International Journal of Mathematical Analysis and Modelling

(Formerly Journal of the Nigerian Society for Mathematical Biology)

Volume 8 Issue 2, 2025

ISSN (Print): 2682 - 5694
ISSN (Online): 2682 – 5708
www.tnsmb.org

Numerical approach to optimal control problems governed by delay differential equations

A.S. Afolabi*

Abstract

This research presents a comprehensive framework for solving optimal control problems (OCPs) governed by delay differential equations (DDEs) with mixed constraints. For this more complex case of DDE-constrained problems, where analytical solutions are generally unavailable, a numerical methodology is developed. The approach involves discretizing the cost functional and the DDE constraints using Simpson's rule and the Crank-Nicolson method, respectively. The resulting finite-dimensional, constrained optimization problem is then transformed into an unconstrained form using the Augmented Lagrangian Method (ALM). This formulation gives rise to an associated operator that facilitates the application of the Extended Conjugate Gradient Method (ECGM) to efficiently solve the problem. Numerical experiments demonstrate that the solutions obtained from the proposed ALM-ECGM scheme are in strong agreement with analytical solutions for ODE cases and outperform existing methods for DDE cases. A rigorous convergence analysis confirms the scheme's effectiveness, accuracy, and linear convergence rate, establishing it as a robust computational tool for this challenging class of OCPs.

Keywords and phrases: optimal control problems; delay differential equations; augmented Lagrangian method; extended conjugate gradient method; mixed constraints

1 Introduction

The pursuit of optimality is central to scientific and engineering advancement. Optimization, at its core, is the mathematical discipline concerned with determining the best possible outcome—minimizing effort or maximizing benefit—under given constraints. It involves identifying extrema of a function representing system performance and underpins decisions in engineering design, economics, and scientific modeling. The introduction of constraints significantly complicates the search for an optimum, as feasible solutions may shift, local minima may emerge, and the global solution may lie on the boundary of the feasible region. Consequently, the development of robust algorithms for constrained optimization remains a central challenge in applied mathematics [1, 2].

Optimal Control Theory extends the calculus of variations to dynamic systems, determining state and control trajectories that minimize a cost functional over time. An Optimal Control Problem (OCP) consists of a cost functional, control and state variables, and a dynamical system, often expressed through differential equations. Foundational principles such as Pontryagin's Maximum Principle provide necessary conditions for optimality by linking the system dynamics with adjoint equations and the Hamiltonian, where adjoint variables serve as Lagrange multipliers representing the sensitivity of the cost functional to state variations [2].

Constraints on control or state variables introduce both theoretical and numerical complexities, often precluding analytical solutions. Numerical approaches are thus indispensable. Historically, penalty function methods converted constrained problems into unconstrained forms by penalizing constraint violations, but these methods can become ill-conditioned when penalty parameters grow large. The Augmented Lagrangian Method (ALM), or method of multipliers, improves upon this by introducing

*E-mail: asafolabi@futa.edu.ng; Department of Mathematical Sciences, Federal University of Technology, Akure, Ondo State, Nigeria

explicit Lagrange multiplier estimates and quadratic penalty terms, thereby achieving convergence without severe ill-conditioning. The global convergence and robustness of ALM were rigorously demonstrated by [3], who applied it successfully to degenerate constraint problems.

Efficient numerical implementation of optimization subproblems often employs iterative solvers such as the Conjugate Gradient Method (CGM), known for its fast convergence and low memory requirements [4, 5, 6, 7, 8, 9]. Early work by [10] extended the Fletcher–Reeves conjugate gradient method to OCPs, converting constrained problems into unconstrained ones via penalty functions and achieving faster convergence than steepest descent. Similarly, [11] established theoretical proofs for the penalty function approach in equality- and inequality-constrained OCPs, validating it through numerical examples. Building on this foundation, the studies of [8, 12, 13, 14] demonstrated the successful integration of discretization techniques and penalty-based CGM solvers, including for OCPs constrained by DDEs. Their results confirmed the linear convergence and computational efficiency of CGM-based frameworks compared to existing benchmarks. Further improvements were reported by [15] and [16], who introduced extended and modified gradient-based CGM variants that enhanced convergence rates for complex control problems.

The Augmented Lagrangian framework was effectively applied to quadratic OCPs constrained by ordinary differential equations by [12], confirming ALM's computational efficiency and its advantage over pure penalty approaches. Its capability to handle constraints without ill-conditioning makes it particularly attractive for large-scale control problems. The practical relevance of such methods is evident in applications to epidemiological models: [17] used optimal control theory to design multi-control strategies for Tungiasis, demonstrating strong effects on disease transmission, while [18] applied a similar approach to a deterministic dengue model using real data from Cape Verde, deriving optimal strategies for disease control through cost functional weighting.

Despite substantial progress in OCP research, the systematic application of the Augmented Lagrangian Method to problems constrained by Delay Differential Equations (DDEs) remains limited. DDEs describe systems whose present evolution depends on past states, as seen in population dynamics, biological systems, and engineering processes. Their infinite-dimensional nature complicates both theory and computation. Although penalty-based CGM techniques have been applied to such systems [14], the superior numerical conditioning and stability properties of ALM suggest it can provide a more robust framework. This study therefore focuses on developing and analyzing numerical schemes that integrate the Augmented Lagrangian Method with gradient-based solvers to address Optimal Control Problems governed by delay dynamics, contributing a computationally stable and efficient approach to this challenging class of systems.

2 Methodology

In this section, OCPs constrained by ordinary or DDEs with mixed constraints and real co-efficients are considered. The necessary conditions for each of the classes of OCPs considered are derived. For this class of problem since the analytical solution does not exist due to the non-differentiability of the delay term, only numerical solutions are generated for this class of problems.

2.1 Necessary conditions for optimal control problems constrained by delay differential equations and inequality constraints

In this subsection, the necessary conditions for different classes of an OCPs constrained by a DDEs and inequality constraints are established. Since the analytical solutions of these classes of problems do not exist, only numerical solutions are generated. The classes considered are:

2.1.1 Necessary Conditions for an optimal control problems with mixed constraints and delay on the state variable

The standard form of a general optimal control problem with mixed constraints and delay on the state variable is given as

$$\begin{aligned} \text{Minimize } I(x,u) = & \int_0^T f(x,u,t)dt \\ \text{Subject to } \dot{x}(t) = & \ h((x-r),u,t), \ x(0) = x_0 \\ & g(x,u,t) \leq 0, \ t \in [0,T] \end{aligned}$$

where $r > 0$ is the delay term, $x(t) \in \Re^n$, $u(t) \in \Re^m$, $f : \Re^n \times \Re^m \times \Re \to \Re$, $h : \Re^n \times \Re^m \times \Re \to \Re^p$, $g : \Re^n \times \Re^m \times \Re \to \Re^r$ are continuously differentiable functions and T denotes the terminal time.

The Hamiltonian function is defined as

$$H(x,u,\lambda,t) = f(x,u,t) + \lambda h((x-r),u,t)$$

where $H : \Re^n \times \Re^m \times \Re^n \times \Re \to \Re$ and $\lambda \in \Re^n$

Similarly, the Lagrangian function is defined as

$$L(x,u,\lambda,\mu,t) = H(x,u,\lambda,t) + \mu g(x,u,t) \tag{1}$$

where $L : \Re^n \times \Re^m \times \Re^n \times \Re^q \times \Re \to \Re$ and the Lagrange multiplier $\mu \in \Re^q$ satisfy the complimentary slackness conditions

$$\mu \geq 0, \ \mu g(x,u,t) = 0$$

The Euler-Lagrange equations for (1) are given as

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = 0 \tag{2}$$

$$\frac{\partial L}{\partial u} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{u}}\right) = 0 \tag{3}$$

$$\frac{\partial L}{\partial \lambda} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\lambda}}\right) = 0 \tag{4}$$

$$\frac{\partial L}{\partial \mu} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mu}}\right) = 0 \tag{5}$$

