

UNIVERSIDAD NACIONAL DE ASUNCIÓN

**Facultad de Ingeniería
Ingeniería Electrónica**



**Desarrollo de una Aplicación Paralela en el Tiempo para la
Simulación y Control de Ecuaciones Diferenciales Parciales
Parabólicas**

Juan Rodrigo Fleitas Figueredo

Diego Herbin Stalder Díaz

San Lorenzo, Paraguay

2010

Miembros del Consejo Directivo

Consejeros Titulares

Prof. Ing. Carlos Dellavedova E. (Decano)

Prof. Ing. Isacio Vallejos Aquino (Vice Decano)

Prof. Ing. Amilcar Troche (Docente)

Prof. Ing. Cesar Sitjar (Docente)

Prof. Ing. Francisco R. Delgado Marquez (Docente)

Prof. Ing. Diógenes Sartorio (Docente)

Prof. Ing. Cirilo Hernáez (Docente)

Prof. Ing. Ramón Pistilli S. (Docente - C.S.U.)

Ing. Pablo Rodriguez (No Docente)

Ing. Gabriel Enrique Fleitas Ferrari (No Docente)

Est. Víctor Omar Morínigo López (Estudiante)

Est. Jorge Manuel Muller Giménez (Estudiante)

Est. Vicente Javier Chaparro Ruíz Díaz (Estudiante)

Consejeros Suplentes

Prof. Ing. Carlos M Montero Volpe (Docente - CSU)

Prof. Ing. María Teresa Pino (Docente)

Ing. Ignacio Daniel Velázquez G. (No Docente)

Ing. Alfredo Javier Coronel Correa (No Docente)

Est. Víctor Daniel Portillo Galván (Estudiante)

Est. Justo Javier Fernández Arguello (Estudiante)

Est. Sergio Daniel Vera Urquhart (Estudiante)

**DESARROLLO DE UNA APLICACIÓN PARALELA EN EL TIEMPO PARA LA SIMULACIÓN Y
CONTROL DE ECUACIONES DIFERENCIALES PARCIALES PARABÓLICAS**

*A nuestros padres, por inculcarnos los valores que nos formaron como personas
por motivarnos a seguir el camino de las ciencias
y por enseñarnos con el ejemplo*

AGRADECIMIENTOS

a Christian Schaerer, por la paciencia, el apoyo y la orientación para la
realización de éste trabajo

a los profesores Cesar Moreira y Miki Saito, por la orientación y el apoyo
continuo durante de nuestra formación

a Hyun Ho Shin, por la ayuda para el buen desarrollo del trabajo

a nuestros amigos y compañeros, por los excelentes momentos compartidos al
pasar por esta Casa de Estudios

AGRADECIMIENTOS INSTITUCIONALES

a la Facultad de Ingeniería de la UNA
al Laboratorio de Computación Científica y Aplicada (LCCA),
por el espacio de trabajo proporcionado

ABSTRACT

In this work we consider the Parareal Algorithm, developed in [1], as a preconditioner for solving the optimal boundary control problem for cooling an electronic circuit. The constitutive equations are obtained by modeling the circuit as a parabolic partial differential equation, where the cooling mechanisms are modeled using Neumann boundary conditions (control variables) [5]. The optimal control problem is obtained by associating to the state and control variables, a linearized quadratic cost function like in [15]. The classical finite element method - (Galerkin)- was implemented for the spatial discretization resulting in a large scale system ODEs [21].

For the implementation of the preconditioner the time domain is divided into coarse time intervals, where the initials conditions are arbitrary. Each of these coarse time intervals are divided into fine time intervals that are solved simultaneously in several processors. The Conjugate Gradient method is used to update the initial conditions and the control variables, the Parareal method is used as a preconditioner to accelerate its convergence.

RESUMEN

Este trabajo implementa el método Parareal, desarrollado en [1] como preconditionador para resolución del problema de control óptimo que involucra el enfriamiento de un circuito electrónico. Las ecuaciones constitutivas se obtienen modelando el circuito como una ecuación diferencial parcial parabólica, donde el mecanismo de enfriamiento es modelado usando condiciones de frontera (variables de control) de Neumann [5]. La formulación de control es obtenida asociando al estado y a las variables de control un funcional de costo lineal-cuadrático [15]. Luego, el problema de control se transforma en un sistema de ecuaciones diferenciales ordinarias aplicando el método de los elementos finitos clásico - (Galerkin)-[21].

Para resolver el problema de control óptimo en forma paralela se realiza una descomposición gruesa del dominio temporal (intervalos grandes) donde se arbitran las condiciones iniciales en cada subdominio. Posteriormente, cada subdominio de la malla gruesa es nuevamente dividido en intervalos menores, determinando en cada subdominio una malla fina que se resuelve simultáneamente en cada procesador. El método Gradiente Conjugado es utilizado para corregir las condiciones iniciales y las funciones de control usando los valores de la solución calculada en cada procesador. Para acelerar la convergencia del método Gradiente Conjugado usamos el método Parareal como preconditionador [1],[15].

ÍNDICE DE CONTENIDO

1. INTRODUCCIÓN	8
1.1. Preliminares	8
1.2. Motivación del Paralelismo	10
1.3. Tendencias de la Ciencia e Ingeniería	12
1.4. Arquitecturas de Cómputo	14
1.4.1. Paso de Mensajes	15
1.5. Antecedentes	16
1.6. Definición del Problema	17
1.7. Objetivos	17
1.7.1. Objetivos Generales	17
1.7.2. Objetivos Específicos	18
2. Problema de control	19
2.1. Formulación matemática	19
2.2. Problema de control óptimo	20

DESARROLLO DE UNA APLICACIÓN PARALELA EN EL TIEMPO PARA LA SIMULACIÓN Y CONTROL DE ECUACIONES DIFERENCIALES PARCIALES PARABÓLICAS

2.3. Condiciones de optimalidad	24
2.3.1. Ecuaciones de estado lineales	25
2.4. Método de solución del problema de control óptimo	27
2.4.1. Discretización temporal	34
3. Algoritmo Parareal	39
3.1. Formulación Matemática	39
3.2. Condiciones de Optimalidad	42
3.3. Método de Solución del Problema de Control Paralelo	44
3.4. Gradiente Conjugado Precondicionado	53
4. Resultados Numéricos	60
4.1. Control óptimo del transitorio térmico de una placa de circuitos electrónicos mediante un sistema de enfriamiento en los bordes	60
4.1.1. Ecuación de conducción de calor con condiciones de fron- tera de Neumann	60
4.1.2. Problema de control óptimo	66
4.1.3. Discretización del dominio espacial (Ω)	67
4.2. Método Gradiente Conjugado	68
4.2.1. Ajuste del controlador	69
4.2.2. Cantidad de Iteraciones vs tolerancia	73
4.2.3. Cantidad de Iteraciones vs cantidad de elementos del do- minio espacial	73

DESARROLLO DE UNA APLICACIÓN PARALELA EN EL TIEMPO PARA LA SIMULACIÓN Y CONTROL DE ECUACIONES DIFERENCIALES PARCIALES PARABÓLICAS

4.3. Método Gradiente Conjugado Precondicionado paralelo	74
4.3.1. Ajuste de la penalización ε	74
4.3.2. Cantidad de Iteraciones vs tolerancia	76
4.3.3. Cantidad de iteraciones vs cantidad de elementos del do- minio espacial	77
4.3.4. Escalabilidad débil	78
4.3.5. Escalabilidad Fuerte	79
5. Conclusiones y Trabajos Futuros	80
5.1. Conclusión	80
5.1.1. Trabajos Futuros	82
A. Apéndice: Problema de control	83
A.1. Condiciones de optimalidad KKT del Lagrangiano	83
A.1.1. Ecuaciones de estado Lineales	88
A.2. Apéndice: Método de solución del problema de control óptimo . .	95
A.2.1. Gradiente conjugado para un problema cuadrático	95
A.2.2. Generación de direcciones conjugadas	99
A.2.3. Cálculo del incremento para la función de control	105
A.2.4. Gradiente del Lagrangiano y del funcional costo	107
A.2.5. Producto de la matriz Hessiana por un vector	111
B. Apéndice: Algoritmo Parareal	114

