injective and defined through the solution of a partial differential equation, J.L. LIONS (1988, 1990) introduced sentinels $s \in F$ such that $(s, Ax)_F$ is sensitive to $x_1 \in E_1$ but insensitive to $x_2 \in E_2$. In this paper, we prove the existence, in the general case of i) a generalized sentinel $(s, \sigma) \in \mathcal{F} \times E$,where $\mathcal{F} \supset F$ with F dense in $\mathcal{F}$, such that for any a priori guess x_0 of x, one has : $< s, Ax >_{\mathcal{F}\mathcal{F}'} + (\sigma, x_0)_E = (c, \hat{x})_E$, where $\hat{x}$ is the least squares estimate of x closest to x_0, and ii) a family of regularized sentinels $(s_n, \sigma_n) \in F \times E$ which converge to (s, σ). Generalized sentinels unify the least square approach (by construction !) and the sentinel approach (when A is injective), and provide a general frame work for the construction of "sentinels with special sensitivity" in the sense of J.L. LIONS (1990)).

Résumé

On s'intéresse à la surveillance d'une fonctionnelle linéaire $(c, x)_E$ d'un vecteur inconnu x d'un Hilbert E, lorsque l'on dispose d'une observation z, dans un Hilbert F, d'un vecteur Ax dépendant linéairement de x par l'intermédiaire d'une application linéaire $A \in \mathcal{L}(E\ ;F)$. Dans le cas où $E = E_1 \times E_2$, $c = (c_1, 0)$ et A est injectif et défini via la résolution d'une équation aux dérivées partielles, J.L. LIONS (1988,1990) a introduit la sentinelle $s \in F$ telle que $(s, Ax)_F$ soit sensible à $x_1 \in E$, mais insensible à $x_2 \in E_2$. Dans l'article qui suit, nous montrons l'existence, dans le cas général, i) d'une sentinelle généralisée $(s, \sigma) \in \mathcal{F} \times E$, où $\mathcal{F} \supset F$ avec F dense dans $\mathcal{F}$, telle que pour tout estimé a priori x_0 de x on ait $< s, Ax >'_{\mathcal{F}\mathcal{F}}$ $+(\sigma, x_0)_E = (c, \hat{x})_E$ où $\hat{x}$ est l'estimé de x au sens des moindres carrés le plus proche de x_0, et ii) d'une famille de sentinelles régularisées $(s_n, \sigma_n) \in F \times E$ qui converge vers (s, σ). Ces sentinelles généralisées unifient les approches par les moindres carrés (par construction !) et par les sentinelles (lorsque A est injectif), et fournissent un cadre général pour la construction de "sentinelles à sensibilité spéciales" au sens de J.L. LIONS (1990).

*INRIA, Domaine de Voluceau-Rocquencourt, BP 105, 78153 Le Chesnay Cédex, France

†CEREMADE, Université Paris Dauphine, 75775 Paris Cédex 16, France

Keywords

Least squares, sentinels, optimal control, regularization, duality.

Mots Clefs

Moindres carrés, sentinelles, contrôle optimal, régularisation, dualité.

1 Introduction

The notion of sentinel was introduced by J.L. Lions (1988) for the monitoring of a few components, say $x_1 \in \mathbf{R}$ and $x_2 \in \mathbf{R}$, of an unknown vector x of some Hilbert space E, the available data being the observation z, in some Hilbert space F, of a vector Ax depending linearly on x through a known operator $A \in \mathcal{L}(E;F)$.

Example :

Let Ω be a regular and bounded domain of $\mathbf{R}^n$; the pressure y in an aquifer located in Ω and produced through two wells at locations S_1 and S_2 is given by :

$$\begin{cases} -\Delta y + y = f_1\delta_{S_1} + f_2\delta_{S_2} & \text{in } \Omega \\ \frac{\partial y}{\partial \nu} = g & \text{on } \partial\Omega \end{cases}$$

where f_1, $f_2 \in \mathbf{R}$, are the production rate of the wells and $g \in L^2(\partial\Omega)$ the (usually unknown) boundary alimentation rate. In case where f_1 and f_2 are not directly accessible to measurement, one may be willing to monitor these two flow rates using a measurement z of the pressure y taken on some observatory $\omega \subset \Omega$, where ω is an open subset of Ω, which usually contains neither S_1 nor S_2. This example can be cast in the previous abstract framework by setting :

$$\begin{aligned} &x_1 = f_1, \quad x_2 = f_2 \\ &E = \mathbf{R} \times \mathbf{R} \times L^2(\partial\Omega), \quad F = L^2(\omega) \\ &A \ : x = (f_1, f_2, g) \in E \rightarrow Ax = y\,|_\omega \in F. \end{aligned}$$

We shall refer througout the introduction to this example to illustrate the various notions of sentinels. ■

If one splits E as $E \equiv \mathbf{R} \times \mathbf{R} \times \tilde{E}$, where $\tilde{E}$ is the space of the uninteresting unknown quantities (in the example $\tilde{E} = L^2(\partial\Omega)$ is the space of the unknown boundary data g), the idea behind sentinels consists in defining a function $S(x_1, x_2, \tilde{x})$ by :

$$(LS - i) \qquad S(x_1, x_2, \tilde{x}) = (s, Ax)_F$$

where $s \in F$ -the "sentinel"- is chosen in such a way that the function S is sensitive to changes in x_1 and x_2 :

$$(LS - ii) \qquad \frac{\partial S}{\partial x_j}(x_1, x_2, \tilde{x}) = \alpha_j \neq 0 \quad j = 1,2$$

but insensitive to the unknown and uninteresting quantities $\tilde{x}$:

$$(LS - iii) \qquad \frac{\partial S}{\partial \tilde{x}}(x_1, x_2, \tilde{x}) = 0.$$

This sentinel function $S(x_1, x_2, \tilde{x})$ can be evaluated easily each time a measurement $z \in F$ of Ax is available by simply computing the scalar product $(s, z)_F$.

In the example, a sentinel s is a function of $L^2(\omega)$ such that, whenever a measure $z \in L^2(\omega)$ of $y\,|_\omega$ is given, the simple calculation of $\int_\omega sz$ yields a number sensitive to the two production rates f_1 and f_2, but insensitive to the boundary data g.

A theory of sentinels (J.L. Lions, 1992) has been developped, which searches for an $s \in F$ which satisfies $(LS - i)$ and $(LS - iii)$ only, and whose non triviality is ensured by requiring that :

$$(LS-iv) \qquad \frac{1}{2} \| s - h_0 \|_F^2 = Min$$

where h_0 is some given a priori guess. In the applications, choosing h_0 irregular enough ensures that $s \not\equiv 0$, so that $(LS-ii)$ will be "generically" satisfied - but it is not included explicitly in the theory.

A more natural way of ensuring the non-triviality of s would be to require that it satisfies the three original properties $(LS-i)$, $(LS-ii)$ and $(LS-iii)$ for given (non zero) sensitivities α_j, $j = 1, 2$, as suggested in J.L. Lions (1990), where an example of such sentinel is given. This approach is the one which has been used in the numerical calculation (Ainseba, 1992, Kernevez, 1993, reintroduce condition $(LS-ii)$ in the objective function and use $h_0 \simeq 0$), and is the one we shall follow. But, instead of using $(LS-i)$, $(LS-ii)$ and $(LS-iii)$ as a definition, we shall make a detour and define our generalised sentinels via least-squares, which was often considered as a (more conventional) competitor to the sentinels.

The least-squares approach consists in :

- First, estimating the whole vector x of unknown quantities by solving :

$$(P) \qquad \text{find } \hat{x} \in E, \text{ close to } x_0, \text{ such that } \frac{1}{2} \| Ax - z \|_F^2 = Min$$

  where $z \in F$ is the measured quantity, and $x_0 \in E$ is an a-priori guess, which is required to ensure uniqueness of $\hat{x}$ in the (usual) case where A is not injective. (In the example, solving (P) amounts to estimate simultaneously the flowrates f_1 and f_2 and the boundary data g, the result here being independant of the a priori guess x_0, as the operator A happens to be injective).

- Second, extracting from $\hat{x}$ the usefull information on the monitored quantities :

$$\hat{x}_1 \in \mathbf{R}, \quad \hat{x}_2 \in \mathbf{R},$$

  (In the example, this amounts to throw away the estimate $\hat{g}$ of the boundary data, and to keep only the estimates $\hat{f}_1$ and $\hat{f}_2$ of the production rates).

When the monitoring has to be repeated many times, the computational advantage is clearly on the side of the sentinels (one scalar product each time opposed to one full optimization run !). When only one variable is being monitored, numerical calculation have shown (Ainseba, 1992) that the two approaches give quantitatively similar results.

We introduce in this paper a more general notion which includes both approaches : if one defines the <u>generalized</u> <u>inverse</u> of A by :

$$(GI) \qquad A^\dagger : (z, x_0) \to \hat{x} \text{ solution of } (P),$$

which is a bounded linear operator on $\mathcal{F}' \times E$, where $\mathcal{F}'$ is the dense subspace of F made of all z for which (P) has a solution, and if one defines a <u>monitoring</u> <u>vector</u> $c \in E$ by :

$$c = (\alpha_1, \alpha_2, 0) \in E = \mathbf{R} \times \mathbf{R} \times \tilde{E} = E',$$

then $c_0 A^+$ is a bounded linear operator from $\mathcal{F}'$ into $\mathbf{R}$, so that there exists always $s \in \mathcal{F}$ and $\sigma \in E$ such that :

$$(GS) \qquad (c, \hat{x})_E = (c, A^\dagger(z, x_0))_E = (s, z)_{\mathcal{F}\mathcal{F}'} + (\sigma, x_0)_E$$

holds for any $z \in \mathcal{F}'$ and $x_0 \in E$. We call <u>the couple</u> <u>(s, σ)</u> <u>a generalized</u> <u>sentinel</u>. It satisfies :

$$\frac{1}{2} \| s \|^2 = Min, \quad \sigma \in KerA.$$

By definition, generalized sentinel always exist (in opposition to the original sentinels defined by (LS)), but they give an information $(c, \hat{x})_E$ which, in general, depends on the chosen a priori guess $x_0 \in E$. However, we shall see that when the operator A is c-injective, i.e. satisfies the <u>sentinel</u> <u>hypothesis</u> :

$$(SH) \qquad Ax = 0 \Longrightarrow (c, x) = 0$$

then $\sigma = 0$, and (GS) implies that :

$$\alpha_1 x_1 + \alpha_2 x_2 = (c, x)_E = (c, \hat{x}) = (s, Ax)_{\mathcal{F}\mathcal{F}'} \text{ for all } x \in E$$

which shows that the sentinel function :

$$S(x_1, x_2, \hat{x}) = (s, Ax)_{\mathcal{F}\mathcal{F}'}$$

associated, in the original sense of J.L. Lions, to the generalized sentinel $s \in \mathcal{F}$ satisfies the original definitions $(S - i$ to $iv)$ of the sentinels (with $h_0 = 0$!). Notice that the monitoring vector $c \in E$ corresponds to the sensitivity vector $(\alpha_1, \alpha_2, 0)$ of the sentinel function S : S is a "sentinel with special sensitivity c" in the terminology of J.L. Lions, 1990, where c can be chosen freely in E.

However, the generalized sentinel s is not in F, but in a larger space $\mathcal{F}$, where $\mathcal{F}$ is a completion of F for the norm :

$$\begin{cases} \left(\| A^* f^I \|_E^2 + \| f^K \|_F^2 \right)^{1/2} \\ f = f^I + f^K, \quad f^I \in \overline{ImA}, \quad f^K \in KerA^*. \end{cases}$$

This space $\mathcal{F}$ is the one introduced by the Hilbert Uniqueness Method (HUM) (J.L. Lions, 1988) in order to give a sense to dual control problems for any target in E.

