ANALYSIS OF CONVERGENCE FOR CONTROL PROBLEMS GOVERNED BY EVOLUTION EQUATIONS INEQUALITY AND EQUALITY CONSTRAINTS WITH MULTIPLIERS IMBEDDED

S.A. OLORUNSOLA

Department of Mathematical Sciences, University of Ado-Ekiti, P.M.B. 5363, Ado-Ekiti, Nigeria.

(*Submitted*: 20 February 2004; *Accepted*: 18 June 2004)

Abstract

The convergence of a scheme to minimize a class of a system of continuous optimal control problems characterized by a system of evolution equations and a system of linear inequality and equality constraints with multiplier imbedding is considered. The result is applied to some problems and the scheme is found to exhibit geometric convergence.

Keywords: Evolution equations, multiplier imbedding, constraints, geometric convergence.

1. Introduction

Recently, Olorunsola (2002) constructed a control operator specifically for a wider class of systems of continuous optimal control problems characterized by a system of evolution equations and a system of linear inequality and equality constraints with multiplier imbedding. He however failed to address the issue of the nature of convergence of this scheme as it applies to the combination of the standard penalty method and the multiplier imbedding algorithm (Glad, 1979). This paper deals with the issue of the convergence of this scheme and a good updating approach as it applies to this scheme.

For the purpose of clarity and completeness, it is therefore relevant at this stage to look at the class of problems dealt with in the above scheme whose convergence is being addressed.

Problem 1.1

$$\text{Minimize } \int_0^T x^T(t)Px(t) + u^T(t)Qu(t)dt \qquad (1.1)$$

$$\text{Subject to } \dot{x}(t) = Rx(t) + Gx(t-r) + Wu(t) \qquad (1.2)$$

$$Cx(t) + Du(t) \geq 0 \qquad (1.3)$$

and

$$Ex(t) + Fu(t) = 0 \qquad (1.4)$$

where $x(o) = x_o$, $t \in [0,T]$, $x(t) = h(t)$ for $t \in [-r,0]$.

$x(t), x(t-r) \in R^n$, $u(t) \in R^m$ are respectively the state variable, delay and the control variables. P and Q are n and m-symmetric positive definite square matrices respectively. The matrices R, G, C and E are any n-square matrices while W, D and F are n x m matrices. In case $m < n$, products of the form Qu(t), Wu(t) …. etc., control vectors u(t) can be made conformable by adjoining n-m zeroes.

Using the combination of the standard penalty function method (Glad, 1979) and the method of the multiplier imbedding Extended conjugate gradient (Olorunsola, 1991), a control operator A for the class of Problem 1.1 was obtained in Olorunsola (2002) as in Theorem 1.1 below.

THEOREM 1.1 The exact control operator that satisfies the given optimization problem 1.1 is given by

$$(AZ)(t)=\begin{bmatrix} a_{11} & a_{12} & \ldots\ldots\ldots\ldots\ldots a_{16} \\ a_{21} & a_{22} & \ldots\ldots\ldots\ldots\ldots : \\ : & & : \\ : & & : \\ a_{61} & \ldots\ldots\ldots\ldots\ldots & \ldots a_{66} \end{bmatrix} \tag{1.5}$$

where $(a_{11}x)(t)=\mu_1\left[x_0(t)-(R+G)\dot{x}_0\right]Sinh(t)+\mu_1\int_0^t\left\{\dot{x}(s)-(R+G)x(s)Cosh(t-s)ds-\right.$

$\int_0^t\left\{[P+\mu_1R^T(R+G)x(s)-\mu_1R^T\{x(s)+\mu_2M^TMx(s)Sinh(t-s)\}ds-\right.$

$\mu_1\int_0^{t-r}\left\{G^T(R+G)x(s+r)Sinh(t-r-s)\right\}ds+\left\{\left[P+\mu_1R^T(R+G)\right]x_0+\mu_2M^TMx_0-\mu_1R^Tx_0(t)+\right.$

$\mu_1G^T(R+G)x(r)-\mu_1G^T\dot{x}(s+r)Cosht+\dfrac{Sinht}{SinhT}\{[P+QR^T(R+G)]x(T)-\mu_1R^T\dot{x}(T)+\mu_2M^TMx(T)+$

$\mu_1G^Tx(T-r)-\mu_1G^Tx(T-r)-\left[(P+\mu_1R^T(R+G)x_0-\mu_1G^Tx_0)\right]+\mu_2M^TMx_0+\mu_1G^T(R+G)x(r)\}CoshT$

$\mu_1\left[\dot{x}(t)-(R+G)x_0\right]SinhT+\int_0^T\left\{P+\mu_1R^T(R+G)x(s)-\mu_1R^T\dot{x}(s)+\mu_2M^TMx(s)\right\}SinhT(T-S)ds$

$+\mu_1\int_0^{T-r}\left\{G^T(R+G)\dot{x}(s+r)\right\}Sinh(T-r-s)ds-\mu_1\int_0^T\left\{\dot{x}(s)-(R+G)x(s)\right\}Cosh(T-s)ds$