Applying equations (2), (3) (4) and (5) on (1), we have

$$\begin{aligned} \frac{\partial f}{\partial x} - \lambda\frac{\partial h}{\partial x} + \mu\frac{\partial g}{\partial x} + \dot{\lambda} = 0 \\ \frac{\partial f}{\partial u} + \lambda\frac{\partial h}{\partial u} + \mu\frac{\partial g}{\partial u} = 0 \\ \frac{\partial f}{\partial \lambda} + h\frac{\partial \lambda}{\partial \lambda} + \mu\frac{\partial g}{\partial \lambda} = 0 \\ \frac{\partial f}{\partial \mu} + \lambda\frac{\partial h}{\partial \mu} + g\frac{\partial \mu}{\partial \mu} = 0 \end{aligned}$$

Hence,

$$\dot{\lambda} = \lambda\tfrac{\partial h}{\partial x} - \tfrac{\partial f}{\partial x} - \mu\tfrac{\partial g}{\partial x} \qquad \textbf{Adjoint Equation} \tag{6}$$

$$\tfrac{\partial f}{\partial u} = -\mu\tfrac{\partial g}{\partial u} - \lambda\tfrac{\partial h}{\partial u} \qquad \textbf{Optimality Condition} \tag{7}$$

$$\dot{x}(t) = h(x,u,t) \qquad \textbf{State Equation} \tag{8}$$

$$\tfrac{\partial H}{\partial \mu} = -g\tfrac{\partial \mu}{\partial \mu} - \lambda\tfrac{\partial h}{\partial \mu} \qquad \textbf{Optimality Condition} \tag{9}$$

The system of first order ordinary differential equations (6) and (8) can be solved concurrently to determine the values of the two constants by applying the boundary conditions. The free-end condition is used to obtain the second boundary condition if only one is provided. Thus, the free-end condition is given as

$$\frac{\partial H}{\partial \dot{x}} = 0 \; or \; \lambda(T) = 0 \qquad \textbf{Transversality Condition}$$

All of the requirements must be satisfied for the solution to be optimal. The solution is not optimal if one or more of the requirements are not met.The optimal control equation is a stationary (static) equation, whereas the state and adjoint (costate) equations are referred to as dynamic equations. The two-point boundary value problem is caused by the boundary conditions. While the state equation evolves in a forward direction, the costate equation does the opposite. The two-point boundary value problems typically require computationally demanding iterative numerical methods. Open loop control solutions are the result of these iterative numerical processes.

2.1.2 Necessary conditions for an optimal control problems with mixed constraints and delay on the control variable

The standard form of a general optimal control problem with mixed constraints and delay on the control variable is given as

$$\begin{aligned} \text{Minimize } I(x,u) = \;& \textstyle\int_0^T f(x,u,t))dt \\ \text{Subject to } \dot{x}(t) = \;& h(x,(u-r),t), \; x(0) = x_0 \\ & g(x,u,t) \leq 0, \; t \in [0,T] \end{aligned}$$

where $r > 0$ is the delay term, $x(t) \in \Re^n$, $u(t) \in \Re^m$, $f : \Re^n \times \Re^m \times \Re \to \Re$, $h : \Re^n \times \Re^m \times \Re \to \Re^p$, $g : \Re^n \times \Re^m \times \Re \to \Re^r$ are continuously differentiable functions and T denotes the terminal time.

The Hamiltonian function is defined as

$$H(x,u,\lambda,t) = f(x,u,t) + \lambda h(x,(u-r),t)$$

where $H : \Re^n \times \Re^m \times \Re^n \times \Re \to \Re$ and $\lambda \in \Re^n$

Similarly, the Lagrangian function is defined as

$$L(x,u,\lambda,\mu,t) = H(x,u,\lambda,t) + \mu g(x,u,t)$$

where $L : \Re^n \times \Re^m \times \Re^n \times \Re^q \times \Re \to \Re$ and the Lagrange multiplier $\mu \in \Re^q$ satisfy the complimentary slackness conditions

$$\mu \geq 0, \; \mu g(x,u,t) = 0$$

The Euler–Lagrange equations presented in 2–5 are employed to derive the optimality conditions and subsequently obtain the numerical solution of the problem using the procedure outlined in 2.1.1.

2.1.3 Necessary conditions for an optimal control problem with mixed constraints and delay on the state and control variables

The standard form of a general optimal control problem with constraints and delays on the state and control variables is given as

$$\begin{aligned} \text{Minimize } I(x,u) = \;& \textstyle\int_0^T f(x,u,t)dt \\ \text{Subject to } \dot{x}(t) = \;& h((x-r),(u-r),t), \; x(0) = x_0 \\ & g(x,u,t) \leq 0 \; t \in [0,T] \end{aligned}$$

where $r > 0$ is the delay term, $x(t) \in \Re^n$, $u(t) \in \Re^m$, $f : \Re^n \times \Re^m \times \Re \to \Re$, $h : \Re^n \times \Re^m \times \Re \to \Re^p$, $g : \Re^n \times \Re^m \times \Re \to \Re^r$ are continuously differentiable functions and T denotes the terminal time.

The Hamiltonian function is defined as

$$H(x,u,\lambda,t) = f(x,u,t) + \lambda h((x-r),(u-r),t)$$

where $H : \Re^n \times \Re^m \times \Re^n \times \Re \to \Re$ and $\lambda \in \Re^n$
Similarly, the Lagrangian function is defined as

$$L(x,u,\lambda,\mu,t) = H(x,u,\lambda,t) + \mu g(x,u,t)$$

where $L : \Re^n \times \Re^m \times \Re^n \times \Re^q \times \Re \to \Re$ and the Lagrange multiplier $\mu \in \Re^q$ satisfy the complimentary slackness conditions

$$\mu \geq 0,\ \ \mu g(x,u,t) = 0$$

Similarly, The Euler–Lagrange equations presented in 2–5 are employed to derive the optimality conditions and subsequently obtain the numerical solution of the problem using the procedure outlined in 2.1.1.

3 Analysis of the model

3.1 Numerical solution of optimal control problems with mixed constraints and delay on the state variable

Consider the quadratic optimal control problem with mixed constraints and delay on the state variable of the form:

$$\text{Minimize } I(x_1,x_2,u) = \int_0^T (px^2(t) + qu^2(t))dt \tag{10}$$

$$\text{Subject to } \dot{x}(t) = ax(t) + bx(t-r) + cu(t) \tag{11}$$

$$dx(t) + eu(t) \leq 0,\ \ x(0) = x_0 \tag{12}$$

where a, b , c, d, e are real constants and $p, q > 0$
In order to use Extended Conjugate Gradient Method to solve (10)-(12), we replace the constrained problem by an unconstrained discretized optimal control problem. Breaking [0,T] into N equal subintervals with knots $0 < t_0 < t_1 < t_2 < t_3 < \ldots < t_N = T$ and $t_j = j\nabla t_j$, we discretize (10) and (11) using fourth order Simson's rule and Cranck-Nilcholson method for the objective function and the differential constraint respectively.

3.1.1 Discretization of the objective function

$$\text{Minimize } I(x,u,t) = \int_0^T (px^2(t) + qu^2(t))dt \tag{13}$$

Equation (13) can be written as

$$I(x,u,t) = \int_0^T px^2(t)dt + \int_0^T qu^2(t)dt$$

where $t_0 = 0$ and $t_f = T$. Recall that

$$h = \frac{T-0}{n-1} = \frac{T}{n-1}$$

Recall from fourth order Simpson's Rule that

$$\int_{t_0}^{t_f} f(x)dx = \frac{2h}{45}(7f(x_0) + 32\sum_{i=1,5,9}^{n-3} f(x_i) + 12\sum_{i=2,6,10}^{n-2} f(x_i) + 32\sum_{i=3,7,11}^{n-1} f(x_i) + 14\sum_{i=4,8,12}^{n-4} f(x_i) + 7f(x_n))$$

This implies that

$$\int_{t_0}^{t_f} px^2(t)dt = \frac{2hp}{45}(\ 7x_0^2 + 32\sum_{i=1,5,9}^{n-3} x_i^2 + 12\sum_{i=2,6,10}^{n-2} x_i^2 + 32\sum_{i=3,7,11}^{n-1} x_i^2 + 14\sum_{i=4,8,12}^{n-4} x_i^2 + 7x_n^2)$$

