GALERKIN PROCEDURES FOR STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS

by

Jocelyn Freitas Bennaton

January 1980

A thesis submitted for the degree of Doctor of Philosophy of the University of London.

Computing and Control Department

Imperial College

ACKNOWLEDGMENTS

This research was made possible by the advice and encouragement of my supervisor, Prof. J.M.C. Clark and by financial support from the "Conselho Nacional de Desenvolvimento Cientifico e Tecnologico - CNPq".

ABSTRACT

This thesis is concerned with stochastic, and non-stochastic, first order linear evolution equations.

The reason for the simultaneous treatment of these topics lies in the fact that a recursive solution for the filtering problem for Markov diffusions can be given either by the stochastic partial differential equation governing the unnormalized conditional density, or by its non-stochastic counterpart, which is a parabolic equation parametrized by the paths of the observation process.

This work embraces both these approaches to the non-linear filtering problem. Convergence results for the Galerkin approximation of the solution, of either the stochastic or the non-stochastic evolution equations, are presented and, for both cases, error estimates of discrete time Galerkin procedures derived. In particular, families of discrete time Galerkin schemes for approximating the solution of the non-linear filtering problem are compared and rates of convergence obtained.

CONTENTS

1 - INTRODUCTION

Although the title of this work makes reference only to stochastic equations, we shall be studying both stochastic and non stochastic linear evolution equations.

It is true that the analysis of stochastic equations contains elements which work in the non stochastic case and, in fact, this happens in the situation we are concerned with. However, in this work, the inclusion of non stochastic evolution equations represents more than a prelude to the stochastic case. The reason for our simultaneous treatment of these topics lies in the relevance they both have in non linear filtering theory.

It is well understood that one way of presenting a recursive solution for the non linear filtering problem for diffusions is by means of the unnormalized density formula, (the Zakai formula, see |54|), which is a stochastic linear evolution equation.

On the other hand, as has been pointed out, (among others, by Clark (| 5 |)), this formula has a non stochastic counterpart parametrized in a convenient way by the sample paths of the observation process. This non stochastic formula is similar to the Fokker-Planck equation for the diffusion under consideration, with the same diffusion coefficients, but with drift and potential coefficients depending on the observation sample paths. In addition, it possesses the special feature of being 'robust' in the sense that its solution is a continuous mapping defined in the sample space of the observation process. Therefore, in practical situations, instead of a given observation sample path, we are allowed to work with suitable approximations belonging to a class dense in the sample space (e.g., functions of bounded variation) without taking the risk of being driven away from the true solution of the filtering problem.

In view of these characteristics we take the point that it is well worth considering the pathwise formula as an alternative and equally important way of representing the solution of the filtering problem, and not merely as a version of the Zakai formula.

A considerable portion of this thesis is devoted to existence and uniqueness results for both non stochastic and stochastic evolution equations, and in this area we follow the work of Lions (|30|, |31|, |32|), and Pardoux (|40|, |41|). However, the inclusion of these results is mainly didactic. The principal purpose of our work is the analysis of Galerkin approximations of the non linear filtering problem.

The duality between the stochastic and the non stochastic representations of the filtering solution is reproduced in the numerical schemes used for its approximation. We can select schemes appropriate to the pathwise formula or, instead, schemes which are suitable for the Zakai formula. As before, both aspects of this duality are equally important, and our intention is to analyse Galerkin schemes both for the non stochastic and for the stochastic representations.

Using a family of implicit Runge-Kutta schemes we show that the corresponding discrete time Galerkin procedure converges, (in the sup norm), to the pathwise solution, for all paths of bounded variation. These schemes, therefore, produce a robust approximation to the filtering solution, in the sense that they are continuous with respect to the observation sample paths, and the approximation converges uniformly in a dense subset of the sample space.

Extensions of the implicit Runge-Kutta schemes, containing terms which are either linear or quadratic in the noise increment, can be used as well. They produce Galerkin approximations that converge uniformly, (in an average sense), to the solution of the Zakai formula. In particular, if sufficient regularity conditions are attained, the standard deviation of the error for the quadratic scheme, converges at a linear rate with respect to the time increment. Judging from what happens for approximations of finite dimensional stochastic differential equations this is the best possible rate of convergence.

As the non linear filtering problem is the 'raison d'etre' of this work we start by presenting in paragraph 1.1, a survey in this subject.

1.1 - The Non Linear Filtering Problem

We start by a general description of the filtering problem.

Suppose the situation where the data concerning an unobservable stochastic process (the signal process) is provided by observation of another stochastic process (the observation process) which is related to the signal in some functional fashion. The question of determining the conditional probability density for the signal process given the observation process constitutes the filtering problem.

Although the filtering problem can be formulated for a wide variety of processes, here we shall be concerned with the case where the signal is a Markov diffusion process in a euclidian space and the observation is a scalar process of the "signal plus white noise" type. Let us be more specific. In relation to some probability space $(\Omega, \mathcal{A}, P)$ let (x,y) denotes the pair signal/observation processes and assume the relation between them being given by the following (Ito's) stochastic differential form:

1. $$dy(t) = h(t,x(t))dt + dw_t \qquad t \in [0,T]$$

where $h \in C([0,T] \times R^n)$ and w_t is a R-valued standard Wiener process.

This formulation of the filtering problem for diffusion process is classical and it is along the lines of that presented by Stratonovich in 1960. In his basic paper (|46|), Stratonovich proposes a stochastic partial differential equation which, under some conditions, represents the dynamics of the conditional density of the signal process. An equivalent result was obtained by Kushner (in |23|) who rederived with some corrections the Stratonovich equation and presented it in terms of Ito integrals.

So, the Kushner-Stratonovich representation for the solution of the non linear filtering problem stands as the first result in a long line of research still being done in this field. Among the subsequent works, a distinctive direction is represented by the search for an extension of the Baye's formula in order to express the density as a functional of the observations. The idea, first proposed by Bucy (in | 3 |) has its complete development in |19| where the authors, Kallianpur and Striebel, presented a precise statement of the formula which generalize a previous one obtained by Wonham (in |52|) for finite state Markov chains.

Although Kallianpur and Striebel's formula is valid for a wide range of situations, especially those regarding estimation problems, it is not useful if a recursive solution is sought for the non linear filtering problem. Solutions having the character of being recursive were, during the sixties, the object of various important papers among which one can select those due to Liptser and Shiryaev (|33|) and Zakai (|54|). In the first, a stochastic differential representation for the solution of the filtering problem is presented for the case where the pair (x,y) is a diffusion process. In the second, under the hypothesis of independence between the signal and the Wiener process in equation 1., the so called unnormalized density formula was derived for the first time, bearing the advantage of being a considerably simpler representation for the solution of the filtering problem for diffusion process. Finally, in |13|, Fujisaki, Kallianpur and Kunita using the innovation process approach introduced by Kailath (|18|), presented a stochastic differential representation for the conditional expectation of the signal process valid for a large range of situations regarding either the signal process or the interdependence between the signal and the observation.

After this brief account of the papers, which are considered classical in non linear filtering theory, let us return to the particular problem we started describing at the beginning of this Introduction.

Regarding the diffusion process x_t, assume that the following stochastic differential form describes its dynamics:

2. $$dx(t) = g(t,x(t))dt + \alpha(t,x(t))dw_t^1$$

where, $g \in C([0,T] \times R^n; R^n)$

$\alpha \in C([0,T] \times R^n; R^{n \times n})$

and w_t^1 is a R^n-valued standard Wiener process.

In equations 1. and 2. suppose

3. $$y(0) = 0 \quad \text{and} \quad x(0) = x_0 ,$$

where x_0 is a random variable.

Suppose that the Wiener processes w_t and w_t^1 are independent and also assume x_0 independent of (w_t, w_t^1).

Consider the stochastic process z_t defined by

4. $$z(t) = -\frac{1}{2}\int_0^t h^2(s,x(s))ds + \int_0^t h(s,x(s))dy_s \qquad t \in [0,T]$$

For the particular class of functions under consideration we can define a new probability measure on the space $(\Omega,\mathcal{A})$ by the following relation:

5. $$d\tilde{P} = \exp(-z(T))dP$$

Write $E,(\tilde{E})$, for the conditional expectation with respect to the measure $P,(\tilde{P})$. If Y_t, $t \in [0,T]$ denotes the σ-algebra generated by $\{y_s : 0 \leq s \leq t\}$ define

6. i) $$\Pi_t(f) = E(f(x_t)/Y_t)$$

ii) $$Q_t(f) = \tilde{E}(f(x_t).\exp(z_t)/Y_t)$$

for all $f \in C(R^n)$, $t \in [0,T]$

By a standard formula relating conditional expectations with respect to equivalent probability measures (see e.g. Kallianpur-Striebel, |19| or Meyer |37|), we have

7. $$\Pi_t(f) = Q_t(f) \,.\, Q_t^{-1}(1) \qquad \text{w.p.1}$$

where the argument 1 denotes the unitary function of $C(R^n)$.

The transformation of probability measure introduced in 5. has some important features. Under the new probability $\tilde{P}$, the observation, y_t, becomes a standard Wiener process independent of the signal process (Girsanov, |15|). This fact can lead us to the Kallianpur-Striebel formula,

8. $$Q_t(f) = \int_W f(\zeta_t) \,.\, \exp(Z_t(\zeta))\mu(d\zeta)$$

where,

$$Z_t(\zeta) = -\frac{1}{2}\int_0^t h^2(s,\zeta_s)ds + \int_0^t h(s,\zeta_s)dy_s$$

$\zeta \in C([0,T];R^n) = W$, W being the sample space for the signal process.

μ is the measure on W induced by the diffusion x.

As we pointed out before, the Kallianpur-Striebel formula gives us a non-recursive representation for the conditional expectation. An alternative and more convenient solution is to express the conditional expectation by means of the Fujisaki-Kallianpur-Kunita formula.

Let L_t denotes the Fokker-Planck operator associated with the diffusion x_t, i.e.,

9. $$L_t u = \frac{1}{2}\sum_{i,j=1}^{n} \frac{\delta^2}{\delta x_i \delta x_j}(a_{i,j}(t,x)u(x)) + - \sum_{i=1}^{n} \frac{\delta}{\delta x_i}(g_i(t,x)u(x))$$

where $[a_{i,j}(t,x)] = \alpha(t,x) \,.\, \alpha^T(t,x)$

The Fujisaki-Kallianpur-Kunita formula under the hypotheses made above, takes the following form:

10. $$d(\Pi_t(f)) = \Pi_t(L_t^* f)dt + (\Pi_t(h_t f) - \Pi_t(h_t)\Pi_t(f))dv_t$$

where $h_t \equiv h(t,.)$, L_t^* is the infinitesimal generator of the diffusion x_t and v_t is the innovation process,

11. $$v(t) = y(t) - \int_0^t \Pi_s(h_s)ds$$

From equation 10. we can derive a recursive representation for the conditional density. So, if $p_t = p(t,x,\omega)$, $(t,x,\omega) \in [0,T] \times R^n \times \Omega$ denotes the conditional probability density of the signal given the observation y_t we can write the Kushner-Stratonovich formula,

12. $$dp_t = L_t p_t dt + (h_t - (h_t,p_t))p_t dv_t$$

where $(.,.)$ denotes the inner product in $L^2(S)$.

Given a suitable initial condition, i.e. the probability density of x_0, equation 12. can give ,under certain conditions, the evolution of the conditional density of the signal and, therefore, it solves the filtering problem. (see e.g. Kushner, |24|) However a better formula can be found, which has the advantage of being linear in the unknown variable. If for the variable $Q_t(f)$ defined in 6. we write

13. $$Q_t(f) = (q_t,f)$$

Then we can deduce the Zakai formula for the unnormalized density,

14. $dq_t = L_t q_t dt + h_t q_t dv_t$

This representation for the solution of the filtering problem has considerable advantages in relation to the previous formulas. It is a simpler formula and, besides, being linear it enlarges the scope vis-a-vis numerical applications.

The concept of unnormalized density and its representation by equation 14. leads us to an alternative form of presenting the solution of the filtering problem under consideration. The idea is to look for non stochastic differential equations parametrized by the paths of the observation process in order to represent the solution of the filtering problem as a continuous function of the sample paths of the observation process. This has been done, for instance, by Clark (in |5|)† and the result is a family of linear partial differential equations, which has the same status as equation 14..

The relation between stochastic differential equations and their non stochastic equivalent representations has been the object of a number of papers and, in particular, some approach the problem by studying stochastic differential forms as the limit of sequences of ordinary differential equations (see e.g. Wong-Zakai, |51|)

A different approach has been adopted by Doss, who, in |11| shows that the solution of a stochastic differential equation is equivalent to the integration of an ordinary differential equation parametrized by the paths of a stochastic process. Here, we shall use his procedure in order to derive the pathwise formula for the solution of the filtering problem for diffusion processes.

† The concept of pathwise solutions has been familiar to the Russian school of probabilists for some time. In particular, we understand that it was used, 'en passant', by Rosovskii in his thesis for the Moscow University in 1972. It also appears in Liptser-Shiryaev, |34|

Let $v(t) = V(t,u)$; $(t,u) \in [O,T] \times L^2(R^n)$, be the solution of the following differential equation in $L^2(R^n)$:

15. $$\frac{d}{dt} v(t) = h_t \, v(t)$$

$$v(O) = u$$

Therefore we can write,

$$V(t,u) = \Phi(t)u$$

where $\Phi(t) = \exp(\int_0^t h_s ds)$

Consider the following ordinary differential equation parametrized by the paths of the process y_t:

16. $$\frac{d}{dt} r(t) = \Phi^{-1}(y(t))\hat{L}(t)\Phi(y(t))r(t)$$

where $\hat{L}(t) = L_t - \frac{1}{2} h_t^2 \quad , \quad t \in [O,T]$.

Using basically Ito's rule of transformations, we can show that the solution of equation 14. can be expressed by means of the following relation:

17. $$q_t = V(y_t, r(t))$$

Therefore, the pathwise formula 15. can represent the solution of the filtering problem for each observed path $y(t)$.

Also, as Clark pointed out, the solution depends continuously on these paths which is important [illegible] for [illegible] numerical applications.

We have presented some of the ways of expressing the solution for the particular non linear filtering problem described here. It can be argued that, in practical cases, the hypothesis we have made concerning the independence between the signal and the observation noise is too restrictive. However, this difficulty can be partially overcome by allowing some dependence between the Wiener processes w_t and w_t^1. In this respect we shall present here the results obtained by Pardoux (in |41|) though similar formulas can be found in Levieux, |28| and Krilov-Rosovskii |22|. We recall that the problem regarding correlation between the signal and the observation noise was also considered in Fujisaki-Kallianpur-Kunita, |13|.

So, instead of assuming independence between w_t and w_t^1, let us suppose that the Wiener process w_t can be expressed by means of the following relation:

18. $$dw_t = \langle \beta^1(t), dw_t^1 \rangle + \beta^2(t) dw_t^2$$

where $\langle .,. \rangle$ denotes the scalar product in R^n, β^1, (β^2), is a continuous R^n, (R)-valued function defined in R^+ and w_t^2 is a R-valued standard Wiener process independent of w_t^1.

In order to guarantee that the above expression is a relation between standard Wiener processes we must assume for all $t \in R^+$,

19. $$\langle \beta^1(t), \beta^1(t) \rangle + (\beta^2(t))^2 = 1$$

Now, consider the following first order differential operator:

20. $$H_t u = - \sum_{i=1} \frac{\delta}{\delta x_i} (b_i(t,x)u(x)) + h(t,x)u(x)$$

where $[b_i(t,x)] = \alpha(t,x).\beta^1(t)$

The formulas we have presented for the recursive solution of the filtering problem can be modified according to assumption 18. In particular, the unnormalized density formula takes now the form,

21. $$dq_t = L_t q_t dt + H_t q_t dy_t$$

(for a precise account of this formula see Pardoux, |41|)

The purpose of this introductory paragraph is to describe in general terms, without proofs, the formulas for the solution of the filtering problem for diffusions. The reason for doing so is to establish the relevance of an analysis of evolution equations presented in the stochastic form 21. (or 14.) and in the pathwise form 16.

These equations are the object of our study in the following sections. With respect to the non linear filtering problem a complete survey of the field can be found in Jazwinski, |17|, Wong, |50| and Liptser-Shiryaev, |34|. In particular, the derivation of the Kushner-Stratonovich formula for partially observed signals, can also be found in Pardoux, |41|. A precise account of the Fujisaki-Kallianpur-Kunita formula is also given by Meyer |37|. Pathwise solutions are considered in greater generality by Davis, |9|.

Here, we have been restricted to the general filtering problem for diffusions in R^n. For (absorbed or reflected) diffusions in subsets $S \subset R^n$ similar formulas can be derived and, in this case, the conditions in the boundary of the domain S define the nature of the diffusion. (see Pardoux |40|

for a precise account on the unnormalized density formulas that correspond to this situation).

2 - BASIC CONCEPTS

The purpose of this section is to present some of the concepts which are in general associated with evolution equations in Hilbert spaces and, in particular, with partial differential equations.

We start in paragraph 2.1 with the introduction of the Sobolev spaces by means of the classical approach using distributions. In paragraph 2.2 we describe the kind of problem we shall be treating through this work and our method of approach to its solution, which is based in the duality between problems and weak forms. It turns out that to this duality there corresponds a duality between linear operators and bilinear forms; this constitutes the subject of the last paragraph of this section.

2.1 - Functional Spaces

Sobolev spaces play a decisive role in partial differential equations and here we shall present a brief account of some of the concepts leading to their definition. We also introduce other functional spaces which will be relevant in the following section. The treatment given here are along the lines of that in Adams |1|, Barros-Neto |2|, and Yosida |53|.

In what follows we reserve the symbol, S, for an open set of a n-dimensional real euclidian space.

If $u \in C(S)$, the space of R-valued continucus functions defined on S, has partial derivatives of order $|\alpha| \geq 0$ we denote by $D^{\alpha}u$ the partial derivative,

$$D^{\alpha}u = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}}$$

where $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ is an n-tuple of non-negative integers and $|\alpha| = \alpha_1 + \alpha_2 + \ldots + \alpha_n$

For $m \geq 0$ we denote by $C^m(S)$ $(C^o(S) = C(S))$, the sets,

$$C^m(S) = \{u \in C(S) : D^\alpha u \in C(S), |\alpha| \leq m\}$$

and by $C^\infty(S)$ we denote the set of "infinitely" continuously differentiable R-valued functions defined on S. In other words, $C^\infty(S) = \bigcap_{0 \leq m} C^m(S)$.

The sets $C^m(S)$ $0 \leq m \leq \infty$ are linear spaces with the usual operations on real-valued functions. In fact, we are able to impose a locally convex topology on them in such way that a sequence $\{u_k\}$ converges to zero if and only if $\{D^\alpha u_k : |\alpha| \leq m\}$ converges uniformly to zero on every compact subset of S. This so called natural topology in $C^m(S)$ is the coarsest one for which the linear maps $D^\alpha : C^m(S) \to C(S)$ for $|\alpha| \leq m$, are continuous.

If $u \in C(S)$, by "support of u" we mean the closure in S of the set $\{x \in S : u(x) \neq 0\}$. For $m \geq 0$ we denote by $C_o^m(S) \subset C^m(S)$ the subset of functions with compact support in S. In particular, it can be shown that $C_o^\infty(S)$ is dense in $L^p(S)$, the space of p-integrable functions on S. As before, the sets $C_o^m(S)$ can be endowed with a locally convex topology in such way that a sequence $\{u_k\}$ converges to zero if and only if there exists a compact set $K \subset S$ such that:

i) support of $u_k \subset K$ for every k

ii) for $|\alpha| \leq m$, $D^\alpha u_k \to 0$ uniformly in K

As it is conventional to write $\mathcal{D}(S)$ for the set $C_0^\infty(\Omega)$ endowed with this topology.

It turns out that a linear functional T defined in $\mathcal{D}(S)$ is continuous if and only if $<T,u_k> \to <T,u>$ whenever $u_k \to u$ in $\mathcal{D}(S)$. This fact enables us to consider the dual of $\mathcal{D}(S)$, $\mathcal{D}'(S)$, which is also a locally convex topological space in such way that a sequence $\{T_k\}$ converges (strongly) to zero if and only if $<T_k,u>$ converges to zero uniformly on every bounded subset of $C_0^\infty(S)$.

The space $L^1_{loc}(S)$, of locally integrable functions on S, can be identified with a subspace of $\mathcal{D}'(S)$. In fact, if $u \in L^1_{loc}(S)$ it can be assigned a distribution $T(u)$ defined by:

1. $$<T(u),v> = \int_S u(x) \cdot v(x)\, dx$$

for all $v \in C_0^\infty(S)$

We can define derivatives of distributions in such a way that it agrees with the conventional derivative, regarding the identification mentioned above. So, if $T \in \mathcal{D}'(S)$ we define the partial derivative $D^\alpha T \in \mathcal{D}'(S)$, by

2. $$<D^\alpha T,u> = (-1)^{|\alpha|} <T,D^\alpha u>$$

for all $u \in C_0^\infty(S)$

Using equation 1. and integration by parts, it can be verified that

$$<T(D^\alpha u),v> = (-1)^{|\alpha|} <T(u),D^\alpha v>$$

for all $v \in C_0^\infty(S)$

As one can see, every distribution has derivatives of all orders and, furthermore, they are independent of the order in which they are taken:

$$\frac{\partial^2 T}{\partial x_j \partial x_k} = \frac{\partial^2 T}{\partial x_k \partial x_j} \qquad j,k = 1,\ldots,n$$

The identification 1. of $L^1_{loc}(S)$ with a subspace of $\mathcal{D}'(S)$ leads us to the concept of "weak derivative". Given $u \in L^1_{loc}(S)$, if there exists a unique (up to a set of measure zero) function $v \in L^1_{loc}(S)$ such that for some multi-index α,

3. $$T(v) = D^\alpha T(u) \quad \text{in} \quad \mathcal{D}'(S)$$

then v is called a weak, or distributional, partial derivative of u. By equation 1. the above weak derivative of u is defined up to a set of measure zero by the following relation:

4. $$\int_S v(x)w(x)\,dx = (-1)^{|\alpha|} \int_S u(x)D^\alpha w(x)\,dx$$

$$\text{for all } w \in C_0^\infty(S)$$

of course, if there exists $D^\alpha u \in L^1_{loc}(S)$ then, (up to a set of measure zero) $D^\alpha u = v$.

Now, consider the set of functions $u \in C^m(S)$ such that, for $1 \leq p < \infty$

5. $$\| u \|_{m,p} = \left(\sum_{|\alpha| \leq m} \int_S (D^\alpha u(x))^p dx \right)^{1/p} < \infty$$

The completion of this set with respect to the norm $\| \cdot \|_{m,p}$ is called the Sobolev space of order (m,p) and it is denoted by $H^{m,p}(S)$. It can be shown (see Adams |1|) that this definition coincides with the following:

$$H^{m,p}(S) = \{u \in L^p(S) : D^\alpha u \in L^p(S) , |\alpha| \leq m\}$$

where $D^\alpha u$ is interpreted as a weak derivative.

We also define $H_o^{m,p}(S)$ as the closure of $C_o^\infty(S)$ in $H^{m,p}(S)$.

In what follows we will be restricted to the case p = 2 where, as it is conventional, the index p is deleted from the notations.

So, the space H^m is a separable Hilbert space with the inner product:

6. $$(.,.)_{H^m(S)} = \sum_{|\alpha| \leq m} (D^\alpha .,.)_{L^2(S)}$$

It turns out that $H^m(R^n) = H_o^m(R^n)$ or, that $C_o^\infty(R^n)$ is dense in $H^m(R^n)$. In general, this result is not true for generic subsets of R^n.

We denote by $H^{-m}(S)$ the dual of $H_o^m(S)$. As $C_o^\infty(S)$ is dense in $H_o^m(S)$ the elements of $H^{-m}(S)$ determine a distribution on S. So, we are able to identify $H^{-m}(S)$ with a subspace of

$\mathcal{D}'(S)$. It can be proved that this subspace is the linear span of the set

$$\{T(D^\alpha u) \; : \; |\alpha| \leq m, \; u \in L^2(S)\}$$

where $D^\alpha u$ is interpreted as a weak derivative.

Some of the concepts introduced here can be extended to H-valued functions where H is a Hilbert space. So, in the following sections we will be often refering to $L^p(S; H)$, $1 \leq p \leq \infty$, the Banach space of (equivalence class of) H-valued functions defined in S such that

7. $$\| u \|_{L^p(S;\, H)} = \left(\int_S \| u(x) \|_H^p \, dx \right)^{1/p} < \infty \; ; \quad 1 \leq p < \infty$$

with the usual modification for $p = \infty$.

We can define the space of distributions on (O,T) with values in H by,

$$\mathcal{D}'(0,T;\, H) = L(\mathcal{D}(0,T);\, H)$$

(see Lions |31| and |32|)

A sequence $\{W_n\}$ converges to W in $\mathcal{D}'(0,T;\, H)$ if and only if $\langle W_n, \psi \rangle \to \langle W, \psi \rangle$ in H for all $\psi \in \mathcal{D}(0,T)$.

If $u \in L^1_{loc}(0,T;\, H)$ we can define the distribution $W(u) \in \mathcal{D}'(0,T;\, H)$ by

8. $$\langle W(u), \psi \rangle = \int_0^T u(t)\, \psi(t) \, dt$$

for all $\psi \in C_0^\infty(0,T)$

Therefore, as we have done before, we can define the derivative of a distribution $W \in \mathcal{D}'(0,T;H)$ by

9. $$\langle \frac{d}{dt} W, \psi \rangle = - \langle W, \frac{d\psi}{dt} \rangle$$

for all $\psi \in C_0^\infty(0,T)$

We can also define as in 3. a weak, or, distributional, derivative of $u \in L^1_{loc}(0,T;H)$ by the relation

10. $$W(\frac{du}{dt}) = \frac{dW}{dt}(u)$$

and therefore, as in 4. and according to 8 and 10, the weak derivative $\frac{du}{dt} \in L^1_{loc}(0,T;H)$ satisfy

11. $$\int_0^T \frac{du}{dt}(t)\ \psi(t)dt = - \int^T u(t)\ \frac{d\psi}{dt}(t)dt$$

for all $\psi \in C_0^\infty(0,T)$

2.2 - Problems and Weak Forms

Suppose it is given a (real) Hilbert space H and taking values in H, a linear operator A(t), depending upon a parameter $t \in (0,T) \subset R$ and with domain $D(A(t)) \subset H$.

Consider the problem of satisfying the following conditions:

1. i) $u(t) \in D(A(t))$, $u'(t) \in H$ for all $t \in (O,T)$ †

ii) $u'(t) + A(t)\, u(t) = f(t) \in H$ for all $t \in (O,T)$

iii) $u(O) = u_o \in H$

This is, perhaps, the simplest evolution problem one can consider relative to a differential equation of first order in the variable t, defined in a Hilbert space H. As might be expected, in order to solve this problem further assumption are necessary. However, at this stage the simple formulation above is sufficient for the objective we have in mind, i.e., to introduce the concept of "weak form".

Consider a subspace V of H. If there exists a function u satisfying 1. we can conclude that this function also verifies:

2. $(u'(t), v) + (A(t)\, u(t), v) = (f(t), v)$

for all $v \in V$, $t \in (O,T)$

where $(.,.)$ denotes the inner product defined in H.

This fact suggests that one can associate with the original problem 1. an alternative formulation represented by statements 1.i), 1.iii) and 2.

Every solution of the original problem is a solution of the alternative formulation although the converse is not, in general, true. So, the alternative formulation is less

† Here, we consider u', the derivative of u, just in a formal way. Of course, in a more rigorous situation, its meaning must be made precise.

restrictive than the original one and, therefore, it is called, appropriately, a weak form for the original problem. Let us extend this concept a little more.

Let $D(A(t)) \cap V \neq \phi$ for all $t \in (0,T)$. Suppose we are given a functional $a(t) = a(t; u, v)$ defined in $(0,T) \times V \times V$ and bilinear in V for each $t \in (0,T)$, such that:

3. $(A(t)\, u, v) = a(t; u, v)$

for all $u \in D(A(t)) \cap V$, $v \in V$, $t \in (0,T)$

If a solution of problem 1. belongs to $D(A(t)) \cap V$ for all $t \in (0,T)$ it also satisfies the equation:

4. $(u'(t), v) + a(t; u(t), v) = (f(t), v)$

for all $v \in V$, $t \in (0,T)$

This fact leads us to consider the problem represented by statements 1.iii), 4. and the following

5. $u(t) \in V$, $u'(t) \in H$ for all $t \in (0,T)$

If $D(A(t)) \subset V$ for all $t \in (0,T)$, the problem 1.iii), 4., 5. is a weak form for the original problem 1. in the sense defined above.

As we shall see in the following sections, a equation like the one in 4. is very well suited for a mathematical treatment. Moreover, if some conditions are imposed on the functional a(t), and on the subspace V, the original problem 1. and the problem 1.iii), 4., 5. are equivalent.

<u>Remark 2.2.1</u> - For a general account of weak forms see Lions |30| and also Necas |38|.

<u>Remark 2.2.2</u> - Following the terminology of Hadamard we say that a problem of the type presented in this paragraph is "well posed" if it admits a unique solution, the solution being continuous with respect to the entries of the problem.

2.3 - Bilinear Forms

Bilinear forms constitute the 'piece de resistance' in the approach we select to study evolution equations. So, in this paragraph we shall present some properties of bilinear forms defined in Hilbert spaces. A general account of what follows can be found in Lions |30| and also in Necas |38|.

As before, let H be a Hilbert space with inner product denoted by (.,.) and norm $|.| = (.,.)^{1/2}$. Let $V \subset H$ be also a Hilbert space and write ((.,.)) and $\|.\|$ for its inner product and norm. Furthermore, suppose,

1. V is dense in H

with the continuous injection.

2. $|v| \leq \|v\|$ for all $v \in V$

Consider now a bilinear form a = a(u,v) defined in

V. Here we make two assumptions. First, we suppose continuity in V, i.e., there exists a constant γ such that

3. $$|a(u,v)| \leq \gamma \|u\| \, \|v\| \qquad \text{for all } u,v \in V$$

Second, we assume the bilinear form to be coercive, i.e., there exists a constant $\sigma > 0$ such that:

4. $$a(u,u) \geq \sigma \|u\|^2 \qquad \text{for all } u \in V$$

We notice that with the above properties the function $a(u,u)^{1/2}$ defined in V is a norm which is equivalent to the original norm $\|\cdot\|$. Furthermore, as a consequence of 3. we can associate with the bilinear form, a, a continuous linear operator $\mathcal{A} \in L(V,V)$ such that:

5. $$a(u,v) = ((\mathcal{A}u,v)) \qquad \text{for all } u,v \in V$$

In view of 4. it can be shown (see Lions |30|) that the operator $\mathcal{A}$ is an isomorphism on V.

Now, for $u \in V$, consider the linear functional:

6. $$v \in V \rightarrow a(u,v)$$

Denote by $D = D(A)$ the set of elements $u \in V$ for which the above linear functional is continuous on V with the topology induced by H. In other words, for all $u \in D$ there exists a constant C, in general depending on u, such that:

7. $|a(u,v)| \leq C|v|$ for all $v \in V$

As V is dense in H, the linear functional 6. can be extended for all $u \in D$ to a continuous linear functional defined in H:

8. $v \in H \rightarrow \bar{a}(u,v)$

Therefore, we can define uniquely a linear operator from $D \subset V$ to H, in general unbounded, such that:

9. $\bar{a}(u,v) = (Au,v)$ $u \in D,\ v \in H$

Now, let $J \in L(H,V)$ be the operator defined by

10. $(u,v) = ((Ju,v))$ $u \in H,\ v \in V$

Consider the problem AB), A'B') and A'B") given by the following statements:

A) $u \in D$

B) $Au = f \in H$

A') $u \in V$

B') $a(u,v) = (f,v)$ all $v \in V$

B") $\mathcal{A}u = Jf$

We have the following proposition:

<u>Proposition 2.3.1</u> - Under the hypotheses 1., 2., 3., 4. the problems AB), A'B') and A'B") are equivalent and admit a unique solution.

Proof of Proposition 2.3.1

By relations 5. and 10. problems A'B) and A'B") are equivalent. On the other hand, problem A'B') is a weak form for the problem AB) and, therefore, a solution for AB) is also a solution for A'B'). But in this case the reverse is also true. In fact if u solves A'B') we conclude that u must belong to D. Therefore:

$$(Au,v) = (f,v) \qquad \text{for all } v \in V$$

and by hypothesis 1., Au = f. The existence of a unique solution follows from the fact that, under the hypotheses made $\mathcal{A}$ is an isomorphism on V. So, problem A'B") admits the unique solution:

$$u = \mathcal{A}^{-1} Jf \in V \qquad \bullet$$

<u>Remark 2.3.1</u> - The proposition 2.3.1 is a version of the well known Lax-Milgram Lemma. For a more extended account of bilinear forms and its relation to linear operators see also Kato |20|.