In the case of the example, the operator A is injective, so that the sentinel hypothesis (SH) is satisfied. Hence we see that there always exists a generalized sentinel $s \in \mathcal{F}$ such that :

$$< s, y \,|_\omega >_{\mathcal{F}\mathcal{F}'} = \alpha_1 f_1 + \alpha_2 f_2,$$

where α_1 and α_2 are any a priori given "sentivities". Of course, this formula is not practical for the computation of $\alpha_1 f_1 + \alpha_2 f_2$, as it requires the evaluation of a duality between $\mathcal{F}$ and $\mathcal{F}'$ instead of a scalar product in $L^2(\omega)$. But we shall see that, replacing the least-square problem (P) by a regularized vection (P_n) one can define a family of regularized sentinel (s_n, σ_n) with $s_n \in L^2(\omega) = F$ and $\sigma_n \in E$ such that (corollary (3.10) belows) :

$$\int_\omega s_n \; y \mid_\omega = \alpha_1 f_1 + \alpha_2 f_2 - (\sigma_n, (f_1, f_2, g))_E$$
$$\sigma_n \to 0 \text{ in } E,$$

and which can be used for the actual estimation of $\alpha_1 f_1 + \alpha_2 f_2$.

The paper is organized in three parts :

In paragraph 2, we give a synthetic presentation of the Hilbert Uniqueness Method in the context of linear least squares, duality and regularization. This section contains known material, or a slight generalization of thereof (see J.L. Lions 1988 a) for the HUM technique, and, among many others Tikhonov and Arsenin 1977, Groetsch 1984, Engl et al 1989, Chavent and Kunisch 1992 for the regularization aspects). This paragraph will introduce the main tools for the next paragraph on sentinels : How to replace the solution of (P) (or of its HUM-relaxed version) by the solution of an (possibly HUM or HUM^2 relaxed) associated dual problem (D^P), and how to approximate the solutions of (P) and (D^P) by regularization.

In paragraph 3, we define the generalized sentinels as suggested above, and investigate their relationship with the original sentinels. We show that generalized sentinels $(s, \sigma) \in \mathcal{F} \times E$ are given by the solution of a HUM-relaxed sentinel dual control problem, with target $c \in E$. We indicate also how to approximate $(s, \sigma) \in \mathcal{F} \times E$ by a sequence of regularized sentinels $(s_n, \sigma_n) \in F \times E$, show how to determine s_n and σ_n through either a dual or a primal regularized control problem, and give convergence results of (s_n, σ_n) to (s, σ).

Finally, we shall illustrate shortly in paragraph 4 the results of paragraph 3 on an example.

2 HUM, Duality and Regularization

Suppose we are given :

$$E, F = \text{ two Hilbert spaces} \tag{2.1}$$

$$A \in \mathcal{L}(E; F) \tag{2.2}$$

We shall use thoughout the paper the following decomposition of E and F :

$$\begin{cases} E = \overline{ImA^*} + KerA \quad , & \text{with } \overline{ImA^*} = KerA^\perp \\ F = \overline{ImA} + KerA^* \quad , & \text{with } \overline{ImA} = KerA^{*\perp}, \end{cases} \tag{2.3}$$

and the according decomposition of any $x \in E$ and $f \in F$ will be denoted by :

$$\begin{cases} x = x^I + x^K \quad , & \text{with } x^I \in \overline{ImA^*}, x^K \in KerA \\ f = f^I + f^K \quad , & \text{with } f^I \in \overline{ImA}, f^K \in KerA^*. \end{cases} \tag{2.4}$$

Given :

$$z \in F, \qquad x_0 \in E \tag{2.5}$$

we consider in this paragraph the resolution of the primal problem :

$$(P) \qquad \text{find } \hat{x} \in E, \text{ close to } x_0, \text{ s.t. } \frac{1}{2} \parallel Ax - z \parallel^2 = Min,$$

which has a solution as soon as the data z satisfies the primal regularity condition :

$$(PR) \qquad z \in ImA + KerA$$

When ImA is closed, then $ImA + KerA^* = F$ and the problem (P) is said to be wellposed.
When ImA is not closed, then $ImA + KerA^*$ is only dense in F, and the problem (P) is said to be illposed. We shall make in the sequel no hypothesis on the closedness of ImA.
The solution $\hat{x}$ of (P), when it exists, is given by :

$$\hat{x} = \hat{x}^I + \hat{x}^K \text{ with } A\hat{x}^I = z^I, \quad \hat{x}^K = x_0^K. \tag{2.6}$$

Similarly, if we are given :

$$z^* \in E, \quad f_0 \in F \tag{2.7}$$

we can consider the resolution of the dual problem :

$$(D) \qquad \text{find } \hat{f} \in F, \text{ close to } f_0, \text{ s.t. } \frac{1}{2} \| A^* f - z^* \|^2 = Min,$$

which has a solution as soon as the data z^* satisfies the dual regularity condition :

$$(DR) \qquad z^* \in ImA^* + KerA$$

According to the closedness of ImA^*, the dual problem can be wellposed or illposed, and its solution $\hat{f}$, when it exists, is given by :

$$\hat{f} = \hat{f}^I + \hat{f}^K \text{ with } A^*\hat{f}^I = z^{*I}, \quad \hat{f}^K = f_0^K. \tag{2.8}$$

Of course, as :

$$ImA \text{ closed} \iff ImA^* \text{ closed}, \tag{2.9}$$

primal and dual problems ar simultaneously wellposed or illposed.
We shall recall, in this paragraph, how the use of the Hilbert Uniqueness Method (HUM) of J.L. Lions (1988) can allow to get rid of the regularity hypothesis (PR) or (DR) required to ensure the existence of solutions to (P) or (D), making thus possible a systematic use of duality to replace the solution of (P) by that of (D) (and conversely !), and how the regularization technique enables us to calculate approximate solutions to these problems. Most of the results of this paragraph are already known, but we have tried to give a synthetic presentation which will serve as a basis for the section on sentinels.
We shall turn the four subspaces occuring in the decomposition (2.3) of E and F into Hilbert spaces by equipping them with the norm induced by E or F :

$$\left\{ \begin{array}{ll} \left(\overline{ImA^*}, \| \|_E\right) & , \ (KerA, \| \|_E) \\ \left(\overline{ImA}, \| \|_F\right) & , \ (KerA^*, \| \|_F) \\ \ldots \text{ are Hilbert spaces} & \end{array} \right. \tag{2.10}$$

We shall also turn $ImA(resp. ImA^*, ImAA^*)$ into an Hilberts space, using the fact that A is one-to-one from $\overline{ImA^*}$ on $ImA(resp. A^*$ is one-to-one from $\overline{ImA}$ on ImA^*, and AA^* is one-to-one from $\overline{ImA}$ on $ImAA^*)$:

$$(2.11)\qquad \begin{cases} (ImA,\ \|\ \|_{ImA}) & ,\ \text{where}\ \|\ Ax^I\ \|_{ImA}=\|\ x^I\ \|_E \\ (ImA^*,\ \|\ \|_{ImA^*}) & ,\ \text{where}\ \|\ A^*f^I\ \|_{ImA^*}=\|\ f^I\ \|_F \\ (ImAA^*,\ \|\ \|_{ImAA^*}) & ,\ \text{where}\ \|\ AA^*f^I\ \|_{ImAA^*}=\|\ f^I\ \|_F \\ \ldots \text{ are Hilbert spaces} \end{cases}$$

The three space in (2.11) are dense respectively in :

$$(2.12)\qquad \begin{cases} (\overline{ImA},\ \|\ \|_F) \\ (\overline{ImA^*},\ \|\ \|_E) \\ (ImA,\ \|\ \|_{ImA}) \text{ and } (\overline{ImA},\ \|\ \|_F), \end{cases}$$

so that one can identify dual spaces of (2.12) to subspaces of the dual spaces of (2.11) (see also figures 2.1 and 2.2) :

$$(2.13)\qquad \begin{cases} ImA^* \subset \overline{ImA^*} \subset (ImA^*)' \\ ImAA^* \subset ImA \subset \overline{ImA} \subset (ImA)' \subset (ImAA^*)' \\ \ldots \text{ all inclusions being continuous and dense.} \end{cases}$$

One checks easily that :

$$(2.14)\qquad \begin{cases} \forall x^I \in \overline{ImA^*} & ,\ \|\ x^I\ \|_{(ImA^*)'}=\|\ Ax^I\ \|_F \\ \forall f^I \in \overline{ImA} & ,\ \|\ f^I\ \|_{(ImA)'}=\|\ A^*f^I\ \|_E \\ & ,\ \|\ f^I\ \|_{(ImAA^*)'}=\|\ AA^*f^I\ \|_F \end{cases}$$

so that :

$$(2.15)\qquad \begin{cases} (ImA^*)' & \text{is a completion of } \overline{ImA^*} \text{ for the norm} & \|\ Ax^I\ \|_F \\ (ImA)' & \text{is a completion of } \overline{ImA} \text{ for the norm} & \|\ A^*f^I\ \|_E \\ (ImAA^*)' & \text{is a completion of } \overline{ImA} \text{ for the norm} & \|\ AA^*f^I\ \|_F \end{cases}$$

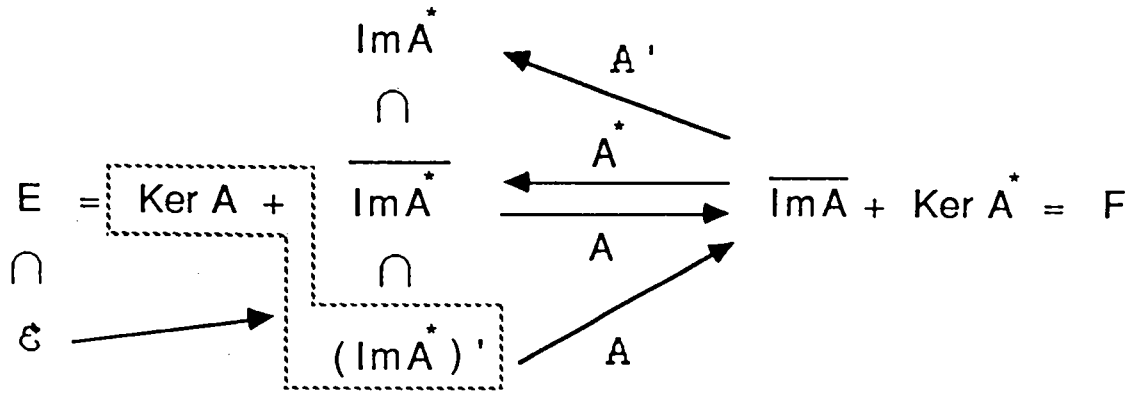

Figure 2.1 The HUM construction for the primal problem.