(1.6)

$$(a_{21}x)(t)=\nu_1W^T(R+G)x(t)-\mu_1W^T\dot{x}(t)+\mu_2N^TMx(t) \tag{1.7}$$

$$(a_{31}x)(t)=\nu_1G^T[(R+G)x(t+r)-\dot{x}(t)]+\mu_2N^TMx(t) \tag{1.8}$$

$$(a_{41}x)(t)=\dot{x}(t)-(R+G)x(t) \tag{1.9}$$

$$(a_{51}x)(t)=\mu_2Mx(t) \tag{1.10}$$

$$(a_{61}x)(t)=0 \tag{1.11}$$

$(a_{12}u)(t)=\mu_1u_0\sinh t-\mu_1\int_0^t Wu(s)Cosh(t-s)ds-\int_0^t\left\{\mu_1R^TWu(s)+\mu_2M^TNu(s)Sinh(t-s)\right\}ds$

$-\mu_1\int_0^{t-r}G^TWu(s+r)Sinh(t-r-s)ds+\left[\mu_1R^TWu_0+\mu_2M^TNu_0+\mu_1G^Tu(r)\right]Cosht+$

$\dfrac{Sinht}{SinhT}\left\{\mu_1G^Tu(T)+\mu_2M^TNu(T)+\mu_1G^TWu(T-r)\right\}-\left[\mu_1R^TWu_0+\mu_2M^TNu_0+\mu_1G^TWu(T)\right]CoshT$

$$-\mu_1 W u_0 SinhT + \int_0^T \{\mu_1 R^T W u(s) + \mu_2 M^T N u(s)\} Sinh(T-s)ds + \mu_1 \int_0^{t-r} G^T W u(s+r) Sinh(t-r-s)ds +$$

$$\mu_1 \int_0^T W u(s) Cosh(T-s)ds \qquad (1.12)$$

$$(a_{22}u)(t) = \{Q + \mu_1 W^T W + \mu_2 N^T N\} u(t) \qquad (1.13)$$

$$(a_{32}u)(t) = \mu_1 G^T W u(t-r) \qquad (1.14)$$

$$(a_{42}u)(t) = -Wu(t) \qquad (1.15)$$

$$(a_{52}u)(t) = \mu_2 N u(t) \qquad (1.16)$$

$$(a_{62}u)(t) = \tfrac{1}{\mu_2} a_{52}u(t) \qquad (1.17)$$

$$(a_{13}h)(t) = \mu_1 G h_0 Sinht - \mu_1 \int_0 Gh(s) Cosh(t-s)ds - \mu_1 \int_0 R^T G h(s) Sinh(t-s)ds + \mu_1 R^T G h_0 Cosht$$

$$\frac{Sinht}{SinhT}\Big(\mu_1\big[R^T G h(r) - R^T G h_0 Coshr - G h_0 Sinht\big] + \mu_1 \int_0 R^T G h(s) Sinh(r-s)ds$$

$$+ \mu_1 \int_0 Gh(s) Cosh(r-s)ds\Big) \qquad (1.18)$$

$$(a_{23}h)(t) = \mu_1 W^T G h(t) \qquad (1.19)$$

$$(a_{33}h)(t) = \mu_1 G^T G h(t) \qquad (1.20)$$

$$(a_{43}h)(t) = -Gh(t) \qquad (1.21)$$

$$(a_{14}y)(t) = -\mu_2 \int_0 y(s) Sinh(t-s)ds - \mu_2 y_0 Cosht - \frac{Sinht}{SinhT}\Big\{\mu_2 y(T) + y_0 Cosht + \mu_2 \int_0^T y(s) Sinh(T-s)ds\Big\}$$

$$(1.22)$$

$$(a_{24}y)(t) = -\mu_2 N y(t) \qquad (1.23)$$

$$(a_{54}y)(t) = -\mu_2 y(t) \qquad (1.24)$$

$$(a_{62}y)(t) = \tfrac{1}{\mu_2} a_{52} y(t) \qquad (1.25)$$

$$(a_{15}\lambda)(t) = -\lambda_0 Sinht + \int_0 \lambda(s) Cosh(t-s)ds + \int_0 \lambda^T(s) R Sinh(t-s)ds - \lambda_0^T R Cosht + \frac{Sinht}{SinhT}\Big(-\lambda(T) + \lambda_0^T R Cosht$$

$$+ \lambda_0 SinhT - \int_0^T \lambda^T(s) R Sinh(T-S)ds - \int_0^T \lambda(s) Cosh(T-s)ds\Big) \qquad (1.26)$$

$$(a_{25}\lambda)(t) = \lambda^T(t) W \qquad (1.27)$$

$$(a_{35}\lambda)(t) = -\lambda^T(t) G \qquad (1.28)$$

$$(a_{16}\rho)(t) = -\int_0^t \rho(s)Sinh(t-s)ds + \rho_0 Cosht + \frac{Sinht}{SinhT}\left\{\rho(T) - \rho_0 CoshT + \int_0^T \rho(s)Sinh(T-S)ds\right\} \quad (1.29)$$

