Control
and Cybernetics

VOL. 9 (1980) No. 4

Convergence of approximations to quadratic optimal control problems with amplitude constrained control

by

KAZIMIERZ MALANOWSKI

Systems Research Institute
Polish Academy of Sciences

A method of estimating the rate of convergence of approximations to quadratic optimal control problems with control subject to amplitude constraints is presented.

In the method the form of the optimal control is exploited.

The obtained general results are appiied to get the estimation of the rate of convergence of Galerkin type approximations to a boundary control problems for a linear parabolic system, with the cost functional depending on terminal state.

1. Introduction

Optimal control problems of distributed systems as a rule can not be solved in an analytical way and necessitate numerical approach.

This in turn requires approximation of initial problem by some finite-dimensional problems of optimization depending on a parameter of discretization.

The approximations should be defined in such a way that the solutions of discretize problems converge to the solution of the initial one, when the parameter of discretization tends to zero.

In some cases of convex optimal control problems it is possible to estimate the rate of this convergence.

This estimation becomes more difficult if control or state of the system is subject to inequality constraints.

The most typical example of such constraints are amplitude constraints of the control function.

In this case to obtain estimations of the rate of convergence of approximations two different approaches have been applied.

One is to use necessary conditions of optimality in variational inequalities form [3]. This approach in general does not allow to obtain optimal estimations.

In the alternative approach the Lagrange formalism is used (cf. [4, 7, 8]). This in turn requires investigation of regularity and convergence of Lagrange multipliers corresponding to control constraints. It is rather complicated procedure.

In this paper another approach is proposed in which the form of optimal control is exploited directly to obtain appropriate estimations. This approach was first applied by W. Hackbusch to investigate convergence of multi-grid method [5].

The general result obtained in Section 1 shows that to estimate the rate of convergence of approximations it is enough to estimate the distance between the optimal solution to the initial problem and its projection on appropriate subspaces as well as the convergence of approximations to the state and adjoint equations.

In these estimations regularity of optimal solutions to the initial problem plays a crucial role.

All investigations are performed for linear-quadratic problem, but can be easily generalized to the case of strongly convex cost functional.

The obtained results are applied in Section 2 and 3 to estimate the rate of convergence of discrete-type Galerkin approximations to boundary control problem of a parabolic system with the cost functional depending on terminal state.

1. Optimal control problem and its approximation

Let U and Z be two Hilbert spaces with inner products and norms denoted by $(\cdot, \cdot)$, $((\cdot, \cdot))$ and $|\cdot|$, $\|\cdot\|$ respectively.

It will be assumed that $U=L^2(\Xi)$ where Ξ is a bounded domain in R^n

On $Z\times U$ there is defined the quadratic functional

$$\hat{J}(z, u)=\frac{1}{2}[\|z-w\|^2+\lambda|u|^2] \tag{1.1}$$

where $w\in Z$ is a given element.

Let

$$z: U \to Z$$

be a continuous affine mapping.

The derivative of z (the linear part of the mapping) will be denoted by q.

Define

$$J: U\to R^1$$

$$J(u)=\hat{J}(z(u), u).$$

We have

$$DJ(u)=q^*(z(u)-w)+\lambda u \tag{1.2}$$

$$(v, D^2 J(u)v)=(v, (q^* q+\lambda I)v)\geq\lambda|v|^2 \qquad \forall u, v\in U \tag{1.3}$$

We introduce a closed, convex subset $V\subset U$:

$$V=\{u\in U=L^2(\Xi) \,\big|\, |u(\xi)|\leq 1 \quad \text{for a.a. } \xi\in\Xi\} \tag{1.4}$$

The following abstract problem of optimization (P) corresponds to quadratic optimal control problems subject to amplitude constraints of the control function:

(P)

find $u^0 \in V$ *such that*

$$J(u^0) = \inf_{u \in V} J(u) \tag{1.5}$$

It is well known (cf.e.g. [9]) that Problem (P) has a unique solution u^0, which can be characterized by the following condition

$$(DJ(u^0), u-u^0) = (q^*(z(u^0)-w)+\lambda u^0, u-u^0) \geqslant 0 \quad \forall u \in V \tag{1.6}$$

It is easy to see that condition (1.6) is equivalent to the following one

$$u^0 - \mathrm{sg}(-DJ(u^0)+u^0) = 0 \tag{1.7}$$

where

$$\mathrm{sg}: U \to U$$

is defined as follows

$$\mathrm{sg}\, f(\xi) = \begin{cases} 1 & \text{if} \quad 1 < f(\xi) \\ f(\xi) & \text{if} \quad -1 \leqslant f(\xi) \leqslant 1 \\ -1 & \text{if} \quad f(\xi) < -1 \end{cases}$$

Note that in case where control constraints are not active, or where in (1.5) V is substituted by U, (1.7) reduces to

$$DJ(u^0) = 0 \tag{1.7a}$$

We consider finite-dimensional approximations to Problem (P).

Let h be a parameter of discretization destinated to tend to zero, and let $U_h \subset U$ and $Z_h \subset Z$ be finite-dimensional subspaces of U and Z respectively depending on h and

$$\overline{\bigcup_{h>0} U_h} = U, \qquad \overline{\bigcup_{h>0} Z_h} = Z.$$

Denote by

$$R_h^U: U \to U_h$$

and

$$R_h^Z: Z \to Z_h$$

operators of orthogonal projections on appropriate finite dimensional subspaces in U and Z.