Similarly, we have

$$\int_{t_0}^{t_f} qu^2(t)dt = \frac{2hq}{45}(\ 7u_0^2 + 32\sum_{i=1,5,9}^{n-3} u_i^2 + 12\sum_{i=2,6,10}^{n-2} u_i^2 + 32\sum_{i=3,7,11}^{n-1} u_i^2 + 14\sum_{i=4,8,12}^{n-4} u_i^2 + 7u_n^2)$$

Hence,

$$\begin{aligned}\int_{t_0}^{t_f} (px^2(t) + qu^2(t))dt = \frac{2h}{45}(\ & 7(px_0^2 + qu_0^2) + 32\sum_{i=1,5,9}^{n-3} (px_i^2 + qu_i^2) + 12\sum_{i=2,6,10}^{n-2} (px_i^2 + qu_i^2) \\ & +32\sum_{i=3,7,11}^{n-1} (px_i^2 + qu_i^2) + 14\sum_{i=4,8,12}^{n-4} (px_i^2 + qu_i^2) + 7(px_n^2 + qu_n^2))\end{aligned}$$

This implies that

$$\begin{aligned}\int_{t_0}^{t_f} (px^2(t) + qu^2(t))dt = \frac{2h}{45}(\ & 7(px_0^2 + qu_0^2) + 32(px_1^2 + qu_1^2 + px_5^2 + qu_5^2 + px_9^2 + qu_9^2 + \ldots + px_{n-3}^2 \\ & +qu_{n-3}^2) + 12(px_2^2 + qu_2^2 + px_6^2 + qu_6^2 + px_{10}^2 + qu_{10}^2 + \ldots + px_{n-2}^2 + qu_{n-2}^2) \\ & +32(px_3^2 + qu_3^2 + px_7^2 + qu_7^2 + px_{11}^2 + qu_{11}^2 + \ldots + px_{n-1}^2 + qu_{n-1}^2) + 14(px_4^2 \\ & +qu_4^2 + px_8^2 + qu_8^2 + px_{12}^2 + qu_{12}^2 + \ldots + px_{n-4}^2 + qu_{n-4}^2) + 7(px_n^2 + qu_n^2))\end{aligned}$$

$$\begin{aligned}\int_{t_0}^{t_f} (px^2(t) + qu^2(t))dt = & \frac{14hp}{45}x_0^2 + \frac{64hp}{45}x_1^2 + \frac{24hp}{45}x_2^2 + \frac{64hp}{45}x_3^2 + \frac{28hp}{45}x_4^2 + \frac{64hp}{45}x_5^2 + \frac{24hp}{45}x_6^2 \\ & + \frac{64hp}{45}x_7^2 + \frac{28hp}{45}x_8^2 + \frac{64hp}{45}x_9^2 + \frac{24hp}{45}x_{10}^2 + \frac{64hp}{45}x_{11}^2 + \frac{28hp}{45}x_{12}^2 + \ldots \\ & + \frac{28hp}{45}x_{n-4}^2 \frac{64hp}{45}x_{n-3}^2 + \frac{24hp}{45}x_{n-2}^2 + \frac{64hp}{45}x_{n-1}^2 + \frac{14hp}{45}x_n^2 + \frac{14hq}{45}u_0^2 \\ & + \frac{64hq}{45}u_1^2 + \frac{24hq}{45}u_2^2 + \frac{64hq}{45}u_3^2 + \frac{28hq}{45}u_4^2 + \frac{64hq}{45}u_5^2 + \frac{24hq}{45}u_6^2 + \frac{64hq}{45}u_7^2 \\ & + \frac{28hq}{45}u_8^2 + \frac{64hq}{45}u_9^2 + \frac{24hq}{45}u_{10}^2 + \frac{64hq}{45}u_{11}^2 + \frac{28hq}{45}u_{12}^2 + \cdots + \frac{28hq}{45}u_{n-4}^2 \\ & + \frac{64hq}{45}u_{n-3}^2 + \frac{24hq}{45}u_{n-2}^2 + \frac{64hq}{45}u_{n-1}^2 + \frac{14hq}{45}u_n^2\end{aligned} \tag{14}$$

Equation (14) can be written as

$$Z^{'}AZ + C^{'}$$

where

$$Z^{'} = (x_1, x_2, x_3, \ldots, x_n, u_0, u_1, u_2, \ldots, u_n),$$

If $g = \frac{14h}{45}$, $k = \frac{64h}{45}$, $l = \frac{24h}{45}$, $m = \frac{64h}{45}$ and $s = \frac{28h}{45}$, then

$$
A = \begin{pmatrix}
kp & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & lp & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & mp & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & sp & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & kp & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & lp & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & mp & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & sp & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ddots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ddots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & kp & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & lp & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & mp & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & sp & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & gp & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & gq & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & kq & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & lq & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & mq & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & sq & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & kq & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & lq & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & mq & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & sq & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ddots & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & kq & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & lq & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & mq & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & sq & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & gq
\end{pmatrix}
$$

and

$$C^{'} = gpx_0^2$$

$$
a_{i,j} = \begin{cases}
kp, & \text{i=j=2,6,10,\dots, n-3;} \\
lp, & \text{i=j=3,7,11,\dots, n-2;} \\
mp, & \text{i=j=4,8,12,\dots, n-1;} \\
sp, & \text{i=j=5,9,13,\dots, n-4;} \\
gp, & \text{i=j=n;} \\
gq, & \text{i=j=n+1;} \\
kq, & \text{i=j=n+2,n+6,n+10,\dots, 2n-3;} \\
lq, & \text{i=j=n+3,n+7,n+11,\dots, 2n-2;} \\
mq, & \text{i=j=n+4,n+8,n+12,\dots, 2n-1;} \\
sq, & \text{i=j=n+5,n+9,n+13,\dots, 2n;} \\
gq, & \text{i=j=2n+1;}
\end{cases}
$$

3.1.2 Discretization of the delay differential equation

3.1.3 Necessary optimality conditions for the DDE-constrained problem

The application of optimality conditions for Delay Differential Equation (DDE)-constrained optimal control problems requires a careful functional analytic treatment. The standard Euler-Lagrange formalism is insufficient due to the presence of the delayed state $x(t-r)$, which makes the problem infinite-dimensional and the dynamics non-Markovian.

We consider the problem of minimizing the cost functional J subject to the DDE dynamics and a path constraint:

$$
\begin{aligned}
\min_{u(\cdot)} \quad & J = \int_0^T L(x(t), u(t), t)dt \\
\text{subject to} \quad & \dot{x}(t) = f(x(t), x(t-r), u(t), t), \quad t \in [0, T], \\
& x(t) = \phi(t), \quad t \in [-r, 0], \\
& g(x(t), u(t)) \leq 0, \quad t \in [0, T],
\end{aligned}
$$

where $\phi(t)$ is the known history function.

To derive the first-order necessary conditions, we define the Hamiltonian $\mathcal{H}$ and augment the cost functional with the dynamics and constraints using Lagrange multipliers $\lambda(t)$ (the costate) and $\mu(t)$ (a Karush-Kuhn-Tucker multiplier).

$$\mathcal{H}(t) = L(x(t), u(t), t) + \lambda(t)^T f(x(t), x(t-r), u(t), t) + \mu(t)^T g(x(t), u(t))$$

A variation of the augmented cost leads to the following necessary conditions for optimality:

1. **State Equation (forward DDE):**

$$\dot{x}(t) = \nabla_\lambda \mathcal{H} = f(x(t), x(t-r), u(t), t), \quad x(t) = \phi(t) \text{ for } t \in [-r, 0].$$

2. **Costate Equation (backward Advanced Differential Equation):** The costate equation is not a simple ODE but an Advanced Differential Equation (ADE), which evolves backward in time:

$$-\dot{\lambda}(t) = \nabla_{x(t)} \mathcal{H}(t) + \chi_{[0,T-r]} \nabla_{x(t)} \mathcal{H}(t+r).$$

More explicitly, this is:

$$-\dot{\lambda}(t) = \frac{\partial L}{\partial x}(t) + \lambda(t)^T \frac{\partial f}{\partial x(t)}(t) + \mu(t)^T \frac{\partial g}{\partial x}(t) + \chi_{[0,T-r]} \lambda(t+r)^T \frac{\partial f}{\partial x(t-r)}(t+r)$$

with the **transversality condition:**

$$\lambda(T) = 0.$$

The term $\lambda(t+r)^T \frac{\partial f}{\partial x(t-r)}(t+r)$ is the "advanced" term, as it depends on the future value of the costate. The indicator function $\chi_{[0,T-r]}$ ensures this term is only present when $t + r \leq T$.