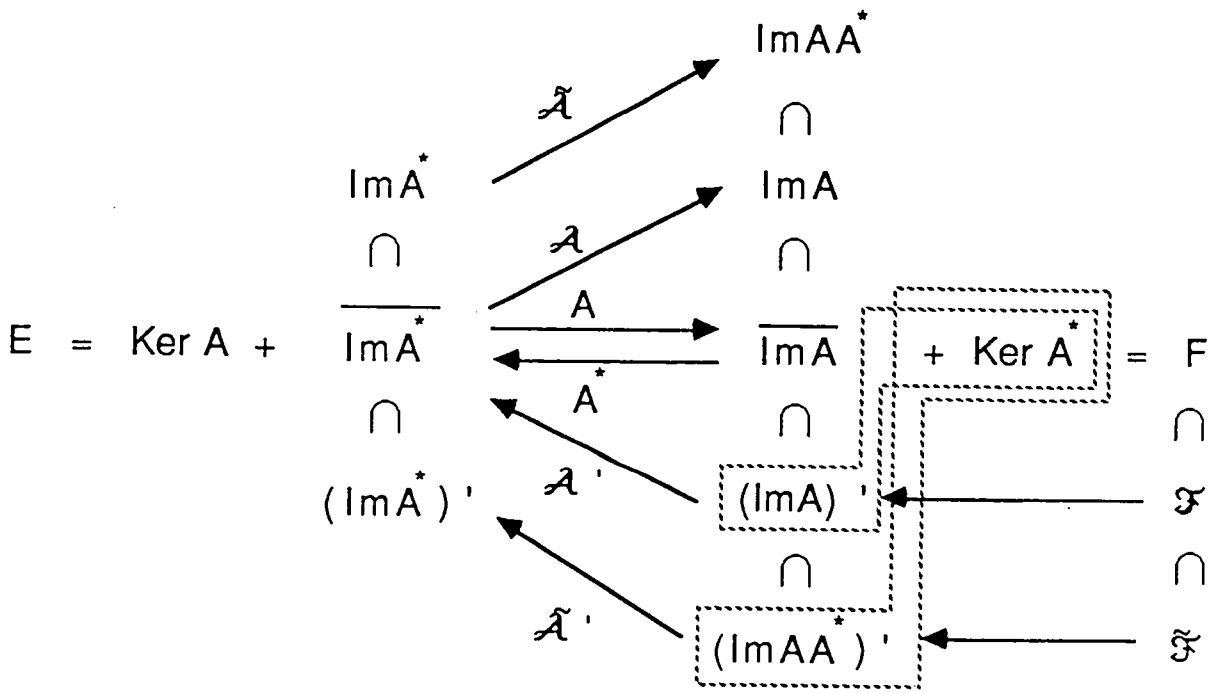

Figure 2.2 The HUM and HUM^2 construction for the dual problem.

It is then natural to define (see figures 2.1 and 2.2) :

$$(2.16)\qquad \begin{cases} \mathcal{A} : \overline{ImA^*} \to ImA & \text{s.t. } \mathcal{A}x^I = Ax^I \\ \tilde{\mathcal{A}} : ImA^* \to ImAA^* & \text{s.t. } \tilde{\mathcal{A}}(A^*f^I) = AA^*f^I \\ \mathtt{A}' : \overline{ImA} \to ImA^* & \text{s.t. } \mathtt{A}'f^I = A^*f^I \end{cases}$$

Obviously, the operators $\mathcal{A}$, $\tilde{\mathcal{A}}$ and $\mathtt{A}'$ are (isometrical) isomorphism, so that $\mathcal{A}'$, $\tilde{\mathcal{A}}'$ and $\mathtt{A}$ are also (isometrical) isomorphism :

$$(2.17)\qquad \begin{cases} \tilde{\mathcal{A}} \in Isom(ImA^*, ImAA^*) \\ \mathcal{A} \in Isom(\overline{ImA^*}, ImA) \\ \mathcal{A}' \in Isom((ImA)', \overline{ImA^*}) \\ \tilde{\mathcal{A}}' \in Isom((ImAA^*)', (ImA^*)') \\ \mathtt{A} \in Isom((ImA^*)', \overline{ImA}) \\ \mathtt{A}' \in Isom(\overline{ImA}, ImA^*) \end{cases}$$

2.1 *HUM* for the primal problem

In its original setting (P), the primal problem has a solution only when $z \in ImA + KerA^*$, which is dense in F. In order to ensure a solution for all $z \in F$, the Hilbert Uniqueness Method replaces the search for $\hat{x} = (\hat{x}^I, \hat{x}^K) \in E$ by the search by for $(\hat{f}^I, \hat{x}^K) \in \overline{ImA} \times KerA$, which is given, for any $z \in F$, by $\hat{f}^I = z^I$ and $\hat{x}^K = x_0^K$ (when z^I happens to be in ImA, then the solution $\hat{x}^I$ of (P) is given by $A\hat{x}^I = z^I$). But this problem is not an extension of (P), as $\overline{ImA}$ is not naturally an overspace of $\overline{ImA^*}$. In fact, one could identify $\overline{ImA^*}$ to a dense subspace of $\overline{ImA}$ using the injection $x^I \to Ax^I$ and keep the (f^I, x^K) unknowns. But the same approach for the dual problem would require to identify $\overline{ImA}$ to a dense subspace of $\overline{ImA^*}$! So we do not pursue in this direction. So the unknown $\hat{f}^I \in \overline{ImA}$ is replaced, using the (isometrical) isomorphism $\mathtt{A}$, by an unknown $x^I \in (ImA^*)'$, which is an overspace of $\overline{ImA^*}$ as we have seen in (2.13).

So we set (see figure 2.1).

$$(2.18)\qquad \begin{cases} \mathcal{E} = (ImA^*)' \times KerA \\ (\mathcal{E} \text{ is a completion of } E \text{ for the norm } (\| Ax^I \|_F^2 + \| x^K \|_E^2)^{1/2}) \end{cases}$$

which satisfies :

$$(2.19)\qquad \begin{cases} \mathcal{E}' \subset E \subset \mathcal{E} \\ \text{continuous and dense inclusions} \end{cases}$$

where :

$$(2.20)\qquad \mathcal{E}' = ImA^* + KerA.$$

We can now extend trivially $\mathtt{A}$ to $\mathcal{E}$ by deciding that :

$$(2.21)\qquad \forall x = (x^I, x^K) \in \mathcal{E} \ : \mathtt{A}x = \mathtt{A}x^I \in \overline{ImA} \subset F$$

so that the <u>*HUM*-relaxed primal problem</u> :

$$(P^{HUM})\qquad \text{find } \hat{x} \in \mathcal{E}, \text{ close to } x_0, \text{ s.t. } \frac{1}{2} \| \mathtt{A}x - z \|^2 = Min$$

has (not surprisingly !) a unique solution for all $z \in F$ and $x_0 \in E$, given by :

$$\hat{x} = (\hat{x}^I, \hat{x}^K) \text{ with } \mathbf{A}\hat{x}^I = z^I, \quad \hat{x}^K = x_0^K. \tag{2.22}$$

2.2 HUM and HUM^2 for the dual problem

Similarly, we can ensure that (D) has a solution for all $z^* \in E$ by defining (see figure 2.2) :

$$\begin{cases} \mathcal{F} = (ImA)' \times KerA^* \\ (\mathcal{F} \text{ is a completion of } F \text{ for the norm } (\| A^* f^I \|_E^2 + \| f^K \|_F^2)^{1/2}) \end{cases} \tag{2.23}$$

which satisfies :

$$\begin{cases} \mathcal{F}' \subset F \subset \mathcal{F}' \\ \text{continuous and dense inclusions} \end{cases} \tag{2.24}$$

where :

$$\mathcal{F}' = ImA + KerA^*. \tag{2.25}$$

We can now extend $\mathcal{A}'$ to $\mathcal{F}$ by :

$$\forall f = (f^I, f^K) \in \mathcal{F} \ : \mathcal{A}' f = \mathcal{A}' f^I \in \overline{ImA^*} \subset E \tag{2.26}$$

so that the <u>HUM-relaxed dual problem</u> :

$$(D^{HUM}) \qquad \text{find } \hat{f} \in \mathcal{F}, \text{ close to } f_0, \text{ s.t. } \frac{1}{2} \| \mathcal{A}' f - z^* \|_E^2 = Min$$

has a unique solution for all $z^* \in E$ and $f_0 \in F$, given by :

$$\hat{f} = (\hat{f}^I, \hat{f}^K) \text{ with } \mathcal{A}' \hat{f}^I = z^{*I}, \hat{f}^K = f_0^K. \tag{2.27}$$

One can even ensure that (D) has a solution for all $z^* \in \mathcal{E}$ by defining :

$$\begin{cases} \tilde{\mathcal{F}} = (ImAA^*)' + KerA^* \\ (\tilde{\mathcal{F}} \text{ is a completion of } F \text{ for the norm } (\| AA^* f^I \|_F^2 + \| f^K \|_F^2)^{1/2}) \end{cases} \tag{2.28}$$

which satisfies :

$$\begin{cases} \tilde{\mathcal{F}}' \subset \mathcal{F}' \subset F \subset \mathcal{F} \subset \mathcal{F} \\ \text{continuous and dense inclusions} \end{cases} \tag{2.29}$$

where :

$$\tilde{\mathcal{F}}' = ImAA^* + KerA^*. \tag{2.30}$$

We can also extend $\tilde{\mathcal{A}}'$ to $\tilde{\mathcal{F}}$ by :

$$\forall f = (f^I, f^K) \in \tilde{\mathcal{F}} \ : \tilde{\mathcal{A}}' f = \tilde{\mathcal{A}}' f^I \in (ImA^*)' \subset \mathcal{E} \tag{2.31}$$

so that the <u>HUM^2-relaxed dual problem</u> :

$$(D^{HUM^2}) \qquad \text{find } \hat{f} \in \tilde{\mathcal{F}}, \text{ close to } f_0, \text{ s.t. } \frac{1}{2} \| \tilde{\mathcal{A}}' f - z^* \|_{\mathcal{E}}^2 = Min$$

has a unique solution for all $z^* \in \mathcal{E}$ and $f_0 \in F$, given by (2.27) with $\mathcal{A}'$ replaced by $\tilde{\mathcal{A}}'$.

2.3 HUM and duality

We show now that, whenever (P) has a solution, it can be obtained by the resolution of an HUM-relaxed dual problem (D^{HUM}) whose target z^* is properly chosen :

<u>Theorem</u> **2.1** *For any a priori guess $x_0 \in E$ and any data $z \in F$ satisfying the primal regularity condition :*

$$(PR) \qquad z \in ImA + KerA^*$$

the unique solution $\hat{x}$ of the primal problem (P) is given by :

$$\hat{x} - x_0 = \mathcal{A}'\hat{f} \tag{2.32}$$

where $\hat{f}$ is the unique solution of the dual problem (D^{HUM}) with target z^* and a priori guess f_0 satisfying :

$$z^* \in E, \quad z^{*I} = \hat{x} - x_0, \quad f_0 \in F, \tag{2.33}$$

which rewrites :

$$(D^{HUM,P}) \quad \textit{find } \hat{f} \in \mathcal{F}, \textit{ close to } f_0, \textit{ s.t. } \frac{1}{2} \| \mathcal{A}'f \|_E^2 - < f, z^I - Ax_0 >_{\mathcal{F}\mathcal{F}'} = Min.$$

<u>Proof</u> :
Let $\hat{f} \in \mathcal{F}$ be the solution of $(D^{HUM,P})$. From (2.27) and (2.33) we see that :

$$\hat{f} = (\hat{f}^I, \hat{f}^K) \text{ with } \mathcal{A}'\hat{f}^I = \hat{x} - x_0, \quad \hat{f}^K = f_0^K$$

which prooves (2.32). Then developping the objective function of (D^{HUM}) yields (up to an additive constant) the objective function of $(D^{HUM,P})$, as :

$$(\mathcal{A}'f, z^*)_E = (\mathcal{A}'f, z^{*I})_E = (\mathcal{A}'f, \hat{x} - x_0)_E = (f, \mathcal{A}\hat{x} - \mathcal{A}x_0)_{\mathcal{F}\mathcal{F}'}. \blacksquare$$

<u>Remark</u> **2.2** *At first glance, replacing the resolution of (P) by that of $(D^{HUM,P})$ does not seem to yield any simplification especially because the objective function of the latter contains $\hat{x}$ or z^I, whose evaluation requires the solution of (P) ! But we shall see in paragraph 2.4 that the regularized versions of (P) and $(D^{HUM,P})$ will have their objective functions defined in term of z only, so that one will have really the choice of solving (P_ε) or $(D_\varepsilon^{HUM,P})$.* $\blacksquare$

Similarly, one can replace the solution of the HUM-relaxed primal problem by that a properly chosen dual problem :

<u>Theorem</u> **2.3** *For any a priori guess $x_0 \in E$ and any dat $z \in F$, the unique solution $\hat{x}$ of the HUM-relaxed problem (P^{HUM}) is given by :*

$$\hat{x} - x_0 = \tilde{\mathcal{A}}'\hat{f} \tag{2.34}$$

where $\hat{f}$ is the unique solution of the HUM^2-relaxed dual problem (D^{HUM2}) with target z^* and a priori guess f_0 satisfying :

$$z^* \in \mathcal{E}, \quad z^{*I} = \hat{x} - x_0, \quad f_0 \in F. \tag{2.35}$$

We do not explicit, for this HUM^2-relaxed dual problem, the objective function in a form similar to $(D^{HUM,P})$: as $\mathcal{E}$ is not identified to its dual, the formula would contian the Riesz isomorphism of $\mathcal{E}$ to $\mathcal{E}'$, and is of no use in itself - the only useful formula will be the regularized formula of the next paragraph.