Let

$$z_h: U_h \to Z_h$$

be continuous affine mappings approximating z.

It is assumed that the usual stability and consistency conditions are satisfied [1].

We shall denote by q_h the derivative (linear part) of z_h.

Define

$$J_h: U_h \to R^1 \qquad J_h(u_h) \triangleq \hat{J}(z_h(u_h), u_h) \tag{1.8}$$

Similarly to (1.2), (1.3) we have

$$DJ_h(u_h) = R_h^U q_h^* (z_h(u_h) - w_h) + \lambda u_h \tag{1.9a}$$

$$(v_h, D^2 J_h(u_h) v_h) = (v_h, (q_h^* q_h + \lambda I) v_h) \geq \lambda |v_h| \ \forall u_h, v_h \in U_h \tag{1.9b}$$

Denote

$$V_h = V \cap U_h = \{u_h \in U_h | |u_h(\xi)| \leq 1 \text{ for a.a. } \xi \in \Xi\} \tag{1.10}$$

Problem (P) is approximated by the family of the following finite-dimensional problems of optimization (P_h):

(P_h)

$$\textit{find } u_h^0 \in V_h \textit{ such that } J_h(u_h^0) = \inf_{u_h \in V_h} J_h(u_h) \tag{1.11}$$

Like in (P) Problems (P_h) have unique solutions. By (1.9) and (1.10) these solutions are uniformly bounded

$$|u_h^0| \leq c \tag{1.12}$$

It will be assumed that

$$\text{sg}\, u_h \in U_h \quad \forall u_h \in U_h. \tag{1.13}$$

Note that condition (1.13) impose some restrictions on the form of subspaces U_h. This condition holds if U_h are spaces of piece-wise constant functions.

If (1.13) is satisfied then u_h^0 can be characterized by the equation analogous to (1.7):

$$u_h^0 - \text{sg}(-DJ_h(u_h^0) + u_h^0) = 0 \tag{1.14}$$

Equations (1.7) and (1.14) will be used to estimate the difference between u^0 and u_h^0.

Denote

$$K: U \to U, \qquad K(u) \triangleq u - \text{sg}(-DJ(u) + u) \tag{1.15a}$$

and

$$K_h: U_h \to U_h, \qquad K_h(u_h) \triangleq u_h - \text{sg}(-DJ_h(u_h) + u_h) \tag{1.15b}$$

Subtracting (1.7) from (1.14) and using the above notation we get

$$K_h(u_h^0) = K(u^0)$$

Adding to both sides of this equation $-K_h(v_h)$, where

$$v_h \triangleq R_h^U u^0 \tag{1.16}$$

we obtain

$$K_h(u_h^0) - K_h(v_h) = K(u^0) - K_h(v_h) \tag{1.17}$$

This equation will be used to estimate $|u^0 - u_h^0|$. First we shall investigate the left hand side of (1.17). To this end let us introduce the function of real parameter s defined by

$$k_h(s) \triangleq K_h(v_h + s(u_h^0 - v_h)) \quad \text{for} \quad s \in [0, 1] \tag{1.18}$$

It is obvious that k_h is a Lipschitz continuous function, therefore it is differentiable almost everywhere and the following equality takes place [12]:

$$K_h(u_h^0)-K_h(v_h)=k_h(1)-k_h(0)=\int_0^1 \frac{dk_h(s)}{ds}\,ds \tag{1.19}$$

Let $s\in[0,1]$ be any point at which k_h is differentiable and denote

$$v_h^s=v_h+s(u_h^0-v_h).$$

Define a linear operator

$$L_h(u_h^0, v_h, s)\colon U_h\to U_h$$

putting

$$L_h(u_h^0, v_h, s)\,u_h(\xi)=\begin{cases} u_h(\xi) & \text{if } |-DJ_h(v_h^s)(\xi)+v_h^s(\xi)|\geqslant 1\\ D^2J_h(v_h^s)\,u_h(\xi) & \text{if } |-DJ_h(v_h^s)(\xi)+v_h^s(\xi)|<1\end{cases} \tag{1.20}$$

for almost all $\xi\in\Xi$.

Note that in case where

$$\operatorname{meas}\{\xi\in\Xi\,|\,|-DJ_h(v_h^s)(\xi)+v_h^s(\xi)|=1\}=0$$

L_h is the Gateaux derivative of K_h at v_h^s.