3. **Stationarity Condition:**

$$\nabla_{u(t)} \mathcal{H} = \frac{\partial L}{\partial u}(t) + \lambda(t)^T \frac{\partial f}{\partial u}(t) + \mu(t)^T \frac{\partial g}{\partial u}(t) = 0, \quad \text{for } t \in [0, T].$$

4. **Complementarity Slackness:**

$$\mu(t) \geq 0, \quad g(x(t), u(t)) \leq 0, \quad \mu(t)^T g(x(t), u(t)) = 0.$$

This coupled system of a forward DDE and a backward ADE, along with the stationarity and complementarity conditions, constitutes the correct necessary conditions for optimality. It properly accounts for the non-anticipative nature of the state and the anticipative nature of the costate caused by the delay.

3.1.4 Discretization of the Optimality System

With the optimality conditions rigorously established, we now proceed to discretize the entire system for numerical solution. We apply the Crank-Nicolson method to both the state and costate equations. We ensure the delay r is a rational multiple of the time step h, i.e., $r = mh$ for an integer m, to ensure all delayed and advanced arguments fall on the discretization grid.

Let N be the number of steps, $h = T/N$, and $t_k = kh$.

- **State DDE Discretization:** The state equation $\dot{x}(t) = ax(t) + bx(t-r) + cu(t)$ is discretized as:

$$\frac{x_{k+1} - x_k}{h} = \frac{1}{2}\left[(ax_{k+1} + bx_{k+1-m} + cu_{k+1}) + (ax_k + bx_{k-m} + cu_k)\right]$$

  Rearranging, we obtain the forward recurrence relation:

$$(2 - ah)x_{k+1} = (2 + ah)x_k + bh(x_{k-m} + x_{k+1-m}) + ch(u_k + u_{k+1}) \tag{15}$$

  or

$$x_{k+1} = Ux_k + V(x_{k-m} + x_{k+1-m}) + W(u_k + u_{k+1})$$

  where $U = \frac{2+ah}{2-ah}$, $V = \frac{bh}{2-ah}$, and $W = \frac{ch}{2-ah}$.

- **Costate ADE Discretization:** The costate equation $-\dot{\lambda}(t) = a\lambda(t)+b\lambda(t+r)+\frac{\partial L}{\partial x}(t)$ (assuming L is solely a function of x and u for simplicity, and the path constraint is handled separately) is discretized backward in time. For $k = N-1, \dots, 0$:

$$-\frac{\lambda_{k+1} - \lambda_k}{h} = \frac{1}{2}\left[(a\lambda_{k+1} + b\lambda_{k+1+m} + l_{k+1}) + (a\lambda_k + b\lambda_{k+m} + l_k)\right]$$

where $l_k = \frac{\partial L}{\partial x}(x_k, u_k)$. Rearranging gives the backward recurrence relation:

$$(2+ah)\lambda_k = (2-ah)\lambda_{k+1} - bh(\lambda_{k+m} + \lambda_{k+1+m}) - h(l_k + l_{k+1}) \tag{16}$$

- **Stationarity Condition Discretization:** The discretized stationarity condition is:

$$\frac{\partial L}{\partial u}(x_k, u_k) + c\lambda_k + \mu_k \frac{\partial g}{\partial u}(x_k, u_k) = 0 \tag{17}$$

- **Path Constraint:** The inequality constraint $dx(t) + eu(t) \leq 0$ is converted to an equality by introducing slack variables z_k^2:

$$dx_k + eu_k + z_k^2 = 0, \quad \text{with } \mu_k \geq 0, z_k^2 \geq 0, \mu_k z_k^2 = 0. \tag{18}$$

The complete discretized problem is a large-scale system of nonlinear equations (15), (16), (17), and (18) for the variables $\{x_k\}, \{\lambda_k\}, \{u_k\}, \{\mu_k\}, \{z_k\}$. This system can be solved using a nonlinear programming solver or a specialized solver for complementarity problems. The matrix formulation presented in the original text represents one part of this larger, coupled system.

We discritize the constraint $\dot{x}(t) = ax(t) + bx(t-r) + cu(t)$ by using Cranck Nilcholson Method and assuming that there exists $x(t_k - r) = x_{k-r}$. This is true for all rational values of r and we ensure that our point of discretization falls exactly on r to avoid off grid points.
We recall from Cranck Nilcholson Method that

$$y_{n+1} = y_n + \frac{h}{2}[f_{n+1} + f_n]$$

This implies that

$$x_{j+1} = ax_j + \frac{h}{2}(ax_{j+1} + bx_{j+1-r} + cu_{j+1} + ax_j + bx_{j-r} + cu_j)$$

$$(2-ah)x_{k+1} = (2a+ah)x_k + bhx_{k-r} + bhx_{k+1-r} + cu_{j+1} + +cu_j$$

$$x_{k+1} = Ux_k + Vx_{k-r} + Vx_{k+1-r} + Wu_k + Wu_{k+1} \tag{19}$$

where $U = \frac{a(2+h)}{2-ah}$,$V = \frac{bh}{2-ah}$ and $W = \frac{ch}{2-ah}$
Considering (19), there exists m such that $m = \frac{r}{h}$, where $h = \Delta t_k$. For $t \in [-r, 0]$, where $x(t-r) = x_{t-r}$ is known to be a constant for each value of r in the interval.
when $k = 0$

$$x_1 - Vx_{-r} - Vx_{1-r} - Wu_0 - Wu_1 = Ux_0 \tag{20}$$

when $k = 1$

$$x_2 - Ux_1 - Vx_{1-r} - Vx_{2-r} - Wu_1 - Wu_2 = 0 \tag{21}$$

when $k = 2$

$$x_3 - Ux_2 - Vx_{2-r} - Vx_{3-r} - Wu_2 - Wu_3 = 0 \tag{22}$$

when $k = 3$

$$x_4 - Ux_3 - Vx_{3-r} - Vx_{4-r} - Wu_3 - Wu_4 = 0$$

when $k = 4$

$$x_5 - Ux_4 - Vx_{4-r} - Vx_{5-r} - Wu_4 - Wu_5 = 0$$

when $k = 5$

$$x_6 - Ux_5 - Vx_{5-r} - Vx_{6-r} - Wu_5 - Wu_6 = 0 \tag{23}$$

$\vdots$

when $k = m - 1$

$$x_m - Ux_{m-1} - Vx_{m-1-r} - Vx_{m-r} - Wu_{m-1} - Wu_m = 0 \tag{24}$$

when $k = m$

$$x_{m+1} - Ux_m - Vx_{m-r} - Vx_{m+1-r} - Wu_m - Wu_{m-1} = 0 \tag{25}$$

when $k = m + 1$

$$x_{m+2} - Ux_{m+1} - Vx_{m+1-r} - Vx_{m+2-r} - Wu_{m+1} - Wu_{m+2} = 0 \tag{26}$$

$\vdots$

when $k = N - 1$

$$x_N - Ux_{N-1} - Vx_{N-1-r} - V_{N-r} - Wu_{N-1} - Wu_N = 0 \tag{27}$$

Equations (20) - (27), can be written in matrix form as

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -W & -W & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -U & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -W & -W & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -U & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -W & -W & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -U & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -W & -W & 0 & 0 & 0 & 0 \\ -V & 0 & 0 & -U & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -W & -W & 0 & 0 & 0 \\ 0 & -V & -V & 0 & -U & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -W & -W & 0 & 0 \\ 0 & 0 & -V & -V & 0 & -U & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -W & -W & 0 \\ 0 & 0 & 0 & -V & -V & 0 & -U & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -W & -W \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \\ x_{m+1} \\ \vdots \\ x_n \\ u_0 \\ u_1 \\ \vdots \\ u_m \\ u_{m+1} \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} Ux_0 + Vx_{-r} + V_{1-r} \\ Vx_{2-r} + V_{1-r} \\ \vdots \\ ex_{m-1-r} \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{28}$$