<u>Remark 2.3.2</u> - Under the hypotheses 1., 2., 3., and 4. it can be

shown that D(A) is dense in V (or H) and that the linear operator A is a isomorphism between D(A) and H when D(A) is endowed with the norm,

$$\|\cdot\|_{D(A)} = (|\cdot|^2 + |A\cdot|^2)^{1/2}$$

<u>Remark 2.3.3</u> - Let us take $H = L^2(S)$ and suppose V is such that

$$H_0^1(S) \subset V \subset H^1(S)$$

For $(i,j) \in \{1,\ldots,n\}$ consider the bilinear form

11. $$a(u,v) = \int_S f(x) D_i u(x) D_j v(x)\,dx$$

defined for all $u,v \in V$, with $f \in L^\infty(S)$.

Fixing $u \in V$ and making v range in $C_0^\infty(S)$ the equation 11. defines a distribution. So, we write

12. $$a(u,v) = \langle T(f.D_i u) , D_j v\rangle$$

where, as in paragraph 2.1 the symbol $\langle .,. \rangle$ denotes the duality between $\mathcal{D}(S)$ and $\mathcal{D}'(S)$ and T(.) denotes, according to relation 2.1.1, the identification between $L^1_{loc}(S)$ and $\mathcal{D}'(S)$.

Recalling the definition of derivative of a distribution we can write,

13. $\langle T(fD_i u), D_j V\rangle = -\langle D_j T(fD_i u), V\rangle$

On the other hand, if $D_j(f.D_i u)$ exists, according to 2.1.3 we can write

14. $D_j T(f.D_i u) = T(D_j(f.D_i u))$

Therefore, comparing 12., 13., and 14 we have

15. $$a(u,v) = -\int_S D_j(f(x)D_i u(x))\ v(x)dx$$

for all $v \in C_0^\infty(S)$.

As $C_0^\infty(S)$ is dense in $L^2(S)$ the equation 15. defines a linear functional in $L^2(S)$ and therefore $D_j(f.D_i u) \in L^2(S)$. We conclude that the linear operator associated with the bilinear form a, has the form

16. $Au = -D_j(fD_i u)$

Also , the domain D(A) is determined by

17. i) $u \in V$, $Au \in H$

ii) $(Au,v) = a(u,v)$ for all $v \in V$

In particular, if $V = H_0^1(S)$ the condition ii) above is always verified since $C_0^\infty(S)$ is also a dense subset of

$H_o^1(S)$ and so, this condition follows from 15.

Remark 2.3.4 - We shall introduce here the concept of k-regularity of a bilinear form.

Suppose we select the Hilbert space V with

$$H_o^m(S) \subset V \subset H^m(S)$$

A bilinear form, a, in V, is said to be k-regular with respect to V, if for all f $\in H^r(S)$, $0 \leq r \leq k$, there exists u $\in H^{2r+m}(S)$ such that

18. $a(u,v) = (f,v)$ for all $v \in V$

The concept of k-regularity, as we shall see, plays a very important role in the situation where the bilinear forms are associated with linear differential operators. In this case this property depends on the coefficients of the differential operator, on the space V and on the regularity of the boundaries of the domain S. (see Lions |30|)

3 - EVOLUTION EQUATIONS

We shall be concerned, in a Hilbert space, with the solution of equations with the following generic form :

$$\frac{du}{dt}(t) + A(t)\,u(t) = f(t)$$

where A(t) is a linear operator, in general unbounded.

Such equations are called Evolution Equations. As the operators A(t) that occur in practical cases are usually partial differential operators, the equations we shall be treating are, in fact, parabolic partial differential equations. Although several methods have been used to study this sort of equation we will be following closely the work of Lions |30|. Our main objective is to derive existence and uniqueness results for the solution of the above equation under special hypotheses, namely, symmetry and differentiability of the principal part of the linear operator A(t). As we shall demonstrate in paragraph 3.4, under these circumstances the above equation can represent the solution of a filtering problem .

In order to show the existence of a solution for the evolution equation, two different techniques will be used. The first one is basically a projection theorem in Hilbert spaces. The second, is the so called Galerkin technique, and its main feature is to present the solution of the evolution equation as the limit of a convergent sequence of weak solutions of the original equation. This is the procedure with which we shall be concerned throughout this work.

The reason for presenting these two techniques is purely didactic. We believe that by presenting an alternative existence proof we introduce an element of comparison for the Galerkin technique.

In paragraph 3.1 we present an existence and existence and uniqueness result for a weak form which, with some manipulations, becomes an existence and uniqueness result for the Evolution Problem introduced in paragraph 3.2. In paragraph 3.3 we present the Galerkin technique. Finally, in paragraph 3.4 we apply the results to the non-stochastic representation of the solution of the filtering problem introduced in paragraph 1.1.

3.1 - A Weak Form

As before, let H, V be two Hilbert spaces with inner product and norm denoted as in paragraph 2.3.

We suppose $V \subset H$ with a continuous injection

1. $|v| \leq \|v\|$ for all $v \in V$

For all $t \in [0,T]$ let $a_j(t) = a_j(t;\ u,\ v)$ $j = 0,1$, be continuous bilinear forms in V such that:

2. $|a_0(t;\ u,\ v)| \leq \gamma_0 \|u\| \ \|v\|$

3. $|a_1(t;\ u,\ v)| \leq \gamma_1 \|u\| \ |v|$

for all $u, v \in V$

for some positive constants γ_0 and γ_1.

We suppose the bilinear form $a_0(t)$ to be symmetric,

4. $a_0(t;\ u,\ v) = a_0(t;\ v,\ u)$

for all $t \in [0,T]$; $u, v \in V$,

and coercive, in the sense that for some $\lambda \in R$ and $\sigma > 0$ the following inequality holds:

5. $$a_0(t; u, u) + \lambda |u|^2 \geq \sigma \| u \|^2$$

for all $u \in V$
$t \in [0,T]$

It turns out that the bilinear form $a(t)$, obtained by adding $a_0(t)$ to $a_1(t)$, also verifies a inequality of the above type. In fact, writing

6. $$a(t) = a_0(t) + a_1(t)$$

$t \in [0,T]$

we have:

$$\sigma \| u \|^2 \leq \lambda |u|^2 + a_0(t; u, u) =$$

$$= \lambda |u|^2 + a(t; u, u) - a_1(t; u, u)$$

So, by hypothesis 3,

$$\sigma \| u \|^2 \leq \lambda |u|^2 + a(t; u, u) + \gamma_1 \| u \| \, |u|$$

Using Cauchy's inequality $p.q \leq p^2 \varepsilon/2 + q^2/2\varepsilon$ with $\varepsilon > \frac{\gamma_1}{2\sigma}$, we have,

$$(\sigma - \frac{\gamma_1}{2\varepsilon}) \|u\|^2 \leq (\lambda + \frac{\gamma_1}{2}) |u|^2 + a(t, u, u)$$

for all $t \in [0,T]$, $u \in V$, which represents a coercivity condition similar to the one in 5.

Therefore, as a consequence of hypotheses 3. and 5. we also write for the bilinear form $a(t)$:

7. $$a(t; u, u) + \lambda|u|^2 \geq \sigma\|u\|^2$$

for all $t \in [0,T]$, $u \in V$
for some $\lambda \in R$, $\sigma > 0$

We also assume the following hypotheses:

8. $a_0(.; u, v) \in C^1(|0,T|); R)$ for all $u,v \in V$

9. $a_1(.; u, v) \in C(|0,T|); R)$ for all $u,v \in V$

10. $|a_0'(t; u, v)| \leq \gamma_0'\|u\| \|v\|$ for all $t \in [0,T]$; $u,v \in V$

where $a_0'(.; u, v)$ represents the derivative of $a_0(.; u, v)$.

Now consider the following problem:

11. i) $u \in L^2(0,T;V)$, $u' \in L^2(0,T;H)$

ii) $(u'(t),v) + a(t;u(t),v) = (f(t),v)$ $v \in V$, $t \in [0,T]$

with $f \in L^2(0,T)$

iii) $u(0) = 0$

where $u' = \frac{du}{dt}$ is taken in distributional sense.

We shall prove the following result:

Theorem 3.1.1 - Assuming hypotheses 1, 2, 3, 4, 5, 8, 9 and 10. the problem 11. admits a unique solution.

Remark 3.1.1 - Before we prove the theorem, let us establish the point that the problem 11. can always be reduced to a case where the coercivity condition 7 holds with $\lambda = 0$.

In fact, under the transformation:

12. $$w(t) = \exp(-\lambda t)u(t) \qquad t \in (0,T)$$

the equation 11.ii) can be replaced by the following equivalent equation:

13. $$(w'(t),v) + a(t;w(t),v) + \lambda(w(t),v) = \\ = \exp(-\lambda t) f(t)$$

where the bilinear form: $a(t;u,v) + \lambda(u,v)$ satisfies inequality 7. with the term in λ deleted. As the transformation 12. doesn't alter the other two statements of the problem 11. we shall, hereafter take inequality 7. with $\lambda = 0$. ●

Proof of uniqueness

If there are u_1 and u_2 solving the problem, their difference, $\Delta u = u_1 - u_2$, satisfies the following equation:

14. $$(\Delta u', v) + a(t; \Delta u, v) = 0 \qquad v \in V, \quad t \in (0,T)$$

Taking $v = \Delta u$ we have

$$(\Delta u', \Delta u) + a(t; \Delta u, \Delta u) = 0 \qquad t \in (0,T)$$

By inequality 7. (with $\lambda = 0$),

$$\frac{1}{2}\frac{d}{dt}|\Delta u|^2 + \sigma\|\Delta u\|^2 \leq 0$$

So,

15. $$\frac{d}{dt}|\Delta u|^2 \leq 0$$

as $\Delta u(0) = 0$, it follows that $\Delta u(t) = 0 \quad t \in (0,T)$ and the uniqueness is proved. ●

Remark 3.1.2 - As one can see by the proof, the solution, if it exists, will be unique, even in the case of a non-homogeneous initial condition. ●

Proof of existence

For $t \in R$ let $b(t)$, $b_o(t)$, $b_1(t)$ be bilinear forms in V such that:

$$b(t) = b_o(t) + b_1(t) \qquad t \in R$$

$$b_j(t) = a_j(0) \qquad t < 0\ ,\ j = 0,1$$

16. $$b_j(t) = a_j(t) \qquad t \in [0,T]\ ,\ j = 0,1$$

$$b_1(t) = a_1(T) \qquad t > T$$

$$b_o(t) = a_o(T) + \varepsilon(1 - \exp(\frac{T-t}{\varepsilon}))\ a_o'(T) \qquad t > T$$

where the parameter $\varepsilon > 0$ is conveniently selected in order to guarantee the existence of positive constants $\tilde{\sigma}$, α, $\tilde{\tilde{\sigma}}$ such that:

17. $$b_o(t;u,u) \geq \tilde{\sigma}\ \|u\|^2$$

18. $$\alpha b_o(t;u,u) - b_o'(t;u,u) \geq \tilde{\tilde{\sigma}}\ \|u\|^2$$

for all $t \in R$, $u \in V$.

As a consequence of the above characterization the bilinear form $b(t)$ is continuous in $V \times V$ for each $t \in R$ and we write,

19. $|b(t;u,v)| \leq \tilde{\gamma}\|u\| \|v\|$ for all $u,v \in V$

We also remark that, by definition,

20. $b_o(.;u,v) \in C^1(R^+)$ for each $u,v \in V$

21 $b_1(.;u,v) \in C(R)$ for each $u,v \in V$

Now let $\tilde{f} \in L^2(R;H)$ be such that,

22. $\tilde{f}(t) = f(t)$ for $t \in (O,T)$

$\tilde{f}(t) = O$ otherwise

With the real valued function h defined by,

23. $h(t) = \exp(-\frac{1}{2}\alpha t)$ $t \in R$

Consider the following auxiliary problem:

24. i) $hw \in L^2(R,V)$, $hw' \in L^2(R,H)$

ii) $\int_R (h(t)w'(t), h(t)\psi'(t)) +$

$+ b(t;h(t).w(t), h(t).\psi'(t))dt =$

(equation 24.ii) - continuation)

$$= \int_R (h(t)\tilde{f}(t),\ h(t)\psi'(t))\,dt,$$

for all V-valued functions ψ such that:

$$h\psi \in L^2(R,V),$$
$$h\psi' \in L^2(R,V),$$
$$\psi(t) = 0 \quad \text{for} \quad t \leq 0.$$

iii) $w(t) = 0$ for $t \leq 0$

The relation between the problem 11. and the problem above is contained in the following Lemma:

Lemma 3.1.1 - If w is a solution of problem 24. its restriction to (O,T) solves problem 11.

Proof of Lemma

For some $v \in V$, $\phi \in \mathcal{D}(R^+)$ the function:

25. $$\left(\int_0^t \phi(s)\,ds\right).v$$

satisfies:

$$h(t).\left(\int_0^t \phi(s)ds\right).v \in L^2(R^+;V),$$

$$h(t).\phi(t).v \in L^2(R^+; V).$$

Therefore, if we choose ψ such that $\psi(t) \equiv (\int_0^t \phi(s)\,ds)\,v$ for $t > 0$ as a test element in 24.ii) we can write,

$$\int_0^\infty (h(t)w'(t), h(t)\phi(t)v) + b(h(t)w(t), h(t)\phi(t)v)dt =$$

$$= \int_0^\infty (h(t)\tilde{f}(t), h(t)\phi(t)v)\,dt$$

As the equation above is true for all $\phi \in \mathcal{D}(R^+)$ we conclude that, almost everywhere,

26. $$(w'(t),v) + b(t;w(t),v) = (\tilde{f}(t),v),$$

for all $v \in V$
$t \in R^+$.

Therefore, the restriction of the function w to the interval $[0,T]$ satisfies all the requirements of problem 11. and the lemma is proved. ●

We now return to the proof of existence. By lemma 3.1.1, this can be done by proving the existence of a solution for problem 24. So, let E be the space of functions $\{w\}$ that verify statement 24.i) and 24.iii). This space can be made into a Hilbert space endowed with the following inner product:

27. $$(w_1, w_2)_E = \int_0^\infty (h(t)w_1(t),\ h(t)w_2(t)) + (h(t)w_1'(t),\ h(t)w_2'(t))\,dt,$$

for all $w_1,\ w_2 \in E$.

Consider the subspace $F \subset E$ of elements $\psi \in E$ such that:

$$h\psi' \in L^2(R,V).$$

Define the following bilinear form on $E \times F$:

28. $$B(w,\psi) = \int_0^\infty (h(t)w'(t),\ h(t)\psi'(t)) + b(t;h(t)w(t),\ h(t).\psi'(t))\,dt.$$

Also, define the following linear functional on F:

29. $$L(\psi) = \int_0^\infty (h(t)\tilde{f}(t),\ h(t).\psi'(t))\,dt.$$

Recalling equation 24.ii) one can observe that the problem 24. is equivalent to the problem of solving the following equation in the Hilbert space E:

30. $$B(w,\psi) = L(\psi) \qquad \text{for all } \psi \in F.$$

In order to establish the existence of a solution for the above equation we shall make use of the following result, which we state here without proof, (the proof can be found in Lions |30|, p. 37).

<u>Lemma 3.1.2</u> - Let E be a Hilbert space and $F \subset E$ a subspace. If B is a bilinear on $E \times F$ such that:

i) $B(.,\psi)$ is continuous for all $\psi \in F$

ii) There exists a constant $C > 0$ such that:

$$B(\psi,\psi) \geq C\|\psi\|_E \qquad \text{for all } \psi \in F$$

Then, if $L(\psi)$ is a continuous linear form on F, there exists a solution to the equation:

$$B(w,\psi) = L(\psi) \qquad \text{for all } \psi \in \phi$$

Let us show that the bilinear form B defined in 28. and the linear form 29. satisfy the requirements of the above Lemma. Equation 28. and inequality 19. give us:

$$|B(w,\psi)| < \int_0^\infty |h(t)w'(t)|\,|h(t)\psi'(t)| \; + $$

$$+ \; \tilde{\gamma}\|h(t)w(t)\| \; \|h(t)\psi'(t)\| \, dt.$$

So, fixing $\psi \in F$, recalling 27. and using Hölder's inequality we have

31. $$|B(w,\psi)| \leq C(\psi)\|w\|_E,$$

where $C(\psi)$ is a constant depending on $\|\psi\|_E$.
On the other hand, using definitions 16. we can write:

32. $$B(\psi,\psi) = \int_0^\infty |h(t)\psi'(t)|^2 + h^2(t)b_o(t;\psi(t),\psi'(t)) + $$

$$+ b_1(t;h(t)\psi(t),h(t)\psi'(t))\,dt$$

$$\text{for all } \psi \in F$$

As $b_o(t;u,v)$ is by definition a symmetric form,

$$b_o(t;\psi(t),\psi'(t)) = \frac{1}{2}\frac{d}{dt} b_o(t;\psi(t),\psi(t)) +$$

$$- \frac{1}{2} b_o'(t;\psi(t),\psi(t)),$$

$$t \in R^+.$$

Substituting in 32 we have,

33. $$B(\psi,\psi) = \int_0^\infty |h(t)\psi(t)|^2 +$$

$$+ \frac{1}{2} h^2(t)(\frac{d}{dt} b_o(t;\psi(t),\psi(t)) - b_o'(t;\psi(t),\psi(t)) +$$

$$+ b_1(t;h(t)\psi(t),h(t)\psi'(t))\,dt,$$

for all $\psi \in F$.

Using integration by parts we deduce the following identity:

$$\frac{1}{2}\int_0^\infty h^2(t)\,\frac{d}{dt}\,b_o(t;\psi(t),\psi(t))\,dt \quad =$$

$$= \alpha\int_0^\infty h^2(t)b_o(t;\psi(t),\psi(t))\,dt,$$

for all $\psi \in F$.

Substituting in 33. we have,

34. $$B(\psi,\psi) = \int_0^\infty |h(t)\psi(t)|^2 + \alpha b_o(t;h(t)\psi(t),h(t)\psi(t)) + - b_o'(t;h(t)\psi(t),h(t)\psi(t)) + + b_1(t,h(t)\psi(t),h(t)\psi'(t))\,dt,$$

for all $\psi \in F$.

Making use of inequalities 18. and 19., we have ,

$$B(\psi,\psi) \geq \int_0^\infty |h(t)\,\psi'(t)|^2 + \tilde{\tilde{\sigma}}\|h(t)\,\psi(t)\|^2 + - \tilde{\gamma}\|h(t)\psi(t)\|\,|h(t)\psi'(t)|\,dt,$$

for all $\psi \in F$.

Using Cauchy's inequality: $pq \le \frac{1}{2\varepsilon} p^2 + \frac{1}{2} \varepsilon q^2$

$$B(\psi,\psi) \ \ge \ \int_0^\infty |h(t)\psi(t)'|^2 + \tilde{\sigma}\|h(t)\psi(t)\|^2 \ +$$

$$- \ \frac{\tilde{\gamma}}{2\varepsilon} \|h(t)\psi(t)\|^2 - \frac{\tilde{\gamma}}{2} \varepsilon |h(t)\psi'(t)|^2 \, dt,$$

for all $\psi \in F$.

Therefore, by a convenient selection of the parameter ε we conclude that there exists a constant $C > 0$ such that:

35. $B(\psi,\psi) \ \ge \ C\|\psi\|_E$ for all $\psi \in F$.

As the linear form L, defined in 29. is continuous, in view of results 31. and 35., we are now able to apply Lemma 3.1.2 to equation 30. So, by this lemma, equation 30. admits a solution and so does problem 24. By Lemma 3.1.1, there exists a solution to problem 11. ●

In the next paragraph we shall see how the result presented in Theorem 3.1.1 can be used in order to obtain an existence and uniqueness result for evolution equations.

Remark 3.1.3 - We have borrowed the technique used in the proof of Theorem 3.1.1 from Lions |30|, where a equivalent result is derived for bilinear forms a(t) which are hermitian and continuously differentiable in relation to t. (Theorem 6.1 , p. 65). Here we have shown that Lion's result is still valid under weaker conditions, i.e., symmetry and differentiability imposed only in the principal part of the bilinear form a(t). As we shall see in paragraph 3.4, this is exactly what happens for evolution equations that arise in non-linear filtering theory.

3.2 - Existence and Uniqueness

In addition to the assumptions made in the last paragraph, let us take

1. V dense in H.

Under the hypotheses made we are now able to associate with the bilinear forms $a_j(t)$, $t \in [0,T]$, $j = 0,1$ a set of linear operators $A_j(t)$ in the sense suggested in paragraph 2.3. So,

2. $$A_j(t) : D(A_j(t)) \subset V \to H \qquad t \in [0,T], \quad j = 0,1$$

where $D(A_j(t))$ denotes the set of all $u \in V$ such that:

$$|a_j(t;u,v)| \leq C|v|, \qquad \text{for all } v \in V, \quad j = 0,1$$

where C is a constant in general depending on u.

In particular, by hypothesis 3.1.3, $D(A_1(t)) = V$ and $A_1(t) \in L(V,H)$ for all $t \in [0,T]$.

We also recall that, by the argument developed in paragraph 2.3, we have:

3. $$a_j(t;u,v) = (A_j(t)u,v),$$

$$\text{for all } u \in D(A_j(t));\ v \in V;\ t \in [0,T], \quad j = 0,1.$$

Let us denote by $A(t)$ the linear operator obtained by adding $A_0(t)$ to $A_1(t)$:

4. $$A(t) = A_0(t) + A_1(t), \qquad t \in [0,T]$$

This operator is the one associated with the bilinear form $a(t)$ and therefore,

5. $$a(t;u,v) = (A(t)u,v)$$

$$\text{for all } u \in D(A_0(t)),\ v \in V,\ t \in [0,T].$$

Consider now the Evolution Problem,

6. i) $u \in L^2(0,T;V)$, $u' \in L^2(0,T;H)$,
$u(t) \in D(A_0(t))$ for all $t \in (0,T)$,

ii) $u'(t) + A(t)u(t) = f(t), \qquad t \in (0,T),$
with $f \in L^2(0,T;H)$,

iii) $u(0) = u_0 \in D(A_0(0))$,

where u' is taken in the distributional sense.

We shall prove the following theorem:

Theorem 3.2.1 - Assuming the hypotheses of Theorem 3.1.1, if V is dense in H, problem 6. above has a unique solution.

Remark 3.2.1 - In other words, Theorem 3.2.1 states that, under certain conditions, equation 6.ii) has a unique solution $u \in L^2(0,T;V)$. Moreover, the derivative, u', is an element of the space $L^2(0,T;H)$.

This result concerning the derivative, is the characteristic of the theorem.

In fact, the existence of a unique solution $u \in L^2(0,T;V)$ for equation 6.ii) can be derived under considerable weaker conditions.

It can be shown (see Lions, Theorem 1.2, p. 102) that if $A(t)$ is a coercive linear operator, $A \in L^\infty(0,T;L(V,V'))$ equation 6.ii) admits a unique solution u such that

i) $u \in L^2(0,T;V)$

ii) $u' \in L^2(0,T;V')$

iii) $u = u_0 \in H$

The objective in this section is to show that, by strengthening the hypotheses relative to the principal part of the operator $A(t)$, we can obtain a stronger result for the derivative. This result can be achieved in the form of a corollary of the general result mentioned above. However, for didactic reasons, we present this result as a theorem. ●

Proof of Theorem 3.2.1

We start by supposing the existence of a function Z such that:

7. $Z \in L^2(0,T;V)$, $Z' \in L^2(0,T;H)$,

$Z(t) \in D(A_0(t))$, for all $t \in [0,T]$.

$$A(t)Z(t) \in L^2(0,T;H), \qquad Z(0) = u_0.$$

Consider the problem:

8. i) $w \in L^2(0,T;V)$, $w' \in L^2(0,T;H)$,

$w(t) \in D(A_0(t))$, for all $t \in [0,T]$.

ii) $w'(t) + A(t)w(t) = g(t)$, $t \in (0,T)$,

with $g(t) = f(t) - A(t)Z(t) - Z'(t)$.

iii) $w(0) = 0$.

We notice that, given the existence of a function Z which verifies the requirements in 7., problems 6. and 8. are equivalent under the tranformation,

9. $u = Z + w$.

Now, consider the equation,

10. $(w'(t),v) + a(t;w(t),v) = (g(t),v)$

for all $v \in V$, $t \in (0,T)$.

By theorem 3.1.1 the weak form 8.i), 8.iii), 10. has a unique solution. Therefore, to prove the theorem it is necessary to show that a solution of the weak form 8.i), 8.iii), 10. is also a solution for problem 8.

In fact, let w be the solution of the weak form. Then, we can write for all $t \in (0,T)$,

11. $a(t;w(t),v) = (g(t) - w'(t),v), \qquad v \in V.$

Using the result of proposition 2.3.1 it follows that:

12. $w(t) \in D(A(t)), \qquad t \in (0,T),$

13. $A(t)w(t) = g(t) - w'(t), \qquad t \in (0,T),$

and therefore w solves the problem 8.

So, to complete the proof of the Theorem we must prove the following,

<u>Lemma 3.2.1</u> - There exists a function Z which verifies requirements 7.

Proof of Lemma 3.2.1

For each $t \in [0,T]$ let $Z(t)$ be the solution of the following equation:

14. $a_0(t;Z(t),v) = (A_0(0)u_0,v), \qquad v \in V.$

By proposition 2.3.1 there exists a unique solution to the above equation satisfying $Z(t) \in D(A_0(t))$ for $t \in [0,T]$. Furthermore,

15. $A(t)Z(t) = A_o(O)u_o + A_1(t)Z(t) \in L^2(O,T;H)$,

16. $Z(O) = u_o$.

Using the coercivity hypothesis 3.1.5 in equation 14. with $v = Z(t)$ as a test element we also have:

17. $\sup_{[O,T]} \|Z(t)\| \leq \sigma^{-1}|A_o(O)u_o|$.

So, to complete the proof, we only need to show that $Z' \in L^2(O,T;H)$. In fact, by 15. we have:

18. $A_o(t+h)\,(Z(t+h) - Z(t)) + (A_o(t+h) - A_o(t))Z(t) = O$,

$$t \in (O,T),$$

and therefore,

19. $a_o(t+h);\ h^{-1}(Z(t+h) - Z(t)),v) =$

$$= -h^{-1}\int_t^{t+h} a_o'(s;Z(t),v)\,ds,$$

$$t \in (O,T).$$

Taking $v = h^{-1}(Z(t+h) - Z(t))$ as a test element, using hypotheses 3.1.5 (with $\lambda = O$), 3.1.10 and relation 17., we have:

21. $$\sigma\|h^{-1}(Z(t+h) - Z(t))\| \leq \gamma_0'\sigma^{-1}|A_0(0)u_0|, \quad t \in (0,T).$$

Therefore, there exists an element $Z'(t) \in V$ for each $t \in (0,T)$ such that as $h \to 0$,

$$h^{-1}(Z(t+h) - Z(t)) \overset{\text{weakly}}{\to} Z'(t).$$

By 21. $Z'(t) \in L^\infty(0,T;V)$ and so, the Lemma is proved. ●

3.3 - The Galerkin Technique

We now present an alternative proof for Theorem 3.2.1, and also derive estimates for the solution of the Evolution Problem 3.2.6. We shall achieve these objectives by using a technique in which the evolution equation is approximated by a sequence of ordinary differential equations.

Let us assume all the hypotheses of paragraph 3.2.. Suppose we are given a family of subspaces V_n, $n = 1,2,\ldots$, such that:

1. $V_n \subset V_m \subset V$ for all $n \leq m$; $n,m = 1,2,\ldots$

2. $\bigcup_n V_n$ is dense in V

In addition, suppose we are able to select from each subspace V_n an element ξ_0^n such that:

3. $\xi_o^n \to u_o$ in V as $n \to \infty$

For each natural number, n, we can, therefore, associate with the Evolution Problem 3.2.6 the following weak form:

4. i) $u_n \in L^2(O,T;V_n)$, $u_n' \in L^2(O,T;V_n)$

ii) $(u_n'(t),v) + a(t;u_n(t),v) = (f(t),v)$

for all $v \in V_n$
$t \in (O,T)$

iii) $u_n(O) = \xi_o^n \in V_n$

In relation to the weak form above we have

Lemma 3.3.1 - For each n = 1,2,... the problem 4. above has a unique solution.

Proof of Lemma 3.3.1

Let the integer N denotes the dimension of the subspace V_n and v_j, j = 1,...,N, a set of linearly independent elements of V_n which constitute a basis in this subspace.

Let M and K(t), $t \in |O,T|$, be N × N matrices, with elements given by:

5. $M_{i,j} = (v_i,v_j)$

6. $K_{i,j}(t) = a(t;v_i,v_j)$ $t \in [O,T]$

for $i,j = 1,\dots,N$

Let $\hat{f}(t) = (\hat{f}_1(t), \hat{f}_2(t), \dots, \hat{f}_N(t))$ be a R^N-valued function with,

$$7. \qquad \hat{f}_i(t) = (f(t), v_i) \qquad t \in (0,T), \quad i = 1,\dots,N.$$

Now, consider the system of N ordinary linear differential equations represented in matrix form by:

$$8. \qquad M.\alpha'(t) + K(t).\alpha(t) = \hat{f}(t) \qquad t \in (0,T)$$

where $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_N)$ is a R^N-valued function.

Since the v_j, $j = 1,\dots,N$ are linearly independent, $\det M \neq 0$. Therefore, the equation above admits a unique solution satisfying the initial condition $\alpha(0) = \alpha_0 \in R^N$ where,

$$9. \qquad M\alpha_0 = ((\xi_0^n, v_1), (\xi_0^n, v_2), \dots, (\xi_0^n, v_N)).$$

Take the function u_n defined by:

$$10. \qquad u_n(t) = \sum_{j=1}^{N} \alpha_j(t) v_j, \qquad t \in (0,T).$$

Simple manipulation shows that u_n, given as above, satisfies equation 4.ii). It also satisfies 4.i) and 4.iii) and, hence, u_n is a solution of the weak form 4. Besides, it is the unique solution, since every solution must have the form 10. and the initial value problem 8., 9. has a unique solution. As the argument is valid for all n, the Lemma is proved. ●

Now, consider the equation 4.ii). By Lemma 3.3.1 we can choose $v = u_n(t)$ as a test element. Substituting in the equation we have,

11. $$(u_n'(t), u_n(t)) + a(t; u_n(t), u_n(t)) = (f(t), u_n(t)),$$
$$t \in (0,T).$$

Using hypothesis 3.1.5 (with $\lambda = 0$),

12. $$\frac{d}{dt} |u_n(t)|^2 + 2\sigma \| u_n(t) \|^2 \leq 2 |f_n(t)| |u_n(t)|,$$
$$t \in (0,T).$$

Integrating over $(0,s)$, $s \in (0,T)$,

13. $$|u_n(s)|^2 + 2\sigma \int_0^s \| u_n(t) \|^2 \, dt \leq |\xi_0^n|^2 + $$
$$+ 2 \int_0^s |f_n(t)| |u_n(t)| \, dt.$$

Making use of Cauchy's inequality: $pq \leq \frac{1}{2\varepsilon} p^2 + \frac{1}{2} \varepsilon q^2$,

14. $$|u_n(s)|^2 + (2\sigma - \frac{1}{\varepsilon}) \int_0^s \| u_n(t) \|^2 \, dt \leq |\xi_0^n|^2 + $$
$$+ \varepsilon \int_0^s |f_n(t)|^2 \, dt.$$

Choosing the parameter ε conveniently and taking into account the hypothesis 3., we can derive, from the inequality 14. the following estimates:

15. $$\int_0^T \|u_n(t)\|^2\, dt \leq C(|u_o|^2 + \int_0^T |f(t)|^2\, dt).$$

16. $$|u_n(s)|^2 \leq C(|u_o|^2 + \int_0^T |f(t)|^2\, dt),$$

where $s \in [0,T]$; $n = 1,2,\ldots$ and C is a constant.

Let us return to equation 4.ii). Taking now $v = u_n'(t)$ as a test element we obtain,

17. $$(u_n'(t),u_n'(t)) + a(t;u_n(t),u_n'(t)) = (f(t),u_n'(t)),$$
$$t \in (0,T).$$

Recalling the composition of the bilinear form a(t), we have,

18. $$|u_n'(t)|^2 + a_o(t;u_n(t),u_n'(t)) =$$
$$= -a_1(t;u_n(t),u_n'(t)) + (f(t),u_n'(t)),$$
$$t \in (0,T).$$

As $a_o(t)$ is symmetric, (hypothesis 3.1.4), we have,

19. $$2|u_n'(t)|^2 + \frac{d}{dt} a_o(t;u_n(t),u_n(t)) = a_o'(t;u_n(t),u_n(t)) +$$

$$- 2a_1(t;u_n(t),u_n'(t)) + 2(f(t),u_n'(t)),$$

$$t \in (0,T).$$

Integrating over $(0,s)$, $s \in (0,T)$,

20. $$2\int_0^s |u_n'(t)|^2\,dt + a_o(s;u_n(s),u_n(s)) =$$

$$= a_o(0;\xi_o^n,\xi_o^n) + \int_0^s a_o'(t;u_n(t),u_n(t))dt +$$

$$- 2\int_0^s a_1(t;u_n(t),u_n'(t)) + (f(t),u_n'(t))dt.$$

Hence, using hypotheses 3.1.2, 3.1.3, 3.1.5 and 3.1.10.