It follows from the definition (1.15) (cf. Appendix) that

$$\frac{dk_h(s)}{ds}=L_h(u_h^0, v_h, s)(u_h^0-v_h). \tag{1.21}$$

From (1.19) and (1.21) we get

$$K_h(u_h^0)-K_h(v_h)=\int_0^1 L_h(u_h^0, v_h, s)\,ds\,(u_h^0-v_h). \tag{1.22}$$

Let us define the linear operator

$$\mathscr{K}_h(u_h^0, v_h)\colon U_h\to U_h, \qquad \mathscr{K}_h(u_h^0, v_h)\triangleq\int_0^1 L_h(u_h^0, v_h, s)\,ds. \tag{1.23}$$

By (1.9b), (1.20) and (1.23) we have

$$(u_h, \mathscr{K}_h(u_h^0, v_h)\,u_h)\geqslant\gamma|u_h|^2 \qquad \forall u_h\in U_h \tag{1.24}$$

where

$$\gamma=\min\{1, \lambda\} \tag{1.24a}$$

Substituting (1.22) and (1.23) to (1.19) we get

$$\mathscr{K}_h(u_h^0, v_h)(u_h^0-v_h)=K(u^0)-K_h(v_h)$$

Let us take inner products of both sides of this equation by $(u_h^0-v_h)$. Using (1.15), (1.24) as well as the norm inequality and the fact that

$$|\operatorname{sg} f-\operatorname{sg}\varphi|\leqslant|f-\varphi|$$

we obtain

$$|u_h^0-v_h|\leqslant\frac{1}{\gamma}|K(u^0)-K_h(v_h)|\leqslant\frac{1}{\gamma}[|u^0-v_h|+|-\mathrm{sg}(-DJ(u^0)+u^0)+$$

$$+\mathrm{sg}(-DJ(v_h)-v_h)|]\leqslant\frac{1}{\gamma}[|DJ(u^0)-DJ_h(v_h)|+2|u^0-v_h|] \tag{1.25}$$

Taking advantage of (1.2) and (1.9) after simple rearrangements we get

$$|DJ(u^0)-DJ_h(v_h)|=[q^*(z(u^0)-w)+\lambda u^0]-[R_h^U q_h^*(z_h(v_h)-w_h)+\lambda v_h]|\leqslant$$

$$\leqslant|q^*(z(u^0)-w)-R_h^U q_h^*(z_h(v_h)-w_h)|+\lambda|u^0-v_h|\leqslant|q^*(z(u^0)-w)+$$

$$-R_h^U q^*(z(u^0)-w)|+|R_h^U[q^*(z(u^0)-w)-q_h^*(z_h(v_h)-w_h)]|+\lambda|u^0-v_h|\leqslant$$

$$\leqslant|q^*(z(u^0)-w)-R_h^U q^*(z(u^0)-w)|+|(q^*-q_h^* R_h^Z)(z(u^0)-w)|+$$

$$+|q_h^* R_h^Z(z(u^0)-w)-q_h^*(z_h(v_h)-w_h)|+\lambda|u^0-v_h| \tag{1.26}$$

Using the fact that $\|q_h^*\|$ is bounded uniformly with respect to h we get

$$|q_h^* R_h^Z(z(u^0)-w(-q_h^*)z_h(v_h)-w_h)|=|q_h^* R_h^Z[z(u^0)-z_h(v_h)]|\leqslant$$

$$\leqslant\|q_h^*\|\,\|z(u^0)-z_h(v_h)\|\leqslant c\|z(u^0)-z_h(v_h)\| \quad *) \tag{1.27}$$

Taking into account that

$$|u_h^0-u^0|\leqslant|u_h^0-v_h|+|u^0-v_h|$$

we finally obtain from (1.25) through (1.27) the following

THEOREM 1.1

If condition (1.13) *is satisfied then the difference between the solution* u^0 *of Problem* (P) *and the solution* u_h^0 *of Problem* (P_h) *can be estimated as follows*

$$|u^0-u_h^0|\leqslant c[|u^0-v_h|+|q^*(z(u^0)-w)-R_h^U q^*(z(u^0)-w)|+$$

$$+|(q^*-q_h^* R_h^Z)(z(u^0)-w)|+\|z(u^0)-z_h(v_h)\|] \tag{1.28}$$ ▼

Note that all terms on the right-hand side of (1.28) contain only optimal solution u^0 of (P) and its projection v_h on U_h.

By properties of the space U_h the first 2 terms tend to zero together with h. If approximation z_h of z is convergent also the other 2 terms converge to zero.

It follows from (1.28) that in estimating of rate of convergence of $|u_h^0-u^0|$ an a priori information on regularity of optimal solution to (P) plays crucial role. Such an information is used in estimation of the distance between given elements and their projections as well as in estimation of the rate of convergence of approximations z_h to z (respectively q_h^* to q^*).

It will be illustrated by an example in next sections.

*) Letter c denotes a generic contstant not necessarily the same in different places

2. Boundary control of parabolic system. Regularity results

Let Ω be a bounded domain (open set) in R^n with properly regular boundary Γ. For the sake of simplicity we shall assume that Γ is of class C^∞. Moreover we assume that locally Ω is situated on one side of Γ.

The following functional space defined on Ω (resp. Γ) will be used in the sequel. For their precize definitions and properties see [10].
$H^0(\Omega)=L^2(\Omega)$—space of measurable functions, square integrable on Ω.
$H^s(\Omega)$—Sobolev space of fractional order s defined on Ω.
Let T be a fixed time. Denote

$$Q=\Omega\times[0, T], \quad \Sigma=\Gamma\times[0, T].$$

On Q (resp. Σ) we define the spaces

$$H^r(0, T; H^s(\Omega))$$

which are Sobolev spaces of order r with respect to t with the range in $H^s(\Omega)$. Denote

$$H^{r,s}(Q)=H^0(0, T; H^r(\Omega))\cap H^s(0, T; H^0(\Omega)).$$

For a sufficiently regular function y defined on Q by

$$y|_\Omega \quad \text{and} \quad y|_\Sigma$$

we denote its traces on Ω and Σ respectively.