$$(a_{26}\rho)(t) = \rho^T(t)N \quad (1.30)$$

$$(a_{34}y)(t) = (a_{44}y)(t) = 0 \quad (1.31)$$

$$(a_{44}\lambda)(t) = (a_{55}\lambda)(t) = (a_{63}\lambda)(t) = (a_{65}\lambda)(t) = 0 \quad (1.32)$$

$$(a_{36}\rho)(t) = (a_{46}\rho)(t) = (a_{56}\rho)(t) = (a_{66}\rho)(t) = 0 \quad (1.33)$$

$$(a_{53}h)(t) = 0 \quad (1.34)$$

Remark 1.1

Although the Scheme reliably converges at low iterations and that the constraints are well satisfied for varying penalty constants μ_1 and μ_2, the nature of its convergence was not addressed. Hence we discuss in this paper the nature of its convergence.

We now state and prove the main results.

2. Convergence of the Scheme

The aim is to prove that the optimal control problem 1.1 exhibits geometric convergence.

Definition 2.1 Let $\{\xi_n\}$ be a sequence of vectors in a Hilbert Space H with limit $\xi^* \in H$

Such that $\{\|\xi_{n+1} - \xi^*\|\}/\|\xi_n - \xi^*\|$ tends to a limit $\gamma < 1$ as $n \to \infty$, then $\{\xi_n\}$ is said to converge geometrically to ξ^* with a convergence ratio γ, as reported by (Ibiejugba, Otunta and Olorunsola, 1992)

We shall recall here for the purpose of clarity and explicitness the various steps of the conjugate gradient algorithm that generates the convergent sequence $\{Z_n(t)\}$ of solutions of Problems 1.1 according to Dipillo, Grippo and Lampariello (1974). The algorithm employs the explicit knowledge of the control operator A developed in Theorem 1.1.

STEP 1 Choose initial values for the conjugate descent algorithm

$$Z_0^t(t) = \{x_0(t), u_0(t), h_0(t), y_0(t), \lambda_0(t), \rho_0(t)\} \qquad Z_0(t) \in H.$$

Compute $P_0 = -g_o$

The remaining members of the sequence are then found as follows:

STEP 2 Update x_o, u_o, h_o, y_o, λ_o, ρ_o, such that

$$\left.\begin{aligned} x_{n+1}(t) &= x_n(t) + \alpha_n P_{x,n} \\ u_{n+1}(t) &= u_n(t) + \alpha_n P_{u,n} \\ h_{n+1}(t) &= h_n(t) + \alpha_n P_{h,n} \\ y_{n+1}(t) &= y_n(t) + \alpha_n P_{y,n} \\ \lambda_{n+1}(t) &= \lambda_n(t) + \alpha_n P_{\lambda,n} \\ \rho_{n+1}(t) &= \rho_n(t) + \alpha_n P_{\rho,n} \end{aligned}\right\} \quad (2.1)$$

where α_n and $P_{\bullet n}$ are the step length and the descent directions respectively.

STEP 3 Update gradients and descent directions using the updating rules:

$$\left.\begin{aligned}
g_{x,n+1} &= g_{x,n} + \alpha_n AP_{x,n}\\
g_{u,n+1} &= g_{u,n} + \alpha_n AP_{u,n}\\
g_{h,n+1} &= g_{h,n} + \alpha_n AP_{h,n}\\
g_{y,n+1} &= g_{y,n} + \alpha_n AP_{y,n}\\
g_{\lambda,n+1} &= g_{\lambda,n} + \alpha_n AP_{\lambda,n}\\
g_{\rho,n+1} &= g_{\rho,n} + \alpha_n AP_{\rho,n}
\end{aligned}\right\} \tag{2.2}$$

where $g_{\cdot,n}$ is the gradient at the n[th] iteration and A is the control operator in Theorem 1.1

$$\alpha_n = \langle g_{\bullet,n} \quad g_{\bullet,n}\rangle_H / \langle P_{\bullet,n} \quad AP_{\bullet,n}\rangle_H$$

$$g_{\bullet,n}^T = (g_{x,n}, g_{u,n}, g_{h,n}, g_{y,n}, g_{\lambda,n}, g_{\rho,n}) \text{ and } p_{\bullet,n}^T = (p_{x,n}, p_{u,n}, p_{h,n}, p_{y,n}, p_{\lambda,n}, p_{\rho,n})$$