2.4 HUM and Regularization

We show in this paragraph how regularization, beside being often a necessity because of the ill conditionned nature of the problems we attempt to solve, will also solve the apparent contradiction mentionned in remark (2.2), as it will lead to the solution of problems (P_n) or (D_n) who require only as entries noise corrupted observations $z_n \in F$ of z.

We shall suppose in this paragraph that the exact observation $z \in F$ is approximated by a sequence $z_n \in F$ of experimental data such that :

$$\| z_n - z \| \leq \delta_n \tag{2.36}$$

Then we choose a sequence ε_n of regularization parameters, satisfying :

$$\varepsilon_n > 0 \tag{2.37}$$

According to Tychonov, we replace, when $z \in ImA + KerA^*$, the problem (P) by <u>a sequence of regularized least squares problems</u> :

$$(P_n) \quad \text{find } \hat{x}_n \in E, \text{ s.t. } \frac{1}{2} \| Ax - z_n \|_F^2 + \frac{\varepsilon_n^2}{2} \| x - x_0 \|_E^2 = Min,$$

and, according to J.L. Lions, we replace the HUM-relaxed dual problem $(D^{HUM,P})$ associated to (P) as in theorem (2.1) by :

$$(D_n^P) \quad \text{find } \hat{f}_n \in F, \text{ s.t. } \frac{1}{2} \| A^* f \|_E^2 - (f, z_n - Ax_0)_F + \frac{\varepsilon_n^2}{2} \| f - f_0 \|_F^2 = Min$$

Notice that, when $z \in F$, the same problem (P_n) will be used as a regularized version of the HUM-relaxed problem (P^{HUM}), and that (D_n^P) will also be used as a regularized version of $(D^{HUM^2,P^{HUM}})$.

Of course, problems (P_n) and (D_n) are equivalent to their Euler equations.

$$(EP_n) \quad \text{find } \hat{x}_n \in E, \text{ s.t. } (A^*A + \varepsilon_n^2 I)\hat{x}_n = A^* z_n + \varepsilon_n^2 x_0$$

and

$$(ED_n^P) \quad \text{find } \hat{f}_n \in F, \text{ s.t. } (AA^* + \varepsilon_n^2 I)\hat{f}_n = z_n - Ax_0 + \varepsilon_n^2 f_0,$$

this latter equation corresponding to the "regularized HUM problems" of J.L. Lions.

The above regularized problems are written only using the original spaces E and F, the original operator A, the a priori guesses $x_0 \in E$ and $f_0 \in F$, and the experimental data $z_n \in F$. The relation between (P_n) and (D_n^P) is given by (compare with theorems 2.1 and 2.3) :

<u>Theorem</u> 2.4 *Let (2.36 (2.37) hold. Then problems (P_n) and (D_n^P) have, for all n, a unique solution each, which are linked by :*

$$\hat{x}_n - x_0 + \varepsilon_n^2 y_{0n} = A^* \hat{f}_n \tag{2.38}$$

where $y_{0n} \in E$ is the solution of :

(2.39) $$\textit{find } y_{0n} \in E, \ s.t. \ \frac{1}{2} \| Ay - f_0 \|_F^2 + \frac{\varepsilon_n^2}{2} \| y \|_E^2 = Min$$

or equivalently of :

(2.40) $$(A^*A + \varepsilon_n^2 I) y_{0n} = A^* f_0.$$

Moreover, the values of the minima in (P_n) and (D_n^P) satisfy :

(2.41) $$Min(P_n) + \varepsilon_n^2 Min(D_n^P) + \frac{\varepsilon_n^4}{2} \left(f_0, Ay_{0n} + 2(\hat{f} - f_0) \right)_F = 0.$$

<u>**Remark**</u> **2.5** *When the dual variable $\hat{f}$ is selected as the minimum norm one (i.e. when $f_0 = 0$), then $y_{0n} = 0$ for all n, and (2.38) and (2.41) rewrite :*

(2.42) $$\hat{x} - x_0 = A^* \hat{f}_n,$$

(2.43) $$Min(P_n) + \varepsilon_n^2 Min(D_n^P) = 0,$$

so that the resolution of (P_n) can simply be replaced by that of (D_n^P), followed by the application of A^*. ■

<u>Proof of theorem (2.4)</u> :

Let $\hat{f}_n$ be the solution of (D_n^P) or (ED_n^P), and define $\tilde{x}_n$ by :

(2.44) $$\tilde{x}_n - x_0 + \varepsilon_n^2 y_{0n} = A^* \hat{f}_n$$

Multiplying left (ED_n^P) by A^* yields :

$$(A^*A + \varepsilon_n^2 I) A^* \hat{f}_n = A^* z_n - A^* A x_0 + \varepsilon_n^2 A^* f_0$$

Substituting the value of $A^* \hat{f}_n$ given by (2.44) yields :

$$A^* A \tilde{x}_n + \varepsilon_n^2 \tilde{x}_n - A^* A x_0 - \varepsilon_n^2 x_0 + \varepsilon_n^2 (A^*A + \varepsilon_n^2 I) y_{0n} = A^* A x_0 + \varepsilon_n^2 A^* f_0$$

i.e. , using (2.40) :

$$A^*(A\tilde{x}_n - z_n) + \varepsilon_n^2 (\tilde{x}_n - x_0) = 0$$

which is the Euler equation (EP_n) of (P_n). Hence $\tilde{x}_n$ is the (unique) solution $\hat{x}_n$ of (P_n) which ends the proof of (2.38). We proove now (2.41) ; we first evaluate $Min(P_n)$:

$$Min(P_n) = \frac{1}{2} \| A\hat{x}_n - z_n \|^2 + \frac{\varepsilon_n^2}{2} \| \hat{x}_n - x_0 \|_E^2$$

and, substituting $\hat{x}_n$ using (2.38) and $AA^* \hat{f}_n$ using (ED_n^P) :

$$Min(P_n) = \frac{\varepsilon_n^4}{2} \| f_0 - \hat{f}_n - Ay_{0n} \|_F^2 + \frac{\varepsilon_n^2}{2} \| A^* \hat{f}_n - \varepsilon_n^2 y_{0n} \|_E^2$$

$$Min(P_n) = \frac{\varepsilon_n^2}{2} \parallel A^* \hat{f}_n \parallel_E^2 + \frac{\varepsilon_n^4}{2} \parallel \hat{f}_n - f_0 \parallel_F^2 + \frac{\varepsilon_n^4}{2} \parallel Ay_{0n} \parallel_F^2 - \varepsilon_n^4 (f_0, Ay_{0n})_F + \frac{\varepsilon_n^6}{2} \parallel y_{0n} \parallel_E^2 .$$

Substituting $\varepsilon_n^2 y_{0n}$ using (2.40) in one factor of the last term yields :

$$Min(P_n) = \frac{\varepsilon_n^2}{2} \parallel A^* \hat{f}_n \parallel^2 + \frac{\varepsilon_n^4}{2} \parallel \hat{f}_n - f_0 \parallel_F^2 - \frac{\varepsilon_n^4}{2} (f_0, Ay_{0n})_F. \tag{2.45}$$

Now we evaluate $Min(D_n^P)$:

$$Min(D_n^P) = \frac{1}{2} \parallel A^* \hat{f}_n \parallel_E^2 - (\hat{f}_n, z_n - Ax_0)_F + \frac{\varepsilon_n^2}{2} \parallel \hat{f}_n - f_0 \parallel^2$$

and, substituting $z_n - Ax_0$ using (ED_n^P) :

$$Min(D_n^P) = -\frac{1}{2} \parallel A^* \hat{f}_n \parallel_E^2 - \frac{\varepsilon_n^2}{2} \parallel \hat{f}_n - f_0 \parallel_F^2 - \varepsilon_n^2 (f_0, \hat{f}_n - f_0)_F. \tag{2.46}$$

Combining (2.45) and (2.46) yields (2.41). ■

We investigate now the convergence of $\hat{x}_n$ and $\hat{f}_n$ towards the solutions $\hat{x}$ and $\hat{f}$ of (P) or (P^{HUM}) and $(D^{HUM,P})$ or $(D^{HUM^2,P^{HUM}})$.

Theorem 2.6 *Let (2.36 (2.37) hold.*

i) *For all n, the K-components of $\hat{x}_n$ and $\hat{f}_n$ satisfy :*

$$\hat{x}_n^K = \hat{x}^K, \text{ with } \hat{x}^K = x_0^K \tag{2.47}$$

$$\varepsilon_n^2 \hat{f}_n^K = \varepsilon_n^2 \hat{f}^K + z_n^K, \text{ with } \hat{f}^K = f_0^K. \tag{2.48}$$

ii) *If moreover :*

$$z \in F \text{ and } \delta_n \to 0, \quad \varepsilon_n \to 0 \tag{2.49}$$

Then :

$$\varepsilon_n \hat{x}_n \to 0 \text{ in } E \tag{2.50}$$

$$A\hat{x}_n^I \to z^I \text{ in } F (\text{hence } \hat{x}_n \to \hat{x} \text{ in } \mathcal{E}) \tag{2.51}$$

$$\varepsilon_n y_{0n} \to 0 \text{ in } E \tag{2.52}$$

$$Ay_{0n} \to f_0^I \text{ in } F \tag{2.53}$$

$$\varepsilon_n^2 \hat{f}_n \to z^K \text{ in } F \tag{2.54}$$

$$AA^* \hat{f}_n^I \to z^I - Ax_0 \text{ in } F (\text{hence } A^* \hat{f}_n^I \to \hat{x} - x_0 \text{ in } \mathcal{E} \text{ and } \hat{f}_n^I \to \hat{f}^I \text{ in } \tilde{\mathcal{F}}) \tag{2.55}$$

iii) *If moreover :*

$$z \in ImA + KerA^* \text{ and } \delta_n/\varepsilon_n \to 0 \tag{2.56}$$

Then :

$$\hat{x}_n \to \hat{x} \text{ in } E \tag{2.57}$$

$$\| A\hat{x}_n - z^I \|_F = 0(\varepsilon_n) \tag{2.58}$$

$$A^* \hat{f}_n^I \to \hat{x} - x_0 \text{ in } E(\text{hence } \hat{f}_n^I \to \hat{f}^I \text{ in } \mathcal{F}) \tag{2.59}$$

iv) *If moreover :*

$$z \in ImAA^* + KerA^* \text{ and } \delta_n = 0(\varepsilon_n^2) \tag{2.60}$$

Then :

$$\| \hat{x}_n - \hat{x} \|_E = 0(\delta_n^{\frac{1}{2}}) = 0(\varepsilon_n) \tag{2.61}$$

$$\| A\hat{x}_n - z^I \|_F = 0(\delta_n) = 0(\varepsilon_n^2) \tag{2.62}$$

$$\hat{f}_n^I \to \hat{f}^I \text{ in } F \tag{2.63}$$

Proof :

i) Formula (2.47) results immediately from the projection of (EP^n) on $KerA$, and (2.48) from the projection of (ED_n^P) on $KerA^*$.