In the sequel we shall need the following known results ([10] vol. 2, p. 10) concerning regularity of traces:

LEMMA 2.1
Let $y\in H^{2q,q}(Q)$
then

$$y|_\Omega\in H^{2q-1}(\Omega) \quad \textit{for} \quad q>\frac{1}{2}$$

$$y|_\Sigma\in H^{2(q-\frac{1}{4}),q-\frac{1}{4}}(\Sigma) \quad \textit{for} \quad q>\frac{1}{4}$$

and the mappings

$$y\dashrightarrow y|_\Omega \quad \textit{and} \quad y\dashrightarrow y|_\Sigma$$

are continuous in respective topologies. ▼

By $\mathscr{L}(X; Y)$ we shall denote the space of linear, continuous operators from X into Y.

Consider the system described in the cylinder Q by the following parabolic equation (state equation):

$$\frac{\partial y(x, t)}{\partial t}-Ay(x, t)=0 \tag{2.1}$$

where

$$Ay(x) \triangleq \sum_{i,j=1}^{n} \frac{\partial}{\partial x_j}\left(a_{ij}(x)\frac{\partial y(x)}{\partial x_i}\right) - a_0(x)\,y(x), \tag{2.1a}$$

the functions $a_{ij}(\cdot)=a_{ji}(\cdot)$ are properly regular (for the sake of simplicity of class C^∞) and there exists such a constant $\rho_0>0$ that

$$\sum_{i,j=1}^{n} a_{ij}(x)\,\xi_i\,\xi_j \geqslant \rho_0 \sum_{i=1}^{n} \xi_1^2 \qquad \forall x\in\Omega, \quad \forall \xi_i, \xi_j \in R^1 \tag{2.1b}$$

For (2.1) the Neumann type boundary conditions are satisfied:

$$\frac{\partial y(\sigma,t)}{\partial \eta_A} \triangleq \sum_{i,j=1}^{n} a_{ij}(\sigma)\frac{\partial y(\sigma,t)}{\partial \sigma_j}\cos(\eta,\sigma_i)=g(\sigma,t) \quad \text{in } \Sigma \tag{2.2}$$

where η is the unit outward normal to Γ and g is a properly regular function defined on Σ.

Moreover the initial condition

$$y(x,0)=y^p(x) \quad \text{in } \Omega \tag{2.3}$$

is satisfied where y^p is a properly regular function defined on Ω.
In the sequel it will be assumed that

$$y^p \in H^{3/2}(\Omega). \tag{2.3a}$$

The solution of (2.1)–(2.3) will be understood in the weak sense (cf. [10]) i.e. as the properly regular function for which the follwoing identity holds

$$\left(\frac{dy(t)}{dt}, \varphi\right) + a(y(t),\varphi) = \langle g(t), \varphi\rangle \quad \forall \varphi\in H^1(\Omega) \text{ for a.a. } t\in[0,T] \tag{2.4}$$

along with (2.3).
Where

$$a(y,\varphi) \triangleq \int_\Omega \left[\sum_{i,j=1}^{n} a_{ij}(x)\frac{\partial y(x)}{\partial x_i}\,\frac{\partial \varphi(x)}{\partial x_j} + a_0(x)\,y(x)\,\varphi(x)\right]dx \tag{2.5}$$

$(\cdot,\cdot)$ — denotes inner product in $L^2(\Omega)$ and $\langle\cdot,\cdot\rangle$— inner product in $L^2(\Gamma)$ extended by continuity to $H^{-1/2}(\Gamma)\times H^{1/2}(\Gamma)$.

We shall need the following lemma concerning existence and regularity of boundary value problem (2.1)–(2.3), which is a particular case of results presented in [10] (vol. 2, pp. 69 and 84)

LEMMA 2.2

Let $\quad s\in[-1/4,\, 1/4).$

If $\quad g\in H^{2(s+\frac{1}{4}),\, s+\frac{1}{4}}(\Sigma)$

$\quad y^p\in H^{2s+1}(\Omega)$

then there exists a unique solution

$$y \in H^{2(s+1),s+1}(Q)$$

of (2.1)–(2.3) *which continuously depends on* g, y^p. ▼

Remark 2.3

Results similar to those in Lemma 2.1 are also true for $s \notin [-1/4, 1/4)$, but the case $s < -1/4$ is not interesting for our applications while for $s > 1/4$ additional compatibility conditions must be satisfied, which usually are not met in boundary control problems. ▼

We shall assume that the space of control

$$U = H^{0,0}(\Sigma) \tag{2.6}$$

and we put

$$g(t) = Cu(t) \tag{2.6a}$$

where

$$C \in \mathscr{L}(H^r(\Sigma), H^r(\Sigma)) \quad 0 \leqslant r \leqslant 1\,. \tag{2.6b}$$

The set V of admissible control is defined as in (1.4):

$$V = \{u \in H^{0,0}(\Sigma) \,|\, |u(\sigma, t)| \leqslant 1 \quad \text{for a.a.} \quad (\sigma, t) \in \Sigma\}\,. \tag{2.7}$$