$$(AP_n)(t) = \begin{pmatrix}
(a_{11}p_{x,n})(t) + (a_{12}p_{u,n})(t) + (a_{13}p_{h,n})(t) + (a_{14}p_{y,n})(t) + (a_{15}p_{\lambda,n})(t) + (a_{16}p_{\rho,n})(t)\\
(a_{21}p_{x,n})(t) + (a_{22}p_{u,n})(t) + (a_{23}p_{h,n})(t) + (a_{24}p_{y,n})(t) + (a_{25}p_{\lambda,n})(t) + (a_{26}p_{\rho,n})(t)\\
(a_{31}p_{x,n})(t) + (a_{32}p_{u,n})(t) + (a_{33}p_{h,n})(t) + (a_{34}p_{y,n})(t) + (a_{35}p_{\lambda,n})(t) + (a_{36}p_{\rho,n})(t)\\
:\\
:\\
(a_{61}p_{x,n})(t) + (a_{62}p_{u,n})(t) + (a_{63}p_{h,n})(t) + (a_{64}p_{y,n})(t) + (a_{65}p_{\lambda,n})(t) + (a_{66}p_{\rho,n})(t)
\end{pmatrix}$$

$$= \begin{pmatrix} AP_1(t)\\ AP_2(t)\\ :\\ :\\ AP_6(t)\end{pmatrix} \tag{2.3}$$

Setting

$$\left.\begin{aligned}
J_{x,n} &= \nabla_{x,n} J(x_n(t), u_n(t), h_n(t), y_n(t), \lambda_n(t), \rho_n(t), \mu_1, \mu_2)\\
J_{u,n} &= \nabla_{u,n} J(x_n(t), u_n(t), h_n(t), y_n(t), \lambda_n(t), \rho_n(t), \mu_1, \mu_2)\\
J_{h,n} &= \nabla_{h,n} J(x_n(t), u_n(t), h_n(t), y_n(t), \lambda_n(t), \rho_n(t), \mu_1, \mu_2)\\
J_{y,n} &= \nabla_{y,n} J(x_n(t), u_n(t), h_n(t), y_n(t), \lambda_n(t), \rho_n(t), \mu_1, \mu_2)\\
J_{\lambda,n} &= \nabla_{\lambda,n} J(x_n(t), u_n(t), h_n(t), y_n(t), \lambda_n(t), \rho_n(t), \mu_1, \mu_2)\\
J_{\rho,n} &= \nabla_{\rho,n} J(x_n(t), u_n(t), h_n(t), y_n(t), \lambda_n(t), \rho_n(t), \mu_1, \mu_2)
\end{aligned}\right\} \tag{2.4}$$

where

$$J(x_n(t), u_n(t), h_n(t), y_n(t), \lambda_n(t), \rho_n(t), \mu_1, \mu_2) = \int_0^T x^T(t)Px(t) + u^T(t)Qu(t)dt +$$

$$\mu_1 \int_0^T \{[x(t) - Rx(t) - Gx(t-r) - Wu(t)]^T [x(t) - Rx(t) - Gx(t-r) - Wu(t)]\}dt +$$

$$\mu_2 \int_0^T \{[Mx(t)+Nu(t)-y(t)]^T [Mx(t)+Nu(t)-y(t)]\}dt + \int_0^T \lambda^T(t)[x(t)-Rx(t)-Gx(t-r)-Wu(t)]+$$

$$\int_0^T \rho^T(t)[Mx(t)+Nu(t)-y(t)]dt \qquad (2.5)$$

and $\nabla_{\bullet,n} J$ is the gradient of any compliment at the n[th] step. We obtain the following relations

$$p_{x,n} = -\int_0 \nabla_{x,n} J(x_n(t),u_n(t),h_n(t),y_n(t),\lambda_n(t),\rho_n(t),\mu_1,\mu_2)ds$$

or more generally $p_{\bullet,n} = -\int_0 \nabla_{\bullet,n} J(x_n(t),u_n(t),h_n(t),y_n(t),\lambda_n(t),\rho_n(t),\mu_1,\mu_2)ds$ for $P_{x,n}$, $P_{u,n}$, $P_{h,n}$, $P_{y,n}$, $P_{\lambda,n}$, and $P_{\rho,n}$.

Therefore

$$(AP_1)(t) = \mu_1\left[J_{x,0}(t)-(R+G)p_{x,0}\right]Sinht + \mu_1 \int_0 \{p_{x,n}(s)-(R+G)J_{x,n}(s)Cosh(t-s)ds$$

$$-\int_0 \{[P+\mu_1 R^T(R+G)J_{x,n}(s) - \mu_1 R^T\{p_{x,n}(s)+\mu_2 M^T MJ_{x,n}(s)Sinh(t-s)ds]$$

$$-\mu_1 \int_0^{t-r} \{G^T(R+G)J_{x,n}(s+r)\sinh(t-r-s)\}ds + ([p+\mu R^T(R+G)J_{x,0} + \mu_2 M^T MJ_{x,0}(t)$$

$$-\mu_1 R^T J_{x,0}(0) + \mu_1 G^T(R+G)J_{x,n}(r) - \mu_1 G^T P_{x,n}(s+r)Cosht + \frac{Sinht}{SinhT}\{[P+QR^T(R+G)]J_{x,n}(T)$$

$$-\mu_1 R^T P_{x,n}(T) + \mu_2 M^T MJ_{x,n}(T) + \mu_1 G^T J_{x,n}(T-r) - \mu_1 G^T J_{x,n}(T-r) -$$