ii) Let (2.49) hold, and choose $\varepsilon > 0$. On can find $\tilde{x} \in E$ such that :

$$\begin{cases} A\tilde{x}^I = \tilde{z}^I \text{ with } \| z^I - \tilde{z}^I \|_F \leq \varepsilon/4 \\ \tilde{x}^K = x_0^K \end{cases} \tag{2.64}$$

Then by definition of $\hat{x}_n$ one has :

$$\| A\hat{x}_n - z_n \|_F^2 + \varepsilon_n^2 \| \hat{x}_n - x_0 \|_E^2 \leq \| A\tilde{x} - z_n \|_F^2 + \varepsilon_n^2 \| \tilde{x} - x_0 \|_E^2$$

which rewrites as :

$$\begin{aligned} \| A\hat{x}_n - \tilde{z}^I \|_F^2 + \varepsilon_n^2 \| \hat{x}_n - \tilde{x} \|_E^2 \quad \leq \quad & \| A\tilde{x} - z_n \|_F^2 + \varepsilon_n^2 \| \tilde{x} - x_0 \|_E^2 \\ + \quad & \| A\hat{x}_n - \tilde{z}^I \|_F^2 + \varepsilon_n^2 \| \hat{x}_n - \tilde{x} \|_E^2 \\ - \quad & \| A\hat{x}_n - z_n \|_F^2 - \varepsilon_n^2 \| \hat{x}_n - x_0 \|_E^2 \end{aligned}$$

Using the identity $a^2 + b^2 - (a+b)^2 = -2ab$ and the fact that $A\tilde{x} = \tilde{z}^I$ yields :

$$\| A\hat{x}_n - \tilde{z}^I \|_F^2 + \varepsilon_n^2 \| \hat{x}_n - \tilde{x} \|_E^2 \leq 2(A\hat{x}_n - \tilde{z}^I, z_n - \tilde{z}^I)_F + 2\varepsilon_n^2(\hat{x}_n - \tilde{x}, x_0 - \tilde{x})_E$$

which rewrites :

$$\begin{aligned} \| A\hat{x}_n - \tilde{z}^I \|_F^2 + \varepsilon_n^2 \| \hat{x}_n - \tilde{x} \|_E^2 \leq 2(A\hat{x}_n - \tilde{z}^I, z_n - z + z^I - \tilde{z}^I)_F \\ + 2\varepsilon_n^2(\hat{x}_n - \tilde{x}, x_0 - \tilde{x})_E \end{aligned} \tag{2.65}$$

Using the Cauchy Schwarz inequality and (2.36) yields :

$$\| A\hat{x}_n - \tilde{z}^I \|_F^2 + \varepsilon_n^2 \| \hat{x}_n - \tilde{x} \|_E^2 \leq 2 \| A\hat{x}_n - \tilde{z}^I \|_F \left(\delta_n + \frac{\varepsilon}{4}\right) + 2\varepsilon_n^2 \| \hat{x}_n - \tilde{x} \|_E \| x_0 - \tilde{x} \|_E$$

which is of the form (with obvious notations) :

$$a^2 + b^2 \leq 2a\alpha + 2b\beta.$$

Hence :

$$(a-\alpha)^2 + (b-\beta)^2 \leq \alpha^2 + \beta^2 \leq (\alpha+\beta)^2$$

and finally :

$$a \leq 2\alpha + \beta \tag{2.66}$$

$$b \leq \alpha + 2\beta. \tag{2.67}$$

From (2.67) we see that :

$$\varepsilon_n \| \hat{x}_n - \tilde{x} \|_E \leq \delta_n + \frac{\varepsilon}{4} + \varepsilon_n \| x_0 - \tilde{x} \|$$

which implies that :

$$\varepsilon_n \| \hat{x}_n \|_E \leq \varepsilon_n \| \tilde{x} \| + \delta_n + \frac{\varepsilon}{4} + \varepsilon_n \| x_0 - \tilde{x} \|$$

so that :

$$\varepsilon_n \| \hat{x}_n \|_E \leq \varepsilon \text{ for } n \text{ large enough,}$$

which proves (2.50).
Then from (2.66) we obtain :

$$\| A\hat{x}_n - \tilde{z}^I \|_F \leq 2(\delta_n + \frac{\varepsilon}{4}) + \varepsilon_n \| x_0 - \tilde{x} \|_E$$

which implies that :

$$\| A\hat{x}_n - z^I \|_F \leq 2\delta_n + \frac{3\varepsilon}{4} + \varepsilon_n \| x_0 - \tilde{x} \|_E$$

so that :

$$\| A\hat{x}_n - z^I \|_F \leq \varepsilon \text{ for } n \text{ large enough}$$

which proves (2.51).

Then (2.52) (2.53) correspond exactly to the convergence results (2.50) (2.51) applied to problem (2.39).

We plug now (2.38) in (ED_n^P) to obtain :

$$A\left[\hat{x}_n - x_0 + \varepsilon_n^2 y_{0n}\right] + \varepsilon_n^2 \hat{f}_n = z_n - Ax_0 + \varepsilon_n^2 f_0 \tag{2.68}$$

Passing to the limit in (2.68) when $n \to \infty$ yields, using (2.51) (2.36) :

$$\varepsilon_n^2 \hat{f}_n \to z - z^I = z^K$$

which is (2.54).

But passing to the limit directly in (ED_n^P) yields :

$$AA^* \hat{f}_n \to z - Ax_0 - z^K = z^I - Ax_0$$

which is (2.55).

iii) Let us suppose now that (2.56) holds. So we can choose in (2.64) $\tilde{x} = \hat{x}$, $\tilde{z}^I = z^I$, so that (2.65) rewrites :

$$\begin{aligned} &\| A\hat{x}_n - z^I \|_F^2 + \varepsilon_n^2 \| \hat{x}_n - \hat{x} \|_E^2 \leq \\ &\quad 2(A\hat{x}_n - z^I, z_n - z)_F + 2\varepsilon_n^2 (\hat{x}_n - \hat{x}, x_0 - \hat{x})_F \end{aligned} \tag{2.69}$$

Applying to (2.69) the same technique as we did to (2.65) yields :

$$\begin{aligned} \| A\hat{x}_n - z^I \|_F &\leq 2\delta n + \varepsilon_n \| x_0 - \hat{x} \| \\ \varepsilon_n \| \hat{x}_n - \hat{x} \|_E &\leq \delta_n + 2\varepsilon_n \| x_0 - \hat{x} \| \end{aligned}$$

which proves (2.58).

Let now $\varepsilon > 0$ be given. As $\hat{x} - x_0 \in \overline{ImA^*}$, there exists $g \in F$ such that :

$$\hat{x} - x_0 = A^* g + e \text{ with } \| e \| \leq \varepsilon/3. \tag{2.70}$$

Plugging this formula in the last term of (2.69) yields :

$$\begin{aligned} \| A\hat{x}_n - z^I \|_F^2 + \varepsilon_n^2 \| \hat{x}_n - \hat{x} \|_E^2 \leq 2(A\hat{x}_n - z^I, z_n - z - \varepsilon_n^2 g)_F \\ -2\varepsilon_n^2 (\hat{x}_n - \hat{x}, e)_E. \end{aligned} \tag{2.71}$$

Using the Cauchy Schwarz inequality and (2.36) (2.70) yields :

$$\begin{aligned} \| A\hat{x}_n - z^I \|_F^2 + \varepsilon_n^2 \| \hat{x}_n - \hat{x} \|_E \leq 2 \| A\hat{x}_n - z^I \| (\delta_n + \varepsilon_n^2 \| g \|) \\ +2\varepsilon_n^2 \| \hat{x}_n - \hat{x} \| \varepsilon/3 \end{aligned} \tag{2.72}$$

which, as in the part ii) of the proof, implies that :

$$\| A\hat{x}_n - z^I \|_F \leq 2(\delta_n + \varepsilon_n^2 \| g \|) + \varepsilon_n . \varepsilon/3 \tag{2.73}$$

$$\varepsilon_n \| \hat{x}_n - \hat{x} \|_E \leq \delta_n + \varepsilon_n^2 \| g \| + 2\varepsilon_n . \varepsilon/3. \tag{2.74}$$

As we have supposed in (2.56) that $\delta_n/\varepsilon_n \to 0$, then we obtain from (2.74) :

$$\| \hat{x}_n - \hat{x} \| \leq \varepsilon \text{ for } n \text{ large enough,}$$

which proves (2.57). Then (2.59) results from (2.57) and (2.38).

iv) Let us now suppose that (2.60) holds. Then we can satisfy (2.70) with $e = 0$ for some $g \in F$, so that (2.73) (2.74) rewrite :

$$\begin{aligned} \| A\hat{x}_n - z^I \|_F &\leq 2(\delta_n + \varepsilon_n^2 \| g \|) \\ \varepsilon_n \| \hat{x}_n - \hat{x} \|_E &\leq \delta_n + \varepsilon_n + \varepsilon_n^2 \| g \| \end{aligned}$$

which implies (2.62) and (2.63) using the hypothesis that $\delta_n = 0(\varepsilon_n^2)$ made in (2.60). ■

3 Least Squares and Generalized Sentinels

3.1 Definition of a generalized sentinel

Suppose one is in charge of monitoring the quantity $(c, x)_E \in \mathbf{R}$ where :

$$c \in E \text{ is a given vector ("monitoring vector")} \tag{3.1}$$

and $x \in E$ is a vector of unknown quantities. The only information for this monitoring is available through a known linear operator :

$$A \in \mathcal{L}(E; F), \tag{3.2}$$

and consists in an observation z of Ax. We suppose that there is no measurement or model error :

(3.3) $$z \in ImA + KerA^*$$

and postpone the handling of noisy measurements $z_n \in F$ to the paragraph 3.3 on regularized sentinels.

The most reasonnable thing to do, at least from a conceptual point of view, for the monitoring of $(c, x)_E$, consists in :

i) Computing the best least-squares estimate $\hat{x}$ of x by solving the problem :

$$(P) \quad \text{find } \hat{x} \in E, \text{ close to } x_0, \text{ s.t. } \frac{1}{2} \| Ax - z \|_F^2 = Min$$

where an a priori guess :

(3.4) $$x_0 \in E$$

has been used to select, when A is not injective, one solution of (P).

ii) Calculating :

(3.5) $$(c, \hat{x}_E) = \text{ best estimate of } (c, x)_E$$

This approach is however computationally expensive, as the whole vector x of unknown quantities has to be estimated (by solving (P)) each time one estimation of the real number $(c, x)_E$ (often one single component of x !) is required.

But using the least-squares and HUM construction of paragraph 2, one sees that problem (P) has a unique solution $\hat{x}$ depending in a linear and continuous manner on $z \in ImA + KerA^*$ and $x_0 \in E$. This defines a mapping :

(3.6) $$A^\dagger : (z, x_0) \to \hat{x}$$

called the <u>pseudo inverse of A</u>, which satisfies :

(3.7) $$A^\dagger \in \mathcal{L}(\mathcal{F}' \times E \,; E)$$

where (cf.paragraph 2)

(3.8) $$\mathcal{F}' = ImA \times KerA^*$$

With this pseudo inverse, the estimation of (c, x) by steps i) + ii) rewrites simply as :

(3.9) $$(c, \hat{x})_E = (c, A^\dagger(z, x_0))_E.$$

But :

(3.10) $$c \circ A^\dagger \in \mathcal{L}(\mathcal{F}' \times E \,; \mathbf{R}) = \mathcal{F} \oplus E$$

so that there exists :

(3.11) $$s \in \mathcal{F}, \quad \sigma \in E$$

such that :

$$(c, \hat{x})_E = (s, z)_{\mathcal{F}\mathcal{F}'} + (\sigma, x_0)_E \tag{3.12}$$

Once the elements $s \in \mathcal{F}$ and $\sigma \in E$ have been computed, the estimation of $(c, \hat{x})_E$ can be performed at a moderate computational cost by (3.12) as many times as needed. This leads us to the :

Definition 3.1 *We shall call generalized sentinel for the monitoring of (c, x) from the observation of Ax any $s \in \mathcal{F}$ and $\sigma \in E$ such that (3.12) holds for any $z \in ImA + \overline{KerA^*}$ and any $x_0 \in E$. Such a sentinel always exists and is unique under the sole hypothesis (3.1) and (3.2).*

This definition is slightly more general than the original definition of J.L. Lions (1988), with which it will coincide when A is c-injective, i.e. when :

$$Ax = 0 \text{ implies } (c, x)_E = 0 \tag{3.13}$$

as we shall see in the next paragraph.