21. $$2\int_0^s |u_n'(t)|^2\,dt + \sigma\|u_n(s)\|^2 \leq \gamma_o\|\xi_o^n\|^2 +$$

$$+ \int_0^s \gamma_o'\|u_n(t)\|^2\,dt + 2\int_0^s \gamma_1\|u_n'(t)\|\|u_n(t)\|dt +$$

$$+ 2\int_0^s |f(t)|\,|u_n'(t)|\,dt.$$

Using twice the Cauchy's inequality $pq = \frac{1}{2\varepsilon}p^2 + \frac{1}{2}\varepsilon q^2$ and rearranging terms,

$$22. \qquad (2 - \frac{1}{\varepsilon_1} - \frac{1}{\varepsilon_2}) \int_0^s |u_n'(t)|^2 dt + \sigma \| u_n(s) \|^2 \leq \gamma_0 \| \xi_0^n |^2 +$$

$$+ (\gamma_0' + \varepsilon_1 \gamma_1) \int_0^s \| u_n(t) \|^2 dt + \varepsilon_2 \int_0^s |f(t)|^2 dt.$$

Choosing the parameters ε_1, ε_2 conveniently, using hypothesis 3. and the previous estimate 15. we are able now to obtain the following estimates:

$$23. \qquad \int_0^T |u_n'(t)|^2 dt \leq C(\|u_o\|^2 + \int_0^T |f(t)|^2 dt),$$

$$24. \qquad \|u_n(s)\|^2 \leq C(\|u_o\|^2 + \int_0^T |f(t)|^2 dt,$$

where $s \in [0,T]$; $n = 1,2,\ldots$ and C is a constant.

Let us examine our position so far. We have obtained four estimates concerning the solution of the problem 4., namely, inequalities 15., 16, 23. and 24.. Inequality 15. suggests that, as n varies, the solution u_n of the problem 4. ranges in a bounded subset of the space $L^2(O,T;V)$. Also, inequality 23. suggests that the derivative u_n' ranges in a bounded subset of $L^2(O,T;H)$. Therefore we may extract from $\{u_n\}$ and $\{u_n'\}$ weak convergent sequences $\{u_m'\}$ and $\{u_m'\}$ such that:

$$25. \qquad u_m \to w \in L^2(O,T;V) \quad \text{weakly}$$

26. $u'_m \to Z \in L^2(0,T;H)$ weakly.

Using conventional arguments involving weak convergence and derivatives in distributional sense one can show that,

27. $Z = w'$,

where the derivative w' is taken in distributional sense.

Naturally we are expecting the function w defined by 25., 26. and 27. to be a solution for the Evolution Problem 3.2.6. In fact, this is the case.

Let us start by fixing some arbitrary natural number n_1. Consider the equation 4.ii) for $n > n_1$ with validity restricted to $V_{n_1} \subset V_n$. Multiplying both sides of the equation by $\psi(t)$ where $\psi \in C^1([0,T])$ with $\psi(T) = 0$, we obtain the following equation:

28. $$(u'_n(t), v\,\psi(t)) + a(t;u_n(t), v\psi(t)) = (f(t), v\psi(t)),$$

$$\text{for all } v \in V_{n_1},\quad t \in (0,T),\quad n > n_1.$$

Integrating over $(0,T)$ and using integration by parts in order to eliminate the derivative of u_n, we have,

29. $$-\int_0^T (u_n(t), v\psi'(t)) + a(t;u_n(t), v\psi(t))\, dt =$$

$$= (\xi_0^n, v\psi(0)) + \int_0^T (f(t), v\psi(t))\, dt,$$

for all $v \in V_{n_1}$,
$n > n_1$.

But by 25. there exists a subsequence $\{u_m : m > n_1\}$ converging weakly to w. So, recalling hypothesis 3. and passing to the limit the equation 29., we obtain,

30. $$-\int_0^T (w(t), v\,\psi'(t)) + a(t;w(t),v\psi(t))\,dt = $$
$$= (u_o, v\psi(0)) + \int_0^T (f(t), v\psi(t))\,dt,$$

for all $v \in V_{n_1}$.

Choosing $\psi \in \mathcal{D}(0,T)$ we have,

31. $$\int_0^T ((w'(t),v) + a(t;w(t),v))\psi(t)\,dt =$$
$$= \int_0^T (f(t),v)\psi(t)\,dt,$$

for all $v \in V_{n_1}$.

As the above is valid for all $\psi \in \mathcal{D}(0,T)$ we can write:

32. $$(w'(t),v) + a(t;w(t),v) = (f(t),v),$$

for all $v \in V_{n_1}$,
$t \in (0,T)$.

In this relation the index n_1 is fixed arbitrarily, and so, by hypothesis 1., we have,

33. $$(w'(t),v) + a(t;w(t),v) = (f(t),v),$$

$$\text{for all } v \in V, \quad t \in (0,T).$$

By hypothesis 3.2.1, V is dense in H. So, using Proposition 2.3.1., we deduce,

34. $$w'(t) + A(t)w(t) = f(t), \qquad t \in (0,T).$$

which is the equation 3.2.6.ii).

With respect to the initial condition, we observe that, multiplying both sides of equation 33. by $\psi(t)$ where $\psi \in C^1(|0,T|)$ with $\psi(T) = 0$ and integrating over $(0,T)$, we obtain after using integration by parts,

35. $$-\int_0^T (w(t),v\psi'(t)) + a(t;w(t),v\psi(t))\,dt =$$

$$= (w(0),v\psi(0)) + \int_0^T (f(t), v\,\psi(t))\,dt,$$

$$\text{for all } v \in V.$$

Comparing with 30. we have,

36. $$(w(0),v)\psi(0) = (u_0,v)\psi(0),$$

$$v \in V_{n_1}.$$

Again, as n_1 is arbitrary and V is dense in H we conclude

$$w(0) = u_o$$

Therefore w is indeed a solution for the Evolution Problem 3.2.6. As this solution must be unique (by. for instance, an argument similar to the one presented in the proof of Theorem 3.1.1), we have proved again Theorem 3.2.1.●

Remark 3.3.1 - The technique used in this paragraph in order to show the existence of a solution for the Evolution Problem 3.2.6 is due to Galerkin who introduced the method for elliptic equations. For parabolic and hyperbolic equations the technique was introduced respectively, by Green and Faedo (see Lions |30| for bibliographical references).

An important aspect of the Galerkin technique lies in the fact that it provides us with estimates for the solution of the Evolution Problem 3.2.6. In fact, recalling estimates 15., 16., 23. and 24., we are able to write for the solution, u, the following inequalities:

37. i) $\|u\|_{L^2(0,T;V)} \leq C\xi_1$

ii) $\|u\|_{L^\infty(0,T;H)} \leq C\xi_1$

iii) $\|u'\|_{L^2(0,T;H)} \leq C\xi_2$

iv) $\|u\|_{L^\infty(0,T;V)} \leq C\xi_2$

where C is a constant depending only on $\sigma, \gamma_o, \gamma_o'$ and γ_1, and

$$\xi_1 = |u_o| + \|f\|_{L^2(O,T;H)}$$

$$\xi_2 = \|u_o\| + \|f\|_{L^2(O,T;H)} .$$

In particular, estimates 37.i) and 37.iii) are sufficient to guarantee that u is a (almost surely) continuous function from $[O,T]$ to H. (see Lions |31| p. 102). ●

<u>Remark 3.3.2</u> - We have shown that the sequence $\{u_n\}$ of solutions of the problem 4. admits a weakly convergent subsequence to the solution of the Evolution Problem 3.2.6.. In fact this convergence is strong.

Considering equations 3.2.6.ii) and 4.ii), we can deduce the following identity:

38. $$(u'(t) - u_m'(t),v) + a(t;u(t) - u_m(t),v) =$$

$$= (f(t),\tilde{v}) - \{(u_m'(t),\tilde{v}) + a(t;u_m(t),\tilde{v})\},$$

for all $v = \hat{v} + \tilde{v} \in V$, with $\hat{v} \in V_m$ and $t \in (O,T)$.

Taking $v = u(t) - u_m(t)$ as a test element we can identify $\hat{v} = -u_m(t)$ and $\tilde{v} = u(t)$. Therefore, using inequality 3.1.7 (with $\lambda = 0$), equation 38. yields:

39. $$\frac{d}{dt}|u(t) - u_m(t)|^2 + \sigma\|u(t) - u_m(t)\|^2 \leq$$

$$\leq (f(t),u(t)) - \{(u_m'(t),u(t)) + a(t;u_m(t),u(t))\},$$

$$t \in (O,T).$$

Integrating over $(0,s)$ for $s \in [0,T]$ we have,

40. $$|u(s) - u_m(s)|^2 + \sigma \int_0^s \|u(t) - u_m(t)\|^2 dt \leq$$

$$\leq |u_0 - \xi_0^n|^2 + \int_0^s (f(t),u(t)) - \{(u_m'(t),u(t) +$$

$$+ a(t;u_m(t),u(t))\} dt.$$

By hypothesis 3., as $\{u_m\} \to u$, weakly, the right side of the above inequality tends to zero as $m \to \infty$. Therefore the subsequence $\{u_m\}$ converges strongly to u in $L^\infty(0,T;H)$ or $L^2(0,T;V)$. ●

Remark 3.3.3 - We have presented two procedures for showing the existence of solution for evolution equation. As we mentioned before, we have borrowed these procedures from Lions (|30| and |31|). Alternative techniques of achieving similar results can be found in Ladyzenskaya (|27|) (for parabolic equations) and in Showalter (|43|).

3.4 - An Application to the Filtering Problem

Here, we shall apply the results derived in the last paragraphs to the non-linear filtering problem introduced in paragraph 1.1. The object of our investigation, is, therefore, the pathwise representation for the filtering solution.

Let S be an open domain in R^n and take $H = L^2(S)$, $V = H_0^1(S)$.

Using the notation presented in paragraph 1.1, let us start by making the following hypotheses:

1. $a_{i,j} \in C^1(0,T;L^\infty(S))$,

$D_j a_{i,j}$, $D_{i,j} a_{i,j} \in C(0,T;L^\infty(S))$,

g_i, $D_i g_i \in C(0,T;L^\infty(S))$,

for all i,j = 1,...,n. We recall that $[a_{i,j}(t,x)] = \alpha(t,x).\alpha^T(t,x)$ and $[g_i(t,x)]$ represent, respectively, the diffusion matrix and the drift vector for the diffusion process 1.1.2..

We also assume that for some $\sigma > 0$,

2. $\langle r,[a_{i,j}(t,x)]r\rangle \geq \sigma\langle r,r\rangle$,

for all $r \in R^n$,
$(t,x) \in [0,T] \times S$,

where $\langle .,.\rangle$ denotes the scalar product in R^n.

Here, we shall be concerned with the case where the function h, introduced in 1.1.1, is invariant in time. We write,

3. $h(t,x) = h(x)$.

Assume that,

4. h, $D_i h$, $D_{i,j} h \in L^\infty(S)$,

for all i,j = 1,...,n.

Consider the bilinear form $a_o(t)$, $t \in [0,T]$, defined in $H_0^1(S)$ by,

5. $$a_o(t;u,v) = \frac{1}{2} \sum_{i,j=1}^{n} \int_S a_{j,i}(t,x) D_j u(x) D_i v(x)\, dx,$$

$$u,v \in H_0^1(S),$$
$$t \in [0,T].$$

Using an argument similar to that developed in the Remark 2.3.3, we find that the operator,

6. $$A_o(t)u = -\frac{1}{2} \sum_{i,j=1}^{n} D_i(a_{i,j}(t,.)D_j u),$$

$$t \in [0,T],$$

represents the linear operator associated with the bilinear form $a_o(t)$.

Consider the first order differential operator $B(t)$, defined by,

7. $$B(t)u = \sum_{i=1} D_i((-\frac{1}{2} \sum_{j=1} (D_j a_{j,i}(t,.)) + g_i(t,.))u).$$

We can write,

8. $$A_o(t) + B(t) = -L_t,$$

where L_t, $t \in [0,T]$ is the Fokker-Planck operator introduced in 1.1.9.

We shall now make some manipulations involving the operators presented above and, for economy of notation we will delete the arguments of the functions.

Let $y \in C([0,T])$. By conventional manipulation of derivatives, we can write,

9. $$\exp[-hy]A_o \exp[hy] = A_o + yB_o + y^2 c_o,$$

where B_o represents a first order differential operator and c_o is a multiplicative factor. We have,

10. $$B_o u = -\frac{1}{2} \sum_{i=1}^{n} D_i((\sum_{j=1}^{n} a_{j,i} D_j h)u) + -\frac{1}{2} \sum_{i=1}^{n} (\sum_{j=1}^{n} a_{i,j} D_j h) D_i u$$

11. $$c_o = -\frac{1}{2} \sum_{i,j=1}^{n} a_{i,j} (D_i h) D_j h \text{ .}$$

Using the same manipulation on the operator B defined in 7., we write,

12. $$\exp[-hy]B \exp[hy] = B + yc_1,$$

where,

13. $$c_1 = \sum_{i=1}^{n} (-\frac{1}{2} \sum_{j=1}^{n} (D_j a_{j,i}) + g_i) D_i h$$

Define the bilinear form $a_1(t)$, $t \in [0,T]$, by the following relation:

14. $$a_1(t;u,v) = ((B(t) + \frac{1}{2}h^2)u,v) + $$
$$+ y(t)((B_0(t) + c_1(t))u,v) + y^2(t)(c_0(t)u,v),$$
$$\text{for all } u,v \in H_0^1(S), \quad t \in [0,T],$$

and by $A_1(t)$, denote the operator associated with $a_1(t)$. We have,

15. $$A_1 = B + \frac{1}{2}h^2 + y(B_0 + c_1) + y^2c_0 .$$

Therefore, with the bilinear form,

16. $$a(t) = a_0(t) + a_1(t) \qquad t \in [0,T]$$

is associated an operator $A(t)$, $t \in [0,T]$ of the form,

17. $$A = -L + \frac{1}{2}h^2 + y(B_0 + c_1) + y^2c_0 =$$
$$= \exp[-hy]\,(-L + \frac{1}{2}h^2)\,\exp[hy] .$$

But this is exactly the differential operator that appears in the pathwise formula 1.1.16. (for h invariant in time). Therefore this equation can be rewritten here in the form,

18. $$u'(t) + A(t)u(t) = 0$$

On the other hand, under hypotheses 1.,...,4., one can easily show that the bilinear forms $a_0(t)$ and $a_1(t)$ satisfy conditions 3.1.2, 3.1.3, 3.1.4, 3.1.5†, 3.1.8, 3.1.9 and 3.1.10. Hence, according to Theorem 3.2.1, the evolution equation 18. has a unique solution $u \in L^2(0,T;H_0^1(S))$ such that,

$$u' \in L^2(0,T;L^2(S))$$

$$u(0) = u_0 \in D(A_0(0))$$

Moreover, recalling the estimates presented in section 3.3.,(see Remark 3.3.1), we can state the following theorem:

<u>Theorem 3.4.1</u> - Under hypotheses 1.,...,4., equation 18. has a unique solution,

$$u \in L^\infty(0,T;H_0^1(S)) \cap C(0,T;L^2(S))$$

satisfying $u(0) = u_0 \quad D(A_0(0))$.

---

† In particular, it can be show that by means of a suitable transformation of the original equation (see Remark 3.1.1) the coercivity condition 3.1.7 for the bilinear form a(t) holds independently of y. The reason for this fact is the quadratic form (in y) of A(t).

we have,

$$u' \in L^2(0,T;L^2(S)).$$

As a consequence of the theorem above, we can derive an existence and uniqueness result for the pathwise solution 1.1.16 of the filtering problem for diffusions in R^n. It suffices to take $S = R^n$ and assume the initial condition $r(0) = q_0$ as an element of $D(A_0(0))$. As the sample paths of the observation process are continuous functions, we deduce from Theorem 3.4.1 that, under hypotheses, 1.,...,4., the pathwise formula 1.1.16 has a unique solution,

19. $$r \in L^\infty(0,T;H^1(R^n)) \cap C(0,T;L^2(R^n)),$$

$$r' \in L^2(0,T;L^2(R^n)),$$

for all initial conditions q_0 such that $A_0(0)q_0 \in L^2(R^n)$.

<u>Remark 3.4.1</u> - We have assumed $V = H_0^1(S)$. In other words, we have been concerned with the initial value problem under Dirichlet boundary conditions associated with equation 18. For Neumann boundary conditions, we can use the same procedure as before with $V = H^1(S)$. (see Lions |30|, chapter VI, for a precise account on this situation).

In the context of the filtering problem, Dirichlet boundary conditions imposed on equation 1.1.16, correspond to the filtering problem for a diffusion absorbed by the boundary. In this case, Theorem 3.4.1 can be used straightaway. (see Pardoux |40|, for the stochastic equations governing the unnormalized conditional density of absorbed diffusions).

Remark 3.4.2 - We mentioned in paragraph 1.1 that the principal characteristic of the pathwise solution is its robustness. This means that the solution of the pathwise formula 1.1.16. is a continuous function defined on the sample space of y_t. Here, we shall present this fact in a more precise form.

Consider equation 18. with initial condition $u_o \in D(A_o(0))$.. Writing $u(t,y) = u(t)$ and $A(t,y) = A(t)$ in order to indicate the dependence on the parameter $y \in C([0,T])$, we can derive from 18. the following evolution equation:

20. $$w'(t) + A(t,y_1)w(t) = f(t),$$

where for $y_1, y_2 \in C([0,T])$,

21 $$w = u(.,y_1) - u(.,y_2),$$

22. $$f = -(y_1 - y_2)(B_o + c_1)u(.,y_2) + \\ - (y_1 - y_2)(y_1 + y_2)c_o u(.,y_2).$$

From Theorem 3.4.1 we can deduce that $f \in L^2(0,T;H)$. Therefore, the evolution equation 20. has exactly the form of the equations we have investigated in paragraphs 3.2 and 3.3. So, we can use the results of paragraph 3.3 in order to estimate its solution. Recalling Remark 3.3.1, we can write from 3.3.37.i) the following inequality:

23. $$\| w \|_{L^2(0,T;V)} \leq C \| f \|_{L^2(0,T;H)} .$$

On the other hand, from 22. we have,

24. $$\int_0^T |f(t)|^2 dt = \int_0^T (|y_1 - y_2| \; |(B_o + c_1)\, u\,(t,y_2|)^2 dt \; +$$

$$+ \int_0^T (|y_1 - y_2||y_1 + y_2||c_o u(t,y_2)|)^2 dt .$$

Again, from Theorem 3.4.1, there exists a constant C such that, for all $t \in [0,T]$,

$$|(B_o + c_1)u(t,y_2)|^2 < C,$$

$$|c_o u(t,y_2)|^2 < C .$$

Taking into account this fact and substituting in 24., we have from 23. the following inequality:

$$\|w\|_{L^2(0,T;V)} \leq C \left(\int_0^T (y_1 - y_2)^2 (1 + (y_1 + y_2)^2) dt \right)^{1/2}.$$

Hence, as y_1, y_2 are continuous functions, we can write,

25. $$\|u(.,y_1) - u(.,y_2)\|_{L^2(0,T;V)} \leq C\|y_1 - y_2\|_{L^2(0,T)} .$$

So, under the hypotheses of Theorem 3.4.1, the solution $u(t,y)$ of equation 18., is a continuous function from $C([0,T]) \subset L^2(0,T)$ to $L^2(0,T;V)$. ●

Remark 3.4.3 - In this paragraph we have investigated the pathwise formula 1.1.16. under the hypothesis that h is invariant in time. As a consequence of this condition, we have a polynomial form for the operator $A(t)$, in terms of powers of y (equation 17.). If h depends continuously on t we can obtain a similar form for the operator $A(t)$. In this case, instead of functions of the form $yD_i h$, we have $\int_0^y D_i h dt$, and similar results can be derived if we also assume $D_i h$, $D_{i,j} h$ belongs to $C(0,T;L^\infty(S))$.

4 - GALERKIN APPROXIMATIONS TO EVOLUTION EQUATIONS

In this section we present a family of unconditionally stable discrete time Galerkin schemes to approximate the solution of the evolution equation introduced in the last section. The kind of numerical procedure with which we shall be concerned has been largely used in relation with parabolic equations, and estimates for the error of approximation under differentiability conditions are very well known. Our objective here is to derive such estimates under weaker differentiability hypotheses.

In paragraph 4.1 we present the class of Implicit Runge-Kutta schemes which will receive our attention in this work. In paragraph 4.2 we derive some properties leading mainly to the stability of the schemes. In paragraph 4.3 an estimate for the error of approximation is deduced and, finally, in paragraph 4.4 we apply the results to the numerical approximation for the non-stochastic representation of the solution of the filtering for diffusion process presented in paragraph 1.1.

4.1 - Discrete Time Galerkin Methods

The Galerkin technique presented in paragraph 3.3 gives a procedure for approximating the solution of equation 3.2.6.ii) by solving a sequence of ordinary differential equations. It is this fact that inspires us to develop the discrete-time methods which we shall now present.

We assume the hypotheses made in section 3.

Therefore V and H are Hilbert spaces satisfying hypotheses 3.1.1 and 3.2.1. The symbols $(.,.)$, $(|.|)$, and $((.,.))$, $(\|.\|)$, denote the inner product, (norm), in H and V respectively.

The objects $a_j(t)$, $j = 0,1$, and $a(t)$, for $t \in [0,T]$ are bilinear forms defined in V, satisfying hypotheses 3.1.2, 3.1.3, 3.1.4, 3.1.5 (and, consequently 3.1.7, both taken here with $\lambda = 0$ according to Remark 3.1.1), 3.1.8, 3.1.9 and 3.1.10.

In addition to the hypothesis 3.1.3, concerning the upper bounds for the bilinear form $a_1(t)$, we also assume,

1. $$|a_1(t;u,v)| \leq \gamma_1 |u| \; \|v\|$$

$$u,v \in V$$
$$t \in [0,T]$$

Furthermore, we suppose that there exists a real valued function $z(t,s)$ defined in $[0,T] \times [0,T]$ such that

2. i) $z(t,s) \geq 0$

ii) $z(t,s) \to 0$ when $(t - s) \to 0$

iii) $|a_1(t;u,v) - a_1(s;u,v)| \leq z(t,s) \|u\| \; |v|$,

for all $u,v \in V$,
$t,s \in [0,T]$.

Throughout this section $\{0 = t_0 < t_1, < \ldots < t_N = T\}$ is a partition of the interval $[0,T]$.

We will use extensively the following notation for increments:

$$3. \quad \Delta f(t,s) = f(t) - f(s) \qquad t,s \in [0,T]$$

$$\Delta f_k(t) = f(t) - f(t_k) \qquad t \in [0,T],\ k = 0,1,\ldots,N$$

$$\Delta f_k = f(t_{k+1}) - f(t_k) \qquad k = 0,1,\ldots,N-1$$

$$\Delta_k = t_{k+1} - t_k \qquad k = 0,1,\ldots,N-1$$

for every function f defined in $[0,T]$.

Let $\mathcal{V} \subset V$ be a finite dimensional subspace.

In this section we shall be concerned with numerical procedures with the following iterative form:

$$4. \quad \frac{U_{k+1} - U_k}{\Delta_k} + \mathcal{G}_k U_k = 0,$$

$$k = 0,1,\ldots,N-1,$$

where, for $k = 0,1,\ldots,N$, $U_k \in \mathcal{V}$ and $\mathcal{G}_k \in L(\mathcal{V},\mathcal{V})$.

In order to be more specific we must determine the linear operator in the general form above.

So, for $k = 0,1,\ldots,N$, let $\mathcal{G}_k$ be such that:

$$5. \quad \mathcal{G}_k U_k + \sum_{j=1}^{r} \rho_j \beta_j = 0,$$

where, for all $j = 1,\ldots,r$, $\rho_j \in R$ and $\beta_j = \beta_j^k$ is an element of $\mathcal{V}$ verifying the following equation:

6. $$(\beta_j, v) + a(\tau ; U_k + \Delta_k \sum_{i=1}^{r} \rho_{i,j} \beta_i, v) = 0$$

Here $\tau = \tau_k \in [t_k, t_{k+1}]$ and $\rho_{i,j} \in R$ for $i,j = 1,\ldots,r$

With this characterization, the scheme 4. defines an r-stage implicit Runge-Kutta discretization method for the equation 3.2.6.ii) (with $f = 0$). This class of numerical procedures was introduced and studied by Butcher | 4 |. It has been widely used in connection with ordinary differential equations where, for a suitable choice of the parameters $\rho = \{\rho_i, \rho_{i,j}\}$ it produces unconditionally stable methods and convergent approximations (see, e.g. Stetter | 44 |). It has been also used in order to obtain approximations for the solution of parabolic equations. For instance, a one-stage scheme was used in Douglas | 12 | and Wheeler | 49 | for a non linear parabolic equation. In | 55 | Zlamal employs for a linear equation, invariant in time, a generic r-stage scheme with parameters obtained by means of Gaussian quadrature formulas.

It can be shown that the order of accuracy of the implicit Runge-Kutta schemes is directly related to the number of stages and, also, to the order of differentiability in time of the functions involved. Here, as the bilinear form $a_1(t)$ is, in general, non differentiable, there is little point in using a high order scheme and hence we shall concentrate on the one-stage case. So, we take equations 5. and 6. with $r = 1$. Making $\rho_1 = 1$, $\rho_{1,1} = \rho > 0$ and bringing the definition of the operator $\mathcal{G}_k$ into equation 4., we can rewrite our numerical scheme in the following, more recognizable, form:

7. $$\left(\frac{U_{k+1} - U_k}{\Delta_k}, v\right) + a(\tau ; \rho U_{k+1} + (1-\rho) U_k, v) = 0$$

$k = 0,1,\ldots,N-1$

Now, in the family of schemes represented above we can identify the Crank-Nicholson method when $\rho = 0.5$ and, with $\rho = 1$, the Implicit Backward method. These two methods are classical in the literature about discrete-time Galerkin procedures (see e.g. Strang |45|).

It is worthwile remarking here that the schemes presented above provide us with numerical procedures to approximate the solution of the ordinary differential equation 3.3.8 (with $f = 0$).

4.2 - Properties of the Numerical Schemes

Let $L(t)$, $t \in [0,T]$ be a family of linear operators from $\mathcal{V}$ to $\mathcal{V}$ defined by the following relation:

$$1. \qquad a(t;u,v) = (L(t)u,v)$$

$$\text{for all } u,v \in \mathcal{V}, \quad t \in [0,T]$$

These are well defined continuous linear operators in a finite dimensional subspace. Furthermore, by the coercivity condition 3.1.7 (with $\lambda = 0$) it follows the existence of the operators in the form $(I + k\,L(t))^{-1}$ where I is the identity operator, $k \geq 0$ and $t \in [0,T]$.

We are able to rewrite the numerical scheme proposed in the introduction of this section in a more compact form. In fact, using the definition 1. in equation 4.1.6, we have for the operator $\mathcal{G}_k$ introduced in 4.1.5 the following form:

2. $$\mathcal{G}_k = (I + \Delta_k \rho L(\tau))^{-1} L(\tau)$$

$$k = 0,1,\ldots,N,$$

where, we recall, $\rho > 0$, $\tau \in [t_k, t_{k+1}]$. So, from 4.1.4 the approximating elements U_k are given by,

3. $$U_{k+1} = (I - \Delta_k \mathcal{G}_k) U_k,$$

$$k = 0,1,\ldots,N.$$

Observe that the behaviour of the scheme is dictated by the operator $(I - \Delta_k \mathcal{G}_k)$.

Therefore writing $|||\cdot|||$ for the natural norm in $L(\mathcal{V},\mathcal{V})$ with $\mathcal{V}$ endowed with the topology of the space H, we introduce the following,

<u>Proposition 4.2.1</u>- Assuming the coercivity condition 3.1.7 (with $\lambda = 0$) the following estimates hold (independently of $\mathcal{V}$):

4. i) $$|||I - \Delta_k \mathcal{G}_k||| \leq \max\left\{ \left|\frac{1 - \Delta_k(1-\rho)\sigma}{1 + \Delta_k \rho\sigma}\right| , \left|\frac{1-\rho}{\rho}\right| \right\}$$

ii) $$|||I - \Delta_k \mathcal{G}_k||| \leq 1 \qquad \text{for} \qquad \rho \geq 0.5$$

iii) For $\rho > 0.5$, there exist constants ∂, $h_o > 0$ such that,

$$|||I - \Delta_k \mathcal{G}_k||| \leq \exp(-\partial \Delta_k)$$

for all $\Delta_k \leq h_o$,

$k = 0,1,\ldots,N.$

Proof of Proposition 4.2.1

Let $u = (I - \Delta_k \mathcal{G}_k)z$. Using the definition of the operator $\mathcal{G}_k$, given by equation 2., we have,

$$(u - z,v) + \Delta_k \rho\, a(\tau;\, \rho u + (1-\rho)z,v) = 0,$$

for all $v \in \mathcal{V}$.

Taking $v = \rho u + (1-\rho)z$ as a test element and using the coercivity condition, we have,

5. $$(u - z,\rho u + (1-\rho)z) + \Delta_k\rho\sigma\, \|\rho u + (1-\rho)z\|^2 \leq 0.$$

Recalling 3.1.1 and rearranging terms,

$$((1 + \Delta_k\rho\sigma)u - (1 - \Delta_k(1-\rho)\sigma)z,\ \rho u + (1-\rho)z) \leq 0.$$

Denoting: $$q = \frac{1 - \Delta_k(1-\rho)\sigma}{1 + \Delta_k\rho\sigma}, \quad r = \frac{1-\rho}{\rho},$$

the inequality 5. yields,

$$|u|^2 - qr|z|^2 - (q - r)(u,z) \leq 0.$$

Using now Schwartz inequality we have,

$$|u|^2 - |q-r|\ |u||z| - q.r|z|^2 \leq 0 .$$

Considering the above as a quadratic inequality in $|u|$ we conclude after conventional manipulations that:

$$|u| \leq \frac{1}{2}(|q-r| + |q+r|)|z| = (\max\{|q|,|r|\})\ |z|$$

and so item i) of the proposition is proved.

Item ii) follows from item i) if we take into account the premise,

$$|1-x| \leq 1 \quad \text{if and only if} \quad 0 \leq x \leq 2$$

and the fact that we can write,

$$|q| = \left|1 - \frac{\Delta_k\sigma}{1+\Delta_k\rho\sigma}\right| \quad \text{and} \quad |r| = |1 - \frac{1}{\rho}| .$$

Item iii) follows from previous items and the fact that if $\rho > 0.5$ it is always possible to find $c_1 > 0$ such that

$$|r| \leq |q| < 1 \quad \text{for} \quad 0 < \Delta_k \leq c_1 .$$

But $|q| \leq \exp(-\partial\Delta_k)$ for some $\partial > 0$ and $\Delta_k \leq c_2$. Making $h_o = \min(c_1,c_2)$ the thesis follows. ●

As a direct consequence of the result above, we can derive stability properties for the scheme 3. So, using a conventional terminology (see e.g. Stetter |44|) we can say that the scheme 3. is unconditionally stable for $\rho \geq 0.5$ and asymptotically stable for $\rho > 0.5$.

We recall that the meaning of these terms lies in the fact that if $X_k \in \mathcal{V}$, $k = 0,1,\ldots,N$, verify,

6. $$X_{k+1} = (I - \Delta_k \mathcal{G}_k)X_k + \Delta_k \zeta_k,$$

$$k = 0,1,\ldots,N-1,$$

$$X_0 = 0,$$

where $\zeta_k \in \mathcal{V}$, we deduce for $\rho \geq 0.5$,

$$\sup_k |X_k| \leq \sup_k |\zeta_k|,$$

which, roughly, means that "small perturbations" in the scheme produce "small displacements" from the initial condition. In the case $\rho > 0.5$ one can verify that the output of the scheme will exhibit a decrescing exponential pattern.

<u>Remark 4.2.1</u> - The first bound in the item i) of Prop. 4.2.1 namely $|q|$, is the usual and unique bound found in connection with an ordinary differential equation. Here, in general, the second bound, $|r|$, is dominant for $\rho \leq 0.5$.