$$[(P+\mu_1 R^T(R+G)J_{x,0}(0) - \mu_1 G^T J_{x,0}(0) + \mu_2 M^T MJ_{x,0}(0) + \mu_1 G^T(R+G)J_{x,n}(r)]CoshT$$

$$+\mu_1\left[p_{x,n}(t)-(R+G)J_{x,0}(0)\right]SinhT + \int_0^T \{p+\mu_1 R^T(R+G)J_{x,n}(s) - \mu_1 R^T p_{x,n}(s) +$$

$$\mu_2 M^T MJ_{x,n}(s)\}Sinh(T-S)ds + \mu_1 \int_0^{t-r} \{G^T(R+G)P_{x,n}(s+r)\}Sinh(T-r-s)ds -$$

$$+\mu_1 \int_0^T \{P_{x,n}(t)-(R+G)J_{x,n}(s)\}Cosh(T-S)ds + \mu_1 j_{U,0}(0)\sinh t - \mu_1 \int_0 WJ_{u,n}(s)Cosh(t-s)ds$$

$$-\int_0 \{\mu_1 R^T WJ_{u,n}(s) + \mu_2 M^T NJ_{u,n}(s)Sinh(t-s)\}ds - \mu_1 \int_0^{t-r} G^T WJ_{u,n}(s+r)Sinh(t-r-s)ds$$

$$+\left[\mu_1 R^T WJ_{u,0}(0) + \mu_2 M^T NJ_{u,0}(0) + \mu_1 G^T J_{u,n}(r)\right]\cosh t + \frac{Sinht}{\sinh T}\left(\mu_1 G^T J_{u,n}(T) + \mu_2 M^T NJ_{u,n}(t)\right.$$

$$+\mu_1 G^T WJ_{u,n}(T-r) - \left[\mu_1 R^T WJ_{u,0}(0) + \mu_2 M^T NJ_{u,0}(0) + \mu_1 G^T WJ_{u,n}(T)\right]\cosh T - \mu_1 WJ_{u,0}(0)\sinh T$$

$$+\int_0^T \{\mu_1 R^T WJ_{u,n}(s) + \mu_2 M^T WJ_{u,n}(s)Sinh(T-s)\}ds - \mu_1 \int_0^{T-r} G^T WJ_{u,n}(s+r)Sinh(t-r-s)ds$$

$$+\mu_1 \int_0^T WJ_{u,n}(s)\cosh(T-s)ds + \mu_1 GJ_{h,0}(0)Sinht - \mu_1 \int_0 GJ_{h,n}(s)Cosh(t-s)ds - \mu_1 \int_0 R^T GJ_{h,n}(s)$$

$$Sinh(t-s)ds + \mu_1 R^T GJ_{h,0}(0)\cosh t + \frac{Sinht}{SinhT}\{\mu_1\left[R^T GJ_{h,n}(r) - R^T GJ_{h,0}(0)\cosh r - GT_{h,0}(0)\sinh t\right]$$

$$+\mu_1 \int_0 R^T GJ_{h,n}(s)\sinh(r-s)ds + \mu_1 \int_0 GJ_{h,n}(s)Cosh(r-s)ds\} - \mu_2 \int_0 J_{y,n}(s)Sinh(t-s)ds$$

$$-\mu_2 J_{y,0}(0)\cosh t - \frac{Sinht}{SinhT}\{\mu_2 J_{y,n}(T) + J_{y,0}(0)Cosht + \mu_2 \int_0^T J_{y,n}(0)Sinh(T-s)ds\} - J_{\lambda,0}(0)Sinht$$

$$+\int_0 J_{\lambda,0}(0)\sinh t + \int_0 J_{\lambda,n}(s)Cosh(t-s)ds + \int_0 J^T_{\lambda,n}(s)R\sinh(t-s)ds - J^T_{\lambda,0}(0)RCosht +$$

$$\frac{Sinht}{SinhT}\{-J_{\lambda,n}(T) + J^T_{\lambda,0}RCosht + J_{\lambda,0}(0)\sinh T - \int_0^T J^T_{\lambda,n}(s)R\sinh(T-s)ds - \int_0^T J_{\lambda,n}(s)Cosh(T-s)ds\}$$

$$-\int_0 J_{p,0}(s)Sinh(t-s)ds + J_{p,0}(0)Cosh + \frac{Sinht}{SinhT}\{J_{p,n}(T) - J_{p,0}(0) + \int_0 J_{p,n}(s)Sinh(T-s)ds\}$$