Equation (28) can be written as

$$GZ = K$$

where G is a block matrix of dimension $N \times (2N + 1)$ with representation $G = (C \vdots E)$ where $[C]$ is an $N \times N$ sparse matrix with principal diagonal elements $[C]_{ij} = 1$, lower diagonal elements -U for every i, j such that $i = j + 1$ and -V for every i, j such that $i = j + 4$. $[E]$ is $N \times (N + 1)$ bidiagonal matrix with $[E]_{ij} = -W$ and upper diagonal elements $-W$ for every i, j such that $j = i + 1$. $[Z]$ is also a column vector of dimension $(2N + 1) \times 1$. The column vector $[K]$ is of order $n \times 1$ with entries given by $[K]_{11} = Ux_0 + Vx_{-r} + V_{1-r}$, $[K]_{12} = Vx_{2-r} + V_{1-r}$, . . ., $[K]_{N1} = ex_{m-1-r}$

3.1.5 Conversion of the inequality constraint to an equality constraint

We convert the inequality constraint $dx(t) + eu(t) \leq 0$ to an equality constraint by introducing a vector of nonnegative slack variables $s = (s_0, s_1, \ldots, s_N)$, where $s_k \geq 0$ for all k. This implies that the problem becomes:

$$dx_k + eu_k + s_k = 0, \quad \text{with } s_k \geq 0, \quad \text{for } k = 0, 1, \ldots, N \tag{29}$$

Equation (29) can be written in matrix form as:

$$\begin{pmatrix} d & 0 & \cdots & 0 & e & 0 & \cdots & 0 \\ 0 & d & \cdots & 0 & 0 & e & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d & 0 & 0 & \cdots & e \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \\ u_0 \\ u_1 \\ \vdots \\ u_N \end{pmatrix} + \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_N \end{pmatrix} = \begin{pmatrix} -dx_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

This can be written compactly as:

$$LZ + S = N \tag{30}$$

where $S = (s_1, s_2, \ldots, s_N)^T$ is the vector of nonnegative slack variables.

The complementarity conditions for the inequality constraint become:

$$\mu_k \geq 0, \quad s_k \geq 0, \quad \mu_k s_k = 0, \quad \text{for } k = 1, \ldots, N$$

This formulation maintains the quadratic nature of the optimization problem while properly handling the inequality constraints through standard Karush-Kuhn-Tucker conditions.

3.2 Numerical solution of optimal control problems with mixed constraints and delay on the state and control variables

Consider the quadratic optimal control problem with mixed constrained and delays on the State and Control Variables of the form:

$$\text{Minimize } I(x_1, x_2, u) = \int_0^T (px^2(t) + qu^2(t))dt \tag{31}$$

$$\text{Subject to} \dot{x}(t) = ax(t) + bx(t-r) + cu(t) + du(t-r) \tag{32}$$

$$ex(t) + fu(t) \leq 0, \; x(0) = x_0 \tag{33}$$

a, b , c, d, e, f are real constants and $p, q > 0$

In order to use Extended Conjugate Gradient Method to solve (31)-(33), we replace the constrained problem by an unconstrained discretized optimal control problem. Breaking [0,T] into N equal sub-intervals with knots $0 < t_0 < t_1 < t_2 < t_3 < \ldots < t_N = T$ and $t_j = j\nabla t_j$, we discretize (31) and (32) using fourth order Simson's rule and Cranck-Nilcholson method for the objective function and the constraints respectively.

3.2.1 Discretization of the Objective Function

The discritized form of the objective function given by problem (31) is as presented by equation (14).

3.2.2 Discretization of the delay differential equation

We discritize the constraint $\dot{x}(t) = ax(t) + bx(t-r) + cu(t) + du(t-r)$ by using Cranck Nilcholson Method and assuming that there exists $x(t_k - r) = x_{k-r}$ and $u(t_k - r) = u_{k-r}$. This is true for rational values of r and we ensure that our point of discretization falls exactly on r to avoid off grid points.
We recall from Cranck Nilcholson Method that

$$y_{n+1} = y_n + \frac{h}{2}[f_{n+1} + f_n]$$

This implies that

$$x_{j+1} = ax_j + \frac{h}{2}(ax_{j+1} + bx_{j+1-r} + cu_{j+1} + du_{j+1-r} + ax_j + bx_{j-r} + cu_j + du_{j-r})$$

$$(2-ah)x_{k+1} = (2a+ah)x_k + bhx_{k+1-r} + chu_{k+1} + dhu_{k+1-r} + bhx_{k-r} + chu_k + dhu_{k-r}$$

$$x_{k+1} = Ux_k + Vx_{k-r} + Vx_{k+1-r} + Wu_k + Wu_{k+1} + Xu_{k-r} + Xu_{k+1-r} \tag{34}$$

where $U = \frac{a(2+h)}{2-ah}$,$V = \frac{bh}{2-ah}$, $W = \frac{ch}{2-ah}$ and $X = \frac{dh}{2-ah}$
Considering (34), there exists m such that $m = \frac{r}{h}$, where $h = \Delta t_k$. For $t \in [-r, 0]$, where $x(t-r) = x_{t-r}$ and $u(t-r) = u_{t-r}$ are known to be constants for each value of r in the interval.
when $k = 0$

$$x_1 - Vx_{-r} - Vx_{1-r} - Wu_0 - Wu_1 - Xu_{-r} - Xu_{1-r} = Ux_0 \tag{35}$$

when $k = 1$

$$x_2 - Ux_1 - Vx_{1-r} - Vx_{2-r} - Wu_1 - Wu_2 - Xu_{1-r} - Xu_{2-r} = 0 \tag{36}$$

when $k = 2$

$$x_3 - Ux_2 - Vx_{2-r} - Vx_{3-r} - Wu_2 - Wu_3 - Xu_{2-r} - Xu_{3-r} = 0 \tag{37}$$

when $k = 3$

$$x_4 - Ux_3 - Vx_{3-r} - Vx_{4-r} - Wu_3 - Wu_4 - Xu_{3-r} - Xu_{4-r} = 0 \tag{38}$$

when $k = 4$

$$x_5 - Ux_4 - Vx_{4-r} - Vx_{5-r} - Wu_4 - Wu_5 - Xu_{4-r} - Xu_{5-r} = 0 \tag{39}$$

when $k = 5$

$$x_6 - Ux_5 - Vx_{5-r} - Vx_{6-r} - Wu_5 - Wu_6 - Xu_{5-r} - Xu_{6-r} = 0 \tag{40}$$

$\vdots$

when $k = m-1$

$$x_m - Ux_{m-1} - Vx_{m-1-r} - Vx_{m-r} - Wu_{m-1} - Wu_m - Xu_{m-1-r} - Xu_{m-r} = 0 \tag{41}$$

when $k = m$

$$x_{m+1} - Ux_m - Vx_{m-r} - Vx_{m+1-r} - Wu_m - Wu_{m+1} - Xu_{m-r} - Xu_{m+1-r} = 0 \tag{42}$$

when $k = m+1$

$$x_{m+2} - Ux_{m+1} - Vx_{m+1-r} - Vx_{m+2-r} - Wu_{m+1} - Wu_{m+2} - Xu_{m+1-r} - Xu_{m+2-r} = 0 \tag{43}$$

$\vdots$

when $k = N-1$

$$x_N - Ux_{N-1} - Vx_{N-1-r} - Vx_{N-r} - Wu_{N-1} - Wu_N - Xu_{N-1-r} - Xu_{N-r} = 0 \tag{44}$$

Equations (35) - (44), can be written in matrix form as

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -W & -W & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -U & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -W & -W & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -U & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -W & -W & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -U & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -W & -W & 0 & 0 & 0 & 0 \\ -V & 0 & 0 & -U & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -W & -W & 0 & 0 & 0 \\ 0 & -V & -V & 0 & -U & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -W & -W & 0 & 0 \\ 0 & 0 & -V & -V & 0 & -U & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -W & -W & 0 \\ 0 & 0 & 0 & -V & -V & 0 & -U & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -W & -W \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \\ x_{m+1} \\ \vdots \\ x_n \\ u_0 \\ u_1 \\ \vdots \\ u_m \\ u_{m+1} \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} Ux_0 + Vu_{-r} + Vu_{1-r} \\ Vu_{2-r} + Vu_{1-r} \\ \vdots \\ eu_{m-1-r} \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{45}$$