3.2 Characterization of Generalized Sentinels

Suppose that we have decomposed the monitoring vector $c \in E$ on the two orthogonal subspaces $KerA$ and $\overline{ImA^*}$:

$$c = c^K + c^I, \text{ with } c^K \in KerA,\ c^I \in \overline{ImA^*} \tag{3.14}$$

Then the following characterization holds :

Theorem 3.2 *Under the general hypothesis (3.1) (3.2), $(s, \sigma) \in \mathcal{F} \times E$ is a generalized sentinel for the monitoring of $(c, x)_E$ from the observation of Ax if and only if :*

$$s \in (ImA)', \quad \sigma \in KerA \tag{3.15}$$

and :

$$\mathcal{A}'s + \sigma = c \tag{3.16}$$

Proof :
We show first that conditions (3.15) (3.16) are necessary. We shall use the notations :

$$\begin{array}{lll} z = z^I + z^K, & \text{with } z^I \in ImA, & z^K \in KerA^* \\ x_0 = x_0^K + x_0^I, & \text{with } x_0^K \in KerA, & x_0^I \in \overline{ImA^*} \\ \hat{x} = \hat{x}^K + \hat{x}^I, & \text{with } \hat{x}^K \in KerA, & \hat{x}^I \in \overline{ImA^*} \end{array}$$

etc....
Then the mapping $c \circ A^\dagger : (z, x_0) \to (c, \hat{x})$ can be decomposed into four steps :
Step 1 : project z on to $\overline{ImA}$:

$$z \to z^I \in ImA$$

Step 2 : invert A :

$$z^I \to \hat{x}^I \in \overline{ImA^*} \text{ such that } A\hat{x}^I = z^I$$

Step 3 : choose $\hat{x}$ close to x_0 :

$$x_0 \to \hat{x}^K = x_0^K$$

Step 4 : gather the result :

$$\hat{x} = \hat{x}^K + \hat{x}^I$$

This prooves that $\hat{x}$, and hence $(c, \hat{x})_E$, are independant of the components z^K of z on $KerA^*$ and x_0^I of x_0 on $\overline{ImA^*}$. But by definition of the sentinel, one has :

$$(c, \hat{x})_E = (s, z)_{\mathcal{F}\mathcal{F}'} + (\sigma, x_0)_E \tag{3.17}$$

which implies that :

$$s \in (ImA)', \quad \sigma \in KerA, \tag{3.18}$$

which proves (3.15). Now using (3.18), one can rewrite (3.17) as :

$$(c, \hat{x})_E = (s, z^I)_{\mathcal{F}\mathcal{F}'} + (\sigma, x_0^K)_E$$

and, using the above steps 2 and 3 :

$$\begin{aligned} (c, \hat{x})_E &= (s, A\hat{x}^I)_{\mathcal{F}\mathcal{F}'} + (\sigma, \hat{x}^K) \\ &= (\mathcal{A}'s, \hat{x}_I)_E + (\sigma, \hat{x}_K)_E \end{aligned}$$

i.e.

$$(\mathcal{A}'s + \sigma - c, \hat{x})_E = 0 \tag{3.19}$$

But (3.17) holds for all $z \in ImA + KerA^*$ and all $x_0 \in E$. From steps 1 thru 4 we see that this implies that (3.19) holds for all $\hat{x} \in E$! This prooves (3.16).

One checks then easily that the reciprocal property holds : if (s, σ) satisfies (3.15) (3.16), then (s, σ) is a sentinel. ■

This characterization suggests immediately a constructive approach for the determination of a sentinel :

Corollary 3.3 *Under hypothesis (3.1) (3.2), the generalized sentinel $(s, \sigma) \in \mathcal{F} \times E$ for the monitoring of $(c, x)_E$ from the observation of Ax is given by resolution of the (HUM-relaxed) sentinel dual control problem :*

$$(D_s) \quad \textit{find } s \in \mathcal{F}, \textit{ close to } 0, \textit{ s.t. } \frac{1}{2} \| \mathcal{A}'f - c \|_E^2 = Min$$

followed by a residual calculation :

$$\sigma = c - \mathcal{A}'s \tag{3.20}$$

or, equivalently, by the resolution of the dual problem associated to (D_s) by theorem (2.3), which is the (HUM^2-relaxed) sentinel primal control problem :

$$(P_s) \quad \textit{find } r \in \tilde{\mathcal{E}}, \textit{ close to } 0, \textit{ s.t. } \frac{1}{2} \| \tilde{\mathtt{A}} x - s \|^2_{\mathcal{F}} = Min$$

or, equivalently, of its Euler equation :

$$(EPs) \quad \textit{find } r \in \tilde{\mathcal{E}} \textit{ s.t. } r^K = 0, \quad \tilde{\mathtt{A}} r^I = s,$$

followed by :

$$\sigma = c - \mathcal{A}' \tilde{\mathtt{A}} r \tag{3.21}$$

(This corollary is a simple rewriting of theorems (3.2) and (2.1)).

Remark 3.4 *Of course, the primal sentinel control problem (P_s) in the above form does not seem very operational, as it is expressed in term of the solution s of (D_s) ! But we shall see in the next paragraph that, for regularized sentinels, (P_s) can be expressed in term of c only, and hence will actually define an alternative way for the calculation of s and σ.* ■

We see from Corollary (3.3) that a sentinel can be computed by solving a (HUM-relaxed) control problem for the adjoint operator of A, whose target is the monitoring vector c. The residual of this control problem yields that part σ of the sentinel which indicates the dependance of the estimated quantity (c, x) on the a priori guess x_0. When this residual is zero, then $\sigma = 0$ and the sentinel $(s, 0)$ coincides with the sentinel in the original sense of J.L. Lions (1988), up to the fact it is in $\mathcal{F}$, not in F :

Corollary 3.5 *Let $(s, \sigma) \in \mathcal{F} \times E$ be a sentinel for the monitoring of $(c, x)_E$ from the observation of $Ax \in F$, to which we associate the sentinel function :*

$$\begin{cases} S \ : E \times E \to \mathbf{R} \\ (x, x_0) \to S(x, x_0) = (s, Ax)_{\mathcal{F}\mathcal{F}'} + (\sigma, x_0)_E, \end{cases} \tag{3.22}$$

which satisfies :

$$S(x, x_0) = (c, x)_E + (\sigma, x_0^K - x^K)_E. \tag{3.23}$$

Then, under any of the three following equivalent hypotheses :

$(S-1)$ *A is c-injective , i.e. $Ax = 0$ implies $(c, x)_E = 0$,*
$(S-2)$ *$c \in \overline{ImA^*}$,*
$(S-3)$ *$\sigma = 0$,*

the sentinel function $S(x, x_0)$ satisfies :

$$S(x, x_0) = (c, x)_E \textit{ for all } x \in E \tag{3.24}$$

so that $s \in \mathcal{F}$ is a sentinel in the sense of J.L. Lions (1988).

We skip the proof of this corollary, which follows immediately from theorem (3.2).

If one wants to match completely the original definition of the sentinels by J.L. Lions, as recalled in the introduction, one needs simply i) to decompose E into $E_1 \times E_2 \times \tilde{E}$, according to the unknowns x_1 and x_2 with respect to which one wants the sentinel function to be sensitive, and $\tilde{x}$ which is made of the uninteresting unknowns, and ii) to choose $c = (\alpha_1, \alpha_2, 0) \in E$ as monitoring vector. Then, under any of the hypothesis (S), the sentinel function satisfies :

$$S(x, x_0) = (s, \mathcal{A}x)_{\mathcal{F}\mathcal{F}'} = \alpha_1 x_1 + \alpha_2 x_2 \tag{3.25}$$

which satisfies :

$$\frac{\partial S}{\partial x_1}(x, x_0) = \alpha_1, \quad \frac{\partial S}{\partial x_2}(x, x_0) = \alpha_2, \quad \frac{\partial S}{\partial \tilde{x}}(x, x_0) = 0 \tag{3.26}$$

and hence is a sentinel in the original sense of J.L. Lions.

3.3 Regularized sentinels

As in practice the control problem (P) used for the definition of the pseudo inverse $A^\dagger$ is illposed, it is necessary to replace it by regularized control problems : given a sequence

$$\varepsilon_n > 0, \quad \varepsilon_n \to 0 \tag{3.27}$$

of regularization parameters, we define a sequence of regularized control problems :

$$(P_n) \quad \text{find } \hat{x}_n \in E \text{ s. t. } \frac{1}{2} \| Ax - z_n \|^2 + \frac{\varepsilon_n^2}{2} \| x - x_0 \|^2 = Min$$

which have, for each n, a unique solution $\hat{x}_n$ as soon as :

$$z_n \in F, \quad x_0 \in E. \tag{3.28}$$

Hence (Pn) defines an <u>approximate pseudo inverse</u> $\underline{A_n^\dagger}$ of $A^\dagger$:

$$A_n^\dagger \ : z_n \in F, \ x_0 \in E \to \hat{x}_n \in E \tag{3.29}$$

which satisfies :

$$A_n^\dagger \in \mathcal{L}(F \times E \ ; E), \tag{3.30}$$

so that :

$$c \circ A_n^\dagger \in \mathcal{L}(F \times E \ ; \mathbf{R}). \tag{3.31}$$

This shows, as in paragraph (3.1), that there exists $s_n \in F$ and $\sigma_n \in E$ such that :

$$(c, \hat{x}_n) = (s_n, z_n)_F + (\sigma_n, x_0)_E \tag{3.32}$$

Hence we are led to the :

<u>Definition</u> 3.6 *Given $\varepsilon_n > 0$, we call <u>regularized sentinel for the monitoring of $(c, x)_E$ from the observation of Ax</u> any $s_n \in F$ and $\sigma_n \in E$ such that (3.32) holds for any $z_n \in F$ and any $x_0 \in E$. This regularized sentinel always exists and is unique, under the sole hypothesis (3.1) (3.2) and (3.27).*

Then it results immediately from the part iii) of theorem (2.6) applied to (P) and (P_n) that one has the :

Corollary 3.7 *Let hypothesis (3.1) (3.2) (3.27) hold.*
Then for any z and any sequence of observation z_n satisfying :

$$z \in ImA + KerA^* \tag{3.33}$$

$$z_n \in F, \ \| z_n - z \|_F \leq \delta_n, \ \text{with } \delta_n/\varepsilon_n \to 0, \tag{3.34}$$

one has :

$$(s_n, z_n)_F + (\sigma_n, x_0)_E = (c, \hat{x}_n) \to (c, \hat{x}) = (s, z)_{\mathcal{F}\mathcal{F}'} + (\sigma, x_0)_E \tag{3.35}$$

where :

$\hat{x}_n$ *is the regularized least square estimate of x, given by (P_n)*
$\hat{x}$ *is the least square estimate of x, given by (P).*

We turn now to the determination of s_n and σ_n :