(2.5)

$$(AP_2)(t) = \mu_1 W^T(R+G)J_{x,n}(t) - \mu_1 W^T P_{x,n}(t) + \mu_2 N^T MJ_{x,n}(t) + [Q + \mu_1 \|W\| + \mu_2 \|N\|]J_{u,n}(t) + \mu_1 W^T GJ_{h,n}(t) - \mu_2 NJ_{y,n}(t) + J_{\lambda,n}(t)W + J_{\rho,n}(t)W \quad (2.6)$$

$$(AP_3)(t) = \mu_1 W^T(R+G)J_{x,n}(t) - \mu_1 W^T P_{x,n}(t) + \mu_2 N^T MJ_{x,n}(t) + \mu_1 G^T WJ_{u,n}(t-r) + \mu_1 J_{h,n}(t)\|G\| - J_{\lambda,n}(t)G \quad (2.7)$$

$$(AP_4)(t) = P_{x,n}(t) - (R+G)J_{x,n}(t) - WJ_{u,n}(t) - G^T J_{h,n}(t) \quad (2.8)$$

$$(AP_5)(t) = \mu_2 MJ_{x,n}(t) + \mu_2 NJ_{u,n}(t) - \mu_2 J_{y,n}(t) \quad (2.9)$$

$$(AP_6)(t) = NJ_{x,n}(t) - \mu_2 J_{y,n}(t) \quad (2.10)$$

The main result is stated below in Theorem 2.1.

REMARKS 2.1

The following remarks are in order as we state and prove theorem 2.1

(i). Z* is the optimal value required in Problem 1.1

(ii). The expression $\|Z_{n+1} - Z^*\|\|Z_n - Z^*\|^{-1}$ is the convergence ratio of the terms of the sequence {Zn(t)} in Hilbert Space H.

(iii). According to Hasdorff (1976), the general quadratic functional in the Hilbert space H to be minimized is given by $\varphi(z) = F_0 + \langle a, z\rangle_H + \frac{1}{2}\langle Z, AZ\rangle_H$ where A is the symmetric positive definite n-square matrix. If $F_0 = \langle a, z\rangle_H = 0$. We just have $\varphi(z) = \langle Z, AZ\rangle_H$. A is the control operator of Theorem 1.1.

(iv). Note that $\{P_{z,n}\}$ are conjugate with respect to the linear operator A

i.e. $\langle P_{z,n}, AP_K\rangle_H = 0 \quad n < k$

THEOREM 2.1

The sequence $\{Z_n(t)\}$ of elements generated in solving Problem 1.1 using the explicit knowledge of the control operator A constructed in Theorem 1.1 converges geometrically, to Z*(t) with ratio γ given by

$$\gamma = 1-\beta \text{ where } \beta = \frac{1}{\| Z_0 \|}\max.\| AZ_n \|^3 \left(< AP_{z,n}, P_{Z,n} >_H\right)^{-1}$$

Proof: Let $\varphi(z) = \langle z - z^*, AZ\rangle_H$ and $Z^T(t) = (x(t), u(t), h(t), y(t), \lambda(t), \rho(t))$ optimality condition demands that AZ*(t) = 0.

In general for Z, Z_n, ∈ H

$$\varphi(z_n) = \langle Z_n - Z^*, A(Z_n - Z^*)\rangle = < Z_n, AZ_n > - < Z^*, AZ_n > \tag{2.11}$$

and

$$\varphi(z_{n+1}) = \langle Z_n + \alpha_n p_{z,n} - Z^*, A(Z_n + \alpha_n p_{z,n} - Z^*)\rangle$$

$$=< Z_n, AZ_n > + \alpha_n < Z_n, AP_{z,n} > + \alpha_n < P_{z,n}, AZ_n > + \alpha_n^z < P_{z,n}, AP_{z,n} >$$

$$- < Z^*, AZ_n > - \alpha_n < Z^*, AP_{z,n} > \tag{2.12}$$

From 2.11 and 2.12

$$\varphi(Z_n) - \varphi(Z_{n+1}) = \frac{\| AZ_n \|^2}{< P_{z,n}, AP_{z,n} >} x \frac{\varphi(Z_n)}{< AZ_n, Z_n >} \tag{2.13}$$

Since A is a Self-adjoint operator, (Rickart, 1960; Olorunsola, 2002) AZ* = 0

Again $< AZ_n, Z_n > = < AZ_n, Z_0 > + \sum_{k=0}^{N-1} \alpha_k < AZ_n, P_{z,k} >, \quad but \quad \sum_{k=0}^{N-1} \alpha_k < AZ_n, P_{z,k} > = 0 \quad \forall n \neq k$

Hence $<AZ_n, Z_n> = <AZ_n, Z_0>$ (2.14)

A is a bounded operator implies that there exist numbers $\bar{m} > 0$ *and* $\bar{M} \in R$ such that for every $z \in H$

$$\bar{m} \| Z - Z^* \|^2 \leq \| A(Z - Z^*) \|^2 \leq \bar{M} \| Z - Z^* \|^2 \tag{2.15}$$

Hence $\langle AZ_n, Z_n \rangle \leq \| AZ_n \| \| Z_0 \|$ (2.16)

Substituting (2.16) in (2.13) we obtain

$$\varphi(Z_n) - \varphi(Z_{n+1}) = \frac{\| AZ_n \|^2 .\varphi(Z_n)}{< P_{z,n}, AP_{z,n} > \| AZ_n \| \| Z_0 \|}$$

Hence $\varphi(Z_{n+1}) \leq \left(1 - \frac{\| AZ_n \|^3}{< P_{z,n}, AP_{z,n} > \| Z_0 \|}\right).\varphi(Z_n)$ (2.17)

But by (2.15), $\varphi(Z_n) \geq \bar{m} \| Z_n - Z^* \|^2$ so that (2.17) becomes

$$(\| Z_{n+1} - Z \|^2)(\| Z_n - Z^* \|^2)^{-1} \leq \frac{\varphi(Z_{n+1})}{\varphi(Z_n)} \leq 1 - \frac{\| AZ_n \|^3}{< P_{z,n}, AP_{z,n} >} \| Z_0 \|^{-1}$$

Hence the result.