DESARROLLO DE UNA APLICACIÓN PARALELA EN EL TIEMPO PARA LA SIMULACIÓN Y CONTROL DE ECUACIONES DIFERENCIALES PARCIALES PARABÓLICAS

B.1. Condiciones de KKT del Lagrangiano	114
B.2. Calculo del $\nabla \mathcal{J}$	122
B.3. Calculo de Hd_k^i	128
B.4. Matriz de Precondicionamiento	133
C. Apéndice: Discretización espacial	141
C.1. Ecuación de conducción de calor con condiciones de frontera de Neumann	141
C.1.1. Formulación Variacional	141
C.1.2. Discretización por el método de Galerkin	144
C.2. Funcional de costo	146
Bibliografía	149

ÍNDICE DE FIGURAS

1.1. Rendimiento de los Super-computadores.	9
1.2. Proyecciones de Rendimiento de los Super-computadores.	10
1.3. Etapas de la simulación. Fuente: Elaboración propia.	14
1.4. Funcionamiento del MPI. Fuente: Elaboración propia.	15
2.1. Problema de Control. Fuente: Elaboración propia.	20
2.2. Diferentes tipos de planteamiento del problema de control óptimo. Fuente: [22].	21
3.1. Temperatura calculada en forma Paralela.	54
3.2. Estado Adjunto en Paralelo. Fuente: Elaboración Propia.	55
3.3. Estados malla gruesa. Fuente: Elaboración Propia.	57
4.1. Ejemplo de aplicación: Control de temperatura de circuitos elec- trónicos. Fuente: Elaboración Propia.	62
4.2. Caso: $s = 1$, (a) Función de control en 1, 2, 3, (b) Temperatura en 4, 5, 6, (c) Temperatura [K] p/t=10[s]. Fuente: Elaboración Propia.	70

DESARROLLO DE UNA APLICACIÓN PARALELA EN EL TIEMPO PARA LA SIMULACIÓN Y CONTROL DE ECUACIONES DIFERENCIALES PARCIALES PARABÓLICAS

4.3. Caso: $s = 10000$, (a) Función de control en 1, 2, 3, (b) Temperatura en 4, 5, 6, (c) Temperatura [K] p/t=10[s]. Fuente: Elaboración Propia.	70
4.4. Caso: $s = 100000$, (a) Función de control en 1, 2, 3, (b) Temperatura en 4, 5, 6, (c) Temperatura [K] p/t=10[s]. Fuente: Elaboración Propia.	70
4.5. Caso: $q = 0,1$, (a) Función de control en 1, 2, 3, (b) Temperatura en 4, 5, 6, (c) Temperatura [K] p/t=10[s]. Fuente: Elaboración Propia.	72
4.6. Caso: $q = 1$, (a) Función de control en 1, 2, 3, (b) Temperatura en 4, 5, 6, (c) Temperatura [K] p/t=10[s]. Fuente: Elaboración Propia.	72
4.7. Caso: $q = 10$,(a) Función de control en 1, 2, 3, (b) Temperatura en 4, 5, 6, (c) Temperatura [K] p/t=10[s]. Fuente: Elaboración Propia.	72
4.8. Caso: $\varepsilon = 10^{-1}$, (a) Función de control en 1, 2, 3, (b) Temperatura en 4, 5, 6, (c) Temperatura [K] p/t=10[s]. Fuente: Elaboración Propia.	75
4.9. Caso: $\varepsilon = 10^{-3}$, (a) Función de control en 1, 2, 3, (b) Temperatura en 4, 5, 6, (c) Temperatura [K] p/t=10[s]. Fuente: Elaboración Propia.	75
4.10.Caso: $\varepsilon = 10^{-4}$, (a) Función de control en 1, 2, 3, (b) Temperatura en 4, 5, 6, (c) Temperatura [K] p/t=10[s]. Fuente: Elaboración Propia.	76
4.11.Norma del gradiente del funcional en función a la cantidad de iteraciones. Fuente: Elaboración Propia.	77

ÍNDICE DE TABLAS

4.1. Número de iteraciones del Método Gradiente Conjugado para un valor fijo de $r = 1, q = 0$	69
4.2. Norma del gradiente del funcional al modificar las relaciones $r/const$ y $s/const$	71
4.3. Ajuste de q	71
4.4. Cantidad de Iteraciones vs tolerancia	73
4.5. Influencia de la cantidad elementos del dominio espacial	74
4.6. Ajuste de ε	75
4.7. Cantidad de Iteraciones vs tolerancia	77
4.8. Cantidad de iteraciones vs cantidad de elementos del dominio espacial	78
4.9. Escalabilidad débil	78
4.10. Escalabilidad Fuerte.	79

CAPÍTULO 1

INTRODUCCIÓN

1.1. Preliminares

El avance de las tecnologías de diseño de los microprocesadores ha hecho que las velocidades de reloj de los mismos se incrementen rápidamente, de unos 25 MHz (Intel 80486) en 1989 a 3.8 Ghz (Intel Pentium 4) en 2005 [6]. En los últimos años los planes para diseño de microprocesadores de 4 o más GHz fueron cancelados y se les dio prioridad a los proyectos para fabricar procesadores con varios núcleos [7]. Consecuentemente los procesadores actuales son capaces de ejecutar múltiples instrucciones en un ciclo de reloj, lo cual se traduce en un incremento de la tasa máxima de operaciones de punto flotante por segundo (FLOPS) [3].

El interés en la computación paralela se ha incrementado debido a la demanda continua de capacidad de cálculo (memoria y procesamiento), necesaria para el modelado, simulación numérica de problemas en ciencias e ingeniería (por ejemplo [8]), procesamiento de datos [10], aplicaciones gráficas y visualización de ambientes virtuales en tiempo real [<http://www.paraview.org/>]

En la siguiente figura observamos el aumento de la capacidad de procesamiento que han tenido los super-computadores durante los últimos 15 años:

DESARROLLO DE UNA APLICACIÓN PARALELA EN EL TIEMPO PARA LA SIMULACIÓN Y CONTROL DE ECUACIONES DIFERENCIALES PARCIALES PARABÓLICAS

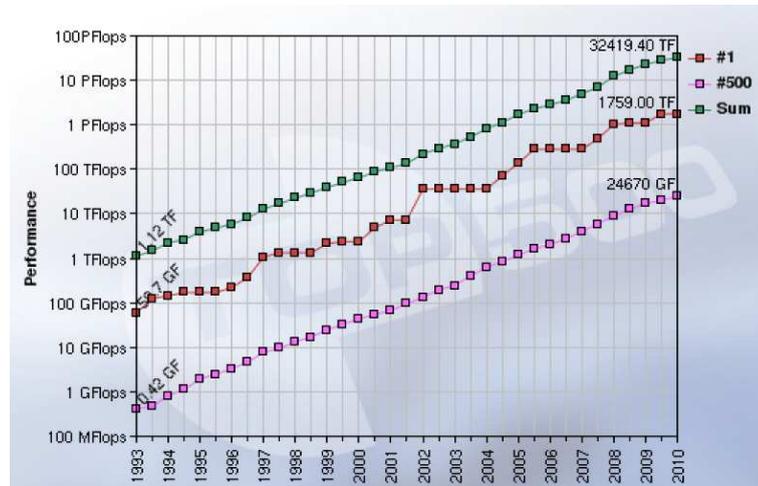


Figura 1.1: Rendimiento de los Super-computadores. Fuente:[11]

Las series referenciadas en el gráfico como #1, #500 y #Sum corresponden al tasa máxima de operaciones (FLOPS) del super-computador con mayor capacidad de procesamiento del mundo, del super-computador ubicado en la posición 500, y la suma de las tasas de operaciones de los 500 computadores respectivamente.

Éste incremento de la capacidad de procesamiento es debido al desarrollo de arquitecturas de sistemas distribuidos. La arquitectura de sistemas distribuidos consiste en utilizar un conjunto de recursos separados físicamente y conectados entre sí por una red de comunicaciones. Ésto permite obtener mayores recursos que un solo computador y consecuentemente mayor capacidad de procesamiento.