Theorem 3.8 *Under the general hypothesis is (3.1) (3.2) and (3.27), the regularized sentinels (s_n, σ_n) are given by the solution of either of :*

i) *the regularized sentinel dual control problem :*

$$(D_{s,n}) \quad \text{find } s_n \in F, \ s.t. \ \frac{1}{2} \| A^* s - c \|_E^2 + \frac{\varepsilon_n^2}{2} \| s \|_F^2 = Min$$

or, equivalently, of its Euler equation :

$$(ED_{s,n}) \quad (AA^* + \varepsilon_n^2 I) s_n = Ac,$$

followed by a residual calculation :

$$\sigma_n = c - A^* s_n \tag{3.36}$$

ii) *the regularized sentinel primal control problem :*

$$(P_{s,n}) \quad \text{find } r_n \in E, \ s.t. \ \frac{1}{2} \| Ar \|_F^2 - (c, r)_F + \frac{\varepsilon_n^2}{2} \| r \|_E^2 = Min$$

or, equivalently, of its Euler equation :

$$(EP_{s,n}) \quad (A^* A + \varepsilon_n^2 I) r_n = c,$$

followed by :

$$s_n = Ar_n, \quad \sigma_n = \varepsilon_n^2 r_n. \tag{3.37}$$

Proof : If we define :

$$\begin{aligned}
&F_n = F \times E, \text{ equipped with the norm} \\
&\| (f,x) \|_{F_n} = (\| f \|_F^2 + \varepsilon_n^2 \| x \|_E^2)^{1/2} \\
&A_n = (A, I) \in \mathcal{L}(E\,;F_n) \\
&\tilde{z} = (z, x_0) \in F_n
\end{aligned}$$

then we can rewrite the regularized problem (P_n) as :

$$(\tilde{P}_n) \quad \text{find } \hat{x}_n \in E \text{ s.t. } \frac{1}{2} \| A_n x - \tilde{z} \|_{F_n}^2 = Min,$$

where of course :

$$(3.38) \qquad \begin{cases} KerA_n = \{0\} \\ ImA_n \text{ is closed, so that } \mathcal{F}_n' = F_n = \mathcal{F}_n \\ A_n^* = A^* \oplus \varepsilon_n^2 I \text{ is onto.} \end{cases}$$

Then the sentinel for the monitoring of $(c, \hat{x}_n)$ from the observation of $A_n x = (Ax, x)$ in F_n is, according to the definition (3.1), the (unique) vector $((\tilde{s}_n, r_n), 0) \in F_n \times E$ such that, for any $(z_n, x_0) \in F_n$ one has :

$$(c, \hat{x}_n) = ((\tilde{s}_n, r_n), (z_n, x_0))_{F_n}$$

i.e. by definition of the scalar product in F_n :

$$(3.39) \qquad (c, \hat{x}_n) = (\tilde{s}_n, z_n)_F + \varepsilon_n^2 (r_n, x_0)_E$$

Comparing (3.39) to the definition (3.32) of the regularized sentinel (s_n, σ_n) shows already that :

$$(3.40) \qquad s_n = \tilde{s}_n$$

$$(3.41) \qquad \sigma_n = \varepsilon_n^2 r_n$$

But we know from theorem (3.2) that the sentinel $((\tilde{s}_n, r_n), 0) \in F_n \times E$ satisfies :

$$(3.42) \qquad (\tilde{s}_n, r_n) \in \overline{ImA_n}, \quad 0 \in KerA = \{0\}$$

$$(3.43) \qquad A_n^*(\tilde{s}_n, r_n) = c$$

Using (3.40) and the fact that ImA is closed, we see from (3.42) that :

$$(3.44) \qquad s_n = Ar_n.$$

Using (3.43) and the formula for A_n^* given in (3.38), we see from (3.43) that :

$$(3.45) \qquad A^* s_n + \varepsilon_n^2 r_n = c.$$

Plugging (3.44) into (3.45) gives :

$$A^*Ar_n + \varepsilon_n^2 r_n = c, \tag{3.46}$$

which is the Euler equation $(EP_{s,n})$. Then multiplying (3.46) left by A and using (3.44) gives :

$$AA^*s_n + \varepsilon_n^2 s_n = Ac, \tag{3.47}$$

which is the Euler equation $(ED_{s,n})$. ■

We can now apply theorem (2.6) on the convergence of regularized solutions to the case where (P) is replaced by (D_s) and the dual (D^P) of (P) is replaced by the dual (P_s) of (D_s). As the sentinel problem (D_s) is a HUM-relaxed problem (the target c is in E, not in $ImA^* + KerA$!) the results of section ii) of theorem (2.6) apply with the following correspondance :

$$\begin{array}{lll} E & \text{replaced by} & F \\ A & \text{replaced by} & A^* \\ z_n = z & \text{replaced by} & c \\ \delta_n & \text{replaced by} & 0 \\ x_0 & \text{replaced by} & 0 \\ f_0 & \text{replaced by} & 0 \\ \hat{x}_n & \text{replaced by} & s_n \\ y_{0n} & \text{replaced by} & 0 \\ \hat{f}_n & \text{replaced by} & r_n, \end{array} \tag{3.48}$$

which gives the :

Theorem 3.9 *Under the general hypothesis (3.1) (3.2) and (3.27), the regularized sentinels $(s_n, \sigma_n) \in F \times E$ converge to the generalized sentinel $(s, \sigma) \in \mathcal{F} \times E$ in the following sense :*

$$s_n^K = s^K = 0 \text{ for all } n \tag{3.49}$$

$$\varepsilon_n s_n^I \to 0 \text{ in } F \tag{3.50}$$

$$s_n \to s \text{ in } \mathcal{F} (\text{i.e. } A^*s_n^I \to c^I \text{ in } E \text{ and } s_n^K \to 0 \text{ in } F) \tag{3.51}$$

$$\sigma_n^K = \sigma = c^K \text{ for all } n \tag{3.52}$$

$$\sigma_n^I \to 0 \text{ in } E \tag{3.53}$$

and the solution $r_n \in E$ of $(P_{s,n})$ is related to the solution $r \in \tilde{\mathcal{E}}$ of (P_s) by :

$$r_n^K = c^K/\varepsilon_n^2 \text{ for all } n (\text{hence } r_n^K \text{ is unbounded in } E \text{ and } \tilde{\mathcal{E}}) \tag{3.54}$$

$$r_n^I \to r^I \text{ in } \tilde{\mathcal{E}} (\text{i.e. } A^*Ar_n^I \to c^I \text{ in } E). \tag{3.55}$$

<u>If moreover the sentinel hypothesis (S) is satisfied</u> (i.e. A is c-injective) then $c^K = 0$ and :

$$\sigma_n^K = 0 \textit{ for all } n \tag{3.56}$$

$$r_n^K = 0 \textit{ for all } n \tag{3.57}$$

$$r_n \to r \textit{ in } \tilde{\mathcal{E}} (\textit{i.e. } A^* A r_n^I \to c^I \textit{ in } E \textit{ and } r_n^K \to 0 \textit{ in } E) \tag{3.58}$$

<u>Corollary</u> **3.10** *Under the general hypothesis (3.1) (3.2) and (3.27), the regularized sentinel function $S_n(x, x_0)$ defined by :*

$$S_n(x, x_0) = (s_n, Ax)_F + (\sigma_n, x_0)_E \tag{3.59}$$

satisfies :

$$\begin{aligned} S_n(x, x_0) &= (c, x)_E + (\sigma_n, x_0 - x)_E \\ &= S(x, x_0) + (\sigma_n^I, x_0^I - x^I)_E \end{aligned} \tag{3.60}$$

<u>Proof</u> : Formula (3.60) follows immediately from (3.36) of theorem (3.8), (3.52) of theorem (3.9) and from the property (3.23) of the sentinel function S. ■

From (3.23) we see that the estimate of $(c, x)_E$ given by the sentinel (s, σ) is exact up to an error proportionnal to the component of $x_0 - x$ on the kernel of A (this term vanishing when A happens to be c-injective). Then (3.60) shows that the use of a regularized sentinel introduces an additional error, proportional to the component of $x_0 - x$ on the orthogonal of the kernel of A (of course, this term does not vanish when A is c-injective !).

In all cases, using the best available estimate x_0 of the unknown x (even a very rough one, using for example the order of magnitude of the unknowns x) will yield an improved estimate of $(c, x)_E$, compared to the one obtained with $x_0 = 0$.

4 One Example

We illustrate the results of paragraph (3) on an example taken from Ainseba (1992). It concerns the monitoring of the pollution in a river we denote by :

$$\Omega \subset \mathbf{R}^2, \quad 0 < T \in \mathbf{R} \tag{4.1}$$

the part of the river under investigation, and the duration of the available informations. Then the concentration $y(x, t)$ of the chemical responsible for the pollution is given by :

$$\left\{ \begin{array}{ll} \frac{\partial y}{\partial t} + V.\nabla y - D_1 \frac{\partial^2 y}{\partial x_1^2} - D_2 \frac{\partial^2 y}{\partial x_2^2} + Ky & = f_1(t)\delta(x - x_S^1) + f_2(t)\delta(x - x_S^2) \\ & \ldots \text{in } \Omega \times]0T[\\ y(x, 0) = y_0(x) & \text{in } \Omega (\text{at } t = 0) \\ y = g & \text{on } \partial\Omega_D \times]0T[\\ \frac{\partial y}{\partial \nu_D} = h & \text{on } \partial\Omega_N \times]0T[\end{array} \right. \tag{4.2}$$

where (see figure 4.1) :

$$(4.3)\quad \begin{cases} V \text{ is the smooth velocity field of the river} \\ D_1, D_2 \text{ are smooth dispersion coefficients} \\ K \text{ is a smooth reaction coefficient} \\ x_S^1, x_S^2 \in \Omega \text{ are the location of two pollution sources} \\ f_1, f_2 \in L^2(0T) \text{ are the intensity of these sources} \\ y_0 \in L^2(\Omega) \text{ is the initial concentration of pollutant} \\ g \in L^2(0T\ ; H^{-\frac{1}{2}}(\partial\Omega_D)) \text{ is the concentration of pollutant on} \\ \text{the upstream boundary } \partial\Omega_D \text{ of the river} \\ \frac{\partial y}{\partial \nu_D} = D_1 \frac{\partial y}{x_1}\nu_1 + D_2 \frac{\partial y}{x_2}\nu_2, \text{ where } \vec{\nu} = (\nu_1, \nu_2) \text{ is the outer unit normal to } \Omega. \\ h \in L^2(0T\ ; H^{-\frac{3}{2}}(\partial\Omega_N)) \text{ is the flux of pollutant caused by dispersion on} \\ \text{the lateral and downstream boundary } \partial\Omega_N. \end{cases}$$

We shall make the hypothesis that :

$$(4.4)\quad \begin{cases} \text{all coefficients and right hand sides in (4.2) are} \\ \text{known, except } f_1, f_2, y_0, g, h, \end{cases}$$

and that there exists an observatory :

$$(4.5)\qquad \omega \subset \Omega, \quad \omega \text{ open}$$

on which a measurement :

$$(4.6)\qquad z \in L^2(\omega \times]0T[)$$

of y is available. Before specifying which quantity we want to monitor, we cast the example in the abstract framework of paragraph (3) by setting :

$$(4.7)\qquad E = L^2(0T) \times L^2(0T) \times L^2(\Omega) \times L^2(0T\ ; H^{-\frac{1}{2}}(\partial\Omega_D)) \times L^2(0T\ ; H^{-\frac{3}{2}}(\partial\Omega_N))$$

$$(4.8)\qquad F = L^2(\omega \times]0T[)$$

$$(4.9)\qquad A\ : (f_1, f_2, y_0, g, h) \in E \to y\,|_{\omega\times]0T[} \in F \text{ where } y \text{ is solution of (4.2)}$$

Theorem 4.1 *The operator A is injective. Hence its transposed A^*, defined by :*

$$(4.10)\qquad A^* w = \left(q\,|_{x=x_S^1}, q\,|_{x=x_S^2}, q\,|_{t=0}, -\frac{\partial q}{\partial \nu_D}\,|_{\partial\Omega_D}, q\,|_{\partial\Omega_N} \right),$$

where q is the solution of the adjoint system :

$$(4.11)\quad \begin{cases} -\frac{\partial q}{\partial t} - \nabla(qV) - D_1 \frac{\partial^2 q}{x_1^2} - D_2 \frac{\partial^2 q}{x_2^2} + Kq & = \chi_\omega w \ \ldots \ in\ \Omega \times]0T[\\ q(x,T) = 0 & in\ \Omega (at\ t = T) \\ q = 0 & on\ \partial\Omega_D \times]0T[\\ qV.\nu + \frac{\partial q}{\partial \nu_D} = 0 & on\ \partial\Omega_N \times]0T[\end{cases}$$

has a dense image.