Equation (45) can be written as

$$GZ = K$$

where G is a block matrix of dimension $N \times (2N+1)$ with representation $G = (C \vdots E)$ where $[C]$ is an $N \times N$ sparse matrix with principal diagonal elements $[C]_{ij} = 1$, lower diagonal elements -U for every i, j such that $i = j+1$ and -V for every i, j such that $i = j+4$. $[E]$ is $N \times (N+1)$ bidiagonal matrix with $[E]_{ij} = -W$ and upper diagonal elements $-W$ for every i, j such that $j = i+1$. $[Z]$ is also a column vector of dimension $(2N+1) \times 1$. The column vector $[K]$ is of order $n \times 1$ with entries given by $[K]_{11} = Ux_0 + Vu_{-r} + Vu_{1-r}$, $[K]_{12} = Vu_{2-r} + Vu_{1-r}$, . . ., $[K]_{N1} = eu_{m-1-r}$

3.2.3 Conversion of the inequality constraint to an equality constraint

We convert the inequality constraint $ex(t) + fu(t) \leq 0$ to an equality constraint by introducing nonnegative slack variables $s = (s_0, s_1, \ldots, s_N)$, where $s_k \geq 0$ for all k. This gives:

$$ex_k + fu_k + s_k = 0, \quad \text{with } s_k \geq 0, \quad \text{for } k = 0, 1, \ldots, N \tag{46}$$

Equation (46) can be written in matrix form as:

$$\begin{pmatrix} e & 0 & \cdots & 0 & f & 0 & \cdots & 0 \\ 0 & e & \cdots & 0 & 0 & f & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e & 0 & 0 & \cdots & f \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \\ u_0 \\ u_1 \\ \vdots \\ u_N \end{pmatrix} + \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_N \end{pmatrix} = \begin{pmatrix} -ex_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

This can be written compactly as:

$$LZ + S = N \tag{47}$$

where $S = (s_1, s_2, \ldots, s_N)^T$ is the vector of nonnegative slack variables.

The corresponding complementarity conditions are:

$$\mu_k \geq 0, \quad s_k \geq 0, \quad \mu_k s_k = 0, \quad \text{for } k = 1, \ldots, N.$$

3.3 Reformulation of the constrained problem into an unconstrained optimal control framework

Now, by parameter optimization, the discretized optimal control problem becomes a large sparse quadratic programming problem written as:

$$\begin{aligned} &\text{Minimize} \quad I(Z) = Z^T AZ + C^T Z \\ &\text{Subject to} \quad GZ = K \\ &\text{and} \quad LZ + S = N, \quad S \geq 0 \end{aligned}$$

Hence, the unconstrained form of the original problem using Augmented Lagrangian Method, $L_{\mu,\lambda}(Z, S, \mu, \lambda) : \Re^{n+m+1} \rightarrow \Re$, becomes:

$$\begin{aligned} L_{\mu,\lambda}(Z, S, \mu, \lambda) = Z^T AZ + C^T Z + \frac{1}{2}\mu_i \|GZ - K\|^2 + \lambda_i^T (GZ - K) \\ + \frac{1}{2}\mu_i \|LZ + S - N\|^2 + \gamma_i^T (LZ + S - N) \end{aligned}$$

Expanding and collecting like terms, we have:

$$\begin{aligned} L_{\mu,\lambda}(Z,S,\mu,\lambda) &= Z^TAZ + C^TZ + \frac{1}{2}\mu_i(GZ-K)(GZ-K)^T + \lambda_i^T(GZ-K) \\ &+ \frac{1}{2}\mu_i(LZ+S-N)(LZ+S-N)^T + \gamma_i^T(LZ+S-N) \end{aligned}$$

$$\begin{aligned} L_{\mu,\lambda}(Z,S,\mu,\lambda) &= Z^TAZ + C^TZ + \frac{1}{2}\mu_i(G^TGZ^TZ - 2K^TGZ + K^TK) + \lambda_i^T(GZ-K) \\ &+ \frac{1}{2}\mu_i(Z^TL^TLZ + 2Z^TL^T(S-N) + (S-N)^T(S-N)) \\ &+ \gamma_i^T(LZ+S-N) \end{aligned}$$

$$\begin{aligned} L_{\mu,\lambda}(Z,S,\mu,\lambda) &= Z^T\left(A + \frac{\mu_i}{2}G^TG + \frac{\mu_i}{2}L^TL\right)Z \\ &+ \left[C^T - \mu_i(K^TG + N^TL) + \mu_i(S-N)^TL + \lambda_i^TG + \gamma_i^TL\right]Z \\ &+ \frac{\mu_i}{2}(S-N)^T(S-N) + \gamma_i^T(S-N) + \left(C^T + \frac{\mu_i}{2}K^TK - \lambda_i^TK\right) \end{aligned}$$

Now, by parameter optimization, the discretized optimal control problem becomes a large sparse quadratic programming problem written as:

$$\begin{aligned} \text{Minimize} \quad & I(z) = Z^TAZ + C^T \\ \text{Subject to} \quad & GZ = K \\ \text{and} \quad & LM = N \end{aligned}$$

Hence, the unconstrained form of the original problem using Augumented Lagrangian Method, $L_{\mu,\lambda}(Z,\mu,\lambda) : \Re^{n+m+1} \to \Re$, becomes

$$\begin{aligned} \text{Minimize } L_{\mu,\lambda}(Z,\mu,\lambda) = \quad & Z^TAZ + C^T + \frac{1}{2}\mu_i\|GZ-K\|^2 + \lambda_i^T(GZ-K) + \\ & \frac{1}{2}\mu_i\|LZ-N\|^2 + \lambda_i^T(LZ-N) \end{aligned}$$

Expanding and collecting like terms, we have

$$\begin{aligned} L_{\mu,\lambda}(Z,\mu,\lambda) = \quad & Z^TAZ + C^T + \frac{1}{2}\mu_i(GZ-K)(GZ-K)^T + \lambda_i^T(GZ-K) \\ & \frac{1}{2}\mu_i(LZ-N)(LZ-N)^T + \lambda_i^T(LZ-N) \end{aligned}$$

$$\begin{aligned} L_{\mu,\lambda}(Z,\mu,\lambda) = \quad & Z^TAZ + C^T + \frac{1}{2}\mu_i(G^TGZ^TZ - 2K^TGZ + K^TK) + \lambda_i^T(GZ-K) \\ & \frac{1}{2}\mu_i(L^TLZ^TZ - 2N^TLZ + N^TN) + \lambda_i^T(LZ-N) \end{aligned}$$

$$\begin{aligned} L_{\mu,\lambda}(Z,\mu,\lambda) = \quad & Z^T(A + \frac{\mu_i}{2}G^TG + \frac{\mu_i}{2}L^TL)Z - \mu_i(K^TG + N^TL)Z + \\ & \lambda_i^T((G+L)Z - K - N) + (C^T + \frac{\mu i}{2}(K^TK + N^TN)) \end{aligned} \tag{48}$$

Equation (48) is a programming problem which can be solved via Conjugate Gradient Method, where

$$A_{\mu_i} = A + \frac{\mu_i}{2}G^T G + \frac{\mu_i}{2}L^T L \text{ with dimension}(2M+1)\times(2M+1) \quad (49)$$

$$B^T = -\mu_i(K^T G + N^T L)Z \text{ with dimension } 1\times(2M+1) \quad (50)$$

$$C = C^T + \frac{\mu i}{2}(K^T K + N^T N) + \lambda_i^T((G+L)Z - K - N) \quad (51)$$

$$\text{with dimension } 1\times 1$$

Hence, by problem (49)-(51), the penalized matrix A_{μ_i} is positive definite for sufficiently large penalty parameter μ_i.

3.4 Convergence Analysis

The following theorem establishes the linear convergence rate of the proposed ALM-ECGM scheme for the discretized DDE-constrained optimal control problem.