Además en la siguiente figura se ilustra las previsiones en cuanto al aumento de la demanda de capacidad de procesamiento en los siguientes diez años [11]:

DESARROLLO DE UNA APLICACIÓN PARALELA EN EL TIEMPO PARA LA SIMULACIÓN Y CONTROL DE ECUACIONES DIFERENCIALES PARCIALES PARABÓLICAS

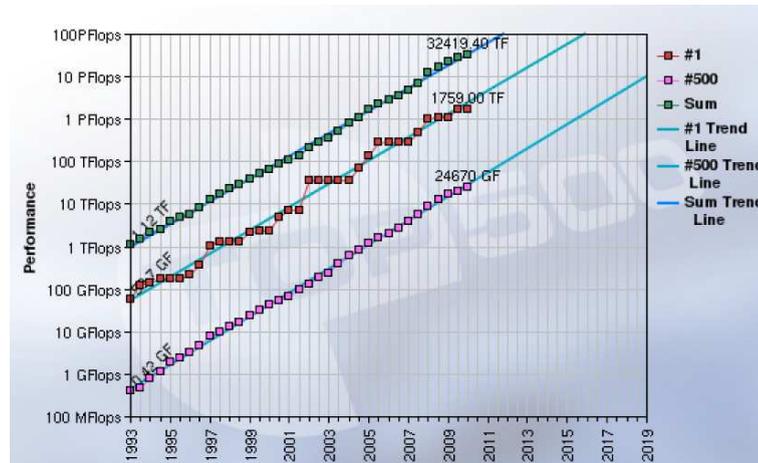


Figura 1.2: Proyecciones de Rendimiento de los Super-computadores.

Fuente:[11]

El aumento en la demanda de la capacidad de procesamiento se debe al desarrollo de aplicaciones en plataforma distribuida, que utilizan al máximo éstos recursos y cuyos requerimientos se incrementan con el transcurso del tiempo.

1.2. Motivación del Paralelismo

La física, la ciencia y la ingeniería, estudian las leyes de la naturaleza que son expresadas matemáticamente mediante ecuaciones diferenciales parciales (EDP). Las soluciones analíticas de éstas no siempre pueden ser obtenidas o son muy difíciles de ser encontradas, por lo cual, generalmente, se recurre a aproximaciones numéricas de la solución. Las aproximaciones son obtenidas a partir de algoritmos numéricos que sustituyen el problema original continuo por uno discreto y establecen el orden de ejecución de las operaciones aritméticas

DESARROLLO DE UNA APLICACIÓN PARALELA EN EL TIEMPO PARA LA SIMULACIÓN Y CONTROL DE ECUACIONES DIFERENCIALES PARCIALES PARABÓLICAS

para obtener un resultado aproximado. Para tener una buena precisión, éstos métodos necesitan un elevado costo computacional en términos de tiempo de cálculo y memoria [12], [13].

Un ejemplo de ello es el control óptimo de la temperatura de una placa de circuitos electrónicos abordado en [20], el cual involucra la resolución de un sistema de ecuaciones lineales de gran porte $3 \times 2600 \times 400 = 3120000$ incógnitas, donde 2600 es la cantidad de puntos considerados sobre la placa y 400 es la cantidad de intervalos tiempo considerados. La resolución de éste sistema demanda gran cantidad de memoria y capacidad de cálculo. Para mejorar la precisión de los resultados numéricos se debe aumentar la cantidad de puntos considerados, lo cual se traduce en un incremento del tamaño del problema, consecuentemente la demanda de recursos.

La demanda de recursos, se refiere a capacidad de memoria y velocidad de cálculo. Al aumentar el tamaño del problema se puede sobrepasar la capacidad de memoria que puede ser proveída por un solo computador y/o los tiempos de cálculo pueden llegar a ser inviables. Entonces la resolución de problemas de gran porte, mediante aplicaciones secuenciales se encuentra limitada.

La velocidad de cálculo de un computador esta limitada por la velocidad de transmisión de datos (velocidad de transmisión de cable de cobre 9 cm/nanosegundos). En consecuencia el aumento de las velocidades requiere mayor proximidad de los elementos de procesamiento. Además la miniaturización está lle-

gando a niveles moleculares, lo cual nos indica que los límites físicos en cuanto a escala de integración están siendo alcanzados [4].

Los algoritmos paralelos son aquellos que pueden ser ejecutados simultáneamente en varias unidades de procesamiento, donde cada unidad resuelve una parte del problema, para finalmente unir todas las partes y ofrecer un resultado. Su implementación se hace posible mediante la utilización de plataformas computacionales distribuidas. Éstas constituyen una colección de computadores separados físicamente y conectados entre sí mediante una red de comunicaciones. Como un determinado problema es dividido en partes menores y resuelto en forma simultánea, el tiempo de procesamiento es menor. Esto constituye una ventaja competitiva frente a los algoritmos secuenciales.

Entonces podemos decir que los motivos para desarrollar una aplicación paralela son la necesidad de resolver problemas de gran envergadura en menor tiempo y con mayores precisiones.

1.3. Tendencias de la Ciencia e Ingeniería

La ciencia y la ingeniería han pasado por un cambio a nivel de investigación, así como también a nivel de desarrollo y tecnología. El viejo paradigma de prueba y error, ha sido reemplazado por el de simular y analizar [14].

Un ejemplo de ello es la nanotecnología, en la cual se están realizando simulaciones numéricas que involucran cálculos computacionales de gran porte,

en lugar de construir prototipos de prueba [14]. Así como también aplicaciones aeroespaciales, síntesis de nuevos materiales, preparación de nuevos medicamentos, dinámica de fluidos, microfabricación de chips, entre otras.

Para realizar las simulaciones, primeramente se obtiene el modelo físico. En muchas ocasiones se da el caso que modelos físicos distintos tienen la misma formulación matemática, consecuentemente al desarrollar algoritmos que permitan encontrar soluciones aproximadas para un tipo de EDP estamos resolviendo problemas de diferentes disciplinas.

Las soluciones aproximadas se obtienen discretizando las EDP, esta discretización consta de dos partes. La primera es la discretización espacial, en la que se convierten las EDP en Ecuaciones Diferenciales Ordinarias (EDO). La segunda es la discretización temporal, que consiste en convertir las EDO en Ecuaciones Algebraicas (EA).

Finalmente para obtener los resultados numéricos, las ecuaciones algebraicas son resueltas mediante uno o varios computadores. Para resolver en varios computadores es necesario dividir el problema en partes, que pueden no ser totalmente independientes. En este caso es necesario una comunicación (paso de mensajes) entre los procesos.

En la siguiente figura se ilustra el esquema a seguir en el proceso de realizar una simulación.

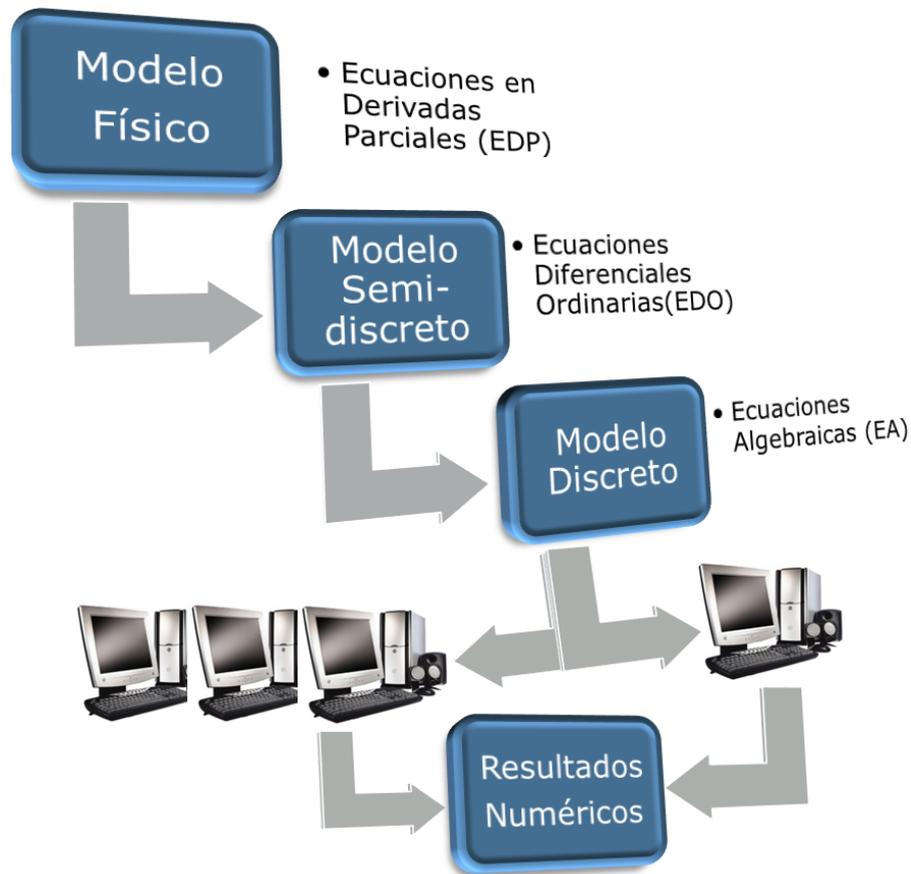


Figura 1.3: Etapas de la simulación. Fuente: Elaboración propia.