As the space Z of output we take

$$Z = H^0(\Omega)$$

and we put $z(u) = y(T; u)$, where $y(\cdot, u)$ denotes the solution of (2.1)–(2.3) along with (2.6a)

Hence the functionals $\hat{J}(z, u)$ and $J(u)$ take on the form

$$\hat{J}(z, u) = \frac{1}{2}\left[\|z - w\|^2_{H^0(\Omega)} + \lambda \|u\|^2_{H^{0,0}(\Sigma)}\right] \tag{2.8}$$

$$J(u) = \hat{J}(z(u), u) = \frac{1}{2}\left[\|y(T; u) - w\|^2_{H^0(\Omega)} + \lambda \|u\|^2_{H^{0,0}(\Sigma)}\right] \tag{2.9}$$

It will be assumed that

$$w \in H^{3/2}(\Omega)\,. \tag{2.9a}$$

Now we can formulate Problem (P^1) of optimal control

(P^1)

find $u^0 \in V$ *such that*

$$J(u^0) = \inf_{u \in V} J(u)\,.$$ ▼

Problem (P^1) is a special case of Problem (P) hence it has a unique solution which can be characterized by (1.6) or (1.7).

Using results of Section 1 we shall investigate convergence of some finite dimensional approximations to (P^1).

To this and we shall need informations on regularity of the solution to (P^1). This informations will be obtained by analyzing (1.7).

The gradient of the functional (2.9) takes on the form

$$DJ(u)=q^*(y(T;u)-w)+\lambda u=C^* p(u)|_\Sigma+\lambda u \tag{2.10}$$

where $p(u)$ is the solution of the following adjoint equation ([9] p. 139).

$$\frac{\partial p(u)}{\partial t}+Ap(u)=0 \tag{2.11}$$

$$\frac{\partial p(u)}{\partial \eta_A}=0 \tag{2.11a}$$

$$p(T;u)=y(T;u)-w \tag{2.11b}$$

From (1.7), and (2.10) we get

$$u^0=\max\left\{-1, \min\left\{1, -\frac{1}{\lambda}C^* p(u^0)|_\Sigma\right\}\right\} \tag{2.12}$$

Since $u^0 \in H^{0,0}(\Sigma)$ substituting $s=-1/4$ and taking into account (2.3a) by Lemma 2.2 we get

$$y(u^0)\in H^{3/2,3/4}(Q). \tag{2.13}$$

Hence by Lemma 2.1

$$y(T;u^0)\in H^{1/2}(\Omega) \tag{2.14}$$

By (2.8a), (2.11), (2.14) and Lemma 2.2 we have

$$p(u^0)\in H^{3/2,3/4}(Q) \tag{2.15}$$

and by Lemma 2.1

$$p(u^0)|_\Sigma \in H^{1,1/2}(\Sigma) \tag{2.16}$$

Condition (2.6b) together with (2.16) yield

$$C^* p(u^0)|_\Sigma \in H^{1,1/2}(\Sigma). \tag{2.17}$$

On the other hand it is known [11] that:

$$\text{if} \quad f\in H^{r,s}(Q)$$

then

$$\varphi \triangleq \max\{1, f\}\in H^{p,q}(Q) \tag{2.18}$$

where $p=\min\{1, r\}$, $q=\min\{1, s\}$.
Therefore (2.12) together with (2.17) imply

$$u^0 \in H^{1,1/2}(\Sigma) \tag{2.19}$$

Taking advantage of (2.19) and reiterating the whole process with $s=1/4-\varepsilon$ we obtain the following:

THEOREM 2.4

Let $y^p, w \in H^{3/2}(\Omega)$, *and condition* (2.6b) *be satisfied then*

$$y(u^0), p(u^0) \in H^{5/2-2\varepsilon, 5/4-\varepsilon}(Q) \tag{2.20}$$

$$u^0 \in H^{1, 1-\varepsilon}(\Sigma) \tag{2.21}$$ ▼

Note that the limiting factor in the obtained regularity of $y(u^0)$ and $p(u^0)$ is not the regularity of u^0, but compatibility condition (cf. [10] vol. 2, p. 69) Due to this fact we can not use condition (2.21) to repeat again the process (2.13)–(2.17).

3. Boundary control of parabolic system. Approximation

To introduce finite dimensional approximations to Problem (P^1) first we must define subspaces U_h and Z_h.

To this end in the domain Ω we introduce a family of grids depending on the parameter h.

On these grids we define spaces $G_h^0(\Omega)$ and $G_h^1(\Omega)$ of piece-wise constant and piece-wise linear finite elements respectively (cf. [2]).

Let

$$R^i_{h,\Omega}: H^0(\Omega) \to G^i_h(\Omega) \qquad i=0, 1$$

denote the operators of orthogonal projection on appropriate spaces.

It is assumed that the grids are introduced in such a way that

$$\|v - R^i_{h,\Omega} v\|_{H^\alpha(\Omega)} \leqslant ch^{\beta-\alpha} \|v\|_{H^\beta(\Omega)} \qquad \forall v \in H^\beta(\Omega) \tag{3.1}$$

where

$$\alpha \in [0, i], \qquad \beta \in [\alpha, i+1], \qquad i=0, 1.$$

Moreover for $G_h^1(\Omega)$ the following inverse property [2] is satisfied

$$\|v\|_{H^1(\Omega)} \leqslant ch^{\beta-1} \|v\|_{H^\beta(\Omega)} \qquad \forall v \in G_h^1(\Omega) \tag{3.2}$$

where $0 < \varepsilon \leqslant \beta \leqslant 1$.