Proof : If $y \equiv 0$ on $\omega\times]0T[$, from the theorem of Mizohata (Saut and Scheurer, 1987), we see that y is identically zero on $(\Omega - \{x_S^1\} - \{x_S^2\})\times]0T[$, hence $y = 0$ a.e. on $\Omega\times]0T[$, which implies that f_1, f_2, y_0, g and h are zero, so that A is injective. The formula (4.10) (4.11) are given by a classical calculations. ■

This implies immediately, through corollary (3.5).

Corollary 4.2 *For any monitoring vector $c \in E$, there exists a unique generalized sentinel $s \in \mathcal{F}$ such that :*

$$(4.12) \qquad \begin{cases} S(f_1, f_2, y_0, g, h) = (s, y\,|_{\omega\times]0T[})_{\mathcal{F}\mathcal{F}'} = (c, (f_1, f_2, y_0, g, h))_E \\ \text{for all } f_1, f_2, y_0, g, h. \end{cases}$$

It is given by the solution of a dual control problem :

$$(4.13) \qquad \text{find } s \in \mathcal{F}, s \text{ of minimum norm, such that } J(s) = Inf_{w\in F} J(w)$$

where :

$$(4.14) \qquad J(w) = \frac{1}{2} \parallel A^* w - c \parallel_E^2,$$

and satisfies :

$$(4.15) \qquad J(s) = 0$$

For example, if one wants to monitor the total amount of pollutant injected by the source 1 during the observation period $[0T]$, one simply chooses :

$$(4.16) \qquad c = (1_{L^2(0T)}, 0, 0, 0, 0) \in E$$

so that :

$$(4.17) \qquad \begin{cases} (s, y\,|_{\omega\times]0T[})_{\mathcal{F}\mathcal{F}'} = \int_0^T f_1(t)\, dt \\ \text{for all } f_1, f_2, y_0, g, h \end{cases}$$

where the sentinel $s \in \mathcal{F}$ is given by (4.13) with :

$$(4.18) \qquad \begin{cases} J(w) = \int_0^T (q(x_S^1, t) - 1)^2\, dt + \int_0^T q(x_S^2, t)^2\, dt + \int_\Omega q(x, 0)^2\, dx \\ \quad + \int_0^T \parallel \frac{\partial q}{\partial \nu_D}(t) \parallel^2_{H^{\frac{1}{2}}(\partial\Omega_D)}\, dt + \int_0^T \parallel q(t) \parallel^2_{H^{\frac{3}{2}}(\partial\Omega_N)}\, dt \end{cases}$$

Remark 4.3 *One can also monitor simultaneously the two pollution sources by choosing :*

$$(4.19) \qquad c = (1_{L^2(0T)}, 1_{L^2(0T)}, 0, 0, 0)$$

in which case :

$$(4.20) \qquad \begin{cases} (s, y\,|_{\omega\times]0T[})_{\mathcal{F}\mathcal{F}'} = \int_0^T f_1(t)\, dt + \int_0^T f_2(t)\, dt \\ \text{for all } f_1, f_2, y_0, g, h. \end{cases}$$

The sentinel s is then sensitive to changes in the intensivity of any of the two sources, and insensitive to changes in any other unknown quantity. ■

Of course, the ideal situation described by corollary (4.2) is a kind of lure, caused by the infinite speed of propagation associated to the Laplacian operator ! This situation is completely lost after discretization : if we denote by :

$$E_h \subset E,\ F_h \subset F \quad A_h \in \mathcal{L}(E_h, F_h) \tag{4.21}$$

the finite dimensional spaces and operator which replace E, F, A defined in (4.7) (4.8) (4.9) once a reasonnable discretization of the state equation (4.2) has been chosen, then A_h has many zero, or very small, singular values, so that it is, by far, no more injective. Hence the corresponding <u>discretized</u> <u>sentinel</u> $(s_h, \sigma_h) \in F_h \times E_h$ will contain a <u>non</u> <u>zero</u> <u>term</u> <u>σ_h</u> giving the influence of the a priori guess chosen for f_1, f_2, y_0, g and h.

The adjoint A_h^* of A_h can be determined by a classical discrete adjoint state calculation, whose result we can write symbolically as :

$$A_h^* w_h = \left(q_h|_{x=x_S^1}, q_h|_{x=x_S^2}, q_h|_{t=0}, -\frac{\partial q_h}{\partial \nu_D}\ |_{\partial\Omega_D}, q_h|_{\partial\Omega_N} \right) \tag{4.22}$$

where q_h is the solution of the discrete adjoint state equation with right-hand side w_h.

Because of the poor conditionning of the dual control problem for the determination of (s_h, σ_h), one has to use regularization :

<u>**Corollary**</u> **4.4** *For any $\varepsilon > 0$, there exists a <u>discretized-regularized</u> <u>sentinel</u> <u>$(s_{h,\varepsilon}, \sigma_{h,\varepsilon}) \in F_h \times E_h$</u> such that :*

$$\begin{aligned} &\left(s_{h,\varepsilon}, y_h\ |_{\omega\times]0T[}\right)_{F_h} + (\sigma_{h,\varepsilon}, x_{0h})_{Eh} = \\ &\int_0^T f_{1h}(t)\ dt + (\sigma_{h,\varepsilon}, x_h - x_{0h})_{Eh} \end{aligned} \tag{4.23}$$

for all $x_h = (f_{1h}, f_{2h}, y_{0h}, g_h, h_h) \in E_h$
and for all a priori guesses $x_{0h} = (f_{1,0,h}, f_{2,0,h}, y_{0,0,h}, g_{0,h}, h_{0,h}) \in E_h$.
The <u>observation</u> <u>related</u> <u>part</u> <u>$s_{h\varepsilon} \in F_h$</u> of the sentinel is given by :

$$s_{h\varepsilon} \in F_h, \quad J_{h,\varepsilon}(s_{h,\varepsilon}) = Inf_{w_h \in F_h} J_{h,\varepsilon}(w_h) \tag{4.24}$$

where :

$$\begin{aligned} J_{h,\varepsilon}(w_h) = &\int_0^T \left(q_h\ |_{x=x_S^1} - 1\right)^2\ dt + \int_0^T \left(q_h\ |_{x=x_S^2}\right)^2\ dt + \int_\Omega (q_h\ |_{t=0})^2\ dx \\ &+ \int_0^T \|\ \tfrac{\partial q_h}{\partial \nu_D}\ |_{\partial\Omega_D}\|^2\ dt + \int_0^T \|\ q_h\ |_{\partial\Omega_N}\|^2\ dt + \varepsilon^2 \int_0^T \int_\omega |\ w_h\ |^2\ dx\ dt. \end{aligned} \tag{4.25}$$

Then <u>the</u> <u>a priori</u> <u>guess</u> <u>related</u> <u>part</u> $\sigma_{h,\varepsilon} \in E_h$ of the sentinel is given by :

$$\left\{ \begin{array}{l} \sigma_{\varepsilon,h} = -\left(q_h\ |_{x=x_S^1} - 1, q_h\ |_{x=x_S^2}, q_h\ |_{t=0}, \frac{\partial q_h}{\partial \nu_D}\ |_{\partial\Omega_D}, q_h\ |_{\partial\Omega_N}\right) \\ \textit{where } q_h \textit{ is the solution of the discrete adjoint} \\ \textit{system with right-hand side } s_{h,\varepsilon}. \end{array} \right. \tag{4.26}$$

When $\varepsilon \to 0$, one has :

$$s_{h,\varepsilon} \in Im A_h \to s_h \in Im A_h \ in\ F_h \tag{4.27}$$

$$\sigma_{h,\varepsilon} \to \sigma_h \in Ker A_h \ in\ E_h \tag{4.28}$$

The formula (4.24) (4.25) for the determination of $s_{h,\varepsilon}$ coincide with the ones which have been used so far for the numerical determination of sentinels (Ainseba, 1992, Kernevez 1993), but the interpretation of the residual (4.26) as yielding the influence of the a priori guess seems to be new.

One could also probably proove that when $h \to 0$, $s_h \to s$ in $\mathcal{F}$ and $\sigma_h \to 0$ in E under reasonnable assumption, but we have not pursued in this direction.

Acknowledgement

I am indebted to Jacques Henry (Institut National de Recherche en Informatique et Automatique, Rocquencourt, France) for many fruitful discussions, Jean-Pierre Kernevez (Université Technologie de Compiègne, France) and Jacques-Louis Lions (Collège de France).

References

[1] B.E. Ainseba. *Controlabilité Exacte, Identifiabilité, Sentinelles.* PhD thesis, UTC Compiègne, Décembre 1992.

[2] G. Chavent and K. Kunisch. Convergence of Tikhonov Regularization for Constrained Illposed Inverse Problems. *Submitted to Journal of Approximation Theory*, July 1992. Rapport INRIA 1729.

[3] H.W. Engl, K. Kunisch, and A. Neubauer. Convergence Rates for Tikhonov Regularization of Non Linear Illposed Problems. *Inverse Problems*, 5:523–540, 1989.

[4] C.W. Groetsch. *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind.* Pitman, 1984.

[5] J.P. Kernevez. Des Sentinelles pour Surveiller l'Environnement. Livre à paraître.

[6] J.L. Lions. *Contrabilité exacte, Perturbations et Stabilisation de Système Distribués.* Masson, 1988. Tomes 1 et 2.

[7] J.L. Lions. Sur les Sentinelles des Systèmes Distribués. CRAS, Tome 307, 1988. Paris.

[8] J.L. Lions. Sentinels with Special Sensitivity. In El Jai et M. Amouroux ed., editor, *IFAC Symposia Series N^0 3*, pages 1–4. Pergamon Press, 1990,.

[9] J.L. Lions. *Sentinelles pour les Systèmes à Données Incomplètes.* Masson, 1992.

[10] J.C. Sant and B. Scheurer. Unique Continuation for Some Evolution Equations. *Journal of Differential Equations*, 66:118–139, 1987.

[11] A.N. Tikhonov and V.Y. Arsenin. *Solution of Illposed Problems.* Wiley, New York, 1977.

Unité de Recherche INRIA Rocquencourt
Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 LE CHESNAY Cedex (France)

Unité de Recherche INRIA Lorraine Technopôle de Nancy-Brabois - Campus Scientifique
615, rue du Jardin Botanique - B.P. 101 - 54602 VILLERS LES NANCY Cedex (France)
Unité de Recherche INRIA Rennes IRISA, Campus Universitaire de Beaulieu 35042 RENNES Cedex (France)
Unité de Recherche INRIA Rhône-Alpes 46, avenue Félix Viallet - 38031 GRENOBLE Cedex (France)
Unité de Recherche INRIA Sophia Antipolis 2004, route des Lucioles - B.P. 93 - 06902 SOPHIA ANTIPOLIS Cedex (France)

EDITEUR
INRIA - Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 LE CHESNAY Cedex (France)

ISSN 0249 - 6399

RR.1932