1.4. Arquitecturas de Cómputo

La Computación distribuida es un modelo para resolver problemas utilizando computadoras interconectadas mediante una infraestructura de red.

Existen tres tipos de arquitecturas de memoria para implementar cálculos paralelos: Memoria Compartida (Shared), Memoria Distribuida (Distributed) y Híbrida (Shared/Distributed). Para utilizar estas arquitecturas existen varios modelos de programación paralela tales como: Shared Memory, Data Parallel, Message

Passing, Hybrid. Cada modelo es una abstracción de la arquitectura de hardware y memoria lo que implica que estos modelos no necesariamente se corresponden con el tipo de máquina o arquitecturas de memoria [4].

1.4.1. Paso de Mensajes

Los Procesos intercambian datos a través de mensajes de comunicación explícitos, cada proceso tiene memoria, datos locales y globales. Message Passing Interface (MPI) es el estándar que define la sintaxis y la semántica de las funciones contenidas en una biblioteca de paso de mensajes diseñada para ser usada en programas que exploten la existencia de múltiples procesadores.

El MPI provee un conjunto de librerías de modo que el programador diseñe su aplicación paralela, sin la necesidad de conocer el hardware sobre el cual se ejecutará la misma. En la siguiente figura se muestra un esquema del funcionamiento del MPI [14].

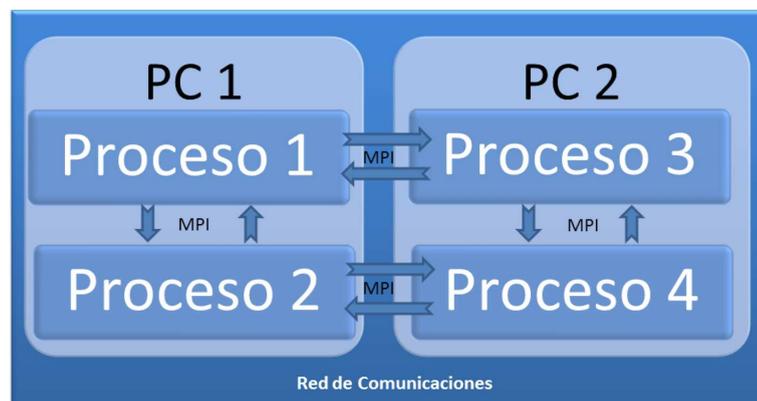


Figura 1.4: Funcionamiento del MPI. Fuente: Elaboración propia.

1.5. Antecedentes

La mayoría de los métodos desarrollados para resolver las ecuaciones diferenciales que envuelven la variable temporal son puramente secuenciales. Esto es debido a la naturaleza secuencial intrínseca del tiempo. En términos prácticos significa que se necesita calcular la solución de un determinado paso de tiempo para obtener la solución al paso de tiempo siguiente. Por esta razón, no se ha prestado mucha atención a los algoritmos paralelos en el tiempo [2].

El punto de partida para este trabajo es el algoritmo Parareal, primeramente presentado por Lions, Maday y Turinici en 2001 [1]. Éste propone usar una descomposición del dominio temporal, utilizando una malla fina y otra gruesa en el tiempo. La malla gruesa divide al dominio en varios subdominios y la malla fina es la discretización de éstos subdominios. Éstas dos mallas son combinadas con un esquema predictor-corrector. La malla gruesa y el predictor corrector son actualizados de manera estrictamente secuencial. Cada uno de éstos subdominios son resueltos simultáneamente usando una plataforma computacional paralela. Se realizan iteraciones del esquema predictor-corrector hasta la convergencia.

Otro antecedente es la publicación de Mathew, Sarkis y Schaerer en 2007 [15]. En el que se describen algoritmos iterativos para la solución de problemas de control óptimo lineales-cuadráticos de gran escala que involucren ecuaciones diferenciales parciales parabólicas.

Además del trabajo final de grado del Ing. Carlos Sauer presentado en 2009

[20]. Donde se modela el control óptimo de temperatura de un circuito electrónico sujeto a fuentes internas de calor. Las ecuaciones constitutivas se modelan como una ecuación diferencial parcial parabólica, donde el mecanismo de enfriamiento es modelado utilizando condiciones de frontera de Neumann (funciones de control). El problema de control óptimo es obtenido asociando un funcional de costo lineal cuadrático a la temperatura y a las funciones de control.

1.6. Definición del Problema

Este trabajo implementa un algoritmo iterativo para resolver un problema de control óptimo lineal cuadrático con tiempo de establecimiento finito, cuyo modelo físico es descrito mediante una ecuación en derivadas parciales parabólicas. El dominio temporal es dividido en varios sub-dominios, para ser resuelto en forma paralela mediante el método del Gradiente Conjugado. Para acelerar la convergencia, se utiliza como preconditionador el método Parareal propuesto en [1].

Finalmente se aplica el algoritmo al problema planteado en el TFG de Carlos Sauer [20] y se muestran los resultados numéricos obtenidos.

1.7. Objetivos

1.7.1. Objetivos Generales

- Desarrollar una aplicación, paralela en el tiempo que permita la simulación y control de ecuaciones diferenciales parciales parabólicas.

1.7.2. Objetivos Específicos

- Formular el problema de control óptimo para la ecuación de conducción de calor.
- Discretizar el dominio espacial mediante el método de Galerkin.
- Discretizar el dominio temporal utilizando la formulación variacional propuesta por J. L. Lions.
- Desarrollar un algoritmo que permita resolver el sistema de ecuaciones resultante en forma paralela.
- Estudiar los sistemas de computación distribuida.
- Implementar el algoritmo.

CAPÍTULO 2

PROBLEMA DE CONTROL

2.1. Formulación matemática

Los procesos físicos, mecánicos, químicos, y de ingeniería, son modelados utilizando ecuaciones diferenciales. Al sistema físico lo denominamos planta y a las ecuaciones que lo modelan, ecuaciones de estado. Éstas últimas establecen las dependencias entre las variables de estado $\underline{z} \in \mathbb{R}^{\hat{q}}$ (variables que deben conocerse para determinar completamente el estado y el funcionamiento de la planta) y las variables de control $\underline{v} \in \mathbb{R}^{\hat{p}}$. Estudiamos el comportamiento de la planta en un intervalo de tiempo $[t_o, t_f]$, asumiendo que de las condiciones iniciales $z_o^h \in \mathbb{R}^{\hat{q}}$ (estado inicial de funcionamiento de la planta) son conocidas.

En este trabajo las ecuaciones de estado son representadas por:

$$\begin{aligned}\dot{\underline{z}} &= f(\underline{z}, \underline{v}, t), \quad \text{para } t_o \leq t \leq t_f \\ \underline{z}(t_o) &= z_o^h,\end{aligned}\tag{2.1}$$

donde, $\underline{v} \in \mathbb{R}^{\hat{p}} \times [t_o, t_f]$, $\underline{z} \in \mathbb{R}^{\hat{q}} \times [t_o, t_f]$ y $f : \mathbb{R}^{\hat{q} \times \hat{p}} \times [t_o, t_f] \rightarrow \mathbb{R}^{\hat{q}} \times [t_o, t_f]$.

Las ecuaciones de estado pueden ser ecuaciones diferenciales lineales o no lineales dependiendo de las propiedades del operador f . En particular en (2.1), las funciones de control \underline{v} son variables libres denominadas variables de control,

ya que al modificarlas se modifica la solución \underline{z} de (2.1).