The rational function q has the form of a Padé approximation for the exponential function $\exp(-\sigma\Delta_k)$ with maximum order of accuracy of 3 in the case $\rho = 0.5$. It seems that this fact is responsible for most of the properties of the scheme regarding stability and convergence. •

Now let $R(t)$, $t \in [0,T]$ be a family of linear operators from V to $\mathcal{V}$ defined by the equation:

7. $$a_o(t;u,v) = a_o(t; R(t)\,u,v),$$

$$\text{for all } u \in V,\ v \in \mathcal{V},\ t \in [0,T].$$

By the coercivity condition 3.1.5 (with $\lambda = 0$) it follows that the operator $R(t)$ is well defined for all $t \in [0,T]$. Furthermore $R(t) \in L(V,\mathcal{V})$, $R(t)\cdot R(t) = I$ for all $t \in [0,T]$ and so $R(t)$ is a projection operator. We also have,

8. $$\| u - R(t) \|^2 \leq \sigma^{-1}\, a_o(t;\, u - R(t)u,\, u - R(t)u) =$$

$$= \sigma^{-1}\, a_o(t;\, u - R(t)u,\, u - v) \leq$$

$$\leq \sigma^{-1}\, \gamma_o \| u - R(t)u \|\, \| u - v \|$$

$$\text{for all } u \in V,\ v \in \mathcal{V},$$

and as a consequence the following lemma can be stated:

<u>Lemma 4.2.1</u> - Under hypotheses 3.1.2 and 3.1.5 (with $\lambda = 0$) we have,

9. $$\| u - R(t)u \| = \sigma^{-1}\, \gamma_o \inf_{v \in \mathcal{V}} \| u - v \| .$$

The operator $R(t)$ is usually called the Ritz projection w.r.t $a_0(t)$ and $\mathcal{V}$ (see e.g. Strang-Fix |45|). In what follows we denote,

$$10. \qquad \tilde{R}(t) = I - R(t), \qquad t \in [0,T].$$

Our objective in this section is to derive estimates for the error of the approximation when we elect the family $U_k \in \mathcal{V}$, given by 3., as representative of the solution u of the Evolution Problem 3.2.6 (with $f = 0$). In other words, we are interested in the element,

$$11. \qquad \left(u(t_k) - U_k\right) \in V \qquad k = 0,1,\ldots,N,$$

or, using the definition 10. above,

$$12. \qquad u(t_k) - U_k = e_k + \tilde{R}(t_k)\, u(t_k)$$

$$k = 0,1,\ldots,N,$$

where e_k is the error in the subspace $\mathcal{V}$. That is:

$$13. \qquad e_k = R(t_k)\, u(t_k) - U_k,$$

$$k = 0,1,\ldots,N.$$

Now, let ϕ_k, $k = 0,1,\ldots,N$ be defined by,

$$14. \qquad \phi_k = R(t_{k+1})\, u(t_{k+1}) - R(t_k)\, u(t_k) + \Delta_k \mathcal{G}_k R(t_k)\, u(t_k)$$

Subtracting equation 3. from the above and rearranging terms, we can write,

15. $$e_{k+1} = (I - \Delta_k \mathcal{G}_k) e_k + \phi_k,$$

$$k = 0,1,\ldots,N-1 .$$

We observe that, roughly, the "size of the error" is directly related to the "size" of the variable ϕ_k, and so we can expect this variable to play a decisive role in the convergence of the method. In numerical analysis terminology, the variable ϕ_k is said to describe the consistency of the method, and we expect this variable to tend to zero as N tends to infinity. (see Stetter |44| for a general account on stability + consistency leading to convergence of numerical methods).

4.3 - An Abstract Error Estimate

According to Proposition 4.2.1, in order to guarantee unconditionally stable schemes we assume for now on

1. $$\rho \geq 0.5$$

Using the definition of the operator $\mathcal{G}_k$, given in 4.2.2. we can rewrite equation 4.2.15 in the following form:

2. $$e_{k+1} - e_k + \Delta_k L(\tau)(\rho e_{k+1} + (1-\rho) e_k) =$$

$$= (I + \Delta_k \rho L(\tau)) \phi_k,$$

$$k = 0,1,\ldots,N .$$

After recalling the definition 4.2.14 we have,

3. $$e_{k+1} - e_k + \Delta_k L(\tau)(\rho e_{k+1} + (1-\rho)e_k) =$$

$$= \Delta R u_k + \Delta_k \rho L(\tau)\, \Delta R u_k + \Delta_k L(\tau)\, R u(t_k),$$

$$k = 0,1,\ldots,N,$$

where $Ru(t) \equiv R(t)u(t)$ for all $t \in [0,T]$.

Using now definition 4.2.1 and some manipulation we can write,

4. $$(e_{k+1} - e_k, v) + \Delta_k\, a(\tau;\, \rho e_{k+1} + (1-\rho)e_k, v) =$$

$$= (\Delta u_k, v) - (\Delta \tilde{R} u_k, v) + \Delta_k a(\tau; \rho u(t_{k+1}) +$$

$$+ (1-\rho)u(t_k), v) - \Delta_k a(\tau;\, \tilde{R}u(t_{k+1}) +$$

$$+ (1-\rho)\, \tilde{R}u(t_k), v),$$

$$\text{for all } v \in \mathcal{V},$$

$$k = 0,1,\ldots,N-1,$$

where, according to 4.2.10, $\tilde{R}u(t) = u(t) - Ru(t)$ for all $t \in [0,T]$.

After the equation 3.2.6.ii) (with $f = 0$), we are able to write the following identity:

5. $$(\Delta u_k, v) + \Delta_k a(\tau; \rho u(t_{k+1}) + (1-\rho)\, u(t_k), v) =$$

$$= \int_{t_k}^{t_{k+1}} a(\tau; \rho \Delta u_k - \Delta u_k(s), v)\, ds \; +$$

$$+ \int_{t_k}^{t_{k+1}} a(\tau; u(s), v) - a(s; u(s), v)\, ds,$$

$$\text{for all } v \in V,$$

$$k = 0, 1, \ldots, N-1.$$

Taking this identity into account we can write equation 4. as follows:

6. $$(e_{k+1} - e_k, v) + \Delta_k a(\tau; \rho e_{k+1} + (1-\rho) e_k, v) =$$

$$= \int_{t_k}^{t_{k+1}} a(\tau; \rho \Delta u_k - \Delta u_k(s), v)\, ds \; +$$

$$+ \int_{t_k}^{t_{k+1}} a(\tau; u(s), v) - a(s; u(s), v)\, ds \; +$$

$$- (\Delta \tilde{R} u_k, v) \; +$$

$$- \Delta_k a(\tau; \rho \tilde{R} u(t_{k+1}) + (1-\rho) \tilde{R} u(t_k), v),$$

$$\text{for all } v \in \mathcal{V},$$

$$k = 0, 1, \ldots, N-1.$$

We observe that the equation above has a suitable form for manipulations in order to estimate e_k. The reason for that is the fact that it concentrates "small objects" in its terms in the right side. But before we continue, let us make a supplementary hypothesis in order to simplify the next steps.

Assume the principal part of the bilinear form $a(t)$, i.e., $a_o(t)$, to be invariant in time,

7. $$a_o(t) = a_o .$$

Remark 4.3.1 - Although the results we shall obtain in this section depend on the above condition, it does not constitute a fundamental hypothesis and equivalent results can be derived if the bilinear form $a_o(t)$ is sufficiently smooth in relation to the variable time. ●

Therefore, the Ritz projection is also invariant in time and recalling its definition in 4.2.7, equation 6. becomes:

8. $$(e_{k+1} - e_k, v) + \Delta_k a(\tau; \rho e_{k+1} + (1-\rho) e_k, v) =$$

$$= \int_{t_k}^{t_{k+1}} a_o(\rho \Delta u_k - \Delta u_k(s), v) ds \ +$$

$$+ \int_{t_k}^{t_{k+1}} a_1(\tau; u(s), v) - a_1(s; u(s), v) ds \ +$$

$$- \ (\Delta \tilde{R} u_k, v) \ +$$

$$- \ \Delta_k a_1(\tau; \rho \tilde{R} u(t_{k+1}) + (1-\rho) \tilde{R} u(t_k), v),$$

for all $v \in \mathcal{V}$,

$k = 0, 1, \ldots, N-1.$

Now consider the following identity regarding inner products:

$$\frac{1}{2}|e_{k+1}|^2 - \frac{1}{2}|e_k|^2 - (e_{k+1} - e_k, \rho e_{k+1} + (1-\rho)e_k) =$$

$$= (e_{k+1} - e_k, (\frac{1}{2} - \rho)e_{k+1} + (\frac{1}{2} - (1-\rho))e_k) =$$

$$= (\frac{1}{2} - \rho)|e_{k+1} - e_k|^2 .$$

Recalling hypothesis 1, we can write,

9. $$\frac{1}{2}|e_{k+1}|^2 - \frac{1}{2}|e_k|^2 \leq (e_{k+1} - e_k, \rho e_{k+1} + (1-\rho)e_k).$$

Returning to equation 8. we select $v = \rho e_{k+1} + (1-\rho)e_k$ as a test element. Taking into account the inequality 9, the coercivity condition 3.1.7 and hypotheses 3.1.2, 4.1.1 and 4.1.2 the following inequality holds:

10. $$\frac{1}{2}|e_{k+1}|^2 - \frac{1}{2}|e_k|^2 + \Delta_k \sigma \| \rho e_{k+1} + (1-\rho)e_k \|^2 \leq$$

$$\leq \int_{t_k}^{t_{k+1}} \gamma_o \| \rho \Delta u_k - \Delta u_k(s) \| \, \| \rho e_{k+1} + (1-\rho)e_k \| \, ds \; +$$

(equation 10. - continuation)

$$+ \int_{t_k}^{t_{k+1}} z(\tau,s) \; \|u(s)\| \; |\rho e_{k+1} + (1-\rho)\, e_k| \, ds \; +$$

$$+ \; |\Delta \tilde{R} u_k| \; |\rho e_{k+1} + (1-\rho)\; e_k| \; +$$

$$+ \; \Delta_k \gamma_1 |\rho \tilde{R} u(t_{k+1}) + (1-\rho)\; \tilde{R} u(t_k)| \; \|\rho e_{k+1} + (1-\rho) e_k\|,$$

$$k = 0,1,\ldots,N-1.$$

Using Cauchy's inequality, $pq \le 0{,}5\; p^2/\varepsilon \; + \; 0.5\varepsilon q^2$, for every term in the right side we have,

11. $$\frac{1}{2}|e_{k+1}|^2 - \frac{1}{2}|e_k|^2 - \Delta_k \sigma \; \|\rho\, e_{k+1} + (1-\rho)\; e_k\|^2 \; \le$$

$$\le \int_{t_k}^{t_{k+1}} \frac{\gamma_0}{2\varepsilon_1} \; \|\rho \Delta u_k - \Delta u_k(s)\|^2 \; ds \; +$$

$$+ \; \Delta_k \frac{\gamma_0}{2} \varepsilon_1 \; \|\rho\, e_{k+1} + (1-\rho) e_k\|^2 \; +$$

$$+ \int_{t_k}^{t_{k+1}} \frac{1}{2\varepsilon_2} \, z^2(\tau,s) \; \|u(s)\|^2 \, ds \; +$$

$$+ \; \Delta_k \frac{1}{2} \varepsilon_2 \; |\rho e_{k+1} + (1-\rho) e_k|^2 \; +$$

(equation 11. - continuation)

$$+ \frac{1}{2\varepsilon_3} |\Delta\tilde{R}u_k|^2 + \frac{1}{2}\varepsilon_3 |\rho e_{k+1} + (1-\rho)e_k|^2 +$$

$$+ \Delta_k \frac{\gamma_1}{2\varepsilon_4} |\rho\tilde{R}u(t_{k+1}) + (1-\rho)\tilde{R}u(t_k)|^2 +$$

$$+ \frac{\Delta_k\gamma_1}{2}\varepsilon_4 \|\rho e_{k+1} + (1-\rho)e_k\|^2,$$

$$k = 0,1,\dots,N-1.$$

Consider that, for $\rho \geq 0.5$, the following inequality holds:

$$|\rho e_{k+1} + (1-\rho)e_k|^2 \leq 2\rho^2\left(|e_{k+1}|^2 + |e_k|^2\right).$$

Now in the inequality 11. choose $\varepsilon_1 = \varepsilon_4 = 2\sigma/(\gamma_o + \gamma_1)$, $\varepsilon_2 = 1/4\rho^2$, $\varepsilon_3 = \Delta_k/4\rho^2$. We have, after rearranging terms,

12. $$\frac{1}{2}|e_{k+1}|^2 - \frac{1}{2}|e_k| \leq \frac{1}{2}\Delta_k|e_{k+1}|^2 + \frac{1}{2}\Delta_k|e_k|^2 + \frac{1}{2}\psi_k,$$

$$k = 0,1,\dots,N-1,$$

where,

13. $$\psi_k = \frac{\gamma_o(\gamma_o+\gamma_1)}{2\sigma}\int_{t_k}^{t_{k+1}} \|\rho\,\Delta u_k - \Delta u_k(s)\|^2\, ds + $$

$$+ 4\rho^2 \int_{t_k}^{t_{k+1}} z^2(\tau,s)\,\|u(s)\|^2\, ds +$$

(equation 13. - continuation)

$$+ \quad 4\rho^2 (\Delta_k^{-1} \; |\Delta \tilde{R} u_k|^2 \quad +$$

$$+ \quad \frac{\gamma_1(\gamma_o + \gamma_1)}{2\sigma} \, \Delta_k |\rho \tilde{R} u(t_{k+1}) \; + \; (1-\rho) \tilde{R} u(t_k)|^2 .$$

To obtain an estimate for the quantity $|e_k|^2$ independent of the remaining terms of the set, we need a version of Gronwall's Lemma appropriate to sequences. Here, we shall make use of the following

<u>Lemma 4.3.1</u> - Let X_k, $k = 0,1,\ldots,N$ be a sequence of real number such that,

14. $$X_{k+1} \leq (1 + h_k) X_k \; + \; \psi_k ,$$

where $h_k \geq 0$ and $\psi_k \in R$.

Then, for all $k = 0,1,\ldots,N$,

15. $$X_k \leq \exp \left[\sum_{o}^{N-1} h_j \right] (X_o \; + \; \sum_{j=0}^{N-1} \psi_j) .$$

Proof of Lemma 4.3.1

The thesis follows taking into consideration that $1 + h_k \leq \exp(h_k)$ and substituting X_k, $X_{k-1}, \ldots,$ X_o into equation 14. ●

Consider inequality 12. again. Multiplying both sides by 2 and rearranging terms under the assumption $\Delta_k < 1$, $k = 0,1,\ldots,N-1$, we have,

$$16. \quad |e_{k+1}|^2 \leq \frac{1+\Delta_k}{1-\Delta_k} |e_k|^2 + \frac{1}{1-\Delta_k} \psi_k =$$

$$= (1 + \frac{2\Delta_k}{1-\Delta_k}) |e_k|^2 + \tilde{\psi}_k ,$$

$$k = 0,1,\ldots,N-1.$$

Applying Lemma 4.3.1 to this inequality, we can write,

$$17. \quad |e_k|^2 \leq \exp \left| 2 \sum_{1}^{N-1} \frac{\Delta_j}{1-\Delta_j} \right| (|e_o|^2 + \sum_{j=0}^{N-1} \tilde{\psi}_j),$$

$$k = 0,1,\ldots,N-1.$$

Let us manipulate this inequality in order to obtain a final estimate suitable for the application we have in mind. So, returning to the expression 13., we can write for each of its terms in the right side, the following set of inequalities:

$$18. \quad \int_{t_k}^{t_{k+1}} \| \rho \Delta u_k + \Delta u_k(s) \|^2 ds \leq 2\rho^2 \Delta_k \| \Delta u_k \|^2 + 2 \int_{t_k}^{t_{k+1}} \| \Delta u_k(s) \|^2 ds$$

$$19. \quad \int_{t_k}^{t_{k+1}} z^2(\tau,s) \| u(s) \|^2 ds \leq \sup_{[0,T]} (\|u(t)\|^2) \cdot \int_{t_k}^{t_{k+1}} z^2(\tau,s) ds$$

20. $\Delta_k^{-1}\ |\Delta\tilde{R}u_k|^2 = \Delta_k^{-1}\ |\tilde{R}(\Delta u_k)|^2 =$

$$= \Delta_k^{-1}\ \left| \int_{t_k}^{t_{k+1}} \tilde{R}u'(s)\, ds \right|^2 \leq \int_{t_k}^{t_{k+1}} |\tilde{R}u'(s)|^2\, ds.$$

21. $\Delta_k |\rho\tilde{R}u(t_{k+1}) + (1-\rho)\tilde{R}u(t_k)|^2 =$

$$= \Delta_k\ |\rho\tilde{R}\Delta u_k + \tilde{R}u(t_k)|^2 \leq$$

$$\leq 2\rho^2\Delta_k^2 \int_{t_k}^{t_{k+1}} |\tilde{R}u'(s)|^2\, ds + \Delta_k|\tilde{R}u(t_k)|^2,$$

$$k = 0,1,\ldots,N-1.$$

By Lemma 4.2.1 and in view of the previous estimates for the solution of our Evolution Problem, given in Remark 3.3.1, the set of inequalities above makes sense.

Define h, the mesh of the partition of the interval $|0,T|$, by,

22. $h = \sup\{\Delta_k : k = 0,1,\ldots,N-1\}.$

Substituting inequalities 18.,...,21. into estimate 17. and rearranging terms, we can write

23. $$|e_k|^2 \le C\left\{ |e_o|^2 + \sum_{j=0}^{N-1} \left\{ h\|\Delta u_j\|^2 + \int_{t_j}^{t_{j+1}} \|\Delta u_k(s)\|^2\, ds \; + \right.\right.$$

$$+ \sup_{[0,T]} \|u(t)\|^2 \int_{t_j}^{t_{j+1}} z^2(\tau,s)\, ds \; +$$

$$\left.\left. + \int_{t_j}^{t_{j+1}} |\tilde{R}u'(s)|^2\, ds + h|\tilde{R}u(t_j)|^2 \right\} \right\},$$

for all $k = 0,1,\ldots,N$, $h \le h_o < 1$, where C is a positive constant depending only on the parameters γ_0, γ_1, σ, ρ and T.

We leave here the inequality 23. as a priori estimate for the error $|e_k|$, without further manipulations in its right side. It is our purpose to proceed in this way in paragraph 4.4 when a practical situation is analysed.

Remark 4.3.2 - Instead of identity 5. the following relation could have been written:

$$(\Delta u_k, v) + \Delta_k a(\tau;\rho u(t_{k+1}) + (1-\rho)\, u(t_k), v) =$$

$$= (\Delta u_k - \Delta_k u'(\tau), v) +$$

$$+ \Delta_k a(\tau;\rho u(t_{k+1}) + (1-\rho)u(t_k) - u(\tau), v),$$

for all $v \in V$,
$k = 0,1,\ldots,N-1$.

It turns out that, under differentiability conditions, this identity is more convenient to be manipulated in order to generate "small" terms in the final estimate. In fact, for the Crank-Nicholson scheme, i.e., $\tau = t_k + 0.5\,\Delta_k$, $\rho = 0.5$, we have,

$$\Delta u_k - \Delta_k . u'(\tau) = \frac{1}{2}\int_{t_k}^{\tau} (s - t_k)^2\, u^{(3)}(s)\, ds + \frac{1}{2}\int_{\tau}^{t_{k+1}} (s - t_{k+1})^2\, u^{(3)}(s)\, ds,$$

and,

$$\rho u(t_{k+1}) + (1-\rho)u(t_k) - u(\tau) = \frac{1}{2}\int_{t_k}^{\tau} (s - t_k)u^{(2)}(s)ds + \frac{1}{2}\int_{\tau}^{t_{k+1}} (s - t_{k+1})u^{(2)}(s)\, ds .$$

If the solution, u, of the Evolution Problem is sufficiently smooth we are able to produce terms of order Δ_k^2 in the final estimation for the error in the Crank-Nicholson case. As we are interested in a more general Evolution Problem, where the second derivative of the solution may not exist, we cannot take advantage of this fact. (Compare Remark 4.2.1 and see e.g. Wheeler |49| and Wilson |48| for the Crank-Nicholson method). ●

Remark 4.3.2 - The restrition $\Delta_k < 1$, which enables us to produce the estimate 17., does not constitute a intrinsic property of the scheme. It is only a consequence of the particular selection of values for the parameters ε_2 and ε_3 in equation 11. So, estimates like the inequality 23., must hold whatever the restrition, $h < h_o \in R$, imposed.

4.4 - An Approximation to the Filtering Solution

Here we shall bring the non linear filtering problem into the framework of this section. In other words, we will be concerned with approximating the pathwise solution 1.1.16 by means of the scheme introduced in 4.2.3.

Let $H = L^2(S)$ and $V = H_o^1(S)$, S being a bounded subset of R^n.

Consider the bilinear forms $a_j(t)$, $j = 0,1$ and $a(t)$, $t \in [0,T]$ introduced in 3.4.5, 3.4.14 and 3.4.16.

As we showed in paragraph 3.4, under hypotheses 3.4.1,...,3.4.4, these bilinear forms satisfy the hypotheses of Theorem 3.2.1. They also satisfy the supplementary conditions 4.1.1 and 4.1.2 introduced in the beginning of this section.

In fact, using integration by parts, equation 3.4.14 yields,

1. $$a_1(t;u,v) = (u,(B^*(t) + \frac{1}{2} h^2)v) + y(t)(u,(B_o^*(t) + c_1(t))v) + y^2(t)(u,c_o(t)v),$$

$$\text{for all } u,v \in H_o^1(S), \quad t \in [0,T],$$

where $B^*(t)$ and $B_o^*(t)$, $t \in [0,T]$, are first order differential operators with the form,

2. $$B^*u = \sum_{i=1}^{n} \left(\frac{1}{2} \sum_{j=1} (D_j a_{i,j}) - g_i\right) D_i u,$$

3. $$B_o^*u = \frac{1}{2} \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{i,j} D_j h\right) D_i u + $$
$$+ \frac{1}{2} \sum_{i=1}^{n} D_i \left(\left(\sum_{j=1}^{n} a_{i,j} D_j h\right) u\right).$$

From equation 1., one can easily show that condition 4.1.1 is satisfied.

Concerning the supplementary condition 4.1.2 we can write, from equation 3.4.14, the following relation:

4. $$a_1(t;u,v) - a_1(s;u,v) = \left(\Delta\left[B + \frac{1}{2} h^2\right] (t,s)u,v\right) +$$
$$+ (y(t) - y(s)) ((B_o(t) + c_1(t))u,v) +$$
$$+ y(s)(\Delta|B_o + c_1|(t,s)u,v) +$$
$$+ (y^2(t) - y^2(s))(c_o(t)u,v) +$$
$$+ y^2(s)(\Delta c_o(t,s)u,v),$$

for all $u,v \in H_o^1(S)$,
$t \in [0,T]$.

Now, if in addition to the hypotheses 3.4.1,...,3.4.4 we assume $D_j a_{i,j}$, $D_{i,j} a_{i,j}$, g_i, $D_i g_i$ belong to $C^1(0,T;L^\infty(S))$ for i,j = 1,...,n, then, from 4., we can deduce the following inequality:

5. $$|a_1(t;u,v) - a_1(s;u,v)| \leq \gamma_1'(|t - s| + |y(t) - y(s)|)\|u\| \, |v|,$$

for all $u,v \in H_o^1(S)$, $t \in [0,T]$,

for some positive constant γ_1' depending on the upper bounds of $y \in C([0,T])$ and, as well, the upper bounds of the first derivative in time of $a_{i,j}$, $D_j a_{i,j}$, $D_{i,j} a_{i,j}$, g_i and $D_i g_i$. Therefore, condition 4.1.2 is also satisfied with,

6. $$z(t,s) = \gamma_1'(|t - s| + |y(t) - y(s)|),$$

$t,s \in [0,T]$.

We shall now specify our approximation subspace $\mathcal{V} \subset V$.

In the beginning of this section we have described $\mathcal{V}$ as a finite dimensional subspace. Here, we improve this characterization by selecting the approximation subspace $\mathcal{V}$ as belonging to a family of subspaces of "finite element" type. This family will now be defined.

Let S be a bounded open set of R^n, $d \in (0,1)$ and r,m positive integers with $r < m$.

We denote by $\mathcal{V}(d,r,m)$ a finite dimensional subspace of $H^r(S) \cap H_o^1(S)$ with the following,

<u>Approximation Property</u>: For all non negative i,j such that,

$$i \leq r,$$
$$i \leq j \leq m,$$

there exists a constant K, independent of d and j, such that,

7. $$\|u - v\|_{H^i(S)} \leq Kd^{j-i}\|u\|_{H^j(S)},$$

for all $v \in \mathcal{V}(d,r,m)$, $u \in H^j(S) \cap H^1_0(S)$.

In order to complement the above definition we state, without proof, the following

<u>Lemma 4.4.1</u> - Let, a, be a coercive and bounded bilinear form defined on $H^1_0(S)$. Assume a* is O-regular on $H^1_0(S)$.†

Then, if $u \in H^p(S) \cap H^1_0(S)$, $p \geq 1$, we have

8. $$\|u - Ru\|_{L^2(S)} \leq Kq^q\|u\|_{H^q(S)},$$

where R is the Ritz projection w.r.t the bilinear form a and the subspace $\mathcal{V}(d,r,m)$; $q = \min(p,m)$; K is a constant independent of d and q.

---

† $a^*(u,v) = a(v,u)$; see the definition of k-regularity in Remark 2.3.4.

Remark 4.4.1 - The result in the lemma above is due to Nitsche (it can be found in Wheeler |49|). We observe that this result complements the approximation property 7. In fact, by Lemma 4.2.1, we have under the condition of Lemma 4.4.1 the following inequality,

$$\|u - Ru\|_{L^2(S)} \leq \sigma^{-1}\gamma_0 \|u-v\|_{H^1(S)},$$

$$\text{for all } v \in \mathcal{V}(d,r,m),$$

By the approximation property 7., we deduce that,

9. $$\|u - Ru\|_{L^2(S)} \leq \sigma^{-1}\gamma_0 K d^{q-1} \|u\|_{H^q(S)}.$$

Comparing the above inequality with 8. we see that the latter presents an extra factor d in the right side. This is a significant improvement because the exponent of d in the above expression can indicate the order of the approximation suggested in its right side. In general, for "finite element" spaces the parameter d represents the maximum diameter of the elements composing the domain S.

We also remark that similar results can be found if we take $V = H^1(S)$. (see Wheeler |49|). ●

Remark 4.4.2 - It is not our intention to present a general account of approximation subspaces of finite element type. For the purposes we have in mind, it is sufficient to record here the possibility of constructing a family of subspaces with the approximation property above. Further information can be found in the literature concerned with finite-element method (e.g, Douglas |12|, Strang-Fix |45|, Wheeler |49|, and in the Wilson-Nickell original paper |48|). ●

We are now in position to estimate the error of approximating the solution of equation 3.4.18 by means of the numerical scheme 4.2.3.

But before we proceed in this direction, in order to validate the use of the estimate 4.3.23, we must first assume the bilinear form $a_o(t)$ to be invariant in time. As we pointed out before (see Remark 4.3.1) this assumption is not restrictive. The character of our final result will not be spoiled by assuming smooth time variability of $a_o(t)$ and here, hypothesis 3.4.5 w.r.t $a_{i,j}$, is sufficient to achieve this smoothness.

We also would like to use the result of Lemma 4.4.1 in order to obtain a faster order of convergence in terms of the parameter d which measure the "discretization" in the space. So, we assume $a_o = a_o^*$ to be O-regular in $H_o^1(S)$.

To avoid confusion, let us recall the hypotheses that we have gathered so far. For the functions $a_{i,j}$, g_i and h, for i,j = 1,...,n, we have,

10. $a_{i,j},\ D_j a_{i,j},\ D_{i,j} a_{i,j} \in L^\infty(S)$,

$g_i,\ D_i g_i \in C^1(O,T;L^\infty(S))$,

$h,\ D_i h,\ D_{i,j} h \in L^\infty(S)$.

We also have (from 3.4.2) the coercivity condition,

11. $\langle r,[a_{i,j}]r\rangle \geq \sigma\langle r,r\rangle$,

for all $r \in R^n$, $x \in S$.

12. a_o is O-regular in $H^1_o(S)$.

With respect to the scheme 4.2.3 we take

13 $\vartheta = \vartheta(d,r,m)$ and $\rho \geq 0.5$.

Now, denoting,

$$|\Delta^h_y| = \sup_{[0,T]} \{ |y(t) - y(s)| : |t - s| \leq h\},$$

we can state the following result:

<u>Theorem 4.4.1</u> - Under the hypotheses 10.,...,13. if the solution of equation 3.4.18 satisfies,

$$u, u' \in L^\infty(0,T; H^p(S) \cap H^1_o(S)), \; p \geq 1$$

then the following estimate holds:

14. $$\sup_k |u(t_k) - U_k|^2 \leq C\{|Ru_o - U_o|^2 +$$

$$+ |\Delta^h_y|^2 \sup_{[0,T]} (\|u(t)\|^2_{H^p(S)}) + h^2 \sup_{[0,T]} (\|u(t)\|_{H^p(S)}) +$$

$$+ d^{2q}(\sup_{[0,T]} (\|u(t)\|^2_{H^q(S)}) + \sup_{[0,T]} (\|u'(t)\|^2_{H^q(S)}))\},$$

where q = min(p,m) and C is a constant independent of p, q, h and d.

Proof of Theorem 4.4.1

First we write according to 4.2.12,

15. $$|u(t_k) - U_k|^2 \leq 2|e_k|^2 + 2|\tilde{R}u(t_k)|,$$

$$k = 0,1,\dots,N.$$

To prove the theorem it suffices to use estimate 4.3.23 under the assumptions of this paragraph. Observe that, with respect to the terms in the right side of 4.3.23, we can write the following set of inequalities:

16. $$\|\Delta u_j\|^2 \leq \left\| \int_{t_j}^{t_{j+1}} u'(s)ds \right\|^2 \leq h^2 \sup_{[0,T]} \|u'(t)\|^2,$$

$$\int_{t_j}^{t_{j+1}} \|\Delta u_k(s)\|^2 ds \leq h^3 \sup_{[0,T]} \|u'(t)\|^2,$$

$$\int_{t_j}^{t_{j+1}} z^2(\tau,s)ds = \gamma_1' \int_{t_j}^{t_{j+1}} (|\tau - s|^2 + |y(\tau) - y(s)|^2)ds \leq$$

$$\leq \gamma_1' h^3 + \gamma_1' h \sup_{[0,T]} \{|\Delta y(t,s)| : |t-s| < h\},$$

$$\int_{t_j}^{t_{j+1}} |\tilde{R}u'(s)|^2 ds \leq hk^2 d^{2q} \sup_{[0,T]} \|u'(t)\|^2_{H^q(S)}$$

$$h|\tilde{R}u(t_j)|^2 \leq hk^2 d^{2q} \sup_{[0,T]} \|u(t)\|^2_{H^q(S)} .$$

In the last two inequalities we have used mainly the result of Lemma 4.4.1 under the hypothesis of O-regularity of our symmetric bilinear form a_o.

Substituting these inequalities in 4.3.23 and using the result in 15. we obtain the estimate 14. •

Remark 4.4.3 - Theorem 4.4.1 shows that the Galerkin scheme 4.2.3 provides us with a numerical procedure for approximating the solution of the Dirichlet problem associated with the equation 3.4.18. In other words, we are approximating the solution of the pathwise formula 1.1.16 defined in the cylinder $[0,T] \times S \subset R \times R^n$, with homogeneous condition on the boundary of the bounded domain S. As we mention before (see Remark 3.4.1) this situation corresponds to a filtering problem for diffusions absorbed by the boundary of S. If Newman boundary conditions are imposed on the pathwise formula, we start by taking $V = H^1(S)$ and then, a similar technique of analysis leads to a result equivalent to Theorem 4.4.1..

The discrete time Galerkin numerical procedure 4.2.3 has been widely used in connection with parabolic equations. Results concerning its rates of convergence are very well known for "smooth in time" differential operators. The purpose of our study is to analyse the procedure under weaker conditions with respect to the time variability of the "secondary" part of the differential operator. In other words, what distinguishes our study from the classical works about Galerkin approximations (e.g. Douglas-Dupont |12|) is our assumption with respect to the function y which, in the pathwise formula 1.1.16, represents the observation sample paths. Here, we take y as a continuous function. The result is that the procedure still converges and, under this condition, the rate of convergence is dictated by the modulus of continuity of the function y. From estimate 14., selecting $U_0 = Ru_0$, we can write,

$$\sup_k |u(t_k) - U(t_k)| \leq C(|\Delta_y^h| + h + d^q).$$

We observe that the procedure converges for all sample paths of bounded variation. The convergence is uniform over families of sample paths that satisfy a uniform Holder condition,

$$|\Delta_y^h| \leq kh^\alpha , \quad 0 < \alpha < 1 .$$

In this case, the order of convergence (w.r.t. h) has the same value as the Hölder coefficient α.

In |5|, Clark has shown that the pathwise solution for filtering problem for Markov chains admits a discrete approximation (Euler scheme) that converges uniformly with a rate depending on the modulus of continuity of the observation sample paths. Here, we have extended this result to diffusion processes.

5 - STOCHASTIC EVOLUTION EQUATIONS

The objective in this section is to examine the stochastic counterpart of the evolution equations studied in section 3, namely equations in the following stochastic differential form:

$$du(t) + A_0(t)u(t)dt + A_1(t)u(t)dw_t = f(t)dt$$

where $A_0(t)$ and $A_1(t)$ represent linear operators in a Hilbert space, which are in general unbounded.