Let $G_h^0(\Gamma)$ and $G_h^1(\Gamma)$ denote the spaces of traces of $G_h^0(\Omega)$ and $G_h^1(\Omega)$ respectively on Γ.

can be shown [2] that conditions similar to (3.1) hold:

$$\|v - R^i_{h,\Gamma} v\|_{H^0(\Gamma)} \leqslant ch^\alpha \|v\|_{H^\alpha(\Gamma)}, \qquad v \in H^\alpha(\Gamma), \qquad \alpha \in [0, 1] \tag{3.3}$$

where $R^i_{h,\Gamma}$ $(i=0, 1)$ denote orthogonal projections on $G_h^i(\Gamma)$.

To discretize the functions define on Q the interval of control $[0, T]$ is divided into $T/\tau = N(\tau)$ subintervals of the length τ.

It is assumed that there exist such constants $0 < c_1 \leqslant c_2$ that

$$c_1 h^2 \leqslant \tau \leqslant c_2 h^2 \tag{3.4}$$

Hence h can be considered as the only parameter of discretization.

On the cylinder $Q=\Omega\times(0, T)$ we define a family of functions piece-wise constant on intervals $[j\tau, (j+1)\tau)$:

$$Y_h^{i,0}(Q)=\{y_h | y_h(t)=y_h(j\tau) \quad \text{for} \quad t\in[j\tau,(j+1)\tau), \quad y_h(j\tau)\in G_h^i(\Omega)\} \tag{3.5}$$

By

$$R_{h,Q}^{i,0}: H^{0,0}(Q)\to Y_h^{i,0}(Q) \quad (i=0, 1) \tag{3.6a}$$

$$R_{h,\Sigma}^{i,0}: H^{0,0}(\Sigma)\to Y_h^{i,0}(\Sigma) \quad (i=0, 1) \tag{3.6b}$$

we denote operators of orthogonal projections on appropriate spaces.

Condition (3.3) and definition of $R_{h,\Sigma}^{i,0}$ imply:

$$\|v-R_{h,\Sigma}^{i,0}v\|_{H^{0,0}(\Sigma)}\leqslant c\,[h^\beta+\tau^\gamma]\,\|v\|_{H^{\beta,\gamma}} \quad \forall v\in H^{\beta,\gamma}(\Sigma), \quad \beta, \gamma\in[0, 1]. \tag{3.7}$$

We put

$$U_h=Y_h^{0,0}(\Sigma), \tag{3.8}$$

$$Z_h=G_h^1(\Omega). \tag{3.9}$$

The state equation (2.4) is approximated by a discrete-time Galerkin scheme. Namely we define a unique function $y_h\in Y_h^{1,0}(Q)$ which satisfies the equation

$$(\nabla y_h(k\tau), \varphi_h)+a\,(y_{h,\theta}(k\tau), \varphi_h)=\langle\bar{g}(k\tau), \varphi_h\rangle \tag{3.10}$$

$$\forall\varphi_h\in G_h^1(\Omega), \quad k=0, 1, \ldots, N(\tau)-1$$

$$y_h(0)=R_{h,\Omega}^1\, y^p \tag{3.10a}$$

where

$$\bar{g}(k\tau)=\frac{1}{\tau}\int_{k\tau}^{(k+1)\tau} g(t)\,dt \tag{3.11}$$

$$\nabla y_h(k\tau)\triangleq\frac{y_h((k+1)\tau)-y_h(k\tau)}{\tau} \tag{3.12a}$$

$$y_{h,\theta}(k\tau)\triangleq\theta y_h((k+1)\tau)+(1-\theta)\,y_h(k\tau) \tag{3.12b}$$

and $a(.,.)$ is given by (2.5).

It is assumed that

$$\theta=\frac{1}{2} \quad \text{(Crank–Nicholson scheme)} \tag{3.13a}$$

or

$$\theta=1 \quad \text{(implicite scheme)} \tag{3.13b}$$

In the sequel we shall use the following results due to Hackbusch (cf. [6] Lemmas 4.1 and 4.2) concerning convergence of solutions of (3.10) to the solution of (2.1)–(2.3)

LEMMA 3.1

Let y be the solution of (2.1)–(2.3) *and* y_h *the solution of* (3.10) *where* (3.4) *and* (3.13) *are satisfied then the following estimates hold*

$$\|y-y_h\|_{H^{2r,0}(Q)} \leqslant Ch^{2(s-r)}\|y\|_{H^{2s,s}(Q)} \tag{3.14}$$

$$\text{for} \quad r\in\left[0,\frac{1}{2}\right], \quad s\in\left(\frac{1}{2},1\right]$$

$$\|y(T)-y_h(T)\|_{H^0(\Omega)} \leqslant Ch^{2s-1}\|y\|_{H^{2s,s}(Q)} \tag{3.15}$$

$$\text{for} \quad s\in\left(\frac{1}{2},1\right].$$

▼

To approximate the control problem considered in Section 2 we put in (3.10)

$$\bar{g}(k\tau)=Cu_h(k\tau) \tag{3.16}$$

The set of admissible control is defined by (2.7) where U is substituted by U_h. The cost functional to be minimize is

$$J_h(u_h)=\hat{J}(z_h(u_h),u_h)=\frac{1}{2}\left[\|y_h(T;u_h)-w\|_{H^0(\Omega)}+\lambda\|u_h\|_{H^{0,0}(Q)}\right]. \tag{3.17}$$

Problem (P^1) is approximated by the family of the following Problems (P_h^1)

(P_h^1)

find $u_h^0\in V_h$ *such that*

$$J_h(u_h^0)=\inf_{u_h\in V_h} J_h(u_h) \tag{3.18}$$

▼

Problems (P_h^1) have unique solutions.