El problema de control consiste en buscar funciones de control \underline{v} que lleven a la planta de un estado inicial \underline{z}_o^h a un estado deseado $\tilde{\underline{y}}(t_f) \in \mathbb{R}^{\hat{q}}$ en un tiempo fijado ($T := t_f - t_o$), y al mismo tiempo, éstas verifiquen las ecuaciones de la planta (2.1). Ésto establece el siguiente esquema de control:

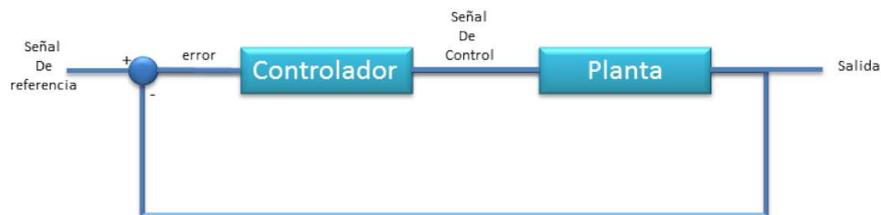


Figura 2.1: Problema de Control. Fuente: Elaboración propia.

2.2. Problema de control óptimo

Si el sistema es controlable, existen infinitas funciones de control \underline{v} que llevan la planta al estado deseado. De forma a encontrar la variable de control óptima en determinado sentido (el superíndice * se utiliza para indicar la solución óptima), en el contexto de este trabajo, planteamos el problema de encontrar \underline{v}^* que lleve la planta del estado $\underline{z}(t_o)$ al estado deseado $\tilde{\underline{y}}$, asociando a la planta un funcional de costo, que debe ser minimizado, satisfaciendo las restricciones de que la solución y la variable de control verifiquen la planta. Es decir planteamos un problema de control óptimo.

Sean \underline{z}^* y \underline{v}^* , las variables de estado y funciones de control óptimas res-

pectivamente, definimos las variaciones de éstos vectores como $\underline{v} = \underline{v}^* + \delta \underline{v}$ y $\underline{z} = \underline{z}^* + \delta \underline{z}$. En la literatura (por ejemplo [22], [23]), es común encontrar cuatro formas de plantear el problema de control, dependiendo de las consideraciones que se hacen con respecto a las variaciones en el tiempo final t_f y el estado final $\underline{z}(t_f)$ (ver siguiente figura).

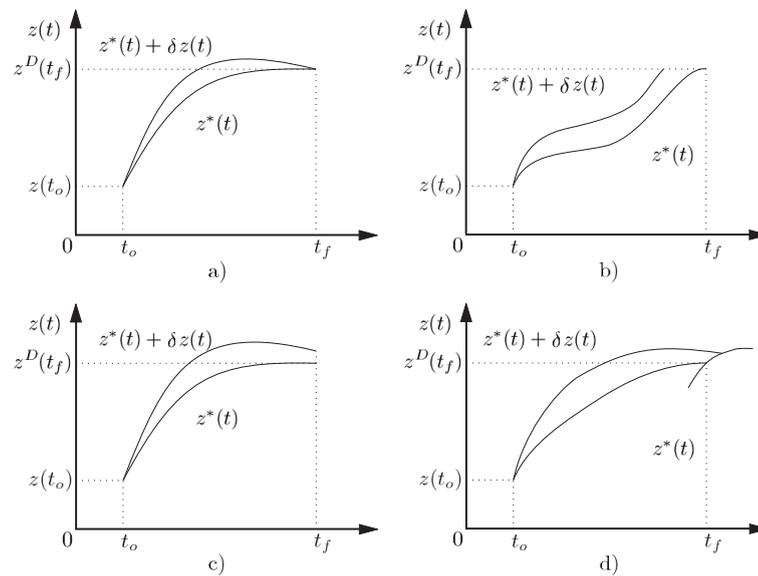


Figura 2.2: Diferentes tipos de planteamiento del problema de control óptimo.

Fuente: [22].

- (a) Sistema con tiempo final fijo y estado final fijo: En este tipo de planteamiento el t_f y $\underline{z}(t_f)$ son fijos, en consecuencia sus variaciones son cero ($\delta \underline{z}(t_f) = \delta t_f = 0$).
- (b) Sistema con tiempo final libre y estado final fijo: En este tipo de planteamiento el t_f es arbitrario o no especificado y $\underline{z}(t_f)$ es fijo, en consecuencia la

variación del estado final es cero ($\delta \underline{z}(t_f) = 0$).

- (c) Sistema con tiempo final fijo y estado final libre: En este tipo de planteamiento el t_f es fijo y $\underline{z}(t_f)$ es arbitrario, en consecuencia la variación del tiempo final es cero ($\delta t_f = 0$).
- (d) Sistema con tiempo final y estado final libres: En este tipo de planteamiento el t_f y $\underline{z}(t_f)$ son arbitrarios.

En particular, en este trabajo utilizamos el planteamiento Tipo (c), para que en un tiempo fijo t_f el estado final del sistema se aproxime a un estado deseado $\tilde{\underline{y}}$.

El funcional de costo $\mathcal{J}_h(\underline{z}, \underline{v}) : \mathbb{R}^{\hat{q} \times \hat{p}} \times [t_o, t_f] \rightarrow \mathbb{R}$, establece un criterio para medir la calidad de las funciones de control \underline{v} y solución $\underline{z}(t)$.

Los siguientes funcionales de costo típicamente encontrados en la literatura [22, 23] son:

- (a) El error cuadrático medio respecto a una solución deseada $\tilde{\underline{y}} \in \mathbb{R}^{\hat{q}} \times [t_o, t_f]$,

$\forall \underline{z} \in \mathbb{R}^{\hat{q}} \times [t_o, t_f]$:

$$\frac{1}{2} \int_{t_o}^{t_f} \|\underline{z} - \tilde{\underline{y}}(t)\|_{M_h}^2 dt = \frac{1}{2} \int_{t_o}^{t_f} (\underline{z} - \tilde{\underline{y}}(t))^T M_h (\underline{z} - \tilde{\underline{y}}(t)) dt, \quad (2.2)$$

donde $M_h \in \mathbb{R}^{\hat{q} \times \hat{q}}$; $M_h^T = M_h$; $M_h > 0$, esto es simétrico y definido positivo.

- (b) El error cuadrático medio en el tiempo final:

$$\frac{1}{2} \|\underline{z}(t_f) - \tilde{\underline{y}}(t_f)\|_{M_h}^2 = \frac{1}{2} (\underline{z}(t_f) - \tilde{\underline{y}}(t_f))^T M_h (\underline{z}(t_f) - \tilde{\underline{y}}(t_f)), \quad (2.3)$$

2.3. Condiciones de optimalidad

Para transformar el problema de minimización con restricciones (2.6), en un problema de minimización irrestricta, definimos el funcional Lagrangiano (\mathcal{L}):

$$\min \mathcal{L}(\underline{z}, \underline{v}, \underline{p}, \underline{\eta}_o) = \mathcal{J}_h(\underline{z}, \underline{v}) + \int_{t_o}^{t_f} \underline{p}^T (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) dt + \underline{\eta}_o^T (\underline{z}(t_o) - \underline{z}_o^h), \quad (2.7)$$

donde $\underline{p} \in \mathbb{R}^{\hat{q}} \times [t_o, t_f]$ y $\underline{\eta}_o \in \mathbb{R}^{\hat{q}}$ conocidos como multiplicadores de Lagrange.

Consecuentemente, minimizar la expresión (2.7), es equivalente a minimizar (2.6). A la solución minimizante de (2.7) la denominamos $\underline{z}^*(t)$ y $\underline{v}^*(t)$.