The relevance of the class of equations above lies in the fact that the solution of the filtering problem for diffusion process admits such representation.

Stochastic evolution equations have received a great amount of attention recently and among the contributions to this field, the work of Pardoux and also Krylov-Rosovskii, are fundamental. Here we shall follow Pardoux |41|.

In paragraph 5.1 we present for random variables in Hilbert spaces, some of the conventional concepts valid for the real case. In paragraph 5.2 we introduce Pardoux's existence and uniqueness proof, which utilize the Galerkin technique presented in paragraph 3.3. It turns out that as in the non-stochastic case, the Galerkin approximation converges strongly to the solution of the stochastic evolution equation. Finally, in paragraph 5.3 the non linear filtering problem is brought into consideration and an existence and uniqueness result is derived.

5.1 - Stochastic Process in Hilbert spaces

We describe some of the usual definitions and results related to the topic above without any intention of giving a complete treatment of the subject. The main idea here is to show that the concepts valid for the real case can be easily extended to more complex spaces.

Our description is along the lines of the treatments given in Curtain-Falb |8|, Doob |10|, Neveu |39| and Scalora |42|.

We start by fixing a probability space $(\Omega, \mathcal{A}, P)$ and a Banach space X with norm denoted by the symbol $\|\cdot\|$.

A X-valued step random variable, x, is a mapping from Ω into X, such that

$$x(\omega) = u_i \quad \text{if} \quad \omega \in A_i \in \mathcal{A}; \quad i = 1,2,\ldots,N,$$

where $\{A_i\}$ is a set of disjunct measurable sets with $\bigcup A_i = \Omega$.

A X-valued random variable, x, is a strongly measurable mapping from Ω into X. We have,

i) There exists a sequence x_n, $n=1,2,\ldots$ of step random variables such that, w.p.1,

$$\| x_n(\omega) - x(\omega) \| \to 0 \quad \text{as} \quad n \to \infty .$$

ii) The set $\{\omega : x(\omega) \in B\} \in \mathcal{A}$ for all Borel set of X.

If x is a step random variable we write,

$$\int_\Omega x(\omega)dP = \sum_{i=1}^{N} u_i \, P(A_i) \in X .$$

A X-valued random variable, x, is said to be integrable if there exists a sequence x_n of step random variable, converging w.p.1 to x, such that,

$$\int_\Omega \| x_n(\omega) - x_m(\omega) \| \, dP \to 0 \quad \text{as} \quad n, m \to \infty .$$

Then, the limit of $\int_\Omega x_n(\omega)dP$ exists and we write,

$$\int_\Omega x(\omega)dP = \lim_{n\to\infty} \int_\Omega x_n(\omega)dP.$$

If x is an integrable random variable we define the expectation of x, E(x), as the element of X such that,

$$E(x) = \int_\Omega x(\omega)dP .$$

We define the space $L^p(\Omega,X)$, $1 \le p \le \infty$ as the space of (equivalent class of) X-valued random variables whose norm is p-integrable. It can be shown that these spaces are Banach under the norm,

$$\| x \|_{L^p(\Omega,X)} = (E(\|x\|^p))^{1/p} \qquad 1 \le p < \infty,$$

with the usual modification for $p=\infty$. As [illegible] we write $L^P(\Omega) = L^P(\Omega,R)$.

Let $\mathfrak{F}$ be a σ-subalgebra of $\mathcal{A}$ and let x be an integrable random variable. The conditional expectation of x relative to $\mathfrak{F}$, $E(x/\mathfrak{F})$, is a X-valued random variable such that,

$$\int_F x(\omega)\,dP = \int_F E(x/\mathfrak{F})\,dP,$$

for all $F \in \mathfrak{F}$.

It can be shown that such a random variable, $E(x/\mathfrak{F})$ is unique w.p.1 and integrable.

If x_1, (x_2), is a X_1, (X_2),-valued random variable, we say that x_1 and x_2 are independent if the sets

$$\{\omega : x_1(\omega) \in B_1\} \quad , \quad \{\omega : x_2(\omega) \in B_2\},$$

are independent for all Borel sets $B_1, (B_2)$ in $X_1, (X_2)$.

It can be shown that if $f_1, (f_2)$ is a Baire function mapping $X_1, (X_2)$ into the real numbers, then $f_1(x_1)$ and $f_2(x_2)$ are independent real random variables.

We also say that a random variable, x, is independent of the σ-algebra $\mathfrak{F} \subset \mathcal{A}$ if the sets F and $\{\omega : x(\omega) \in B\}$ are independent for all $F \in \mathfrak{F}$ and all Borel sets B of X.

If $\mathfrak{F} \subset \mathcal{A}$ is a σ-algebra and f, x and ϕ are respectively R, X_1, $L(X_1,X_2)$ - valued random variables, then the following statements can be proved (Curtain, |7|, |8|).

i) if $x \in L^1(\Omega,X_1)$ then $E(E(x/\mathfrak{F})) = E(x)$

ii) If in addition to i), x is $\mathcal{F}$-measurable then

$$E(x/\mathcal{F}) = x, \qquad \text{w.p.1.}$$

iii) If in addition to ii), $E(|f|.\|x\|) < \infty$ then

$$E(fx/\mathcal{F}) = E(f/\mathcal{F}).x, \qquad \text{w.p.1.}$$

iv) If in addition to i), $E\{\|\phi\| \cdot \|x\|\} < \infty$ and ϕ is $\mathcal{F}$-measurable then

$$E(\phi x/\mathcal{F}) = \phi E(x/\mathcal{F}), \qquad \text{w.p.1.}$$

Consider the interval $[0,T]$. Let $\mathcal{B}$ denotes the σ-algebra of Borel sets in $[0,T]$ and λ the Lebesgue measure. Consider the set $[0,T] \times \Omega$ and let $\mathcal{B} \times \mathcal{A}$ denotes the product σ-algebra and $\lambda \times P$ the corresponding measure (see Neveu, |39|).

We define a X-valued stochastic process as a X-valued random variable in the space $(|0,T| \times \Omega, \mathcal{B} \times \mathcal{A}, \lambda \times P)$.

We remark that, although this definition is less extensive than the usual one (see e.g. Doob |12| and Neveu |39|), it is adequate for the objectives we have in mind.

We shall now present the concept of stochastic integral for X-valued stochastic process. Here, for our purposes, we shall restrict ourselves to the special case where X is a Hilbert space. A more general account can be found in Curtain-Falb |8|.

We start by recalling the definition of a real valued Wiener process.

Let w_t be a R-valued stochastic process, with $w(0) = 0$, defined for $t \geq 0$ and continuous w.p.1.

If there exists an increasing family $\{\mathcal{F}_t\}$ of σ-subalgebra of $\mathcal{A}$ such that,

i) w_t is $\mathcal{F}_t$-measurable

ii) $E(w(t+h) - w(t)/\mathcal{F}_t) = 0$ w.p.1.

iii) $E((w(t+h) - w(t))^2/\mathcal{F}_t) = h$ w.p.1.

for all $t \geq 0$, $h > 0$.

Then w_t is a real valued, $\mathcal{F}_t$-measurable, non-antecipative standard Wiener process on the probability space $(\Omega, \mathcal{A}, P)$.

Now, let H be a Hilbert space with inner product and norm denoted respectively by $(.,.)$ and $|.|$. Assume that the concept of stochastic integral for real-valued processes is already familiar (see e.g. Gikhman-Skorokhod |14|).

Let w_t be a real valued, $\mathcal{F}_t$-measurable, non-anticipative standard Wiener process and $x(t)$, $t \in [0,T]$ be a H-valued stochastic process such that,

i) $E\int_0^T |x(s)|^2 ds < \infty$,

ii) $x(t)$ is $\mathcal{F}_t$-measurable.

For all $\phi \in H'$ (dual of H) the mapping $\phi X(t)$, $t \in [0,T]$ is a real-valued, $\mathcal{F}_t$-measurable, stochastic process such that,

$$\phi X \in L^2(\Omega; L^2[0,T]).$$

Therefore, we can define the stochastic integral of the process ϕx in Ito's sense, i.e.

$$\int_0^T \phi x(s)\,dw_s \in L^2(\Omega), \quad \text{for all } \phi \in H'.$$

and so, along with this, we have defined a linear mapping from H' into $L^2(\Omega)$.

This fact suggests the definition of the stochastic integral of the process x as the element of $L^2(\Omega, H)$ such that,

$$1. \qquad \phi \int_0^T x(s)\,dw_s = \int_0^T \phi x(s)\,dw_s, \qquad \text{for all } \phi \in H'.$$

This definition agrees with the conventional definition of stochastic integrals by means of finite sums. In fact, if $\{0 = t_o < t_1 < \ldots < t_N = T\}$ is a partition of the interval $[0,T]$ and

$$x(t) = x_i \in L^2(\Omega, H), \quad t \in [t_i, t_{i+1}],$$

$$i = 0,1,\ldots N-1.$$

then, it follows from 1.

$$\int_0^T x(s)\,dw_s = \sum_{i=1}^{N} x_i (w(t_{i+1}) - w(t_i))$$

The following items describe some properties of the stochastic integral defined in 1.

The mapping $\int_0^t x(s)dw_s$, $t \in |0,T|$, is an H-valued, $\mathcal{F}_t$-measurable, stochastic process, continuous in t w.p.1, such that,

$$\text{i)} \quad E \int_0^t x(s)dw_s = 0,$$

$$\text{ii)} \quad E \left| \int_0^t x(s)dw_s \right|^2 = E \int_0^t |x(s)|^2 ds .$$

(See also Pardoux |41|).

We can also introduce the concept of stochastic differential forms.

Let $u(t)$, $t \in |0,T|$, be an H-valued stochastic process such that,

$$2. \qquad u(t) - u(0) + \int_0^t f(s)ds + \int_0^t \alpha(s)dw_s = 0,$$

$$t \in [0,T],$$

where f, α are H-valued, $\mathcal{F}_t$-measurable stochastic processes such that,

$$\int_0^T |f(s)|ds < \infty \qquad \text{w.p.1}$$

$$\alpha \in L^2(\Omega;\ L^2(0,T;\ H))$$

Then we can rewrite 2. in the following stochastic differential form:

3. $$du(t) + f(t)dt + \alpha(t)dw_t = 0$$

Finally, we can state a Ito's rule of transformation for our stochastic differential forms. Here we recall the following Lemma which is a particular case of the one presented in Curtain-Falb |7|.

<u>Lemma 5.1.1</u> - (Ito's Lemma) Let the stochastic process u be given by 2 (or 3). Let $\psi \in C([0,T] \times H)$ with

i) $\dfrac{\partial \psi}{\partial t}(t,x) \in C([0,T] \times H)$,

ii) $\dfrac{\partial \psi}{\partial x}(t,x) \in C([0,T] \times H, H')$,

iii) $\dfrac{\partial^2 \psi}{\partial x^2}(t,x) \quad C([0,T] \times H, L(H,H))$.

Then, $Z(t) = \psi(t, u(t))$ is a real valued stochastic process with the following stochastic differential form:

$$dZ(t) = \left\{ \frac{\partial \psi}{\partial t}(t, u(t)) - \langle f(t), \frac{\partial \psi}{\partial x}(t, u(t)) \rangle + \right.$$

$$\left. + \frac{1}{2} \operatorname{tr} \left[(\alpha(t)\alpha^*(t)) . \frac{\partial^2 \psi}{\partial x^2}(t, u(t)) \right] \right\} dt +$$

$$- < \alpha(t), \frac{\partial \psi}{\partial x} (t, u(t))> dw_t \cdot$$

Here, <.,.> denotes the duality between H and H' and tr|.| denotes the trace of the operator indicated within the brackets.

5.2 - The Stochastic Evolution Problem

We shall introduce in this paragraph a basic result on existence and uniqueness for the solution of a Stochastic Evolution Problem. The proof we present is originally due to Pardoux (see |41|) and it makes use of the Galerkin technique we presented in paragraph 3.3. We also show that the Galerkin approximations converge strongly to the solution of the Stochastic Evolution Problem.

Let H, V be separable Hilbert spaces with inner products, (norms), denoted by the symbols $(.,.)$, $(|.|)$, and $((.,.))$, $(\|.\|)$, respectively.

Suppose V is dense in H with a continuous injection

1. $|v| \leq \|v\|$ for all $v \in V$

For $t \in [0,T]$, $a_j(t)$, $j = 0,1$ are bilinear functionals in the space V such that,

2. $a_j(.; u, v) \in L^\infty([0,T])$, $j = 0,1$

$u, v \in V$

3. $|a_0(t; u, v)| \leq \gamma_0 \|u\| \|v\|$

$u, v \in V$
$t \in [0,T]$

4. $|a_1(t;\ u,\ v)| \leq \gamma_1 \| u \| \, |v|,$

$u, v \in V,$

$t \in [0,T].$

By means of the argument presented in paragraph 2.3, we can associate with the bilinear forms $a_j(t)$, $j = 0,1$, linear operators, $A_j(t)$, such that

5. i) $a_j(t;\ u,\ v) = (A_j(t)\ u,\ v),$

ii) $A_j(t) : D(A_j(t)) \to H,$

$u \in D(A_j(t)),\ v \in V,$

$t \in [0,T],$

$j = 0,1.$

Here $D(A_j(t))$ denotes the set of all $u \in V$ such that $a_j(t;\ u,.)$ can be continuously extended to give an element of H'. As a consequence of hypothesis 4., we have $D(A_1(t)) = V$ for all $t \in [0,T]$.

We assume the following coercivity condition:

6. $2a_0(t;\ u,\ u) + \lambda |u|^2 \geq \sigma \| u \|^2 + |A_1(t)u|^2,$

for all $u \in V,$

$t \in [0,T],$

where $\lambda \in R$ and $\sigma > 0$.

Now, let w_t be a real valued $\mathcal{F}_t$-measurable, non-anticipative, standard Wiener process on a probability space $(\Omega, \mathcal{A}, P)$.

Denote by $M^2(0,T;\ V)$ the space of V-valued stochastic processes, x, such that,

7. i) $E\int_0^T |x(t)|^2 dt < \infty$,

ii) $x(t)$ is $\mathcal{F}_t$-measurable.

In this section we shall be concerned with the following Stochastic Evolution Problem:

8. i) $u \in M^2(0,T;\ V) \cap L^2(\Omega;\ C(0,T;\ H))$,

$u(t) \in D(A_0(t))$, $t \in [0,T]$, (w.p.1)

ii) $du(t) + A_0(t)u(t)dt + A_1(t)u(t)dw_t = 0$,

iii) $u(0) = u_0 \in H$.

In relation to this problem the following Theorem can be stated:

<u>Theorem 5.2.1</u> - Under hypothesis 2., 3., 4. and 6. the problem 8. has a unique solution.

This result has been obtained by Pardoux (|41|). Here, we present his proof.

To prove Theorem 5.2.1 we shall make use of the Galerkin technique introduced in paragraph 3.3. So, in order to proceed in this direction we must first bring into consideration the following weak form:

9. i) $u_n \in M^2(0,T;\, V_n) \cap L^2(\Omega;\, C(0,T;\, V_n))$,

ii) $d(u_n(t),\, v) + a_o(t;\, u_n(t),\, v)dt +$

$+ a_1(t;\, u_n(t),\, v)dw_t = 0$,

for all $v \in V_n$,

iii) $u_n(0) = u_o^n \in V_n$,

where V_n, $n = 1,2,\ldots$ is a family of finite dimensional subspaces of V.

Let us denote by P_n, $n = 1,2,\ldots$ the projection operator in H with respect to the subspace V_n.

The following Lemma can be stated:

<u>Lemma 5.2.1</u> - For each $n = 1,2,\ldots$ the problem 9. has a unique solution.

In addition, the following stochastic differential form holds:

10. $d\,|u_n(t)|^2 + 2a_o(t;\, u_n(t),\, u_n(t))dt +$

$-\ (A_1(t)\, u_n(t),\, P_n\, A_1(t)\, u_n(t))dt +$

(equation 10.; continuation)

$$+ \quad 2a_1(t;\ u_n(t),\ u_n(t))dw_t \quad = \quad 0.$$

Proof of Lemma 5.2.1

Let N denotes the dimension of the subspace V_n and v_j V_n, $j = 1,\ldots,N$ a set of linearly independent elements contituting a basis in V_n.

We can write the follwing identity:

11. $$(u,\ v) \quad = \quad <[u],\ M[v]>, \qquad\qquad u,\ v \in V_n.$$

where the symbol <.,.> denotes here the scalar product in R^N, [.] denotes the representation with respect to the basis $\{v_1,\ldots,v_N\}$ and M is an $n \times n$ matrix with,

$$M_{i,j} \;=\; (v_i,\ v_j), \qquad\qquad i,j = 1,\ldots,N.$$

In a similar fashion, we have,

12. $$a_o(t;\ u,\ v) \quad = \quad <[u],\ K(t)[v]>,$$

13. $$a_1(t;\ u,\ v) \quad = \quad <[u],\ R(t)[v]>,$$

$$u,\ v \in V_n,$$
$$t \in [0,T],$$

where K(t) and R(t), $t \in [0,T]$, are $n \times n$ matrices with,

$$K_{i,j}(t) = a_o(t; v_i, v_j)$$

$$R_{i,j}(t) = a_1(t; v_i, v_j),$$

$$i,j = 1,\dots,N.$$

So, equation 9.ii) can be rewritten in the following equivalent matricial form:

15. $$<[u_n(t)], M[v]> + \int_o^t <[u_n(s)], K(s)[v]> ds + \int_o^t <[u_n(s)], R(s)[v]> dw_s = 0,$$

$$\text{for all } v \in V_n, \quad t \in [0,T] \text{ (w.p.1)}.$$

As the matrix M is invertible and symmetric, the following stochastic differential equation is also equivalent to equation 9.ii):

16. $$d[u_n(t)] + M^{-1}K(t)[u_n(t)]dt + M^{-1}R(t)[u_n(t)]dw_s = 0,$$

$$t \in [0,T].$$

But by the theory of finite dimensional Ito's stochastic differential equations, equation 16. has a unique solution (see e.g. Gikhman - Skorokhod, |14|).

$$[u_n(t)] \in M^2(0,T;\ R^N) \cap L^2(\Omega;\ C(0,T;\ R^N)),$$

satisfying the initial condition,

$$[u_n(0)] = [u_o^n].$$

Therefore the first part of Lemma 5.2.1 is proved.

To show the second part of the Lemma we can use the standard Ito's rule of transformation for finite dimensional stochastic differentials. (See e.g. McKean |35|)

From equation 16., we deduce:

$$d([u_n(t)]^T M[u_n(t)]) = \Big\{ -2\, [u_n(t)]^T K(t)\, [u_n(t)] + $$

$$+ \operatorname{tr}\Big[R^T(t)\ [u_n(t)]\, [u_n(t)]^T\ R(t)\ M^{-1}\Big]\Big\}\, dt +$$

$$- 2\, [u_n(t)]^T\ R(t)\, [u_n(t)]\ dw_t .$$

The result follows if we use, in the above equation, relations 11., 12., 13. and the following identity:

$$(A_1(t)u\ ,\ P_n(A_1(t)v)) = \operatorname{tr}[R^T(t)\, [u]\, [v]\ R(t)M^{-1}]$$

$$u,\ v \in V_n$$

$$t \in [0,T]$$

Now, in order to show the above identity we first write for all $u,v \in V_n$, $t \in [0,T]$

17. $$tr\left[R^T(t)[u][v]^T R(t)M^{-1}\right] = \langle R^T(t)[u], M^{-1}R^T(t)[v]\rangle \,.$$

But we also have,

$$\langle [u], R(t)[v]\rangle = a_1(t;\, u,\, v) = (A_1(t)u,\, v) =$$

$$= (P_n(A_1(t)u),\, v) = \langle [P_n(A_1(t)u)], M[v]\rangle,$$

and therefore $[P_n(A_1(t)u)] = M^{-1}R^T[u]$ for all $u \in V_n$. Substituting this relation in 17. we have,

$$tr\left[R^T(t)[u][v]^T R(t)M^{-1}\right] = \langle R^T(t)[u], [P_n(A_1(t)v)]\rangle =$$

$$= (A_1(t)u,\, P_n(A_1(t)\,v)),$$

and so Lemma 5.2.1 is proved. ●

We can now prove Theorem 5.2.1. Before we proceed, let us make the following comment:

Remark 5.2.1 - As before (see Remark 3.1.1), without loss of generality, we can always take $\lambda = 0$ in the coercivity condition 6.. In fact, under the transformation

$$\tilde{u}(t) = \exp(-\lambda t)\, u(t) \qquad t \in [0,T],$$

equation 8.ii) becomes

$$d\tilde{u}(t) + (A_o(t)\ \tilde{u}(t) + \lambda\tilde{u}(t))dt + A_1(t)\ \tilde{u}(t)dw_t = 0,$$

and the corresponding form $a_o(t;\ u,\ v) + \lambda(u,\ v)$ now satisfies 6. with the term in λ deleted. •

Proof of Theorem 5.2.1; Uniqueness

To prove uniqueness we need a representation for the stochastic process $|u(t)|^2$, $t \in [0,T]$, when u satisfies the stochastic evolution equation 8.ii). In order to obtain such representation, we need an Ito's rule of transformation for infinite dimension stochastic processes.We can use either the Ito's Lemma presented in paragraph 5.1 or the Ito's Lemma proved by Pardoux in |41| and the result must, in conformity with equation 10, be valid for the finite dimensional case. In fact, this ...

So, if u solves problem 8. it can be shown that the following stochastic differential form holds:

18. $$d|u(t)|^2 + \{2\, a_o(t;\ u(t),\ u(t)) - |A_1(t)\ u(t)|^2\}\, dt + $$

$$+\ 2\ a_1(t,\ u(t),\ u(t))dw_t \ =\ 0,$$

$$t \in [0,T].$$

Now, suppose u_1 and u_2 solve problem 8. Then, $u = u_1 - u_2$ is also a solution with initial condition $u_o = 0$.

Using the equation 18. above, we can write,

$$|u(t)|^2 + \int_0^t 2\, a_o(s;\, u(s),\, u(s)) - |A_1(s)\, u(s)|^2\, ds +$$

$$+ \; 2\int_0^t a_1(s;\, u(s),\, u(s))\, dw_s \;=\; 0,$$

$$t \in [0,T] \qquad \text{w.p.1.}$$

Taking the expectation and recalling 6. we have,

$$E\, |u(t)|^2 + \sigma\, E \int_0^t \| u(t) \|^2\, dt \;\leq\; 0\, .$$

Therefore, if problem 8. has a solution, this solution must be unique. ●

Proof of Theorem 5.2.1; Existence

Let us assume that in addition to the hypotheses made for the weak form 9. we have,

19. i) $V_n \subset V_m$ for all $n \leq m$, $n,m = 1,2,...$

ii) $\cup V_n$ is dense in V

iii) $u_o^n \to u_o$ in H as $n \to \infty$,

(in other words we are assuming V to be separable).

Using the result of Lemma 5.2.1 we can write,

$$|u_n(t)|^2 + \int_0^t 2a_o(s; u_n(s), u_n(s) +$$

$$- (A_1(s)\ u_n(s), P_n(A_1(s)\ u_n(s)))\ ds =$$

$$= |u_o^n|^2 - 2\int_0^t a_1(s; u_n(s), u_n(s))\ dw_s,$$

$$t \in [0,T] \quad \text{w.p.1.}$$

Taking the expectation on both sides, using Schwartz inequality and the coercivity condition 6., we have

20. $$E|u_n(t)|^2 + \sigma E\int_0^t \| u_n(s)\|^2\, ds \leq |u_o|^2,$$

$$t \in [0,T].$$

Therefore, we can write the following estimate:

21. $$E\int_0^T \| u_n(s)\|^2\, ds \leq |u_o|^2 .$$

It follows that we can extract from the sequence $\{u_n\}$ a weakly convergent subsequence $\{u_\nu\}$ and so, we write,

22. $$u_\nu \to Z \in M^2(0,T;\ V), \quad \text{weakly}.$$

Let $\psi \in C([0,T])$ be such that,

23. i) $\psi' = \dfrac{d\psi}{dt} \in L^2(0,T)$,

ii) $\psi(T) = 0$.

From equation 9.ii), using Ito's rule of transformation and taking into account hypotheses 23., we have the following identity:

$$24. \quad \int_0^T a_o(s;\ u_\nu(s),\ v\psi(s)\ ds + \int_0^T a_1(s;\ u_\nu(s),\ v\psi(s))\,dw_s +$$

$$- \int_0^T (u_\nu(s),\ v\psi'(s)\ ds = (u_o^\nu,\ v\psi(0)),$$

for all $v \in V_{n_1}$,

$\nu \geq n_1$,

where n_1 is some natural number.

Now, let $x \in L^2(\Omega)$ be a random variable. Multiplying both sides of the above equation by x and taking the expectation, we can write,

$$25. \quad E(x.\phi_1) + E(x.\phi_2) + E(x\ .\ \phi_3) = E(x(u_o^\nu,\ v\psi(0)),$$

where for simplicity, by $\phi_i = \phi_i(u_\nu,\ v,\ \psi) \in L^2(\Omega)$ $i = 1,2,3$, we denote, respectively, the terms in the left side of equation 24..

We observe that, for $i = 1,2,3$, the expression $E(x.\phi_i(u_\nu,\ v,\ \psi))$, considered as a function of the variable $u_\nu \in M^2(0,T;\ V)$, defines a continuos linear functional on

$M^2(0,T;\ V)$. Therefore, by 22. we can take the limit of this expression as $\nu \to \infty$, yielding,

$$E(x.\phi_i(u_\nu,\ v,\ \psi)) \ \to\ E(x.\phi_i(Z,\ v,\ \psi)).$$

So, taking into account hypotesis 19.iii), it follows from equation 25,

$$\sum_{i=1}^{3} E(x.\phi_i(Z,\ v,\ \psi)) \ =\ E(x(u_0,\ v\psi)),$$

$$\text{for all } v \in V_{n_1}.$$

As the above identity is valid for all $x \in L^2(\Omega)$ we can conclude that, almost surely,

26. $$\sum_{i=1}^{3} \phi_i(Z,\ v,\ \psi) \ =\ (u_0,\ v\psi(0)),$$

$$v \in V_{n_1}.$$

The index n_1 has been fixed arbitrarily and so, using hypotheses 19., we can extend the validity of the above expression for all $v \in V$.

Assume the function ψ defined by $\psi(t) = \psi(\varepsilon,t)$ where

i) $\psi(\varepsilon,s) = 1 \quad$ for $\quad s \leq t - \varepsilon$

ii) $\psi(\varepsilon,s) = \dfrac{1}{2}\left(1 + \dfrac{1}{\varepsilon}(t - s)\right),$

for $s \in (t-\varepsilon,\ t+\varepsilon)$,

iii) $\psi(\varepsilon,s) = 0$, for $s \geq t + \varepsilon$,

where $\varepsilon > 0$ and $[t - \varepsilon, t + \varepsilon] \subset [0,T]$.

Sustituting in equation 28, with validity extended to all $v \in V$ and recalling the original expressions for $\phi_i(Z, v, \psi)$, $i = 1,2,3$, we have,

27. $$\int_0^T a_0(s;\ Z(s),\ v)\ \psi(\varepsilon,s)\ ds + \int_0^T a_1(s;\ Z(s),\ v)\ \psi(\varepsilon,s)\ dw_s + \frac{1}{2\varepsilon}\int_{t-\varepsilon}^{t+\varepsilon} (Z(s),\ v)\ ds = (u_0,\ v)$$

for all $v \in V$

We can now take the limit of the above expression as $\varepsilon \to 0$ for almost all $t \in (0,T)$, yielding the following identity:

$$(Z(t),\ v) - (u_0,\ v) + \int_0^t a_0(s;\ Z(s),\ v)\ ds + \int_0^t a_1(s;\ Z(s),\ v)\ dw_s = 0$$

for all $v \in V$
$t \in [0,T]$ w.p.1

As V is dense in H, by a standard argument (see paragraph 2.3) we conclude,

$$Z(t) - u_o + \int_o^t A_o(s)\, Z(s)\, ds + \int_o^t A_1(s)\, Z(s)\, dw_s = 0,$$

$$t \in [0,T], \qquad \text{w.p.1.}$$

So, $Z(t)$, $t \in [0,T]$ is w.p.1 equal to a continuous H-valued stochastic process which satisfies the requirements of problem 8. •

Remark 5.2.2 - Inequality 20. also give us an estimate for the solution of problem 8. considered as an element of $L^2(\Omega;\ C(0,T;\ H))$, and this [illegible] the question of the stability of the solution of equation 8.ii. For an account on the asymptotic stability of the second moment of the solution of equation 8.ii., see Haussmann, |16|. •

Remark 5.2.3 - Here, as in the non-stochastic case presented in paragraph 3.3., the solution of the weak form 9. converges strongly to the solution of the stochastic evolution equation 8.ii).

To show this fact we start by writing the identity,

28. $$|u(t) - u_n(t)|^2 = |u(t)|^2 + |u_n(t)|^2 - 2(u(t),\ u_n(t)),$$

$$t \in [0,T],$$
$$n = 1,2,\ldots$$

Recalling the energy formulas 10. and 18. and substituting in the above relation we have,

29. $$|u(t) - u_n(t)|^2 = |u_o|^2 + |u_o^n|^2 - 2(u(t), u_n(t)) + \\ - \int_0^t b(s; u(s), u(s)) + b_n(s; u_n(s), u_n(s))\, ds + \\ -2\int_0^t a_1(s; u(s), u(s)) + a_1(s; u_n(s), u_n(s))dw_s ,$$

$t \in [0,T]$ w.p.1

$n = 1,2,...$

$b(t)$ and $b_n(t)$, $n = 1,2,...$ denote the following bilinear forms on V:

30. i) $b(t, u, v) = 2a_o(t; u, v) - (A_1(t)u, A_1(t)v)$

ii) $b_n(t; u, v) = b(t; u, v) + (A_1(t)u, \tilde{P}_n A_1(t) v)$

where $\tilde{P}_n = (I - P_n)$.

We also have, using the above definitions,

31. $$b(t; u, u) + b_n(t; v, v) = b_n(t; u-v, u-v) + \\ - (A_1(t)u, \tilde{P}_n A_1(t)u) + b_n(t; u, v) + \\ + b_n(t; v, u) .$$

Making use of the above equation in 29. we can write, after some manipulation,

32. $$|u(t) - u_n(t)|^2 + \int_0^t b_n(s; u(s) - u_n(s), u(s) - u_n(s))\,ds =$$

$$= \phi(t; u, u_n) - \int_0^t a_1(s; u(s), u(s)) +$$

$$+ a_1(s; u_n(s), u_n(s))\,dw_s,$$

$$t \in [0,T] \quad \text{w.p.1}$$

$$n = 1,2,\ldots$$

Here,

33. $$\phi(t; u, u_n) = |u_0|^2 + |u_0^n|^2 - 2(u(t), u_n(t)) +$$

$$+ \int_0^t (A_1(s)\,u(s), \tilde{P}_n A_1(s)\,u(s))\,ds +$$

$$- \int_0^t b_n(s; u(s), u_n(s)) +$$

$$+ b_n(s; u(s), u_n(s))\,ds\,.$$

Taking the expectation on both sides of equation 32. and using the coercivity condition 6. transferred to the bilinear form $b_n(t)$, we have,

31. $$E\,|u(T) - u_n(T)|^2 + \sigma E \int_0^T \| u(s) - u_n(s) \|^2 \, ds \leq$$

$$\leq E(\phi(T;\, u,\, u_n)),$$

$$n = 1,2,\ldots$$

But by inequalities 21. and 22. we can select from the sequence $\{u_n\}$ a weakly convergent sequence $\{u_\nu\}$ such that, as $\nu \to \infty$

35. i) $E\,(u(T),\, u_\nu(T)) \to E|u(T)|^2$

ii) $E \int_0^T b(s;\, u_\nu(s),\, u(s))\, ds \to E \int_0^T b(s;\, u(s),\, u(s)))\, ds$.

Besides, by hypotheses 19., we also have as $\nu \to \infty$

36. i) $E \int_0^T (A_1(s)\, u(s),\, \tilde{P}_\nu\, A_1(s)\, u(s)\, ds \to 0,$

ii) $|u_0^\nu|^2 \to |u_0|^2,$

$$\text{iii)} \quad E \int_0^T (A_1(s)\, u(s),\ \tilde{P}_\nu\, A_1(s)\, u_\nu(s))\, ds \to 0$$

Therefore, by equation 33. and relations 35. and 36., as $\nu \to \infty$, we have

$$E\phi(T;\ u,\ u_\nu) \to 2|u_o|^2 - 2\, E|u(T)|^2 + $$
$$- 2\, E \int_0^T b(s;\ u(s),\ u(s))\, ds\,.$$

Comparing with the energy formula 18. we bserve that the right side of the above relation is zero. Therefore, returning to inequality 34 we conclude that, as $\nu \to \infty$

$$u_\nu \to u\ , \quad \text{strongly in } M^2(O,T;\ V)\,.$$

<u>Remark 5.2.4</u> - Let us investigate what happens if in the argument leading to the existence proof of Theorem 5.1.1, we consider stochastic integrals in the Stratonovich's sense (instead of Ito's).