Theorem 3.1. *Consider the discretized quadratic programming problem:*

$$\min_Z J(Z) = Z^T AZ + C^T Z \quad \textit{subject to} \quad GZ = K, \quad LZ = N$$

Under the assumptions that:

1. *A is symmetric positive definite,*
2. *The combined constraint matrix $\begin{bmatrix} G \\ L \end{bmatrix}$ has full row rank,*
3. *The penalty parameter μ_k is chosen sufficiently large and non-decreasing,*

the ALM-ECGM algorithm converges linearly. That is, there exist constants $\eta \in (0,1)$ and $M > 0$ such that:

$$\|Z_k - Z^*\| + \|\lambda_k - \lambda^*\| + \|\gamma_k - \gamma^*\| \le M\eta^k$$

for all sufficiently large k.

Proof. The convergence analysis proceeds in two parts:

Part 1: Outer ALM Convergence. For the quadratic programming problem with linear constraints, the Augmented Lagrangian Method exhibits linear convergence when the penalty parameter μ_k is sufficiently large. The update:

$$Z_{k+1} = \arg\min_Z L_{\mu_k}(Z, \lambda_k, \gamma_k), \quad \lambda_{k+1} = \lambda_k + \mu_k(GZ_{k+1} - K), \quad \gamma_{k+1} = \gamma_k + \mu_k(LZ_{k+1} - N)$$

satisfies the bound:

$$\|Z_{k+1} - Z^*\| + \|(\lambda_{k+1}, \gamma_{k+1}) - (\lambda^*, \gamma^*)\| \le \frac{C}{\mu_k}\left(\|Z_k - Z^*\| + \|(\lambda_k, \gamma_k) - (\lambda^*, \gamma^*)\|\right)$$

for some constant $C > 0$. With μ_k sufficiently large and non-decreasing, this guarantees linear convergence.

Part 2: Inner ECGM Convergence. Each inner subproblem $\min_Z L_{\mu_k}(Z, \lambda_k, \gamma_k)$ is strictly convex quadratic since $A_\mu = A + \frac{\mu}{2}(G^T G + L^T L)$ is positive definite. The Extended Conjugate Gradient Method applied to this subproblem converges linearly with error bound:

$$\|Z^{(m)} - Z^*_{k+1}\|_A \le 2\left(\frac{\sqrt{\kappa(A_\mu)} - 1}{\sqrt{\kappa(A_\mu)} + 1}\right)^m \|Z^{(0)} - Z^*_{k+1}\|_A$$

where $\kappa(A_\mu)$ is the condition number. By terminating the inner iterations with accuracy $\epsilon_k = O(\eta^k)$, the overall ALM-ECGM scheme maintains linear convergence.

The assumptions of Theorem 1 are satisfied for the discretized DDE-constrained problems considered in this work, as A arises from a positive definite cost functional and the discretization preserves constraint linear independence. The numerical results in Tables 1 and 2 demonstrate this linear convergence behavior empirically.

4 Numerical simulations

Example 1.

$$\begin{aligned} \text{Minimum } I(t, x(t), u(t)) = & \int_0^1 (x^2(t) + u^2(t))dt \\ \dot{x} = & \; 5x(t) + x(t-0.3) + 3u(t), \; x(0) = 0, t \in [-0.3, 0] \\ & u(t) - x(t) \leq 0 \end{aligned}$$

Table 1: Constraint Satisfaction and Objective Function Values

S/N	Constraint satisfaction (Proposed Method)	objective function value (Proposed Method)
1	0.1970E-0	0.9875
2	0.3050E-1	1.5281
3	0.3200E-2	1.6166
4	0.3000E-3	1.6260
5	0.3000E-4	1.6260
6	0.3000E-5	1.6260

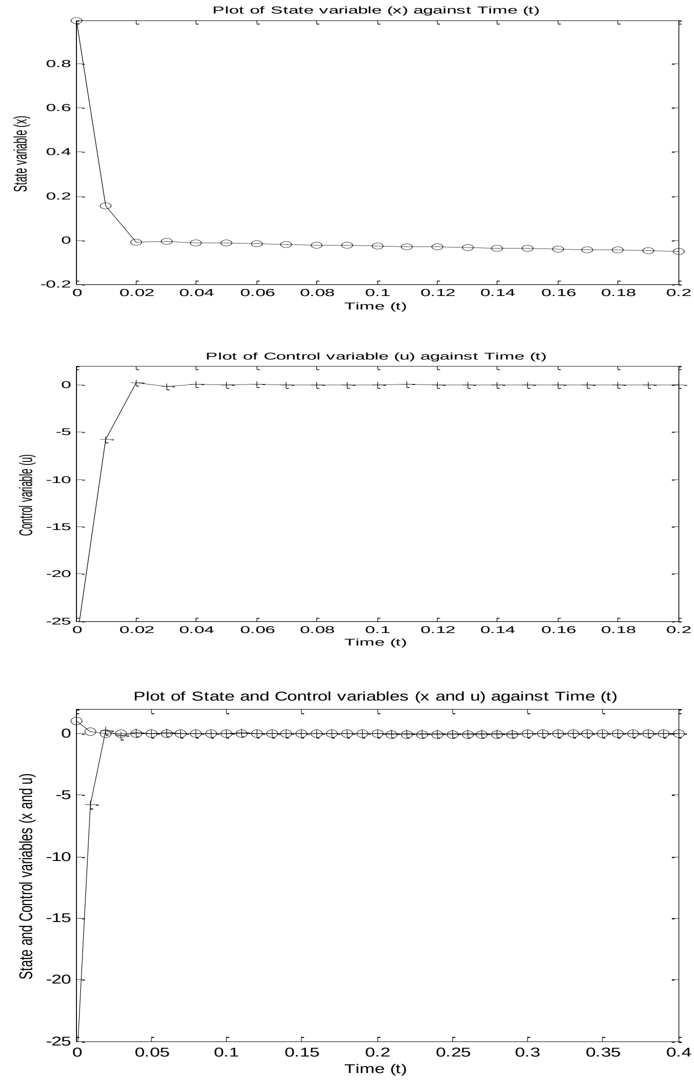

Figure 1: Evolution of the State and Control Variables of the DDE-Constrained Problem for Problem 1

The numerical solution, obtained via the ALM-ECGM scheme, exhibits excellent convergence properties. As detailed in Table 1, the constraint violation, measured as $\max(u(t) - x(t))$, decreases systematically from an initial value of 0.1970 to a negligible magnitude of 0.3000×10^{-5} over the course of the iterations. This monotonic reduction in constraint violation is a hallmark of the Augmented Lagrangian Method's effectiveness, as it enforces constraints without the ill-conditioning associated with pure penalty methods.

Concurrently, the objective functional value converges robustly to a final value of 1.6260. The stabilization of the objective value, coinciding with the high-degree of constraint satisfaction, indicates that the algorithm has reached a feasible and optimal solution. The optimal control $u(t)$ is successfully computed to always lie below the state trajectory $x(t)$, thereby adhering to the problem's physical or design limitation.