La variación total de primer orden $\delta\mathcal{L}$ tiene la forma:

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\underline{v}} \delta\underline{v} + \frac{\partial\mathcal{L}}{\partial\underline{p}} \delta\underline{p} + \frac{\partial\mathcal{L}}{\partial\underline{z}} \delta\underline{z} + \frac{\partial\mathcal{L}}{\partial\underline{\eta}_o} \delta\underline{\eta}_o. \quad (2.8)$$

Las condiciones de optimalidad de Karush-Kuhn-Tucker (KKT) son obtenidas anulando la variación total de primer orden $\delta\mathcal{L}$ al ser evaluada en los valores $\underline{z}^*(t)$ y $\underline{v}^*(t)$, véase [22, 23], o sea:

$$\delta\mathcal{L} = \left. \frac{\partial\mathcal{L}}{\partial\underline{v}} \right|_{\underline{z}^*, \underline{v}^*} \delta\underline{v} + \left. \frac{\partial\mathcal{L}}{\partial\underline{p}} \right|_{\underline{z}^*, \underline{v}^*} \delta\underline{p} + \left. \frac{\partial\mathcal{L}}{\partial\underline{z}} \right|_{\underline{z}^*, \underline{v}^*} \delta\underline{z} + \left. \frac{\partial\mathcal{L}}{\partial\underline{\eta}_o} \right|_{\underline{z}^*, \underline{v}^*} \delta\underline{\eta}_o = 0, \quad (2.9)$$

donde las derivadas del funcional \mathcal{L} respecto a $\underline{z}, \underline{v}, \underline{p}$, son dadas usando la derivada en el sentido de Gâteaux, que se define como:

Definición 2.1 (Derivada de Gâteaux). *Sea un funcional, la derivada del funcional con respecto a la variable $y_i \in \mathbb{R}^{\hat{q}} \times [t_o, t_f]$, en el punto $y_i^* \in \mathbb{R}^{\hat{q}} \times [t_o, t_f]$ a lo largo de la dirección, $\phi \in \mathbb{R}^{\hat{q}} \times [t_o, t_f]$ es dada por:*

$$\frac{D\mathcal{J}_h}{Dy_i} = \left(\frac{\partial\mathcal{J}_h}{\partial y_i}, \phi \right) = \lim_{\epsilon \rightarrow 0} \left(\frac{\mathcal{J}_h(y_i^* + \epsilon\phi) - \mathcal{J}_h(y_i^*)}{\epsilon} \right) = \left. \frac{d\mathcal{J}_h(y_i^* + \epsilon\phi)}{d\epsilon} \right|_{\epsilon=0},$$

donde (\cdot, \cdot) es producto interno en el dominio del funcional.

Utilizando la Definición 2.1 obtenemos las condiciones de optimalidad (2.8) para el Lagrangiano (2.7):

Ecuaciones de estado

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \underline{p}} \Big|_{\underline{z}^*, \underline{v}^*} &= 0 \Rightarrow (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) \Big|_{(\underline{z}^*, \underline{v}^*)} = 0, \quad \text{para } t_o \leq t \leq t_f, \\ \frac{\partial \mathcal{L}}{\partial \underline{\eta}_o} \Big|_{\underline{z}^*, \underline{v}^*} &= 0 \Rightarrow \underline{z}(t_o) = \underline{z}_o^h. \end{aligned} \quad (2.10)$$

Problema Adjunto

$$\frac{\partial \mathcal{L}}{\partial \underline{z}} \Big|_{\underline{z}^*, \underline{v}^*} = 0 \Rightarrow \left(\frac{\partial \mathcal{J}_h}{\partial \underline{z}} \Big|_{(\underline{z}^*, \underline{v}^*)} \right) + \left(\frac{\partial \left(\int_{t_o}^{t_f} \underline{p}^T (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) dt \right)}{\partial \underline{z}} \Big|_{(\underline{z}^*, \underline{v}^*)} \right) = 0. \quad (2.11)$$

Condición de Euler

$$\frac{\partial \mathcal{L}}{\partial \underline{v}} \Big|_{\underline{z}^*, \underline{v}^*} = 0 \Rightarrow \left(\frac{\partial \mathcal{J}_h}{\partial \underline{v}} \Big|_{(\underline{z}^*, \underline{v}^*)} \right) + \left(\frac{\partial \left(\int_{t_o}^{t_f} \underline{p}^T (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) dt \right)}{\partial \underline{v}} \Big|_{(\underline{z}^*, \underline{v}^*)} \right) = 0. \quad (2.12)$$

La deducción de las ecuaciones (2.10), (2.11) y (2.12) se encuentran en el apéndice (A.1).

2.3.1. Ecuaciones de estado lineales

En el contexto de este trabajo el operador f de las ecuaciones de estado (2.1) es considerado lineal, consecuentemente las condiciones de optimalidad pueden re-escribirse como sigue:

Ecuaciones de estado

$$\begin{cases} M_h \dot{\underline{z}} = A_h \underline{z} + B_h \underline{v} + c_h, & \text{para } t_o \leq t \leq t_f, \\ \underline{z}(t_o) = z_o^h, \end{cases} \quad (2.13)$$

donde, $M_h, A_h \in \mathbb{R}^{\hat{q} \times \hat{q}}$; $B \in \mathbb{R}^{\hat{q} \times \hat{p}}$; $c_h \in \mathbb{R}^{\hat{q}}$. Asumimos que las matrices M_h y A_h , son simétricas y definidas positivas.

Luego el problema adjunto y la condición de Euler, definidos en (2.11) y (2.12) respectivamente, pueden ser re-escritos utilizando las ecuaciones de estado lineales (para mas detalles véase (A.1.1)).

Problema adjunto

$$\begin{cases} M_h \dot{\underline{p}} = -A_h^T \underline{p} + q M_h (\underline{z} - \tilde{\underline{y}})(t), & \text{para } t_o \leq t \leq t_f, \\ \underline{p}(t_f) = -s (\underline{z}(t_f) - \tilde{\underline{y}}(t_f)), \end{cases} \quad (2.14)$$

Condición de Euler

$$R_h \underline{v} = B_h^T \underline{p}, \quad \text{para } t_o \leq t \leq t_f. \quad (2.15)$$

Entonces, el Problema de Control Óptimo P1 es equivalente a resolver el sistema de ecuaciones diferenciales ordinarias EDOs determinado por las ecuaciones (2.13), (2.14) y (2.15).

Éste sistema es un problema de valores de frontera en dos puntos (TPBVP), y puede ser transformado a sistema de ecuaciones algebraicas mediante una discretización temporal. Éste puede ser resuelto mediante el método "todo de una vez"(all at once) como en [24] y [20]. Sin embargo, en la medida que se refina la malla el costo computacional de éste método se vuelve prácticamente

inviabile. Sobretudo teniendo en cuenta la limitación de memoria y la capacidad de procesamiento.

2.4. Método de solución del problema de control óptimo

En las secciones anteriores formulamos el problema de control óptimo como un problema de minimización con restricciones. Luego para transformar este problema en un problema de minimización irrestricta introdujimos los multiplicadores de Lagrange (2.7).

En ésta sección presentamos el algoritmo de optimización iterativo denominado Método del Gradiente Conjugado (CG). Éste algoritmo es aplicable a problemas donde se quiere minimizar funcionales del tipo cuadráticos particularmente simétricos y definidos positivos [25]. El funcional de costo \mathcal{J} definido en (2.6) depende del cuadrado de la función de control y del cuadrado del error con respecto a una señal de referencia, en consecuencia es un funcional cuadrático y podemos utilizar el CG para minimizarlo.

Antes de presentar el CG, introducimos los siguientes conceptos fundamentales:

Definición 2.2 (Producto interno). *El producto interno de dos vectores en un espacio vectorial V es una función bilineal, hermítica y definida positiva: $(\cdot, \cdot) : V \times V \rightarrow \mathbb{R}$.*

En el espacio vectorial \mathbb{R}^n el producto interno puede ser expresado mediante

la siguiente expresión $(\mathbf{w}, \mathbf{u}) = \mathbf{w}^T \mathbf{u}$, donde $\mathbf{w}, \mathbf{u} \in \mathbb{R}^n$

Definición 2.3 (Vectores ortogonales). *Dos vectores son ortogonales cuando su producto interno es igual a cero.*

Definición 2.4 (Vectores \mathbf{H} -ortogonales). *Dos vectores son \mathbf{H} -ortogonales cuando su producto interno con respecto a una matriz \mathbf{H} (simétrica y definida positiva) es igual a cero. $(\mathbf{w}, \mathbf{H}\mathbf{u}) = (\mathbf{w}, \mathbf{u})_{\mathbf{H}} = \mathbf{w}^T \mathbf{H}\mathbf{u} = 0$.*

Definición 2.5 (Norma- \mathbf{H}). *La norma de un vector \mathbf{w} respecto de una matriz \mathbf{H} se define como sigue $\|\mathbf{w}\|_{\mathbf{H}} = \sqrt{(\mathbf{w}, \mathbf{H}\mathbf{w})}$, donde \mathbf{H} es simétrica y definida positiva.*