Consider the R^N-valued, stochastic differential form 16. Taking into account the relation between Ito's and Stratonovich integrals (see Stratonovich, |47|), this equation has the following stochastic representation in the Stratonovich's sense:

37. $$d[u_n(t)] + (M^{-1} K^T(t) + \frac{1}{2} (M^{-1} R(t))^2) [u_n(t)]) dt + M^{-1} R(t) [u_n(t)] dw_t = 0,$$

$$t \in [0,T]; \quad (S).$$

Or equivalently,

38. $$d(u_n(t), v) + \{a_o(t; u_n(t), v) + \frac{1}{2} (A_1(t) P_n A_1(t) u_n(t), v)\} dt + a_1(t; u_n(t), v) dw_s = 0,$$

$$\text{for all } v \in V_n$$
$$t \in [0,T] \; ; \; (S) .$$

The equation above is the Stratonovich counterpart of equation 9.ii), and in its derivation we have used 11., 12., 13. and the following relation:

39. $$< (M^{-1}R^T(t))^2 [u], M [v] > = < M^{-1} R^T(t) [u], R(t) [v] > = (A_1(t) (P_n A_1(t) u), v)$$

$$u, v \in V_n$$
$$t \in [0,T]$$

If we suppose $A_1(t) \in L(H, H)$, $t \in [O,T]$, a copy of the existence proof of Theorem 5.2.1 must lead us to the conclusion that there exists a weakly convergent subsequence $\{u_\nu\}$ which converges to the solution of the following evolution equation:

$$40. \quad du(t) + (A_o(t) + \frac{1}{2} A_1^2(t)) u(t) dt + A_1(t) u(t) dw_t = O, \qquad t \in [O,T;] (S).$$

In his paper, Stratonovich gives the rule of transformation between his integral and Ito's integral for finite dimensional integrand process. One must be able to extend this rule to more complex spaces in order to conclude that, in fact, equation 40. is the Stratonovich' version of equation 8.ii).

Now, let us write a weak form for equation 40. equivalent to the equation 9.ii) which is a weak form for 8.ii). It has the Stratonovich differential form,

$$41. \quad d(\tilde{u}_n(t), v) + \{a_o(t; \tilde{u}_n(t), v) - \frac{1}{2} (A_1^2(t) \tilde{u}_n(t), v)\} dt + a_1(t; \tilde{u}_n(t), v) dw_s, \qquad \text{for all } v \in V_n, \quad t \in [O,T] ; (S).$$

where we have written $\tilde{u}_n$ instead of the conventional u_n, to underline the fact that equation 41. above and equation 38. are in general, two different objects. (However, if the subspace V_n is invariant for the operator $A_1(t)$, equations 38. and 41. are equivalent).

Using the same technique used before , one must be able to prove that the sequence of solutions for equation 41. has a weakly convergent sequence which converges to the solution of equation 40.

Therefore we may say that the Stratonovich and Ito's versions of the original evolution equation 8.ii) produce two different weak forms, both convergent.

5.3 - The Non Linear Filtering Problem

In this paragraph we return to the filtering problem introduced in section 1. We shall use the results derived in the previous paragraph in order to produce a existence and uniqueness result for the stochastic parabolic equation 1.1.21. which represents the solution of the filtering problem for partially observed diffusion process.

Let S be an open domain in R^n and take $H = L^2(S)$, $V = H_o'(S)$.

Using the notation presented in paragraph 1.1., denote, $a_o(t)$, $t \in [0,T]$, the bilinear form on $H_o^1(s)$ defined by the following relation:

1. $$a_o(t; u, v) = \frac{1}{2} \sum_{i,j=1}^{n} \int_S a_{j,i}(t,x)\, D_j\, u(x)\, D_i\, v(x)\, dx \; +$$

$$+ \sum_{i=1}^{n} \int_S D_i((-\frac{1}{2} \sum_{j=1} (D_j\, a_{j,i}(t,x) + g_i(t,x)\, u)\, v dx,$$

$$u, v \in H_o^1(S)$$
$$t \in [O,T],$$

We recall that,

$$[a_{i,j}(t,x)] = \alpha(t,x)\ \alpha^T(t,x),$$

is the diffusion matrix and $[g_i]$ is the drift vector for the diffusion 1.1.2..

Let us suppose that for i,j = 1,...,n, the functions,

2. $$a_{i,j}\ ,\ D_j\, a_{i,j}\ ,\ D_{i,j}\, a_{i,j}\ ,\ g_i\ ,\ D_i\, g_j,$$

are elements of the space $C(O,T;\ L^{\infty}(S))$

Using a standard argument (see Remark 2.3.3) we can deduce the linear operator $A_o(t)$, $t \in [O,T]$, associated with the bilinear form $a_o(t)$. We have,

3. $$A_o(t) = -L_t,$$

where L_t denotes the Fokker-Planck operator introduced in 1.1.9..

Define the bilinear form $a_1(t)$, $t \in [O,T]$ by,

4. $$a_1(t;\ u,\ v) = (A_1(t),\ v),$$

$$u, v \in H_o^1(S)$$
$$t \in [O,T],$$

where $A_1(t) = -H_t$ and H_t is the first order differential operator introduced in 1.1.20.

We recall that,

5. $$H_t u = - \sum_{i=1}^{n} \frac{\delta}{\delta x_i} (b_i(t,x)\ u(x)) + h(t,x)\ u(x).$$

Here $[b_i(t,x)] = \alpha(t,x) . \beta^1(t)$ and the functions h, β^1 are parameters of the observation process.

Let us assume that for i = 1,...,n the functions

6. $$b_i,\ D_i b_i,\ h,$$

are elements of the space $C(O,T;\ L^\infty(S))$.

It is very easy to show that under hypotheses 2. and 6. the bilinear forms $a_o(t)$, $a_1(t)$ verify assumptions 5.2.2, 5.2.3 and 5.2.4. In order to have also here the coercivity condition 5.2.6 we assume that for some constant $\sigma > O$

7. $$<r,([a_{i,j}] - [b_i][b_i]^T)r> \ \geq \ \sigma<r,r>,$$

$$\text{for all } r \in R^n$$
$$(t,x) \in |O,T| \times S,$$

where $<.,.>$ denotes the scalar product in R^n,

$[a_{i,j}] = [a_{i,j}(t,x)]$ and $[b_i] = [b_i(t,x)]$.

Consider the observation process introduced in 1.1.1 plus 1.1.18 and 1.1.19. Let $\mathcal{F}_t = \sigma(y(s) : O \leq s \leq t)$.

Consider the following stochastic evolution equation:

8. $$du(t) + A_o(t)dt + A_1(t)d\tilde{w}_t = 0,$$

where $\tilde{w}_t$ is a real-valued, $\mathcal{F}_t$-measurable, non-antecipative, standard Wiener process on the probability space $(\Omega, \mathcal{A}, \tilde{P})$.

According to Theorem 5.2.1, equation 8. has a unique solution u,

$$u \in \tilde{M}^2(0,T;H_o^1(S)) \cap \tilde{L}^2(\Omega;C(0,T;L^2(S)),$$

satisfying $u(0) = u_o \in H_o^1(S)$. (Here, the symbol ~ is used to indicate the dependence with respect to the probability $\tilde{P}$).

It can be shown (see e.g. Pardoux |41|) that under the transformation of probability measure indicated in 1.1.5, the observation process, y(t), becomes a real-valued, $\mathcal{F}_t$-measurable non-anticipative standard Wiener process on $(\Omega, \mathcal{A}, \tilde{P})$. Therefore, equation 8. is equivalent to equation 1.1.21 and so, we have proved the following result:

<u>Theorem 5.3.1</u> - Under hypotheses, 2., 6. and 7. equation 1.1.21 has a unique solution

$$q \in M^2(0,T;H_o^1(S)) \cap L^2(\Omega;C(0,T;L^2(S))$$

satisfying $q(0) = q_o \in H_o^1(S)$.

Here q_o is the density of the law of X_o (see 1.1.3)

Selecting $S = R^n$, the result above enables us to derive a existence and uniqueness result for the filtering problem for partially observed diffusions in R^n. As we mentioned in sections 3. and 4., the assumption $V = H_0^1(S)$, S an open set of R^n, corresponds to the filtering problem for diffusions absorbed by the boundary of S. Selecting $V = H^1(S)$, we shall be able to analyse the case where the diffusion is reflected in an inelastic boundary. (see Pardoux |40|, for both situations). In particular the case $S = R^n$, diffusions in R^n, has been analysed also by Krilov-Rosovskii (|22|) and Levieux (|28|).

<u>Remark 5.3.1</u> - We remark that the coercivity condition 7 is achieved automatically if we assume that for all $(t,x) \in [0,T] \times S$, $r \in R^n$ there exist constant $\sigma > 0$ and $\varepsilon \in (0,1)$ such that,

9. i) $\langle r, |a_{i,j}| r\rangle > \sigma . \varepsilon^{-1} \langle r,r\rangle$

ii) $\langle \beta^1, \beta^1\rangle \leq 1 - \varepsilon$.

In fact, under these conditions we can write,

$$\langle \alpha^T r, \beta^1 (\beta^1)^T \alpha^T r\rangle = (\langle \alpha^T r, \beta^1\rangle)^2 \leq$$

$$\leq \langle \alpha^T r, \alpha^T r\rangle \quad (1 - \varepsilon).$$

Rearranging terms,

$$\varepsilon < r,[a_{i,j}]r> \leq <r, ([a_{i,j}] - [b_i] [b_i]^T)r>,$$

and so, the coercivity condition 7. holds.

We also observe that, recalling hypothesis 1.1.19, condition 9.ii) above is equivalent to the following:

10. $$(\beta^2(t))^2 \geq \varepsilon, \qquad t \in [0,T].$$

Therefore, as the coercivity condition is a crucial assumption in the proof of Theorem 5.3.1, we conclude that condition 10. is an equally crucial condition to the solution of the filtering problem. It means that in the observation process, the proportion of the noise independent of the signal must be positive. (See Pardoux |41| for an extended analysis on this subject). ●

<u>Remark 5.3.2</u> - With respect to the regularity of the solution of equation 1.1.21 one can show that, similarly to what happens for non-stochastic partial differential equations, this regularity depends on how regular are the coefficients and the initial condition associated with the equation.

In Pardoux (|41|) (and also in Krilov-Rosovskii (|21|)) regularity results are presented for the solution of the Cauchy problem for the evolution equation 1.1.21 (i.e. for $S = R^n$ in Theorem 5.3.1). It turns out that, if the functions described in 2. and 6. have bounded partial derivatives (w.r.t. $x \in R^n$) up to order $p \geq 1$ and if $q_0 \in H^p(R^n)$, then equation 1.1.21 admits a unique solution,

$$q \in M^2(O,T;H^{p+1}(R^n)) \cap L^2(O,T;C(O,T;H^p(R^n))).$$

(Theorem 2.1 in Pardoux |41|) .

For the case $S \subset R^n$, similar results can be derived if the boundary of the domain S is sufficiently "smooth". Here, we register a result presented by Pardoux (|40|) where a stochastic equation of the form 1.1.14 (the Zakai equation) is analysed.

Let the boundary of S be of Class C^2.

Take $\beta^1 = O$ in 1.1.21. (In other words, consider equation 1.1.14) If, in addition to hypotheses 2. and 6. we have, for i,j = 1,...,n,

$$a_{i,j} \in C^1([O,T];L^\infty(S))$$

$$D_i h \in C((O,T);L^\infty(S)),$$

then for $q_o \in H_o^1(S)$ the solution q of equation 1.1.21 satisfies

$$q \in M^2(O,T;H^2(S)) \cap L^2(\Omega;C(O,T;H_o^1(S))).$$

(Theorem 2.3 in Pardoux |40|)

Remark 5.3.3 - Consider the case $\beta^1 = O$ in equation 1.1.21. In other words, we are assuming independence between the noise in the observation process and the signal and, in this case, equation 1.1.21 is identical to the Zakai formula 1.1.14. But this equation admits a non-stochastic counterpart, i.e., equation 1.1.16. Therefore, an existence and uniqueness result for equation 1.1.14 can be obtained by means of the results presented in section 3. for (non-stochastic) evolution equations. In particular, if we also assume the function h to be invariant in time, Theorem 3.4.1 and Theorem 5.3.1 are equivalent, (in the sense that both represent an existence and uniqueness result for the Zakai formula).

The concept of the non-stochastic counterpart offers other interesting aspects for investigation. Consider the finite dimensional stochastic equation that constitutes a Galerkin approximation to equation 8. It has the form of equation 5.2.16 but with $w_t = \tilde{w}_t = y(t)$. In addition to the hypotheses made in this paragraph assume, $\beta^1 = 0$ and h, invariant in time. In 5.2.16. these assumptions mean that $R(t) = R = R^T$. A non-stochastic counterpart of 5.2.16 can be obtained using the procedure presented by Doss (|11|). We first write the following equation in V_n:

11. $$\left(\frac{d}{dt} V(t,v_o^n),v\right) = (P_n(hV(t,v_o^n)),v)$$

for all $v \in V_n$

$t \in [0,T]$,

where P_n is the projection on V_n and $V(0,v_o^n) = v_o^n \in V_n$.

Therefore, $[V(t,v_o)] = F(t).[v_o] = \exp(-M^{-1}R)[v_o]$ and a pathwise solution for 5.2.16 has the form,

12. $$\frac{d}{dt}[r_n(t)] + F^{-1}(y(t))\mathcal{M}(t)F(y(t))[r_n(t)] = 0,$$

where $M(t) = M^{-1}K^T(t) + \frac{1}{2}(M^{-1}R)^2$. The relation between 5.2.16 and 12. is given by

13. $u_n(t) = V(y(t), r_n(t))$.

We observe that equation 11. is a Galerkin approximation to equation 1.1.15 (in the sense that they tend to describe the same object as $n \to \infty$). On the other hand, one must be able to prove that the solution of 12. converges to the solution of the pathwise formula 1.1.16. Therefore, equation 12. represents a Galerkin approximation to the pathwise formula 1.1.16. (However, this Galerkin approximation is different from the one obtained when we start with 1.1.16.. So, we have here the same situation as in Remark 5.2.4: equation 1.1.14 and its non-stochastic version 1.1.16 produce two different weak forms both convergent).

6 - GALERKIN APPROXIMATIONS TO STOCHASTIC EVOLUTION EQUATIONS

The objective in this section is to present two families of discrete time Galerkin schemes in order to approximate the solution of stochastic evolution equations. These families are characterized by having terms which are respectively linear and quadratic in the noise increment. With respect to the time increment the schemes in both families are implicit Runge-Kutta of the variety studied in section 4. and, therefore, the methodology used here follows the same pattern as before.

In paragraph 6.1 we introduce a family of linear schemes. Consistency of the numerical method is studied in paragraph 6.2 and in paragraph 6.3 an estimate for the error of approximation is presented. It turns out that if sufficient regularity is attained by the solution of the stochastic evolution equation, the method has a non linear rate of convergence in relation to the discretization in time. In paragraph 6.4 we study a family of quadratic schemes. In this case if stronger regularity conditions hold, the method admits a linear rate of convergence in the time increment. Finally in paragraph 6.4, we bring into consideration the filtering problem for diffusion process.

6.1 - A Numerical Scheme

Basically, we assume the hypotheses of section 5.

So, V and H are Hilbert spaces, V is dense in H and its injection is continuous according to 5.2.1. The symbols $((\cdot,\cdot))$ and $\|\cdot\|$, $(\cdot,\cdot)$ and $|\cdot|$, denote the inner product and norm in V and H respectively.

The objects $a_j(t)$, $j = 0,1$ are bilinear functionals defined in the space V, satisfying hypotheses 5.2.2, 5.2.3, 5.2.4 and 5.2.6, the latter taken with $\lambda = 0$ for reasons given in Remark 5.2.1.

Here, we strengthen hypothesis 5.2.4, by assuming

1. $|a_1(t;u,v)| \leq \gamma_1 |u||v|,$

$$u,v \in V$$
$$t \in [0,T].$$

In other words, the operator $A_1(t)$ introduced in 5.2.5 is now an element of $L(H,H)$.

We also make the following additional hypothesis:

2. $a_j(.,u,v) \in C^1(0,T),$

$$\text{for all } u,v \in V$$
$$j = 0,1$$

Let $\mathcal{V}$ be a finite dimensional subspace of V.

For all $t \in [0,T]$, let $L_j(t)$, $j = 0,1$, be linear operators from $\mathcal{V}$ to $\mathcal{V}$ defined by the following relations:

3. $a_j(t;u,v) = (L_j(t)u,v),$

$$\text{for all } u,v \in \mathcal{V},$$
$$t \in [0,T],$$
$$j = 0,1.$$

Since $\mathcal{V}$ is a finite dimensional subspace, these are well defined continuous linear operators. In particular, by hypothesis 1., we have

$$|L_1(t)u| \leq \gamma_1 |u|, \qquad u \in \mathcal{V}, \quad t \in [0,T].$$

or, equivalently,

4. $$|||L_1(t)||| \leq \gamma_1 \qquad t \in [0,T]$$

independently of the subspace $\mathcal{V}$. Here, the symbol $|||\cdot|||$ stands for the natural norm of $L(\mathcal{V},\mathcal{V})$ when $\mathcal{V}$ is endowed with the $|\cdot|$ norm.

The coercivity condition 5.2.6 (with $\lambda = 0$, see Remark 5.2.1) implies that the operator $L_o(t)$ is invertible and so are the operators of the form $(I + kL_o(t))$ where I is the identity operator and $k \geq 0$. Also, by the continuity of the injection $V \subset H$, the following estimate holds:

5. $$|||(I + kL_o(t)^{-1}||| \leq (1 + k\sigma)^{-1}, \qquad t \in [0,T].$$

Now, let $\{0 = t_o < t_1 < \ldots < t_N = T\}$ be a partition of the interval $[0,T]$ with mesh,

6. $$h = \sup\{|t_{k+1} - t_k| : k = 0,1,\ldots,N-1\}.$$

With respect to this partition, we shall use the same set of notation for increments introduced in 4.1.3.

We shall now present a discrete time stochastic scheme for approximating the stochastic evolution equation 5.2.8.ii).

So, let w_t be the real valued, $\mathcal{F}_t$-measurable non-antecipative standard Wiener process on the probability space $(\Omega, \mathcal{A}, P)$ introduced in paragraph 5.2 and consider the following stochastic scheme:

7. $$U_{k+1} - U_k + \Delta_k \mathcal{G}_k^o U_k + \Delta w_k \mathcal{G}_k^1 U_k = 0,$$

$$k = 0,1,\ldots,N-1,$$

where $U_k \in \mathcal{V}$ and $\mathcal{G}_k^j \in L(\mathcal{V},\mathcal{V})$, $j = 0,1$ are linear operators defined by the following relations:

8. i) $$\mathcal{G}_k^o = (I + \Delta_k \rho L_o(\tau))^{-1} L_o(\tau),$$

ii) $$\mathcal{G}_k^1 = (I + \Delta_k \rho L_o(\tau))^{-1} L_1(t_k),$$

$$k = 0,1,\ldots,N-1$$

with $\rho > 0$ and $\tau = \tau_k \in [t_k, t_{k+1}]$.

Concerning the operators $\mathcal{G}_k^j$, $j = 0,1$; $k = 0,1,\ldots,N-1$ the following Proposition can be stated:

<u>Proposition 6.1.1</u> - Under the hypotheses above the following estimates hold independently of the subspace $\mathcal{V}$:

9. i) $$|||I - \Delta_k \mathcal{G}_k^o||| \leq 1 \quad \text{for} \quad \rho \geq 0.5$$

and, in particular, if $\rho > 0.5$, there exist constants $\delta, h_o > 0$ such that:

$$|||I - \Delta_k \mathcal{G}_k^o||| \leq \exp(-\delta\Delta_k),$$

for all partitions of the interval $[0,T]$ with $h \leq h_o$

ii) $|||\mathcal{G}_k^1||| \leq \gamma_1$, $\quad k = 0,1,\ldots,N-1$.

Proof of Proposition 1

The first part is identical to the thesis of the Proposition 4.1.1 and so, is already proven. The second part follows from inequalities 4. and 5. ●

So, from the above proposition we can affirm that, given an initial condition $U_o \in \mathcal{V}$, the set of iterative equations 7. uniquely defines a sequence U_k, $k = 0,\ldots,N$ of $\mathcal{V}$-valued $\mathcal{F}_k \equiv \mathcal{F}(t_k)$-measurable, random variables.

We can also, as we did in paragraph 4.2, explore some of the stability properties of the scheme 7. In particular, we observe that the expectation of the variables U_k satisfy a scheme identical to the one analysed in section 4. In fact, we can write from equation 7.,

10. $$EU_{k+1} = (I - \Delta_k \mathcal{G}_k^o) EU_k$$

which is identical to equation 4.2.3 and therefore has the same properties regarding stability.

Now, let $R(t)$, $t \in [0,T]$ be the Ritz projection with respect to the bilinear form $a_o(t)$ and the subspace $\mathcal{V}$.

Recalling the definition given in 4.2.7, we can write

11. $$a_o(t;u - R(t)u,v) = 0,$$

for all $u \in V$, $v \in \mathcal{V}$, $t \in [0,T]$.

The coercivity condition imposed on the bilinear form $a_o(t)$ guarantees the existence and uniqueness of such an operator.

The purpose of this section is the estimation of the error of approximating the solution of the stochastic evolution equation 5.8.ii) by means of the set of random variables defined by equation 7. So, in what follows, the object of our attention will be the random variable

$$U_k - u(t_k),$$

$$k = 0,1,\ldots,N,$$

where by, u, we denote the solution of the Stochastic Evolution Problem 5.2.8..

Using the definition 11. above we can write,

12. $$U_k - u(t_k) = e_k + \tilde{R}(t_k)u(t_k),$$

$$k = 0,1,\ldots,N.$$

Here the random variable e_k and the linear operator $\tilde{R}(t)$ are defined by the following relations:

13. $$e_k = U_k - R(t_k)u(t_k)$$

14. $\tilde{R}(t) = I - R(t).$

Now, define the sequence ϕ_k, $k = 1,\ldots,N$ of $\mathcal{U}$-valued, $\mathcal{F}_k$-measurable, random variables by the following relation:

15. $$\phi_{k+1} = R(t_{k+1})\,u(t_{k+1}) - R(t_k)u(t_k) + \Delta_k \mathcal{G}_k^o R(t_k)u(t_k) + \Delta w_k \mathcal{G}_k^1 R(t_k)u(t_k),$$

$$k = 0,1,\ldots,N-1.$$

Subtracting equation 7. from the above, using equation 13. and rearranging terms, we have,

16. $$e_{k+1} - e_k + \Delta_k \mathcal{G}_k^o e_k + \Delta w_k \mathcal{G}_k^1 e_k + \phi_{k+1} = 0,$$

$$k = 0,1,\ldots,N-1.$$

Here, as in paragraph 4.2, the error of the approximation is determined by the variable ϕ_k. So, extending the concept of consistency of a numerical method to this case, we can say that ϕ_k measures the consistency of the method of approximating the solution of the evolution equation 5.8.ii) by means of the scheme 7.

Remark 6.1.1 - The discrete time stochastic scheme 7 can be written in other forms which are, perhaps, more familiar to the reader. So, it can be presented in a "stage" form,

$$U_{k+1} - U_k + \Delta_k \beta_o + \Delta w_k \beta_1 = 0,$$

$$k = 0,1,\ldots,N-1,$$

where $\beta_j \in \mathcal{V}$, $j = 0,1$, are such that:

$$(\beta_o, v) + \Delta_k a_o(\tau; \beta_o, v) + a_o(\tau; U_k, v) = 0,$$

$$(\beta_1, v) + \Delta_k a_o(\tau; \beta_1, v) + a_1(t_k; U_k, v) = 0,$$

$$\text{for all } v \in \mathcal{V}.$$

Alternatively,

$$(U_{k+1} - U_k, v) + \Delta_k a_o(\tau; \rho U_{k+1} + (1-\rho) U_k, v) +$$

$$+ \Delta w_k a_1(t_k; U_k, v) = 0,$$

$$\text{for all } v \in \mathcal{V},$$
$$k = 0,1,\ldots,N-1.$$

We observe that scheme 7. differ from the implicit Runge-Kutta scheme analysed in section 4 only by the term containing the increment in the noise.

Basically, a numerical scheme appropriate to give approximations to the finite dimensional stochastic equation that governs the continuous time Galerkin approximation (equation 5.2.16) can be used in order to produce discrete time Galerkin schemes. For instance, if we take $\rho = 0$ in equation 7., we have the so called Cauchy-Maruyama scheme (McShane, |36|). However, as we pointed out before (section 4) this particular explicit scheme is not appropriate for Galerkin approximations and that is the reason why we assume the parameter ρ to be positive. So, the scheme presented in this paragraph is the natural and simplest extension of the first order Runge-Kutta scheme introduced in section 4.

6.2 - Consistency Properties of the Method

In this paragraph we shall evaluate the consistency of the approximation method proposed in the last paragraph.

Two proposition will be presented with estimates for the random variables ϕ_{k+1} and $E(\phi_{k+1}/\mathfrak{F}_k)$.

We start by considering the equation 6.1.15. Using the definitions of the elements involved it can be rewritten in the following form:

1. $$(\phi_{k+1},v) + \Delta_k \rho a_o(\tau;\phi_{k+1},v) = (\Delta u_k,v) +$$

$$+ \Delta_k a_o(\tau;\rho u(t_{k+1}) + (1-\rho)u(t_k),v) +$$

$$+ \Delta w_k a_1(t_k;u(t_k),v) - (\Delta \tilde{R}u_k,v) +$$

$$- \Delta_k a_o(\tau;\rho\tilde{R}u(t_{k+1}) + (1-\rho)\tilde{R}u(t_k),v) +$$

$$- \Delta w_k a_1(t_k;\tilde{R}u(t_k),v),$$

for all $v \in \mathcal{V}$
$k = 0,1,\ldots,N-1.$

Here, according to 6.1.14, we write $\tilde{R}u(t) = \tilde{R}(t)u(t)$, $t \in [0,T]$.

As u is the solution of the problem 5.2.8, we have,

2. $$(\Delta u_k, v) + \int_{t_k}^{t_{k+1}} a_o(s;u(s),v)ds + \int_{t_k}^{t_{k+1}} a_1(s;u(s),v)dw_s = 0,$$

for all $v \in \mathcal{V}$
$k = 0,1,\ldots,N-1$ w.p.1.

Substituting this identity in expression 1. and rearranging terms, we have,

3. $$(\phi_{k+1},v) + \Delta_k \rho a_o(\tau;\phi_{k+1},v) =$$

$$= \int_{t_k}^{t_{k+1}} a_o(\tau;\rho u(t_{k+1}) + (1-\rho)u(t_k),v) + - a_o(s;u(s),v)ds + \int_{t_k}^{t_{k+1}} a_1(t_k;u(t_k),v) + - a_1(s;u(s),v)dw_s - (\tilde{R}\Delta u_k,v) - \Delta_k a_o(\tau;\rho\tilde{R}u(t_{k+1}) + + (1-\rho)\tilde{R}u(t_k),v) - \Delta w_k a_1(t_k;\tilde{R}u(t_k),v),$$

for all $v \in \mathcal{V}$
$k = 0,1,\ldots,N-1.$

Now, for simplicity, let us strengthen to some extent our hypotheses by supposing the bilinear form $a_o(t)$ is invariant in time,

4. $$a_o(t) = a_o.$$

Remark 6.2.1 - Although our conclusions will be obtained under the above condition, it does not constitute a fundamental hypothesis like those presented in the beginning of this section. If $a_o(t)$ is sufficiently 'smooth' in relation to the variable time, similar results can be obtained. ●

From condition 4., the Ritz projection is also invariant in time and we are able to write,

5. $$a_o(\tilde{R}u,v) = 0$$

for all $u \in V$, $v \in \mathcal{V}$

On the other hand, hypotheses 6.1.1 and 6.1.2 enable us to define the operator A'(t) such that

6. i) $$A_1'(t) = \frac{d}{dt} A_1(t) \in L(H,H) \qquad t \in [0,T]$$

ii) $$|A_1'(t)u| \leq \gamma_1'|u|$$

for all $u \in H$
$t \in [0,T]$

for some constant γ_1'.

So, the following identity can be written:

7. $$\int_{t_k}^{t_{k+1}} a_1(t_k;u(t_k),v) - a_1(s;u(s),v)dw_s =$$

$$= \int_{t_k}^{t_{k+1}} (A_1(t_k)u(t_k) - A_1(s)u(s),v)dw_s =$$

$$= - \int_{t_k}^{t_{k+1}} (A_1(s)\Delta u_k(s),v)dw_s +$$

$$- \int_{t_k}^{t_{k+1}} (\int_{t_k}^{s} A_1'(\xi)u(t_k)d\xi,v)dw_s ,$$

for all $v \in \mathcal{V}$
$k = 0,1,\ldots,N-1$.

Taking 4., 5., and 7. into account and rearranging terms, equation 3. now becomes

8. $$(\phi_{k+1},v) + \Delta_k\rho a_o(\phi_{k+1},v) = \Delta_k\rho a_o(\Delta u_k,v) +$$

$$- \int_{t_k}^{t_{k+1}} a_o(\Delta u_k(s),v)ds - (\int_{t_k}^{t_{k+1}} A_1(s)\Delta u_k(s)dw_s,v) +$$

$$- (\int_{t_k}^{t_{k+1}} (\int_{t_k}^{s} A_1'(\xi)u(t_k)d\xi)dw_s,v) +$$

$$- (\tilde{R}\Delta u_k,v) - \Delta w_k a_1(t_k;\tilde{R}u(t_k),v),$$

for all $v \in \mathcal{V}$
$k = 0,1,\ldots,N-1$ wpl.

Now, choose $v = \phi_{k+1}$ as a test vector in the above equation. Using hypotheses 5.2.3, 5.2.6 (with $\lambda = 0$), 6.1.1 and the Schwartz' inequality, equation 8. yield the following inequality:

9. $$|\phi_{k+1}|^2 + \Delta_k \rho\sigma \|\phi_{k+1}\|^2 \leq \Delta_k \rho \gamma_0 \|\Delta u_k\| \, \|\phi_{k+1}\| + \int_{t_k}^{t_{k+1}} \gamma_0 \|\Delta u_k(s)\| \, \|\phi_{k+1}\| \, ds + \left| \int_{t_k}^{t_{k+1}} A_1(s) \Delta u_k(s) \, dw_s \right| |\phi_{k+1}| + \left| \int_{t_k}^{t_{k+1}} \left(\int_{t_k}^{s} A_1'(\xi) u(t_k) d\xi \right) dw_s \right| |\phi_{k+1}| + |\tilde{R}\Delta u_k| |\phi_{k+1}| + \gamma_1 |\Delta w_k| |\tilde{R} u(t_k)| |\phi_{k+1}|,$$

$k = 0,1,\ldots,N-1,$ wpl.