Example 2.

$$\begin{aligned} \text{Minimum } 0I(t, x(t), u(t)) = & \int_0^1 (x^2(t) + u^2(t))dt \\ \dot{x} = & \; 5x(t) + x(t-0.3) + 3.021u(t),\ x(0) = 1, t \in [-0.3, 0] \\ & u(t) \leq 1 \end{aligned}$$

Table 2: Constraint satisfaction and objective function values

S/N	Constraint satisfaction (Proposed Method)	objective function value (Proposed Method)
1	0.1449E-0	0.5388
2	0.2230E-1	0.8298
3	0.2400E-2	0.8772
4	0.2000E-3	0.8822
5	0.2000E-4	0.8822
6	0.2000E-5	0.8822

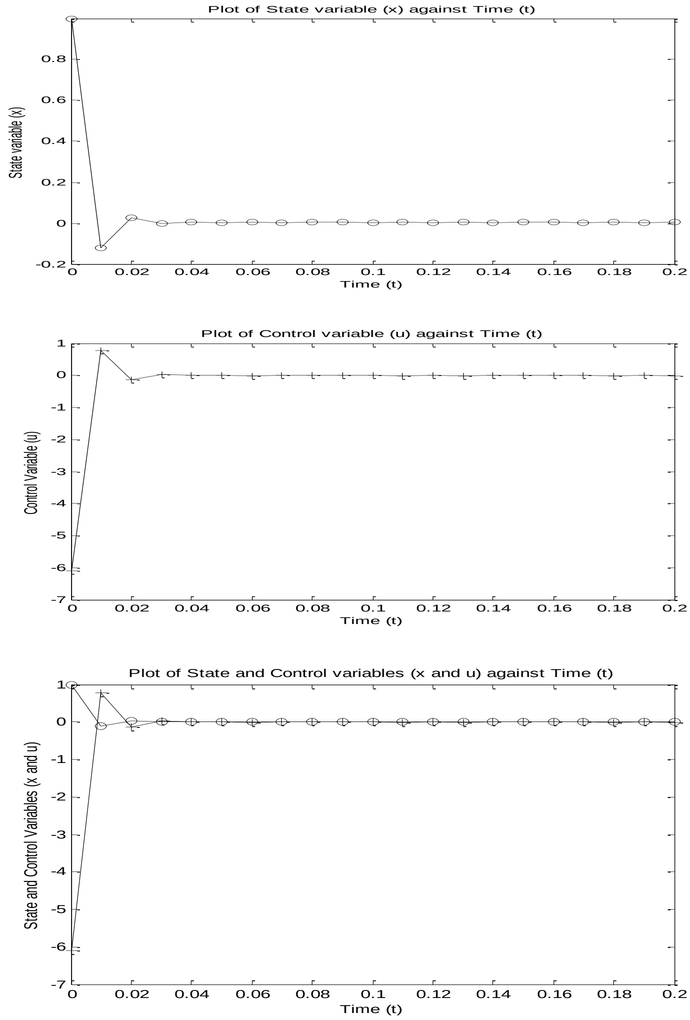

Figure 2: Evolution of the State and Control Variables of the DDE-Constrained Problem for Problem 2

The results for this problem, summarized in Table 2, further confirm the robustness of the proposed method. The constraint violation for the control bound $u(t) \leq 1$ is efficiently driven down from 0.1449 to 0.2000×10^{-5}. This demonstrates the method's capability in handling different types of inequality constraints, whether they are mixed (as in Example 1) or pure control constraints.

The objective functional converges to an optimal value of 0.8822. The higher initial cost compared to Example 1 is a direct consequence of the non-zero initial state $x(0) = 1$, which immediately contributes to the cost integral. The algorithm successfully finds a control strategy that regulates the system from this initial state while strictly respecting the control saturation limit.

5 Discussions and conclusion

This study has successfully developed and validated a novel numerical scheme that integrates the Augmented Lagrangian Method (ALM) with the Extended Conjugate Gradient Method (ECGM) for solving OCPs subject to mixed constraints and governed by DDEs. The research establishes a unified computational framework that effectively handles both ordinary and delay differential constraints alongside inequality limitations.

The results demonstrate the scheme's robust performance and computational efficiency. For problems with ODE constraints, the numerical solutions exhibit excellent agreement with analytical results, validating the proposed discretization and solution methodology. In the more complex case of DDE constraints, the ALM-ECGM scheme outperforms existing methods, producing superior solutions with consistent convergence behavior. The comprehensive convergence analysis confirms the method's linear convergence rate, while the progressive constraint satisfaction and stabilization of objective function values to their optima further substantiate the algorithm's reliability and numerical stability.

The integration of ALM has proven particularly effective in managing mixed constraints without the ill-conditioning issues associated with traditional penalty methods, while the ECGM efficiently solves the resulting large-scale unconstrained optimization subproblems. This research thus provides a powerful and computationally stable approach for addressing challenging OCPs with delayed dynamics and mixed constraints, representing a significant advancement in the numerical treatment of such systems.

References

[1] Polak, E. (1973). An Historical Survey of Computational Methods in Optimal Control. SIAM Review, 15(2), 552–584.

[2] Rao, S. S. (2019). Engineering Optimization: Theory and Practice, 5th Edition. John Wiley and Sons, Hoboken, NJ, USA, pp. 1–864.

[3] Izmailov, A.F., Solodov, M.V. and Uskov, E.I. (2012). Global convergence of augmented lagrangian methods applied to optimization problems with degenerate constraints, including problems with complementarity constraints. SIAM Journal on Optimization, 22(4), 1579–1606.

[4] Afolabi, A.S. and Olotu, O. (2020). Analytical Solutions of Optimal Control Problems with Mixed Constraints. Iconic Research and Engineering Journals, 3(12), 1–9, India.

[5] Afolabi, A.S. (2022). On the Constrained Optimal Control Problems with Mixed Constraints. International Journal of Mathematical Analysis and Modelling, 5(2), 25–35, India.

[6] Afolabi, A.S. and Oyewale, A.A. (2025). On the Solutions of Optimal Control Problems Constrained by Ordinary Differential Equations with Vector-Matrix Coefficients Using FICO Xpress Mosel. Journal of Mathematical Sciences and Optimization: Theory and Application, 11(2), 111–132, Nigeria.

[7] Crankson, M.V., Afolabi, A.S. and Eshun, S. (2023). Comprehensive Semi-Analytical Approach to Quadratic Optimal Control Problems Governed by Ordinary Differential Equations. International Journal of Science and Research, 7(942), 2319–7064, India.

[8] Orji, J.C., Miswanto, and Afolabi, A.S. (2024). Optimal Control Problems Constrained by Ordinary Differential Equations Using Conjugate Gradient Method and FICO Xpress Mosel. International Research and Innovation Journal, 1(1), 134–149, UK.

[9] Oyewale, A.A., Afolabi, A.S. and Dawodu, K.A. (2024). On the Optimal Control Problems Constrained by Ordinary Differential Equations Using Conjugate Gradient Method and FICO Xpress Mosel. Journal of Mathematics and Computational Science, 14(18), 1–25, UK.

[10] Lasdon, L.S., Mitter, S.K. and Warren, A.D. (1967). The Conjugate Gradient Method for Optimal Control Problems. IEEE Transactions on Automatic Control, 12(2), 132–138.

[11] Zboon, R.A., Yadav, S.P. and Mohan, C. (1999). Penalty Method for an Optimal Control Problem with Equality and Inequality Constraints. Indian Journal of Pure and Applied Mathematics, 30(1), 1–14.

[12] Olotu, O. and Akeremale, O.C. (2012). Optimization of Quadratic Constrained Optimal Control Problem Using Augmented Lagrangian Function. Journal of the Nigerian Association of Mathematical Physics, 20, 185–192.

[13] Olotu, O. and Adekunle, A.I. (2010). Analytic and Numeric Solutions of Discretized Constrained Optimal Control Problem with Vector-Matrix Coefficients. AMO – Advanced Modeling and Optimization, 12(1), 119–131.

[14] Olotu, O. (2011). Discretized Scheme Examining Convergence and Geometric Ratio Profiles for Control Problem Constrained by Evolution Equation. Journal of Mathematics and Statistics, 7(2), 116–123.

[15] Olotu, O. and Afolabi, A.S. (2017). On the Optimal Control Problems Constrained by Delay Differential Equation Using Penalty Functions with B-Spline Function. Journal of Mathematical Physics, Benin, Nigeria, 39, 119–128.

[16] Adebayo, K.J., Aderibigbe, F.A. and Akinmuyise, M.F. (2014). On Application Of Modified Gradient to Extended Conjugate Gradient Method Algorithm For Solving Optimal Control Problems. IOSR Journal of Mathematics, 9(5), 30–35.

[17] Kahuru, J., Luboobi, L.S. and Gyekye, Y.N. (2017). Optimal Control Techniques on a Mathematical Model for the Dynamics of Tungiasis in a Community. International Journal of Mathematics and Mathematical Sciences, 2017, Article ID 4804897, 1–19.

[18] Sofia, H. (2014). Optimal Control and Numerical Optimization Applied to Epidemiological Models. Ph.D. Thesis, Department of Mathematics, University of Aveiro, pp. 1–111.