Definición 2.6 (Vectores conjugados respecto a la matriz \mathbf{H}). *Un vector \mathbf{w} es el conjugado de otro \mathbf{u} cuando son \mathbf{H} -ortogonales.*

Definición 2.7 (Gradiente del funcional). *Dado un funcional $f(\underline{u}) : \mathbb{R}^n \rightarrow \mathbb{R}$, se define el gradiente del funcional como $\nabla f(\underline{u}) := \frac{\partial f(\underline{u})}{\partial \underline{u}} = \left[\frac{\partial f(\underline{u})}{\partial u_1} \quad \frac{\partial f(\underline{u})}{\partial u_2} \quad \dots \quad \frac{\partial f(\underline{u})}{\partial u_n} \right]^T$, y puede obtenerse como sigue:*

$$(\nabla f(\underline{u}), \phi) = \frac{Df(\underline{u})}{D\underline{u}} = \lim_{\epsilon \rightarrow 0} \left(\frac{f(\underline{u} + \epsilon \phi) - f(\underline{u})}{\epsilon} \right). \quad (2.16)$$

Definición 2.8 (Matriz Hessiana). *Dado un funcional $f(\underline{u}) : \mathbb{R}^n \rightarrow \mathbb{R}$, se define la matriz Hessiana como sigue:*

$$\mathbf{H} := \nabla^2 f(\underline{u}) = \begin{bmatrix} \frac{\partial^2 f(\underline{u})}{\partial^2 u_1} & \frac{\partial^2 f(\underline{u})}{\partial u_1 \partial u_2} & \dots & \frac{\partial^2 f(\underline{u})}{\partial u_1 \partial u_n} \\ \frac{\partial^2 f(\underline{u})}{\partial u_2 \partial u_1} & \frac{\partial^2 f(\underline{u})}{\partial^2 u_2} & \dots & \frac{\partial^2 f(\underline{u})}{\partial u_2 \partial u_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 f(\underline{u})}{\partial u_n \partial u_1} & \frac{\partial^2 f(\underline{u})}{\partial u_n \partial u_2} & \dots & \frac{\partial^2 f(\underline{u})}{\partial^2 u_n} \end{bmatrix},$$

El Gradiente Conjugado corrige en forma iterativa una función de control arbitraria \underline{v}^i (el superíndice i indica el número de iteraciones que se han realizado). Con ésta obtenemos el estado \underline{z}^i y el estado adjunto \underline{p}^i a partir de las ecuaciones (2.13) y (2.14) respectivamente. Consecuentemente la variación de primer orden del Lagrangiano (2.8) es igual a la variación total del funcional (2.5) (véase apéndice(A.2.4)),

$$\delta\mathcal{L}(\underline{z}^i, \underline{v}^i, \underline{p}^i) = \frac{\partial\mathcal{L}}{\partial\underline{v}} \delta\underline{v} = \nabla J_h \delta\underline{v}, \quad (2.17)$$

note que al proceder de ésta manera, el minimizar \mathcal{L} es equivalente a minimizar J_h , por lo cual de aquí en adelante toda la deducción del algoritmo se basará en el funcional J_h de (2.6).

Mientras la función de control no sea la óptima, la variación (2.17) no será nula, si éste es el caso, entonces existe una función de control \underline{v}^{i+1} tal que:

$$\mathcal{J}_h(\underline{z}^{i+1}(\underline{v}^{i+1}), \underline{v}^{i+1}) < \mathcal{J}_h(\underline{z}^i(\underline{v}^i), \underline{v}^i), \quad (2.18)$$

Para encontrar la función de control \underline{v}^{i+1} con la propiedad (2.18) utilizamos el siguiente esquema de actualización de \underline{v}^i :

$$\underline{v}^{i+1} = \underline{v}^i + \alpha^i \underline{d}^i, \quad (2.19)$$

donde α^i es un escalar y \underline{d}^i es un vector que se denomina dirección de búsqueda.

El valor de α^i en cada iteración, se obtiene de tal manera que el funcional J_h sea mínimo en la dirección \underline{d}^i es decir:

$$\frac{d\left(\mathcal{J}_h|_{(\underline{z}^{i+1}, \underline{v}^{i+1})}\right)}{d\alpha^i} = 0, \quad (2.20)$$

ésta condición se denomina de máximo descenso. El valor del incremento se puede obtener expandiendo el funcional \mathcal{J}_h entorno a \underline{v}^i utilizando la fórmula de Taylor (o con la deducción alternativa (A.2.3)), como sigue:

$$\mathcal{J}_h|_{(\underline{z}^{i+1}, \underline{v}^{i+1})} = \mathcal{J}_h|_{(\underline{z}^i, \underline{v}^i)} + \alpha^i \left(\nabla \mathcal{J}_h|_{(\underline{z}^i, \underline{v}^i)}, \underline{d}^i \right) + \frac{1}{2} (\alpha^i)^2 \left(\underline{d}^i, \nabla^2 \mathcal{J}_h|_{(\underline{z}^i, \underline{v}^i)} \underline{d}^i \right), \quad (2.21)$$

y aplicamos la condición de máximo descenso (2.20), luego tenemos:

$$0 = \left(\nabla \mathcal{J}_h|_{(\underline{z}^i, \underline{v}^i)}, \underline{d}^i \right) + (\alpha^i) \left(\underline{d}^i, \nabla^2 \mathcal{J}_h|_{(\underline{z}^i, \underline{v}^i)} \underline{d}^i \right), \quad (2.22)$$

por tanto:

$$\alpha^i = - \frac{\left(\nabla \mathcal{J}_h^T|_{(\underline{z}^i, \underline{v}^i)}, \underline{d}^i \right)}{\left(\underline{d}^i, H \underline{d}^i \right)}, \quad (2.23)$$

donde $H := \nabla^2 \mathcal{J}_h|_{(\underline{z}^i, \underline{v}^i)}$ es la matriz hessiana del funcional.

Note que al aplicar la condición de máximo descenso (2.20) estamos haciendo que el gradiente del funcional (2.5) sea ortogonal a las direcciones de búsqueda anteriores (véase apéndice (A.2.1)), es decir:

$$\nabla \mathcal{J}_h^T|_{(\underline{z}^{i+1}, \underline{v}^{i+1})} \underline{d}^j = 0 \quad \forall j = 0, 1, \dots, i. \quad (2.24)$$

Para simplificar la nomenclatura al gradiente del funcional (2.5) en la iteración i lo denotaremos como sigue:

$$\underline{g}^i := \nabla \mathcal{J}_h|_{(\underline{z}^i, \underline{v}^i)}, \quad (2.25)$$

El Método Gradiente Conjugado utiliza un conjunto de direcciones de búsqueda ortogonales respecto a la matriz hessiana del funcional [26] (véase apéndice

(A.2.1)), es decir:

$$(\underline{d}^i, H\underline{d}^j) = 0, \quad j = 0, 1, \dots, n, \quad \forall i \neq j, \quad (2.26)$$

A continuación construimos las direcciones de búsqueda utilizando el procedimiento de Ortogonalización de Gram-Schmidt (véase apéndice (A.2.2)), usando la definición 2.2 de producto interno (\cdot, \cdot) .

La primera dirección tomamos como la dirección contraria al gradiente del funcional (2.5), es decir:

$$\underline{d}^0 = -\underline{g}^0, \quad (2.27)$$

y las siguientes direcciones como sigue:

$$\underline{d}^{i+1} = -\underline{g}^{i+1} + \beta^i \underline{d}^i, \quad (2.28)$$

donde β^i es:

$$\beta^i = \frac{(\underline{g}^{i+1}, H\underline{d}^i)}{(\underline{d}^i, H\underline{d}^i)}, \quad (2.29)$$

o equivalentemente mediante la fórmula de Fletcher-Reeves [26]:

$$\beta^i = \frac{(\underline{g}^{i+1}, \underline{g}^{i+1})}{(\underline{g}^i, \underline{g}^i)}. \quad (2.30)$$

Por último, la corrección de las funciones de control se realiza (proceso iterativo (2.19)) hasta satisfacer algún criterio de parada. El criterio de parada que consideramos en este trabajo consiste en evaluar la norma del gradiente del

funcional para verificar si se cumplen las condiciones de optimalidad de primer orden, en el contexto de este trabajo el criterio de parada a utilizar es el siguiente:

$$\frac{\|g^i\|}{\|g^0\|} \leq tol, \quad (2.31)$$

donde tol es un real positivo.