Using Cauchy's inequality, $pq \leq p^2/2\varepsilon + \varepsilon q^2/2$ with $\varepsilon = 2\rho\sigma/\gamma_0(\rho+1)$ for the first and the second terms of the right side and with $\varepsilon = 1/4$ for the remaining terms, we obtain after standard manipulation, the following inequality:

10. $$\frac{1}{2}|\phi_{k+1}|^2 \leq \Delta_k \frac{\gamma_o^2(\rho+1)}{4\rho\sigma} \|\Delta u_k\|^2 + $$

$$+ \frac{\gamma_o^2(\rho+1)}{4\rho\sigma} \int_{t_k}^{t_{k+1}} \|\Delta u_k(s)\|^2 \, ds + $$

$$+ 2 \left| \int_{t_k}^{t_{k+1}} A_1(s)\Delta u_k(s)\, dw_s \right|^2 + $$

$$+ 2 \left| \int_{t_k}^{t_{k+1}} \left(\int_{t_k}^{s} A_1'(\xi) u(t_k) d\xi \right) dw_s \right. + $$

$$+ 2|\tilde{R}\Delta u_k|^2 + 2\gamma_1^2 (\Delta w_k)^2 \; |\tilde{R}u(t_k)|^2,$$

$$k = 0,1,\ldots,N-1 .$$

Taking the expectation on both sides of this inequality, we can write,

11. $$E|\phi_{k+1}|^2 \leq \Delta_k \frac{\gamma_o^2(\rho+1)}{2\rho\sigma} E\|\Delta u_k\|^2$$

$$+ \frac{\gamma_o^2(\rho+1)}{2\rho\sigma} E \int_{t_k}^{t_{k+1}} \|\Delta u_k(s)\|^2 \, ds + $$

$$+ 4 \int_{t_k}^{t_{k+1}} E|A_1(s)\Delta u_k(s)|^2 \, ds + $$

(equation 11. - continuation)

$$+ \; 4 \int_{t_k}^{t_{k+1}} E \left| \int_{t_k}^{s} A_1'(\xi) u(t_k) d\xi \right|^2 ds + 4E|\tilde{R}\Delta u_k|^2 \; +$$

$$+ \; 4\gamma_1^2 E\left(E(\Delta^2 w_k / \mathcal{F}_k)|\tilde{R}u(t_k)|^2\right),$$

$$k = 0,1,\ldots,N-1.$$

Using the estimates 6.1.1 and 6.ii) we can finally write the following inequality:

12. $$E|\phi_{k+1}|^2 \leq \Delta_k \frac{\gamma_o^2(\rho + 1)}{2\rho\sigma} E\|\Delta u_k\|^2 \; +$$

$$+ \frac{\gamma_o^2(\rho + 1)}{2\rho\sigma} \int_{t_k}^{t_{k+1}} E\|\Delta u_k(s)\|^2 ds \; +$$

$$+ \; 4\gamma_1^2 \int_{t_k}^{t_{k+1}} E|\Delta u_k(s)|^2 ds + (\Delta_k)^3 \, 4(\gamma_1')^2 \, E|u(t_k)|^2 \; +$$

$$+ \; 4E|\tilde{R}\Delta u_k|^2 + \Delta_k 4\gamma_1^2 \, E|\tilde{R}u(t_k)|^2,$$

$$k = 0,1,\ldots,N-1, \qquad \text{wpl.}$$

We state this result in the following,

<u>Proposition 6.2.1</u> - Under hypotheses 5.2.3, 5.2.6, 6.1.1, 6.1.2 and 6.2.4 the following estimates holds:

$$13. \quad E|\phi_{k+1}|^2 \leq C\left\{\Delta_k E\|\Delta u_k\|^2 + \int_{t_k}^{t_{k+1}} E\|\Delta u_k(s)\|^2\,ds + (\Delta_k)^3 E|u(t_k)|^2 + E|\tilde{R}\Delta u_k|^2 + \Delta_k E|\tilde{R}u(t_k)|^2\right\},$$

$$k = 0,1,\ldots,N-1.$$

Here C is a positive constant depending only on the parameters ρ, γ_o, γ_1, γ_1' and σ.

Remark 6.2.2 - The inequality 12. shows that $\phi_k \in L^2(\Omega,H)$ for all $k = 1,\ldots,N$ since, by the estimates presented in paragraph 5.2, its right side is finite. Moreover, we shall have $\phi_k \to 0$ in $L^2(\Omega,H)$ as $N \to \infty$. ●

Remark 6.2.3 - In the steps leading to the estimate 12. we have used, implicitly, some of the standard properties of stochastic integrals (in Ito's sense) and Wiener processes which are registered in paragraph 5.1. ●

The result presented in Proposition 6.2.1 enables us to estimate the random variable $E(\phi_{k+1}/\mathcal{F}_k)$, $k = 0,1,\ldots,N-1$ by means of the inequality,

$$14. \quad E(|E(\phi_{k+1}/\mathcal{F}_k)|^2) \leq E|\phi_{k+1}|^2.$$

However, for the purposes we have in mind, the above estimate is not accurate enough. So, we shall now prove the following proposition:

Proposition 6.2.2 - Under the hypotheses of Proposition 6.2.1, the following estimate holds:

15. $$E(|E(\phi_{k+1}/\mathcal{F}_k)|^2) \le C(\Delta_k)^2 \int_{t_k}^{t_{k+1}} E\|\frac{d}{ds}\theta(s;t_k,u(t_k))\|^2 ds + \Delta_k \int_{t_k}^{t_{k+1}} E|\tilde{R}\frac{d}{ds}\theta(s;t_k,u(t_k))|^2 ds,$$

$$k = 0,1,\ldots,N-1.$$

Here C is a positive constant depending only on the parameters ρ, γ_0, σ; $\theta(.;\hat{t},z)$ is a V-valued function defined in $|\hat{t},T|$ and related to the parameters $\hat{t} \in [0,T]$, $z \in V$ by the following initial valued evolution equation:

16. i) $$\frac{d}{dt}\theta(t;\hat{t},z) + A_o\theta(t;\hat{t},z) = 0,$$

ii) $$\theta(\hat{t};\hat{t},z) = z \in D(A_o),$$

Remark 6.2.4 - The result in Proposition 6.2.2 is established by the fact that equation 16. has a unique solution. Although we are not allowed to use the results of section 3. in order to show existence of a solution, (because here we are not supposing the bilinear form a_o with a symmetric principal part), the existence of such a solution can be shown by means of the techniques introduced in that section. Here, we shall not present this proof. Instead, we will make use of a similar result presented in Lions, |31|.

Consider the evolution equation,

17. $$\frac{d}{dt} Z(t;Z_o) + A_o Z(t;Z_o) = 0$$

$$Z(0,Z_o) = Z_o \in H.$$

It can be shown (Lions |31|, Theorem 1.2, p. 102) that the equation above has a unique solution

$$Z(.,Z_o) \in L^2(0,T;V) \cap C(0,T;H).$$

Also, we can write,

18. $$\frac{d}{dt} Z(t;Z_o) = Z(t;-A_o Z_o), \qquad t \in [0,T],$$

for all $Z_o \in D(A_o)$.

Therefore, using this argument in relation to equation 16. we can conclude that,

$$\theta(.;\hat{t},z) \text{ and } \frac{d}{dt}\theta(.;\hat{t},z),$$

are elements of $L^2(\hat{t},T;V) \cap C(\hat{t},T;H)$ for all $z \in V$ such that $z \in D(A_o)$.

(or similar results when A_o depends on time, see Lions |30|, chapter V) ●

Proof of Proposition 6.2.2

Let $\hat{\phi}(z_k)$, $z_k \in D(A_o) \subset V$, $k = 0,1,\ldots,N-1$ be a family of elements belonging to the subspace $\mathcal{V}$, defined by,

$$\text{19.} \qquad \hat{\phi}(z_k) = R\theta(t_{k+1};t_k,z_k) - Rz_k + \Delta_k \mathcal{G}_k^o z_k, \qquad k = 0,1,\ldots,N-1.$$

For simplicity, in the steps hereafter we will delete the argument z by writing

$$\hat{\phi} = \hat{\phi}(z_k), \quad \theta(t) = \theta(t;t_k,z_k).$$

Recalling the definition of the elements involved, equation 17. can be rewritten in the form

$$(\hat{\phi},v) + \Delta_k \rho a_o(\hat{\phi},v) = (\Delta\theta_k,v) + \Delta_k a_o(\rho\theta(t_{k+1}) + (1-\rho)z_k,v) - (\tilde{R}\Delta\theta_k,v),$$

$$\text{for all } v \in \mathcal{V}, \quad k = 0,1,\ldots,N-1.$$

Using equation 16. to evaluate the increment $\Delta\theta_k$ and substituting in the above equation we have after rearranging terms, the following identity:

$$\text{20.} \qquad (\hat{\phi},v) + \Delta_k \rho a_o(\hat{\phi},v) = \Delta_k \rho a_o(\Delta\theta_k,v) + - \int_{t_k}^{t_{k+1}} a_o(\Delta\theta_k(s),v)\,ds - (\tilde{R}\Delta\theta_k,v),$$

$$\text{for all } v \in \mathcal{V}, \quad k = 0,1,\ldots,N-1.$$

Take $v = \hat{\phi}$ as a test vector. Using hypotheses 5.2.3 and 5.2.6 (with $\lambda = 0$) jointly with Schwartz' inequality, equation 20 yields,

21. $$|\hat{\phi}|^2 + \Delta_k \rho\sigma \|\hat{\phi}\|^2 \leq \Delta_k \rho \gamma_o \|\Delta\theta_k\| \, \|\hat{\phi}\| + \int_{t_k}^{t_{k+1}} \gamma_o \|\Delta\theta_k(s)\| \, \|\hat{\phi}\| \, ds + |\tilde{R}\Delta\theta_k| \, |\hat{\phi}|,$$

$$k = 0,1,\ldots,N-1.$$

Apply Cauchy's inequality $pq \leq 0.5\, p^2/\varepsilon + 0.5\varepsilon q^2$ with $\varepsilon = \sigma/\gamma_o$, $\varepsilon = \sigma\rho/\gamma_o$ and $\varepsilon = 1$, respectively, for the terms in the right side of the above equation. After some manipulation we have,

22. $$\frac{1}{2}|\hat{\phi}|^2 \leq \Delta_k \frac{\rho\gamma_o^2}{2\sigma} \|\Delta\theta_k\|^2 + \frac{\gamma_o^2}{2\sigma\rho} \int_{t_k}^{t_{k+1}} \|\Delta\theta_k(s)\|^2 \, ds + \frac{1}{2}|\tilde{R}\Delta\theta_k|^2,$$

$$k = 0,1,\ldots,N-1.$$

Let us write, again for simplicity,

23. $$\theta'(t) = \frac{d}{dt}\theta(t;t_k,z_k).$$

Using Schwartz' inequality we can deduce the following inequalities:

$$24. \quad \text{i)} \quad \|\Delta\theta_k(s)\|^2 \leq \Delta_k \int_{t_k}^{t_{k+1}} \|\theta'(s)\|^2\, ds,$$

$$\text{ii)} \quad |\tilde{R}\Delta\theta_k|^2 \leq \Delta_k \int_{t_k}^{t_{k+1}} |\tilde{R}\theta'(s)|^2\, ds.$$

Substituting 24. in 22. and eliminating the factor 1/2 in the left side we have

$$25. \quad |\hat{\phi}|^2 \leq \Delta_k^2 \frac{\rho\gamma_o^2}{\sigma} \int_{t_k}^{t_{k+1}} \|\theta'(s)\|^2\, ds + $$

$$+ \Delta_k^2 \frac{\gamma_o^2}{\sigma\rho} \int_{t_k}^{t_{k+1}} \|\theta'(s)\|^2\, ds + \Delta_k \int_{t_k}^{t_{k+1}} |\tilde{R}\theta'(s)|^2 ds,$$

$$k = 0,1,\ldots,N-1.$$

Now, consider the $\mathcal{V}$-valued, $\mathcal{F}_k$-measurable, random variable $\hat{\phi}(u(t_k))$, $k = 0,1,\ldots,N-1$, obtained by means of equation 17. when the variable z_k is fixed at $u(t_k)$, the function, u, being the solution of the evolution problem 5.2.8..

We shall show that $\hat{\phi}(u(t_k))$ is a version of the conditional expectation of ϕ_{k+1} with respect to the σ-algebra $\mathcal{F}_k$. In other words,

$$26. \quad E(\phi_{k+1}/\mathcal{F}_k) = \hat{\phi}(u(t_k)), \qquad \text{w.p.1}$$

$$k = 0,1,\ldots,N-1.$$

In order to prove the above relation consider the equation 6.1.15.. Taking the conditional expectation from both sides we have,

27. $$E(\phi_{k+1}/\mathcal{J}_k) = R\,E(u(t_{k+1})/\mathcal{J}_k) - Ru(t_k) + \Delta_k \mathcal{G}_k^o Ru(t_k),$$

$$k = 0,1,\ldots,N-1.$$

Subtracting the above relation from equation 17. we have,

28. $$\hat{\phi}(z_k) - E(\phi_{k+1}/\mathcal{J}_k) = R(\theta(t_{k+1};t_k z_k) - E(u(t_{k+1})/\mathcal{J}_k)) +$$

$$- R(z_k - u(t_k)) + \Delta_k \mathcal{G}_k^o R(z_k - u(t_k)),$$

$$k = 0,1,\ldots,N-1.$$

Taking $z_k = u(t_k)$ in the above equation,

29. $$\hat{\phi}(u(t_k)) - E(\phi_{k+1}/\mathcal{J}_{t_k}) = R(\theta(t_{k+1};t_k,u(t_k)) +$$

$$- E(u(t_{k+1})/\mathcal{J}_k)),$$

$$k = 0,1,\ldots,N-1.$$

Now, compare equation 16.i) with equation 5.2.8.ii). We observe that the following identity can be written:

30. $$\theta(t;\hat{t},u(\hat{t})) - E(u(t)/\mathcal{J}_{\hat{t}}) = \theta(t;\hat{t},0),$$

for all $0 \leq \hat{t} \leq t \leq T$.

But, by the results obtained in section 3 we have,

31. $$\theta(t;\hat{t},0) = 0 .$$

Therefore relation 26. is proved. Using the estimate 25. as an estimate for the conditional expectation, the thesis of Proposition 6.2.2 is demonstrated.

6.3 - An Abstract Error Estimate

We shall now present an estimate for the error of approximation.

From equation 6.1.16, the following inequality can be written:

1. $$\begin{aligned} |e_{k+1}|^2 \leq\ & |||I - \Delta_k \mathcal{G}_k^0|||^2 \, |e_k|^2 + \\ & + (\Delta w_k)^2 |||\mathcal{G}_k^1|||^2 \, |e_k|^2 + |\phi_{k+1}|^2 + \\ & - 2\,\Delta w_k ((I - \Delta_k \mathcal{G}_k^0) e_k, \mathcal{G}_k^1 e_k) + \\ & - 2((I - \Delta_k \mathcal{G}_k^0) e_k, \phi_{k+1}) + 2 \Delta w_k (\mathcal{G}_k^1 e_k, \phi_{k+1}), \end{aligned}$$

$$k = 0,1,\ldots,N-1 .$$

Take the expectation on both sides of this equation. Recalling that e_k is a $\mathcal{F}_k$-measurable random variable and using Schwartz' inequality we have,

2. $$E|e_{k+1}|^2 \le |||I - \Delta_k \mathcal{G}_k^o|||^2 \, E|e_k|^2 + \Delta_k |||\mathcal{G}_k^1|||^2 \, E|e_k|^2 + E|\phi_{k+1}|^2 + 2\,|||I - \Delta_k \mathcal{G}_k^o|||\; E(|e_k|\,|E(\phi_{k+1}/\mathcal{F}_k)|) + 2\,|||\mathcal{G}_k^1|||\; E(|\Delta w_k|\,|e_k|\,|\phi_{k+1}|),$$

$$k = 0,1,\dots,N-1.$$

Now let us suppose that in the scheme 6.1.7 we are taking,

3. $$\rho \ge 0.5.$$

Recalling the estimates in Proposition 6.1.1 we have,

4. $$E|e_{k+1}|^2 \le E|e_k|^2 + \Delta_k \gamma_1^2 E|e_k|^2 + E|\phi_{k+1}|^2 + 2\,E(|e_k|\,|E(\phi_{k+1}/\mathcal{F}_k)|) + 2\gamma_1 E(|\Delta w_k|\,|e_k|\,|\phi_{k+1}|),$$

$$k = 0,1,\dots,N-1.$$

Making use of Cauchy's inequality, $pq \le p^2/2\varepsilon + \varepsilon q^2/2$, with $\varepsilon = (\Delta_k)^{-1}$ and $\varepsilon = (\gamma_1)^{-1}$ in the last two terms respectively, we have,

5. $$E|e_{k+1}|^2 \le E|e_k|^2 + \Delta_k\gamma_1^2 E|e_k|^2 + E|\phi_{k+1}|^2 + \Delta_k E|e_k|^2 + \Delta_k^{-1} E|E(\phi_{k+1}/\mathcal{F}_k)|^2 + \gamma_1^2 E(\Delta^2 w_k|e_k|^2) + E|\phi_{k+1}|^2,$$

$$k = 0,1,\ldots,N-1.$$

After some manipulation inequality 5. yields

6. $$E|e_{k+1}|^2 \le E|e_k|^2 + \Delta_k(2\gamma_1^2 + 1) E|e_k|^2 + 2E|\phi_{k+1}|^2 + \Delta_k^{-1} E|E(\phi_{k+1}/\mathcal{F}_k)|^2,$$

$$k = 0,1,\ldots,N-1.$$

Recalling Lemma 4.3.1, we are able to deduce the following inequality:

7. $$E|e_k|^2 \le \exp\left[\sum_{j=0}^{N-1}(2\gamma_1^2 + 1)\Delta_j\right]\left\{E|e_o|^2 + \sum_{j=0}^{N-1}\left\{2E|\phi_{j+1}|^2 + \Delta_j^{-1} E|E(\phi_{j+1}/\mathcal{F}_j)|^2\right\}\right\},$$

$$k = 0,1,\ldots,N.$$

Now, Propositions 6.2.1 and 6.2.2 enable us to present the final result. Substituting estimates 6.2.13 and 6.2.15 in the inequality 7. above, we have,

8. $$E|e_k|^2 \leq C \left\{ |Ru_o - U_o|^2 + \sum_{j=0}^{N-1} \left\{ h E\|\Delta u_j\|^2 + \int_{t_j}^{t_{j+1}} E\|\Delta u_j(s)\|^2 ds + h^3 E|u(t_j)|^2 + E|\tilde{R}\Delta u_k|^2 + h\, E|\tilde{R}u(t_k)|^2 + h \int_{t_j}^{t_{j+1}} E\left\|\frac{d}{ds}\theta(s;t_j,u(t_j))\right\|^2 ds + \int_{t_j}^{t_{j+1}} E\left|\tilde{R}\frac{d}{ds}\theta(s;t_j,u(t_j))\right|^2 ds \right\} \right\},$$

$$k = 0,1,\ldots,N,$$

where C is a positive constant depending only on ρ, γ_o, γ_1, γ_1', σ and T.

Although the estimate 8. provides us with the means for proving convergence of the numerical method given by the scheme 6.1.7, it does not represent alone, a convergence result. If these results are sought, we need supplementary assumptions.

There we shall now present a set of hypotheses and a convergence result for scheme 6.1.7.

First, let us assume that our bilinear form a_o can be written as a sum of two bilinear forms b_o and b_1, defined on the space V, such that,

9. i) $a_o = b_o + b_1$,

ii) b_o is symmetric,

iii) $B_1 \in L(V,H)$,

where B_j, $j = 0,1$, denotes the linear operator associated with the bilinear form b_j.

With the addition of hypothesis 9. we are now able to use the results of section 3. with respect to the evolution equation 6.2.16. Consider equation 6.2.17. From estimate 3.3.24, we conclude that there exists a constant C such that,

$$\| Z(t;Z_o) \|^2 \leq C \| Z_o \|^2, \qquad t \in [0,T],$$

for all $Z_o \in V$.

Therefore, using relation 6.2.18, we have for the solution of equation 6.2.16 the following estimate,

10. $$\left\| \frac{d}{dt} \theta(t;\hat{t},z) \right\|^2 \leq C \| A_o z \|^2, \qquad 0 \leq \hat{t} \leq t \leq T,$$

for all z such that $A_o z \in V$.

So, let us suppose that for the solution of the problem 5.2.8 we have,

11. i) $A_o u(t) \in V$,

ii) $E\|A_o u(t)\|^2 \leq M < \infty,$

for all $t \in [0,T]$.

By inequality 10., hypotheses 9. and 11. lead us to the conclusion that there exists a constant C such that,

12. $$E\|\frac{d}{dt}\theta(t;\hat{t},u(\hat{t})\|^2 \leq C E\|A_o u(\hat{t})\|^2 < \infty,$$

for all $0 \leq \hat{t} \leq t \leq T$.

On the other hand, using equation 5.2.8.ii) and a standard procedure, hypothesis 11.ii) allow us to conclude that there exists a constant C such that

13. $$E\|\Delta u_k(s)\|^2 \leq C h,$$

for all $s \in [t_k, t_{k+1}]$, $k = 0,1,\ldots,N-1$.

Now, let us consider the approximation subspace, $\mathcal{V}$, where the scheme 6.1.7 is defined.

We suppose that there exists a family of finite dimensional subspaces $\mathcal{V}(d) \subset V$ with $d > 0$ such that, with respect to the bilinear form a_o and the spaces H and V, the following approximation property holds:

14. $$|\tilde{R}u| \leq d\|u\|,$$

for all $u \in V$.

So, selecting $\mathcal{V}$ as a member of the family of subspaces described above,

15. $\vartheta = \vartheta(d)$,

we are able to show the following Theorem:

Theorem 6.3.1 - Under the hypotheses of Proposition 6.2.2 plus hypotheses 9., 11. and 15. the following estimate holds:

16. $$\sup_k (E|u(t_k) - U_k|^2) \leq C \Big\{ |Ru_o - U_o|^2 + h(1 + \sup_{[0,T]} (E\|A_o u(t)\|^2)) + h^2 \sup_{[0,T]} (E|u(t)|^2) + d^2 (1 + \sup_{[0,T]} (E\|u(t)\|^2) + \sup_{[0,T]} (E\|A_o u(t)\|^2)) \Big\},$$

where C is a positive constant.

Proof of Theorem 6.3.1

The proof follows after using inequalities 12., 13., and 14. in the estimate 8. and then substituting in 6.1.12. ●

Remark 6.3.1 - The estimate 16. means that under the conditions of Theorem 6.3.1, a numerical procedure given by the scheme 6.1.17, with $U_o = Ru_o$, will converge to the solution of problem 5.2.8, in the norm,

$$\sup_k \| u(t_k) - U_k \|_{L^2(\Omega,H)} \quad \bullet$$

Here t_k, $k = 0,1,\ldots,N$ are the dividing points of the partition of the interval $|0,T|$. The rate of convergence in the time will be $h^{1/2}$. This is a slow rate of convergence. In paragraph 6.4 we shall present a family of schemes that, under stronger conditions, will converge with a faster rate.

Here, we observe that the crucial hypothesis is stated in 12.. It is possible to interpret these conditions by saying that they represent a certain regularity attained by the solution of problem 5.2.8 and this interpretation has a precise meaning when A_0 is a partial differential operator. We shall return to this situation in paragraph 6.4.

We also remark that the hypothesis concerning the approximation subspace is standard and can be verified for finite-element subspaces (see paragraph 4.3.4).

6.4 - A Quadratic Scheme

In paragraph 6.1 we introduced a simple numerical scheme which is linear in terms of the increment in the noise. We remarked in the end of paragraph 6.3, that the rate of convergence of such a scheme can be disappointingly slow. Here we shall present another scheme which, under suitable conditions, can have a faster rate of convergence.

As has been pointed out by McShane (|36|) and, also Clark (| 6 |), for finite dimensional stochastic differential equations, a higher order of convergence in time can be achieved if, in the numerical scheme, we take into account terms containing powers of the noise increment.

This fact can be understood with an analogy between stochastic and non-stochastic differential equations. Consider the scalar linear differential equation,

$$\frac{du}{dt}(t) = a\,u(t), \qquad a \in R.$$

Therefore $u(t) = \exp(at)u(0)$ and we may say that numerical schemes for the above equation are constructed in order to approximate the exponential $\exp(a\Delta_k)$, where Δ_k is the increment in time (see Remark 4.2.1.).

Now, consider the simplest scalar version of the stochastic equation 5.8.ii). It has the form

$$du(t) = au(t)dt + bu(t)dw_t \;; \quad a,b \in R.$$

So, $u(t) = \exp(at - \frac{1}{2}b^2 t + bw_t)u(0)$, (w.p.1) and therefore, in this case, schemes should be constructed in order to produce approximations to the exponencial $\exp(a\Delta_k - \frac{1}{2}b^2\Delta_k + b\Delta w_t)$.

It is easy to see that, in relation to the above stochastic equation, the scheme introduced in paragraph 6.1 fails to approximate the second term in the exponential and, besides, gives a mediocre approximation to the third term.

Following this line of argument we can produce a more complex scheme, containing a second order term (in the power of the noise increment) which may have a faster rate of convergence. This scheme corresponds to McShane's numerical method (McShane, |36|, p. 205).

In what follows, we shall use the notation introduced in paragraphs 6.1.1, 6.1.2 and 6.1.3. However, we must consider supplementary hypotheses.

First, for simplicity, we also assume the operator $A_1(t)$ to be invariant in time,

1. $A_1(t) = A_1$.

Remark 6.2.1 also applies to the above hypothesis. In other words, hypothesis 1. is not a fundamental hypothesis and basically, the results of this paragraph could be obtained with hypothesis 6.1.2 alone.

We also assume that the linear operator $A_1 \in L(H,H)$ is such that,

2. $$A_1^* v \in V \qquad \text{for all } v \in V,$$

where A_1^* denotes the adjoint of A_1.

In addition to the operators L_j, $\mathcal{G}_k^j$, $j = 0,1$ defined in paragraph 6.1 define linear operators L_2, $\mathcal{G}_k^2 \in L(\mathcal{V},\mathcal{V})$ by the following relations:

3. $$(A_1^2 u,v) = (L_2 u,v), \qquad \text{for all } u,v \in \mathcal{V}.$$

4. $$\mathcal{G}_k^2 = (I + \Delta_k \rho L_o)^{-1} L_2, \qquad k = 0,1,\ldots,N-1.$$

Consider the second order stochastic numerical scheme,

5. $$\tilde{U}_{k+1} - \tilde{U}_k + \Delta_k(\mathcal{G}_k^o + \frac{1}{2}\mathcal{G}_k^2)\tilde{U}_k + {} + \Delta w_k \mathcal{G}_k^1 \tilde{U}_k - \frac{1}{2}(\Delta w_k)^2 \mathcal{G}_k^2 \tilde{U}_k = 0, \qquad k = 0,1,\ldots,N-1,$$

where, here, we use the symbol ~ to differentiate the above scheme from the scheme 6.1.7.

Starting with the above equation we can follow the same pattern of analysis as we did before.

First, we recall a basic identity concerning Wiener processes:

6. $$(\Delta w_k)^2 = \Delta_k + 2\int_{t_k}^{t_s} \Delta w_k(s)\, dw_s,$$

$$k = 0,1,\ldots,N-1.$$

As a consequence of this identity we can write,

7. $$\Delta_k \mathcal{G}_k^2 - (\Delta w_k)^2 \mathcal{G}_k^2 = -2\int_{t_k}^{t_{k+1}} \Delta w_k(s)\, \mathcal{G}_k^2\, dw_s.$$

Therefore, if we want to explore stability properties of the scheme 5., we can start from the fact that, as before (see equation 6.1.10), the expectation of the variables $\tilde{U}_k$ satisfy a scheme identical to the one studied in section 4.. Substituting 7. in 5. we have,

8. $$E\,\tilde{U}_{k+1} = (I - \Delta_k \mathcal{G}_k^o) E\,\tilde{U}_k,$$

$$k = 0,1,\ldots,N-1.$$

We can also write expressions for the error of approximation. So, the counterpart of equation 6.1.16 has now the form,

9. $$\tilde{e}_{k+1} - \tilde{e}_k + \Delta_k(\mathcal{G}_k^o + \frac{1}{2}\mathcal{G}_k^2)\tilde{e}_k + \Delta w_k \mathcal{G}_k^1 \tilde{e}_k + - \frac{1}{2}(\Delta w_k)^2 \mathcal{G}_k^2 \tilde{e}_k + \tilde{\phi}_{k+1} = 0,$$

$$k = 0,1,\ldots,N-1.$$

Here,

10. $$\tilde{\phi}_{k+1} = R\Delta u_k + \Delta_k(\mathcal{G}_k^o + \frac{1}{2}\mathcal{G}_k^2)Ru(t_k) + + \Delta w_k \mathcal{G}_k^1 Ru(t_k) - \frac{1}{2}(\Delta w_k)^2 \mathcal{G}_k^2 Ru(t_k).$$

Now, multiplying both sides of equation 10. by $(I + \Delta_k \rho L_o)$ and using relation 7. we have, after rearranging terms,

11. $$(I + \Delta_k \rho L_o)\tilde{\phi}_{k+1} = R\Delta u_k + \Delta_k L_o(\rho Ru(t_{k+1}) + + (1-\rho)Ru(t_k)) + \Delta w_k L_1 Ru(t_k) + - \int_{t_k}^{t_{k+1}} \Delta w_k(s) L_2 Ru(t_k)\, ds,$$

$$k = 0,1,\ldots,N-1.$$

Using the definition of the operators involved and identity 6.2.2, we can derive the following expression:

12. $$(\tilde{\phi}_{k+1},v) + \Delta_k \rho a_o(\tilde{\phi}_{k+1},v) = \Delta_k \rho a_o(\Delta u_k,v) +$$

$$- \int_{t_k}^{t_{k+1}} a_o(\Delta u_k(s),v)\, ds - \int_{t_k}^{t_{k+1}} (A_1 \Delta u_k(s),v)\, dw_s +$$

$$- \int_{t_k}^{t_{k+1}} \Delta w_k(s)\; (A_1^2 u(t_k),v)\, dw_s +$$

$$- (\tilde{R}\Delta u_k,v) - \Delta w_k(A_1 \tilde{R} u(t_k),v) +$$

$$+ \int_{t_k}^{t_{k+1}} \Delta w_k(s)(A_1^2 \tilde{R} u(t_k),v)\, dw_s,$$

for all $v \in \vartheta$,

$k = 0,1,\ldots,N-1.$

We observe that this expression differs from 6.2.8 only by the terms that contain an integral of the noise increment, and also by the term in 6.2.8 that contains the derivative of $A_1(t)$. This term is "small" in relation to the others and so, hypothesis 1. is justifiable.

Now, consider the following relation:

13. $$(A_1 \Delta u_k(s),v) + \Delta w_k(s)(A_1^2 u(t_k),v) =$$

$$= - \int_{t_k}^{s} a_o(u(\xi),A_1^* v)\, d\xi - \int_{t_k}^{s} (A_1^2 u(\xi),v)\; dw_\xi +$$

(Equation 13. - continuation)

$$+ \int_{t_k}^{s} (A_1^2 u(t_k), v)\, dw_\xi =$$

$$= - \int_{t_k}^{s} a_o(u(\xi), A_1^* v)\, d\xi - \int_{t_k}^{s} (A_1^2 \Delta u_k(\xi), v)\, dw_\xi ,$$

for all $v \in V$

$k = 0, 1, \ldots, N-1.$

(We have used basically identity 6.2.2 and hypothesis 2. in the above derivation)

Substituting in 12. we have,

14. $$(\tilde{\phi}_{k+1}, v) + \Delta_k \rho a_o(\tilde{\phi}_{k+1}, v) = \Delta_k \rho a_o(\Delta u_k, v) +$$

$$- \int_{t_k}^{t_{k+1}} a_o(\Delta u_k(s), v)\, ds + \int_{t_k}^{t_{k+1}} \int_{t_k}^{s} a_o(u(\xi), A_1^* v)\, d\xi\, dw_s +$$

$$+ \int_{t_k}^{t_{k+1}} \int_{t_k}^{s} (A_1^2 \Delta u_k(\xi), v)\, dw_\xi\, dw_s +$$

$$- (\tilde{R} \Delta u_k, v) - \Delta w_k (A_1 \tilde{R} u(t_k), v) +$$

$$- \int_{t_k}^{t_{k+1}} \Delta w_k(s) (A_1^2 \tilde{R}(t_k), v)\, dw_s ,$$

for all $v \in \mathcal{V}$

$k = 0,1,\ldots,N-1$

We observe now how the second order term in the numerical scheme 5. can be used in order to produce faster rate of convergence. By means of relation 13. we have eliminated the third term in the right side of equation 12. which also appears in 6.2.8. This term contributes, in the error estimate 6.3.8, to give a slow rate of convergence of the scheme 6.1.7. Here, as a consequence of the second order term in scheme 5., we have replaced it by higher order terms. However, this is not enough to guarantee a faster order of convergence for the scheme 5. In fact, we observe that the first and the second terms in the right side of 6.2.8 are also responsible for the slow rate of convergence of the method. These terms also appear in equation 14. and so, in this case, we can not make use of the advantages of a second order scheme, unless some additional hypotheses are made.

We already know that the solution of the problem 5.2.8 satisfies,

$$u(t) \in D(A_0(t)) \;, \qquad t \in [0,T],$$

(see section 5.)

Therefore, we can write,

$$a_0(\Delta u_k(s),v) = (A_0 \Delta u_k(s),v),$$

for all $v \in V$, $k = 0,1,\ldots,N-1$, $s \in [0,T]$.