We shall use Theorem 1.1 to estimate

$$\|u^0-u_h^0\|_{H^{0,0}(Q)}.$$

It is easy to see that in this case like in (2.10), (2.11)

$$q_h^*(z)=C^*p_h|_\Sigma \tag{3.19}$$

where $p_h\in Y^{1,0}(Q)$ satisfies the following equation adjoint to (3.10)

$$(\nabla p_h(k\tau),\varphi_h)-a(p_{h,(1-\theta)}(k\tau),\varphi_h)=0$$

$$\forall\varphi_h\in G_h^1(\Omega), \quad k=N(\tau)-1,\dots,1,0 \tag{3.20}$$

$$p_h(N(\tau)\tau)=p_h(T)=z. \tag{3.20a}$$

Note that changing direction of time we can apply Lemma 3.1 to adjoint equations (2.11) and (3.20).

Now we are in position to use Theorem 1.1 to estimate

$$\|u^0-u_h^0\|_{H^{0,0}(\Sigma)}.$$

Taking advantage of (1.16), (2.21). (3.4), (3.7) and (3.8) we obtain

$$\|u^0-v_h\|_{H^{0,0}(\Sigma)}\leqslant c\,[h+\tau^{(1-\varepsilon)}]\|u^0\|_{H^{1,1-\varepsilon}(\Sigma)}\leqslant ch \tag{3.21}$$

By (2.6b), (2.9), (2.10), (2.20) and by Lemma 2.1 we get

$$q^* \left(y(T; u^0) - w\right) \in H^{2-2\varepsilon, 1-\varepsilon}(\Sigma).$$

Hence by (3.4) and (3.7)

$$\|q^* \left(y(T; u^0) - w\right) - R_h^U q^* \left(y(T; u^0) - w\right)\|_{H^{0,0}(\Sigma)} \leqslant ch. \tag{3.22}$$

Note that by (3.9), (3.19) and (3.20)

$$q_h^* (R_h^Z z) = q_h^* (z).$$

Hence by (2.10) and (3.19) we have

$$\|(q - q_h^*) \left(y(T; u^0) - w\right)\|_{H^{0,0}(\Sigma)} \leqslant c \, \|p(u^0)|_\Sigma - p_h(u^0)|_\Sigma\|_{H^{0,0}(\Sigma)} \tag{3.23}$$

Well known theorem of trace (cf. [10] vol. 1, p. 47) implies

$$\|p|_\Sigma\|_{H^{0,0}(\Sigma)} \leqslant c \, \|p\|_{H^{1/2+2\varepsilon,0}(Q)} \qquad \varepsilon > 0.$$

Hence substituting in (3.14) $r = 1/4 + \varepsilon$, $s = 1$ and taking into account (2.20) as well as (3.23) we obtain

$$\|(q^* - q_h^* R_h^Z) \left(y(T; u^0) - w\right)\|_{H^{0,0}(\Sigma)} \leqslant c \, \|p(u^0) - p_h(u^0)\|_{H^{1/2+2\varepsilon,0}(Q)} \leqslant ch^{3/2-2\varepsilon}. \tag{3.24}$$

By stability of Galerkin approximation (3.10) we have

$$\|y(T; u^0) - y_h(T; v_h)\|_{H^0(\Omega)} \leqslant \|y(T; u^0) - y_h(T; u^0)\|_{H^0(\Omega)} +$$
$$+ \|y_h(T; u^0) - y_h(T; v_h)\|_{H^0(\Omega)} \leqslant \|y(T; u^0) - y_h(T; u^0)\|_{H^0(\Omega)} + c \, \|u^0 - v_h\|_{H^{0,0}(\Sigma)}$$

Putting in (3.15) $s = 1$ and taking into account (3.21) we get

$$\|y(T; u^0) - y_h(T; v_h)\|_{H^0(\Omega)} \leqslant ch \tag{3.25}$$

Substituting (3.21), (3.22), (3.24) and (3.25) into (1.28) we finally obtain

THEOREM 3.2
If conditions (3.1) *through* (3.4) *and* (3.13) *are satisfied then*

$$\|u^0 - u_h^0\|_{H^{0,0}(\Sigma)} \leqslant ch$$ ▼

Acknowledgment

The author would like to express his gratitude to Dr. Jan Sokołowski for careful reading of the manuscript and many important comments which helped to improve the text.

References

[1] AUBIN J. P.: Approximation of elliptic boundary-value problems. Wiley–Interscience, New York 1972

[2] BABUŠKA I., AZIZ A. K.: Survey lecture on mathematical foundations of the finite element method. In: The mathematical foundations of the finite element method with applications to partial differential equations — edited by A. K. Aziz. Academic Press, New York 1972.