Table 1(a): Solution Profile for numerical Example 3.1. We compute the objective functional $J(x(t),u(t))$ and the constraint Satisfaction $CSAT(x,u) = \|\varphi(x,u)\|^2$

Penalty Constant μ_1	Penalty Constant μ_2	Iteration number K	Objective Functional $J(x(t),u(t))$	Constraint Satisfaction $\|\varphi(x(t),u(t))\|^2$
10^{-2}	10^{-2}	0	5.0000	25.7825
		1	2.86444	5.98355×10^{-1}
		2	1.942341	5.822259×10^{-1}
		3	1.390102	5.91823×10^{-1}
		4	1.021468	5.54895×10^{-1}
2×10^{-2}	2×10^{-2}	0	5.00000	25.7825
		1	2.988816	5.786152×10^{-1}
		2	2.225503	5.435591×10^{-1}
		3	1.831498	5.224789×10^{-1}
		4	1.58644	5.103435×10^{-1}
		5	1.440511	5.16819×10^{-1}
3×10^{-2}	3×10^{-2}	0	5.00000	25.7825
		1	2.930837	5.06737×10^{-1}
1×10^{-2}	3×10^{-2}	0	5.0000	25.7825
		1	2.90313	5.993069×10^{-1}
		2	2.102618	5.737325×10^{-1}
		3	1.730622	5.572019×10^{-1}
		4	1.615126	5.415913×10^{-1}
2×10^{-2}	3×10^{-2}	0	5.00000	25.7825
		1	2.986242	5.754655×10^{-1}
		2	2.306321	5.301353×10^{-1}
		3	2.171998	4.891628×10^{-1}
1×10^{-2}	4×10^{-2}	0	5.00000	25.7825
		1	2.942877	6.000822×10^{-1}
		2	2.150869	5.677288×10^{-1}
		3	1.904404	5.355533×10^{-1}
2×10^{-2}	4×10^{-2}	0	5.00000	25.7825
		1	2.925442	5.969161×10^{-1}
		2	2.141035	5.629162×10^{-1}
		3	1.899237	5.290906×10^{-1}

Table 1(b): Convergence ratio for numerical Example 3.1

Penalty	Constants	Iteration	Convergence Ratio
μ_1	μ_2	N	$\gamma^2 = 1-\beta$
0.01	0.01	1	0.204197
		2	0.1827247
		3	0.1761002
		4	0.1769987
		5	0.1877945
0.02	0.02	1	0.2042600
		2	0.1827929
		3	0.1759149
		4	0.1773238
		5	0.1907254
0.03	0.03	1	0.2044095
		2	0.1827606
		3	0.1755998
		4	0.1775935
		5	0.193959
0.04	0.04	1	0.2046224
		2	0.1826889
		3	0.1751335
		4	0.17778556
		5	0.193959
0.05	0.05	1	0.2048749
		2	0.1824520
		3	0.17443623
		4	0.1778679
		5	0.200913

3. Numerical Results

Example 3.1 Minimize the following quadratic functional governed by a system of linear of evolution equations and a system of linear inequality and equality constraints.

$$\text{minimize} \quad \int_0^1 x^T(t)Px(t) + u^T(t)Qu(t)dt$$

$$\text{Subject to} \quad \dot{x}(t) = Rx(t) + Gx(t-r) + Wu(t)$$

$$Cx(t) + Du(t) \geq 0$$

and Ex(t) + Fu(t) = 0

where

$$P = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, Q = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, R = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, G = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, W = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$

$$C = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, D = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, E = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}, F = \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix}$$

The solution to this problem is presented in Table 1(a) below. Table 1(b) shows that numerical example 3.1 converges geometrically according as in Theorem 2.1. Table 2(a) shows the solution profile for numerical Example 3.2.