So, the supplementary hypothesis that we need is the following:

15. $E|A_o u(t)|^2 \leq M < \infty, \quad t \in [0,T].$

Now we can return to equation 14.. Choosing $v = \tilde{\phi}_{k+1}$ as a test function, using hypotheses 5.2.4 and 5.2.6 jointly with the Schwartz' inequality we have,

16. $$|\tilde{\phi}_{k+1}| \leq \Delta_k \rho |A_o \Delta u_k| + \int_{t_k}^{t_{k+1}} |A_o \Delta u_k(s)| ds +$$

$$+ \gamma_1^2 \left| \int_{t_k}^{t_{k+1}} \int_{t_k}^{s} |A_o u(\xi)| d\xi dw_s \right| +$$

$$+ \gamma_1^2 \left| \int_{t_k}^{t_{k+1}} \int_{t_k}^{s} |\Delta u_k(\xi)| dw_\xi dw_s \right| +$$

$$+ |\tilde{R} \Delta u_k| + |\Delta w_k| |\tilde{R} u(t_k)| +$$

$$+ \gamma_1^2 \left| \int_{t_k}^{t_{k+1}} |\Delta w_k(s)| |\tilde{R} u(t_k)| dw_s \right|,$$

$$k = 0, 1, \ldots, N-1.$$

We can now estimate $E|\tilde{\phi}_{k+1}|^2$. From equation 16. and using usual properties of Wiener processes and stochastic integrals (see paragraph 5.1) we are able to deduce the following estimate:

17. $$E|\tilde{\phi}_{k+1}|^2 \leq C\left\{ \Delta_k^2 E|A_o \Delta u_k|^2 + \Delta_k \int_{t_k}^{t_{k+1}} E|A_o \Delta u_k(s)|^2 ds + \right.$$

$$+ (\Delta_k)^2 \int_{t_k}^{t_{k+1}} E|A_o u(s)|^2 ds + E|\tilde{R}\Delta u_k|^2 +$$

$$\left. + \Delta_k E|\tilde{R}u(t_k)|^2 + \Delta_k^2 E|\tilde{R}u(t_k)|^2 \right\},$$

$$k = 0,1,\ldots,N-1,$$

where C is a constant depending on ρ and γ_1.

Let us return to equation 9. Using identity 6., equation 9. can be rewritten in the following form:

18. $$\tilde{e}_{k+1} = (I - \Delta_k \mathcal{G}_k^o)\tilde{e}_k - \Delta w_k \mathcal{G}_k^1 \tilde{e}_k +$$

$$+ \left(\int_{t_k}^{t_{k+1}} \Delta w_k(s) dw_s \right) \mathcal{G}_k^2 \tilde{e}_k - \tilde{\phi}_{k+1} .$$

Now we use the same procedure used in paragraph 6.3.. So, apply the operator $E|.|^2$ in both sides of equation 18.. After expanding the right side we obtain the following terms and their estimates:

19. $$E|(I - \Delta_k \mathcal{G}_k^o)\tilde{e}_k|^2 \leq E|\tilde{e}_k|^2,$$

20. $$E|\Delta w_k \mathcal{G}_k^1 \tilde{e}_k|^2 \leq \Delta_k \gamma_1^2 E|\tilde{e}_k|^2,$$

21. $$E\left|\left(\int_{t_k}^{t_{k+1}} \Delta w_k(s)\,dw_s\right)\mathcal{G}_k^2 \tilde{e}_k\right|^2 \le$$

$$\le \gamma_1^4 E\left(E\left(\left|\int_{t_k}^{t_{k+1}} \Delta w_k(s)\,dw_s\right|^2 / \mathcal{F}_k\right)|\tilde{e}_k|^2\right) \le$$

$$\le \gamma_1^4 \Delta_k^2 E|\tilde{e}_k|^2,$$

22. $$E|\tilde{\phi}_{k+1}|^2 \le \zeta,$$

where ζ represents the right side of equation 17.

23. $$E(-2((I - \Delta_k \mathcal{G}_k^0)\tilde{e}_k, \Delta w_k \mathcal{G}_k^1 \tilde{e}_k)) = 0.$$

24. $$E\left(2\left((I - \Delta_k \mathcal{G}_k^0)\tilde{e}_k, \left(\int_{t_k}^{t_{k+1}} \Delta w_k(s)\,dw_s\right)\mathcal{G}_k^2 \tilde{e}_k\right)\right) = 0.$$

25. $$E(-2((I - \Delta_k \mathcal{G}_k^0)\tilde{e}_k, \tilde{\phi}_{k+1})) \le$$

$$\le 2\,E(|\tilde{e}_k||E\tilde{\phi}_{k+1}/\mathcal{F}_k|) \le$$

$$\le \Delta_k E|\tilde{e}_k|^2 + \Delta_k^{-1} E|E\tilde{\phi}_{k+1}/\mathcal{F}_k|^2.$$

26. $$E\left(-2\left(\Delta w_k \mathcal{G}_k^1 \tilde{e}_k\,,\,\left(\int_{t_k}^{t_{k+1}} \Delta w_k(s)\,ds\right)\mathcal{G}_k^2 \tilde{e}_k\right)\right) \le$$

$$\le E\left(\gamma_1^2 |\Delta w_k|^2 |\tilde{e}_k|^2 + \gamma_1^4 \left|\int_{t_k}^{t_{k+1}} \Delta w_k(s)\,ds\right|^2 |\tilde{e}_k|^2\right) =$$

$$= (\gamma_1^2 \Delta_k + \gamma_1^4 \Delta_k^2)\, E|\tilde{e}_k|^2.$$

27. $$E(2(\Delta w_k \mathcal{G}_k^1 \tilde{e}_k\,,\,\tilde{\phi}_{k+1})) \le E(\gamma_1^2 |\Delta w_k|^2 |\tilde{e}_k|^2 + |\tilde{\phi}_{k+1}|^2) =$$

$$= \gamma_1^2 \Delta_k E|\tilde{e}_k|^2 + E|\tilde{\phi}_{k+1}|^2.$$

28. $$E\left(-2\left(\left(\int_{t_k}^{t_{k+1}} \Delta w_k(s)\,dw_s\right)\mathcal{G}_k^2 \tilde{e}_k\,,\,\tilde{\phi}_{k+1}\right)\right) \le$$

$$\le E\left(\gamma_1^4 \left|\int_{t_k}^{t_{k+1}} \Delta w_k(s)\,dw_s\right|^2 |\tilde{e}_k|^2 + |\tilde{\phi}_{k+1}|^2\right) =$$

$$= \gamma_1^4 \Delta_k^2\, E|\tilde{e}_k|^2 + E|\tilde{\phi}_{k+1}|^2,$$

$$k = 0,1,\ldots,N-1.$$

In the derivation of inequalities 19.,...,28. we have used basically hypothesis 6.3.3, the results of proposition 6.1.1 and standard properties of Wiener processes.

So, using the estimates 19.,...,28., we can write,

29. $$E|\tilde{e}_{k+1}|^2 \leq (1 + \Delta_k(3\gamma_1^2 + 1) + \Delta_k^2 3\gamma_1^4)E|\tilde{e}_k|^2 + + 3E|\tilde{\phi}_{k+1}|^2 + \Delta_k^{-1} E|E\tilde{\phi}_{k+1}/\mathcal{J}_k|^2,$$

$$k = 0,1,\ldots,N-1.$$

We observe that, as before we need now an estimate for $|E\tilde{\phi}_{k+1}/\mathcal{J}_k|^2$.

Let us return to equation 10. Taking the conditional expectation in both sides of this equation we have,

30. $$E(\tilde{\phi}_{k+1}/\mathcal{J}_k) = RE(u(t_{k+1})/\mathcal{J}_k) + - Ru(t_k) + \Delta_k \mathcal{G}_k^o Ru(t_k).$$

Comparing 30. with 6.2.27 we have

31. $$E(\tilde{\phi}_{k+1}/\mathcal{J}_k) = E(\phi_{k+1}/\mathcal{J}_k).$$

In other words, the numerical methods that correspond to schemes 6.1.7 and 5. have, almost surely, the same "consistency" at the dividing points of the partition of the interval [O,T] conditioned to the information stored from the previous points.

So, it can be argued that scheme 5. will not produce faster rates of convergence since the conditional expectation of ϕ_{k+1} is also responsible for the slow convergence of the scheme 6.1.7. However, in view of our supplementary assumption 15., the result of Proposition 6.2.2 can be improved.

Consider equation 6.2.16. From Remark 6.2.4 and using a standard procedure (see paragraph 3.4) we have

$$32. \qquad |A_o\theta(t)| = |\tfrac{d}{dt}\theta(t)| = |\theta(t;\hat{t},-A_oz| \leq C|A_oz|,$$

for some constant C. (Here, $\theta(t) = \theta(t;\hat{t},z)$).

Therefore, in equation 6.2.20, we can write

$$33. \qquad a_o(\Delta\theta_k(s),v) = (\Delta A_o\theta_k(s),v) \leq |\Delta A_o\theta_k(s)||v|,$$

for all $v \in \mathcal{V}$

$k = 0,1,\ldots,N-1$

$s \in [0,T]$.

So, instead of inequality 6.2.22 we now have,

$$34. \qquad |\hat{\phi}| \leq \Delta_k\rho|\Delta A_o\theta_k| + \int_{t_k}^{t_{k+1}} |\Delta A_o\theta_k(s)|ds + |\tilde{R}\Delta\theta_k|.$$

The result of Proposition 6.2.2 can now be rewritten,

$$35. \qquad E(|E(\phi_{k+1}/\mathcal{J}_k)|^2) \leq C\,\Delta_k^3 \int_{t_k}^{t_{k+1}} E|A_o\tfrac{d}{ds}\theta(s;t_k,u(t_k))|^2 ds \; +$$

$$+ \;\Delta_k \int_{t_k}^{t_{k+1}} E|\tilde{R}\,\tfrac{d}{ds}\theta(s;t_k,u(t_k))|^2 ds,$$

$k = 0,1,\ldots,N-1,$

where C is a positive constant depending only on ρ.

Therefore, recalling Lemma 4.3.1 and making use of estimates 17. and 35. jointly with identity 31., equation 29. yields the following estimate:

36.
$$E|\tilde{e}_{k+1}|^2 \leq C\Bigg\{|Ru_o - \tilde{U}_o|^2 + \sum_{j=0}^{N-1}\Bigg\{ h^2 E|\Delta A_o u_j|^2 + h\int_{t_j}^{t_{j+1}} E|\Delta A_o u_j(s)|^2 ds + h^2\int_{t_j}^{t_{j+1}} E|A_o u(s)|^2 ds + E|\tilde{R}\Delta u_j|^2 + h\,E|\tilde{R}u(t_j)|^2 + h^2 E|\tilde{R}u(t_j)|^2 + h^2\int_{t_j}^{t_{j+1}} E|A_o \frac{d}{ds}\theta(s;t_j,u(t_j))|^2 ds + \int_{t_j}^{t_{j+1}} E|\tilde{R}\frac{d}{ds}\theta(s;t_j,u(t_j))|^2 ds\Bigg\}\Bigg\},$$

where C is a constant depending only on ρ, γ_1 and T.

A result similar to Theorem 6.3.1 can be derived. In order to proceed in this direction, let us assume hypothesis 6.3.9 concerning the composition of the bilinear form a_o. Consider the evolution equation 6.2.17 in the Remark 6.2.4. From equation 6.2.18 we can write

$$\frac{d^2}{dt^2} Z(t;Z_o) = \frac{d}{dt} Z(t;-A_o Z_o) = Z(t;A_o^2 Z_o), \qquad t \in [0,T],$$

for all $Z_o \in D(A_o^2)$.

So, we conclude that

$$A_o \frac{d}{dt} Z(.,Z_o) = - \frac{d^2}{dt^2} Z(.,Z_o) \in C(0,T;H),$$

for all $Z_o \in D(A_o^2)$.

Transfering this argument to equation 6.2.16 and using estimate 3.3.16 weconclude that there exists a constant C such that,

37. $$|A_o \frac{d}{dt} \theta(t;\hat{t},z)|^2 \leq C|A_o^2 z|^2 < \infty,$$

$$0 \leq \hat{t} \leq t \leq T,$$

for all $z \in D(A_o^2)$.

Here, we need hypotheses which are stronger than those in 6.3.11. So, assume that for the solution of problem 5.2.8 the following conditions hold:

38. i) $A_o A_j u(t) \in H$

ii) $E|A_o A_j u(t)|^2 \leq M < \infty,$

for all $t \in [0,T]$, $j = 0,1$.

Using equation 5.2.8.ii) and a standard procedure hypothesis 38. enable us to conclude that there exists a constant C such that,

39. $E|\Delta A_o u_k(s)|^2 \leq C h,$

for all $s \in [t_k, t_{k+1}]$, $k = 0,1,\ldots,N-1$.

Assuming 6.2.15 for the approximation subspace $\mathcal{V}$ we can now introduce the following theorem:

<u>Theorem 6.4.1</u> - Under the hypotheses of Proposition 6.2.2 plus hypotheses 6.3.9, 6.3.15, 1, 2, 38 the following estimate holds:

40. $$\sup_k E|u(t_k) - \tilde{U}_k|^2 \leq C\Big\{ |Ru_o - \tilde{U}_o|^2 + h^2(1 + \sup_{[0,T]} (E|A_o^2 u(t)|^2)) + d^2(1 + (1+h)\sup_{[0,T]} (E\|u(t)\|^2) + \sup_{[0,T]} (E\|A_o U(t)\|^2))\Big\},$$

where C is a positive constant

Proof of Theorem 6.4.1

We can use inequalities 37., 39. and 6.2.14 in order to estimate the terms in the right side of equation 36. Recalling that,

$$u(t_k) - \tilde{U}_k = u(t_k) - Ru(t_k) + \tilde{e}_k, \qquad k = 0,1,\ldots,N-1,$$

we obtain the result above. ●

Remark 6.4.1 - According to Theorem 6.4.1, the quadratic scheme 5. can produce approximations with errors of order h. This is a considerable improvement with respect to the linear scheme 6.1.7 which converges at a rate $h^{1/2}$. However, to guarantee this fact, a condition stronger than 6.3.11 must be imposed on the solution of problem 5.2.8, namely, hypothesis 38. As we mention before (see Remark 6.3.1) hypotheses like these in 6.3.11 or 38. have a clear interpretation in terms of the regularity of the solution of the stochastic evolution equation when A_o is a partial differential operator. This is the subject of our next paragraph.

6.5 - An Application to the Filtering Problem

We shall now apply the results obtained in the previous paragraphs to the numerical solution of the non linear filtering problem for diffusion introduced in paragraph 1.1. We will be concerned with Galerkin approximations of the solution of the Zakai formula 1.1.14.

Let $H = L^2(S)$, $V = H_0^1(S)$, where S is a bounded subset of R^n.

Consider the stochastic evolution equation 5.3.8. In addition to hypotheses 5.3.2 and 5.3.6 assume

1. $a_{i,j}$, g_i are invariant in time,

$$i,j = 1,\dots,n$$

As a consequence of this hypothesis the bilinear form $a_o(t)$ introduced in 5.3.1 is invariant in time and we are now able to use the estimates presented in paragraphs 6.3 and 6.4. As we pointed out before (see Remark 6.2.1) this hypothesis is not restrictive, and it was only made in order to simplify the steps leading to estimate 6.3.8 and 6.4.30. If $a_{i,j}$ and g_i are of class C^1 with respect to $t \in [0,T]$, similar results hold regarding the error of approximation of the numerical methods with which we are concerned.

We also assume the diffusion matrix to be positive definitive. In other words, for some $\sigma > 0$,

2. $$\langle r,[a_{i,j}]r\rangle \geq \sigma\langle r,r\rangle,$$

for all $r \in R^n$

$x \in S$.

Now, let $\beta^1 = 0$. Equation 5.3.8 (with $\tilde{w}_t \equiv y(t)$) now becomes identical to the Zakai formula 1.1.14.. In particular, hypothesis 6.1.1 is satisfied and the condition 2. above guarantees the coercivity condition 5.2.6. In order to have hypothesis 6.1.2 satisfied we assume,

3. $$h \in C^1(0,T;L^\infty(S)).$$

We observe that now, all the hypotheses made at the beginning of paragraph 6.1 with respect to $a_0(t)$ and $a_1(t)$ are satisfied. Therefore, we can use inequality 6.3.8 in order to estimate the error of approximation of the Galerkin scheme 6.1.7. Before we proceed in this direction we select the approximation subspace $\mathcal{V}$ as an element of the family of subspaces of "finite element" type introduced in paragraph 4.4.. So, in relation to scheme 6.1.7 we assume,

4. $$\mathcal{V} = \mathcal{V}(d,r,m),\ \rho \geq 0.5,\ U_0 = Rq_0$$

where $q_0 \in H_0^1(S)$ is the initial condition for 1.1.14.

In order to make the best use of this family of approximation subspaces (see Lemma 4.4.1) we also assume,

5. a_0 is O-regular in $H_0^1(S)$.

We can now present the following result:

Theorem 6.5.1 - Let conditions 1.,...,5. be satisfied. Assume that for the solution of equation 1.1.14 we have,

6. $$E\,\|Lq_t\|_{H^1(S)} < \infty, \qquad t \in [0,T].$$

Then, for the linear scheme 6.1.7, the following estimate holds:

$$\sup_k \left(\|q(t_k) - U_k\|_{L^2(\Omega,H)}\right) \leq C(h^{1/2} + h + d),$$

where C is a positive constant independent of h and d.

Proof of Theorem 6.5.1

Condition 6.3.9 is satisfied. From 5. and Lemma 4.4.1, condition 6.3.14 is also satisfied. So, the result above follows from Theorem 6.3.1. ●

Remark 6.5.1 - The crucial hypothesis of Theorem 6.5.1 is condition 6. and, as we pointed out before, (Remark 6.3.1), this condition can be interpreted in terms of the regularity of the solution of the stochastic evolution equation. In fact, assume that the coefficients of the Fokker-Planck operator, L_t, have first order bounded partial derivatives and that $E\,\|q_t\|_{H^3(S)} < \infty \quad t \in [0,T]$ for the solution of 1.1.14. It is easy to see that these conditions are sufficient to guarantee hypothesis 6. ●

Now, consider the quadratic scheme introduced in paragraph 6.4. As $\beta^1 = 0$, the operator A_1 in 5.3.8 satisfies 6.4.2. In order to satisfy 6.4.1 we must assume the function h to be invariant in time. So, we take

7. $\qquad h \in L^\infty(S)$

As we remarked before in section 6.4, this hypothesis is made with the intention of simplifying the steps leading 6.4.30. It does not constitute a fundamental condition and, in this case, results similar to 6.4.30 can be obtained by assuming $h \quad C^1(0,T;L^\infty(S))$.

The following result is a consequence of Theorem 6.4.1..

<u>Theorem 6.5.2</u> - Let conditions 1., 2., 4., 5., 7. be satisfied. Assume that for the solution of equation 1.1.14 we have,

8. $$E\,\|L^2 q_t\|_{L^2(S)}\ ,\ E\,\|Lhq_t\|_{L^2(S)} < \infty,$$
$$t \in [0,T].$$

Then, for the quadratic scheme 6.4.5 the following estimate holds:

$$\sup_k \|q(t_k) - \tilde{U}_k\|_{L^2(\Omega,H)} \leq C\{h + dh + d\},$$

where C is a positive constant independent of h and d.

Remark 6.5.2 - As in Theorem 6.5.1, the result depends on the regularity of the solution expressed here by condition 10. We assume that this condition is attained if the coefficients of the Fokker-Planck operator, L_t, have second order bounded partial derivatives, the functions $D_i h$, $D_{i,j} h$ belong to $C(0,T;L^\infty(S))$ and $E\|q_t\|^4_{H^4(S)} < \infty$, $t \in [0,T]$.

Regularity conditions for the solution of stochastic parabolic equations are discussed in Krylov-Rosovskii (|21|), Pardoux (|41|) and Levieux (|28|) (for the case $S = R^n$). In (|40|) Pardoux presents some conditions leading to a result of the type: $q \in M^2(0,T;H^2(S))$ (see Remark 5.3.2.). ●

Theorems 6.5.1 and 6.5.2 represent convergence results for discrete time Galerkin approximations of the solution of the stochastic evolution equation 1.1.14. defined in a cylinder $[0,T] \times S \subset R \times R^n$ under Dirichlet boundary conditions. These results show that, under certain regularity conditions, the linear scheme 6.1.7 produces a numerical approximation that converges at a rate $h^{1/2}$. On the other hand, under stronger regularity conditions, it is possible to obtain a faster rate of convergence by means of the quadratic scheme 6.4.5. In this circumstance, the rate is linear in the time increment. It goes without saying that, under the regularity conditions of Theorem 6.5.1, the quadractic scheme 6.4.5 also produces convergent approximations but, in this case, with a slower rate of convergence ($h^{1/2}$).

We observe that the rate of convergence in the "space discretization" can increase depending on how regular is the solution of the evolution equation (according to Lemma 4.4.1). However, the linear rate of convergence in the time increment achieved by the quadratic scheme can not be improved. We are led to this conclusion by the fact that, with respect to finite dimensional stochastic differential equations, the linear rate is the best possible rate of convergence for numerical procedures that depend on the values of the noise only at the dividing points of the partition of the time interval†. In our case, the numerical schemes can be viewed as schemes for approximating the solution of a finite dimensional equation (the continuous time Galerkin approximation). Therefore, we conclude that the linear rate must be the best possible rate of convergence for discrete time Galerkin approximations.

In |36|, McShane has presented a modified Euler scheme containing quadratic and cubic terms in the noise increment. His scheme converges at a linear rate for a wide class of finite dimensional stochastic differential equations. Here, we have seen that, for stochastic linear evolution equations, we do not need cubic terms in order to achieve the best rate of convergence.

According to Remark 6.5.2, in order to satisfy the regularity of Theorem 6.5.2, (condition 8.) we must include some requirements concerning the regularity of the function h. It is interesting to notice that these requirements are necessary in order to approximate the solution of the non stochastic counterpart of the equation 1.1.14 (see Theorem 4.4.1). As might be expected, schemes which are appropriate to the pathwise formula can be adapted for the approximation of the solution of the stochastic formula (and vice versa, since the relation between the non stochastic and the stochastic formulas is invertible; cf. equation 1.1.17). Also, it seems that the existence of a numerical procedure which converges to the solution of the pathwise formula at a rate $|\Delta_y^h|$, (the modulus of continuity of the observation sample path; see paragraph 4.4), corresponds to the existence of a procedure which converges at a linear rate to the solution of

† This fact has been shown by Clark, in |6|.

the stochastic formula. In |5|, Clark has presented an (Euler) method for approximating the pathwise solution of a filtering problem for Markov chains. It turns out that this scheme also represents an approximation procedure which converges at a linear rate to the solution of the stochastic version of the pathwise formula. Here, this aspect of the numerical schemes is not so evident. This is because, as we pointed out in Remark 5.3.3, the stochastic and the non stochastic formulas have different Galerkin approximations with respect to a given family of subspaces. However, it is not difficult to see that schemes which are appropriate to the pathwise version of the continuous time Galerkin approximation of 1.1.14 (cf. equation 5.3.12) can also produce approximations for the equation 1.1.14. In this case, one must be able to show that these schemes converge at a rate $|\Delta_y^h|$ to the pathwise formula, and at a linear rate to the stochastic formula.

It can be argued that the results of Theorems 6.5.1 and 6.5.2 are too restrictive vis-a-vis the class of filtering problem that satisfy the hypotheses of these theorems. This is so, because: 1) the operator L_t and the function h_t are assumed invariant in time; 2) we are considering only Dirichlet boundary conditions associated with equation 1.1.14.

As we pointed out before, the hypotheses concerning invariance in time can be relaxed. Assuming L_t and h_t of class C^1 one must be able to obtain results that are identical to those in the theorems.

With respect to the Dirichlet boundary conditions, we recall that these conditions are implicit in the assumption $V = H_0^1(S)$. Selecting instead $V = H^1(S)$, one should be able to consider Neumann conditions and again, similar results could be achieved. (In particular, Lemma 4.4.1 could be extended to approximation subspaces of $H^1(S)$; see e.g. Weeler, |49|).

The scope of applications of the results in both theorems can be enlarged, in order to include more complex situations. The conclusions concerning the rate of convergence in the time increment can be assumed as general results valid for discrete time Galerkin approximations of the solution of the filtering problem.

Finally we remark that numerical procedures for approximating the solution of the stochastic evolution equation that governs the unnormalized conditional density, has also been considered by Kushner and Levieux.

In |29|, Levieux has presented a numerical method which is similar to the one produced by our linear scheme (with $\rho = 1$, i.e., the backward implicit scheme). He shows that the method converges strongly in $L^2(\Omega \times (0,T) \times R^n)$ (Theorem IV.2 in |29|)

Kushner's method has a different conception. The basic idea lies in the approximation of the diffusion process by means of Markov chains. It turns out that the filter for the approximating chain converges to the filter for the diffusion. He shows that his method is robust in the sense we have described at the beginning of this work (see Kushner |25| and |26|).

In this work we have presented families of (one stage, Runge-Kutta) discrete time Galerkin procedures which possess the advantages of both Levieux's and Kushner's methods for approximating the solution of the filtering problem for diffusions. Schemes 6.1.7. and 6.4.5. produce approximations which converge uniformly in a L^2 sense and, in particular, scheme 6.4.5. has a maximum order of convergence with respect to the increment in time. On the other hand, schemes which are appropriate for the pathwise solution of the filtering problem (e.g. scheme 4.2.3.) produce robust approximations to the filtering solution.

REFERENCES

|1| ADAMS, R. Sobolev Spaces, Acad. Press, New York, 1975.

|2| BARROS-NETO, J. An Introduction to the Theory of Distributions, Marcel Dekker, New York, 1973.

|3| BUCY, R.C. Nonlinear filtering theory, IEEE Trans. Autom. Control 10, (1965), pp.198-199.

|4| BUTCHER, J.C. Implicit Runge-Kutta processes, Math. Comp. 18, (1964), pp.50-64.

|5| CLARK, J.M.C. The design of robust approximations to the stochastic differential equations of non-linear filtering, in "Communication Systems and Random Process Theory", ed. J.K. Skwirzynski, NATO Advanced Study Inst. Series, Sijthoff and Noordhoff, Alphen aan der Rijn, 1978.

|6| CLARK, J.M.C. and CAMERON, R.J. The maximum rate of convergence of discrete approximations for stochastic differential equations, Proc. International Symposium on Stochastic Differential Equations, Vilnius, 1978.

|7| CURTAIN, R.F. and FALB, P.L. Ito's lemma in infinite dimension, J. Math. Anal. Appl. 31, (1970), pp. 434-440.

|8| CURTAIN, R.F. and FALB, P.L. Stochastic differential equations in Hilbert space, J. Diff. Equations 10, (1971), pp. 412-430.

| 9 | DAVIS, M.H.A. Pathwise solutions and multiplicative functionals in nonlinear filtering, Proc. IEEE CDC Conference, Fort Lauderdale, Florida, 1979.

| 10 | DOOB, J.L. Stochastic Processes, Wiley, New York, 1963.

| 11 | DOSS, H. Liens entre equations differentielles stochastiques et ordinaires, Ann. Inst. H. Poincare XIII, (1977), pp. 99-125.

| 12 | DOUGLAS, J.Jr. and DUPONT, T. Galerkin methods for parabolic equations, SIAM J. Numer. Anal. 7, (1970), pp.576- 626.

| 13 | FUJISAKI, M., KALLIANPUR, G. and KUNITA, H. Stochastic differential equations for the non linear filtering problem, Osaka J. Math. 9, (1972), pp.19-40.

| 14 | GIKHMAN, I.I. and SKOROKHOD, A.V. Introduction to the Theory of Random Processes, Saunders, Philadelphia, 1969.

| 15 | GIRSANOV, I.V. On transforming a certain class of stochastic processes by absolutely continuous substitution of measures, Theory Prob. Appl. V, 3, (1960), pp.285- 301.

| 16 | HAUSSMANN, V.G. Asymptotic stability of the linear Ito equation in infinite dimensions, J. Math. Anal. Appl. 65, (1978), pp.219-235.

| 17 | JAZWINSKI, A.H, Stochastic Processes and Filtering Theory, Acad. Press, New York, 1971.

| 18 | KAILATH, T. An innovations approach to the least square estimation, Part I: linear filtering with additive white noise, IEEE Trans. Autom. Control 13, 6, (1968), pp.646-654.

|19| KALLIANPUR, G. and STRIEBEL, C. Estimation of stochastic systems: arbitrary system process with additive white noise observation errors, Ann. Math. Stat. 39, 3, (1968), pp.785-801.

|20| KATO, T. Perturbation Theory for Linear Operators, Springer-Verlag, Berlin, 1966.

|21| KRYLOV, N.V. and ROSOVSKII, B.L. The Cauchy problem for linear stochastic partial differential equations, Izvestia Akademia Nauk CCCP, Math. Series 41, 6, (1977), pp.1329-1347 (in Russian)

|22| KRYLOV, N.V. and ROSOVSKII, B.L. On the conditional distribution of diffusion processes, Izvestia Akademia Nauk CCCP, Math. Series 42, 2, (1978), pp.356-378 (in Russian)

|23| KUSHNER, H.J. On the differential equations satisfied by conditional probability densities of Markov processes with applications, SIAM J. Control 2, (1964), pp.106-119.

|24| KUSHNER, H.J. Dynamical equations for optimal nonlinear filtering, J. Diff. Equations, 3, (1967), pp.179-190.

|25| KUSHNER, H.J. Probability Methods for Approximations in Stochastic Control and for Elliptic Equations, Acad. Press, New York, 1977.

|26| KUSHNER, H.J. A robust discrete state approximation to the optimal nonlinear filter for a diffusion, Stochastics, Vol.3,n° 2, (1979), pp.75-83.

|27| LADYZENSKAYA, O.A., SOLONNIKOV, V.A. and URAL'CEVA, N.N. Linear and quasilinear equations of parabolic type, Transl. Math. Monogr. 23, Amer. Math. Soc., Providence R.I. (1968).

|28| LEVIEUX, F. Filtrage non-lineaire et analyse fonctionelle, Rapport LABORIA 57, 1974.

|29| LEVIEUX, F. Conception d'algorithmes parallelisables et convergents de filtrage recursif non-lineaire, Rapport LABORIA 235, 1977.

|30| LIONS, J.L. Equations Differentielles Operationelles, Springer-Verlag, Berlin, 1961.

|31| LIONS, J.L. Optimal Control for Systems Governed by Partial Differential Equations, Springer-Verlag, Berlin, 1971.

|32| LIONS, J.L. and MAGENES, E. Problemes aux limites non homogenes, Vol. I,II,III, Dunod, Paris 1968.

|33| LIPTSER, R.S. and SHIRYAEV, A.N. Nonlinear filtering of Markov diffusion processes, Trudy Mat. Inst. Steklov 104, (1968), pp.135-180.

|34| LIPTSER, R.S. and SHIRYAEV, A.N. Statistics of Random Processes, Springer-Verlag, New York,1977.

|35| McKEAN, H.P.Jr. Stochastic Integrals, Acad. Press, New York, 1969.

|36| McSHANE, E.J. Stochastic Calculus and Stochastic Models, Acad. Press, New York, 1974.

|37| MEYER, P.A. Sur un probleme de filtration, Universite de Strasbourg, Seminaire de Probabilites VII, Lecture Notes in Math. 321, Springer-Verlag, Berlin, 1973.

|38| NECAS, J. Les Methodes Directes en Theorie des Equations Elliptiques, Masson, Paris, 1967.

|39| NEVEU, J. Mathematical Foundations of the Calculus of Probability, Holden Day, San Francisco, 1965.

|40| PARDOUX, E. Filtrage de diffusions avec conditions frontieres, in Proc. Journees de Statistique dans le Processus Stochastiques, Lecture Notes in Math. 636, Springer-Verlag, 1978.

|41| PARDOUX, E. Stochastic partial differential equations and filtering of diffusion processes, Stochastics, Vol.3, n° 2, (1979), pp. 127-167.

|42| SCALORA, F.S. Abstract martingale convergence theorems, Pacific J. Math. 11, (1961), pp.347-374.

|43| SHOWALTER, R.E. Hilbert Space Methods for Partial Differential Equations, Pitman, 1977.

|44| STETTER, H.J. Analysis of Discretization Methods for Ordinary Differential Equations, Springer-Verlag, Berlin, 1973.

|45| STRANG, G. and FIX, G. An Analysis of the Finite Element Method, Prentice Hall, New Jersey, 1973.

|46| STRATONOVICH, R.L. Conditional Markov process, Theory Prob. Appl. 5, (1960), pp.156-178.

|47| STRATONOVICH, R.L. A new representation for stochastic integrals and equations, SIAM J. Control 4, 2, (1966), pp.362-371.

|48| WILSON, E.L. and NICKELL, R.E. Application of finite element method to heat conduction analysis, Nuclear Eng. Design 4, (1966), pp.276-286.

|49| WHEELER, M.F. A priori L_2 error estimates for Galerkin approximations to parabolic partial differential equations, SIAM J. Numer. Anal. 10, 4, (1973), pp.723-759.

|50| WONG, E. Stochastic Process in Information and Dynamical Systems, McGraw-Hill, New York, 1971.

|51| WONG, E. and ZAKAI, M. On the convergence of ordinary integrals to stochastic integrals, Ann. Math. Stat. 36, (1965), pp.1560-1564.

|52| WONHAM, W.M. Some applications of stochastic differential equations to optimal nonlinear filtering, SIAM J. Control, Ser.A,2,3,(1964), pp.347-369.

|53| YOSIDA, K. Functional Analysis, Springer- Verlag, Berlin, 1974.

|54| ZAKAI, M. On the optimal filtering of diffusion processes, Z. Wahrscheinlichkeitstheorie verw. Geb. 11, (1969), pp.230-243.

|55| ZLAMAL, M. Unconditionally stable finite element schemes for parabolic equations, in "Topics in Numerical Analysis" Vol. II, ed. J.J.H. Miller, Acad. Press, London, 1975.