[3] FALK R. S.: Approximation of a class of optimal control problems with order of convergence estimates. *J. Math. Anal. Appl.* 44 (1973), 28–47.
[4] GEVECI T.: On the approximation of the solution of an optimal control problem governed by an elliptic equation. *R.A.I.R.O. Analyse Numerique* 13 (1979), 313–328.
[5] HACKBUSCH W.: On the fast solving of parabolic boundary control problems. *SIAM J. Contr. Opt.* 17 (1979), 231–244.
[6] HACKBUSCH W.: Optimal $H^{p,p/2}$ error estimates for a parabolic Galerkin method (to appear).
[7] HOLNICKI P.: Rate of convergence estimates for a discrete-time Galerkin approximation to optimal control problems for hyperbolic systems. *Arch. Aut. Telemech.* 23 (1978), 46–64.
[8] LASIECKA I., MALANOWSKI K.: On discrete-time Ritz-Galerkin approximation of control constrained optimal control problems for parabolic systems. *Control and Cybernetics*, 7 (1978), 21–36.
[9] LIONS J. L.: Contrôle optimal de systèmes gouvernès par des equations aux dèrivèes partielles. Dunod, Paris 1968.
[10] LIONS J. L., MAGENES G.: Problemes aux limites non homogènes et applications, vol. 1, 2. Dunod, Paris 1968.
[11] TARTAR L.: Equations with order preserving properties. MRC Tech. Summary Report 1580, University of Visconsin, Febr. 1976.
[12] VAINBERG M. M.: Variatsionnyi metod i metod monotonnykh opieratorov. Nauka, Moskwa 1972.

Appendix

Proof of (1.21)

It is assumed that at the point $v_h^s = v_h + s\,(u_h^0 - v_h)$ the function $k_h(s)$ is differentiable. Consider the following subsets of Ξ

$$\Xi^+ = \{\xi \in \Xi \,|\, |-DJ_h(v_h^s)(\xi) + v_h^s(\xi)| > 1\},$$

$$\Xi^- = \{\xi \in \Xi \,|\, |-DJ_h(v_h^s)(\xi) + v_h^s(\xi)| < 1\},$$

$$\Xi^0 = \{\xi \in \Xi \,|\, |-DJ_h(v_h^s)(\xi) + v_h^s(\xi)| = 1\}.$$

It follows from definitions (1.15b) and (1.18) that

$$\frac{dk_h(s)}{ds}(\xi) = (u_h^0 - v_h)(\xi) \quad \text{for} \quad \xi \in \Xi^+. \tag{A.1}$$

For $\xi \in \Xi^-$ we have

$$\mathrm{sg}\,(-DJ_h(v_h^s) + v_h^s)(\xi) = -DJ_h(v_h^s)(\xi) + v_h^s(\xi),$$

hence from (1.15b) and (1.18) we get

$$\frac{dk_h(s)}{ds}(\xi) = D^2 J_h(v_h^s)(u_h^0 - v_h)(\xi). \tag{A.2}$$

Note that for $\xi \in \Xi^0$ we have

$$\lim_{\Delta s \to 0+} \frac{k_h(s+\Delta s) - k_h(s)}{\Delta s}(\xi) \neq \lim_{\Delta s \to 0-} \frac{k_h(s+\Delta s) - k_h(s)}{\Delta s}(\xi)$$

unless

$$(u_h^0 - v_h)(\xi) = D^2 J_h(v_h^s)(u_h^0 - v_h)(\xi) \tag{A.3}$$

Since $k_h(s)$ is differentiable condition (A.3) must be satisfied for almost all $\xi \in \Xi^0$ which fields

$$\frac{dk_h(s)}{ds}(\xi) = (u_h^0 - v_h)(\xi) \quad \text{for a.a.} \quad \xi \in \Xi^0. \tag{A.4}$$

(A.1), (A.2) and (A.4) imply (1.21). q.e.d.

Zbieżność aproksymacji dla kwadratowych zadań sterowania optymalnego z ograniczeniami amplitudy funkcji sterującej

Podaje się ogólną metodę oszacowania prędkości zbieżności aproksymacji dla kwadratowych zadań sterowania optymalnego z ograniczeniami amplitudy funkcji sterującej.

W metodzie tej wykorzystuje się bezpośrednio postać sterowania optymalnego.

Otrzymane wyniki ogólne są użyte do oszacowania prędkości zbieżności aproksymacji typu Galerkina dla zadania sterowania brzegowego dla układu opisywanego liniowym równaniem parabolicznym z funkcjonałem jakości zależnym od stanu końcowego.

Сходимость аппроксимации для задач оптимального управления с квадратным функционалом качества и ограничениями по амплитуде управляющей функции

Дается общий метод оценки скорости сходимости аппроксимации для задач оптимального управления с квадратным функционалом качества и ограничениями по амплитуде управляющей функции.

В этом методе непосредственно используется вид оптимального управления.

Полученные общие результаты используются в оценке скорости сходимости аппроксимации для задачи краевого управления в системе описываемой линейным параболическим уравнением, с функционалом качества зависящим от конечного состояния.