Table 2(a): Solution Profile for numerical Example 3.2

Penalty Constant μ_1	Penalty Constant μ_2	Iteration number K	Objective Functional J(x(t),u(t))	Constraint Satisfaction $\|\|\varphi(x(t),u(t))\|\|^2$
0.01	0.01	0	6.00000	22.07000
		1	3.700829	5.14899×10^{-1}
		2	2.620612	5.173527×10^{-1}
		3	1.945335	6.043216×10^{-1}
		4	1.86899	9.016032×10^{-1}
0.02	0.01	0	6.00000	22.06948
		1	3.862853	5.039536×10^{-1}
		2	2.934001	4.368277×10^{-1}
		3	2.323645	4.995465×10^{-1}
		4	1.90697	6.417061×10^{-1}
0.03	0.01	0	6.00000	22.05934
		1	4.192574	4.67285×10^{-1}
		2	3.9332	4.209466×10^{-1}
0.05	0.01	0	6.0000	22.05934
		1	3.927693	4.930718×10^{-1}
		2	3.160065	4.655135×10^{-1}
		3	2.716471	4.547522×10^{-1}
		4	2.375047	4.661923×10^{-1}
0.06	0.01	0	6.00000	22.04826
		1	3.82174	5.26245×10^{-1}
		2	2.497295	5.16670×10^{-1}
		3	1.886065	5.025523×10^{-1}

Table 2(b): Convergence ratio for Example 3.2

Penalty μ_1	Constants μ_2	Iteration N	Convergence Ratio $\gamma^2 = 1 - \beta$
0.01	0.01	1	0.350039
		2	0.308697
		3	0.3088348
		4	0.309716
		5	0.3177035
0.02	0.02	1	0.3467576
		2	0.3427000
		3	0.3053388
		4	0.305913
		5	0.315491
0.03	0.03	1	0.3438357
		2	0.302687
		3	0.302739
		4	0.3023933
		5	0.313749

Example 3.2

$$\text{minimize} \quad \int_0^1 \left\{ (x_1 + 0.5x_2)^2 + 0.75x_2^2 + 2(u_1 + 0.5u_2)^2 + 0.5u_2^2 \right\} dt$$

$$\text{Subject to} \quad \dot{x}_1 + x_1(t - 0.5) = x_1 - x_2 + x_2(t - 0.5) + 2u_1 + 2u_2$$

$$\dot{x}_2 + x_2(t - 0.5) = x_1 + x_2 - x_1(t - 0.5) - u_2$$

$$x_1 + 2x_2 - u_1 \geq 0$$

$$x_2 - 0.5x_1 - 0.5u_2 \geq 0$$

$$x_1 - x_2 + 0.5u_1 = 0$$

$$x_2 + 0.5x_1 - u_1 = 0$$

Again we compute the objective functional J(x,u) i.e. OBJ and the constraint satisfaction CSAT(x,u) = $\|CSAT(x,u)\|^2$

4. Comments and Conclusions

In Table 1(a) the solution to example 3.1 converges for varying penalty parameters in at most five iterations. It is observed that as μ_1 and μ_2 increase converges takes place at the 3rd iteration. The same pattern is observed in Table 2(a) for solution to example 3.2. Convergence takes place in at most five iterations even when μ_1 increases and μ_2 is kept constant.

Table 1(b) and 2(b) confirm that the scheme developed earlier, described and extended in this paper converges geometrically. The general trend is that as the iteration increases the convergence ratio falls rapidly from iteration N =1 to N = 3. For example, In Table 1(b) for $\mu_1 = \mu_2 = 0.01$, N = 1, $\gamma^2 = 0.2049197$ and for N = 3, $\gamma^2 = 0.17610$. γ^2 falls initially and increases as N increases is the general pattern in the convergence profile. In all cases $\gamma^2 < 1$, showing that the scheme exhibits geometric convergence.

For numerical Example 3.1, $\gamma^2 \approx 0.18 < 1$ and for Example 3.2, $\gamma^2 \approx 0.31 < 1$.

Acknowledgement

I am particularly grateful to the anonymous referee whose comments and suggestions have contributed immensely to the quality, explicitness and readability of this paper.

REFERENCES

Di Pillo, G. Grippo, L. and Lampariello, F., 1974. The multiplier method for optimal control problems, Conference on Optimization in Engineering and Economics, Naples, Italy.

Glad, S.T., 1979. A combination of penalty function and multiplier methods for solving Optimal Control problems, Journal of Optimization Theory and applications, Vol. 28 No. 3, Pp. 303-329

Hasdorff, L., 1976. Gradient Optimization and Nonlinear Control, John Wiley and Sons, New York, pp. 45-61.

Ibiejugba, M.A. and Onumayi, P., 1984. A control operator and some of its applications, Journal of Optimization Theory and Applications, vol. 103, 31-47.

Ibiejugba, M.A., Otunta, F. and Olorunsola, S.A., 1992. The role of the multipliers in the multiplier methods - Part 3, Journal of the Nigerian Mathematical Society, Vol. 3, No. 2;

Kreyszig, E., 1978. Introductory Functional Analysis with Applications, John Wiley and Sons, New York, pp. 459-555.

Olorunsola, S.A., 1991. *On a Multiplier Imbedding Extended Conjugate Gradient Method*, Unpublished Ph.D Thesis, University of Ilorin, Nigeria.

Olorunsola, S.A., 2002. An exact operator for the method of the multiplier, Journal of Sciences, Islamic Republic of Iran (to appear).

Rickart, C.E., 1960. General Theory of Banach Algebras, D-van Nostrand Company, Inc. Princeton.