Entonces el Método Gradiente Conjugado (GC) puede resumirse en el siguiente algoritmo:

Algoritmo GC (2.4)

- 1 Sea $\underline{v}^{(0)} \in \mathbb{R}^{\hat{p}} \times [t_0, t^f]$ arbitraria.
- 2 $i = 0$, Calcular el gradiente del funcional $g^{(0)}$
- 3 **mientras** $\frac{\|g^{(i)}\|}{\|g^{(0)}\|} \leq tol$
- 4 **if** $i = 0$ **then**,
- 5 $d^{(i)} = -g^{(i)}$
- 6 **else**
- 7 $d^{(i+1)} = -g^{(i)} + \frac{(g^{(i)}, Hd^{(i)})}{(d^{(i)}, Hd^{(i)})} d^{(i)}$
- 8 **end**
- 9 $\underline{v}^{(i+1)} = \underline{v}^{(i)} + \frac{(g^{(i)}, g^{(i)})}{(d^{(i)}, Hd^{(i)})} d^{(i)}$
- 10 $g^{(i+1)} = g^{(i)} + \alpha^{(i)} Hd^{(i)}$
- 11 $i = i + 1$
- 12 **end**

Gradiente del funcional

El gradiente del funcional (2.5), obtenemos utilizando el método del adjunto [24]. Éste método consiste en obtener el gradiente del funcional utilizando el estado adjunto. Entonces a partir de una función de control arbitraria obtenemos el estado z^i resolviendo las ecuaciones de estado (2.13) y luego el estado adjunto p^i resolviendo las ecuaciones del estado adjunto. Luego podemos calcular el gradiente del funcional mediante la siguiente expresión:

$$\underline{g}^i = rR_h\underline{v}^i - B_h\underline{p}^i \quad \text{para } t_o \leq t \leq t_f, \quad (2.32)$$

donde la obtención de ésta expresión puede encontrarse en el apéndice (A.2.4). Finalmente la obtención del gradiente del funcional a partir de una función de control arbitraria puede expresarse como la solución del sistema sistema de EDOs:

$$\left\{ \begin{array}{l} \left\{ \begin{array}{l} M_h \dot{\underline{z}}^i = A_h \underline{z}^i + B_h \underline{v}^i + c_h, \quad \text{para } t_o \leq t \leq t_f, \\ \underline{z}^i(t_o) = z_o^h, \end{array} \right. \\ \left\{ \begin{array}{l} M_h \dot{\underline{p}}^i = -A_h \underline{p}^i + qM_h(\underline{z}^i - \tilde{y}), \quad \text{para } t_o \leq t \leq t_f, \\ \underline{p}^i(t_f) = -s(\underline{z}^i(t_f) - \tilde{y}(t_f)), \end{array} \right. \\ \underline{g}^i = rR_h\underline{v}^i - B_h\underline{p}^i \quad \text{para } t_o \leq t \leq t_f, \end{array} \right. \quad (2.33)$$

Producto de la matriz Hessiana por la dirección

En el paso 10 para obtener α^i es necesario calcular el producto de la matriz hessiana H por la dirección de búsqueda, para ello utilizamos el siguiente

sistema de ecuaciones diferenciales ordinarias (véase apéndice (A.2.5)):

$$\left\{ \begin{array}{l} \left\{ \begin{array}{l} M_h \dot{\underline{\psi}}^i = A_h \underline{\psi}^i + B_h \underline{d}^i, \quad \text{para } t_o \leq t \leq t_f, \\ \underline{\psi}^i(t_o) = 0, \end{array} \right. \\ \left\{ \begin{array}{l} M_h \dot{\underline{\psi}}_p^i = -A_h \underline{\psi}_p^i + q M_h \underline{\psi}^i, \quad \text{para } t_o \leq t \leq t_f, \\ \underline{\psi}_p^i(t_f) = -s \underline{\psi}^i(t_f), \end{array} \right. \\ \underline{H} \underline{v}^i = r R_h \underline{d}^i - B_h \underline{\psi}_p^i \quad \text{para } t_o \leq t \leq t_f, \end{array} \right. \quad (2.34)$$

donde ψ y ψ_p son funciones auxiliares.

2.4.1. Discretización temporal

Para hallar una solución numérica aproximada mediante una plataforma computacional del sistema de ecuaciones diferenciales (2.13), (2.14) y (2.15) por lo general se realiza una discretización del dominio temporal.

La discretización temporal consiste en dividir el intervalo $[t_o, t_f]$ en \hat{l} sub-intervalos de tiempo iguales, con pasos de tiempo $\tau = (t_f - t_o)/(\hat{l})$. Asociados a éstos sub-intervalos de tiempo, tenemos a los nodos que serán denotados como sigue: $t_l = t_o + l\tau$. Asumimos que las funciones de control, las variables de estado, los valores de referencia y las variables adjuntas son constantes en cada sub-intervalo $(t_{l-1}, t_l]$, $1 < l < \hat{l}$.

A continuación realizamos la aproximación de las derivadas de las variables de estado $\dot{\underline{z}}$, en t_l mediante el método de Euler implícito (Backward-Euler) (véase [19]) y derivadas de las variables adjuntas $\dot{\underline{p}}$, en t_l mediante el método de Euler explícito (Forward-Euler) (véase [19]).

Ecuaciones de estado discretas

La aproximación de la derivadas de las variables de estado \underline{z} en el tiempo t_l , a partir del paso de tiempo τ , se realiza mediante la siguiente expresión:

$$\dot{\underline{z}}(t_l) = \frac{\underline{z}(t_l) - \underline{z}(t_{l-1})}{\tau} = \frac{z_l - z_{l-1}}{\tau} \quad (2.35)$$

Reemplazando la expresión (2.35) en la ecuación (2.13):

$$M_h \left(\frac{z_l - z_{l-1}}{\tau} \right) = A_h z_l + B_h v_l + c_h, \quad \forall 1 < l < \hat{l} \quad (2.36)$$

donde $z_0 = z_0^h$.

Entonces si definimos las matrices $F_1 := M_h - \tau A_h$ y $F_0 := M_h$, podemos obtener los estados $z_l \forall 1 < l < \hat{l}$ utilizando la ecuación (2.36), a partir del estado inicial z_0^h , las perturbaciones c_h y los valores de las funciones de control $v_l := \underline{v}(t_l)$ mediante el siguiente esquema iterativo:

Esquema iterativo para obtener el estado

- 1** $z_0 = z_0^h$
- 2** desde $l = 1$ hasta \hat{l}
- 3** $F_1 z_l = F_0 z_{l-1} + \tau B_h v_l + \tau c_h$
- 4** fin-desde

La aproximación de las derivadas de las variables adjuntas \underline{p} en el tiempo t_l se hace mediante la siguiente expresión:

$$\dot{\underline{p}}(t_l) = \frac{\underline{p}(t_{l+1}) - \underline{p}(t_l)}{\tau} = \frac{p_{l+1} - p_l}{\tau} \quad (2.37)$$

Reemplazando la expresión (2.37) en la ecuación (2.14):

$$M_h \left(\frac{p_{l+1} - p_l}{\tau} \right) = -A_h p_l + q M_h(z_l - \tilde{y}_l), \quad \forall 0 < l < \hat{l} - 1, \quad (2.38)$$

donde $t_i = t_f$ y $p_i = -s(z_i - \tilde{y}_i)$.

Entonces podemos obtener el estado adjunto $p_l \forall 0 < l < \hat{l} - 1$ utilizando la ecuación (2.38), a partir del estado adjunto final $p_{\hat{l}}$, los valores deseados $\tilde{y}_{(l)}$, luego de haber obtenido el estado actual mediante (2.4.1) y los valores de las funciones de control v_l :

Esquema Iterativo para obtener el estado adjunto

- 1 $p_i = -s(z_i - \tilde{y}_i)$
- 2 **desde** $l = \hat{l} - 1$ **hasta** 0
- 3 $F_1 p_l = F_0 p_{l+1} - \tau q M_h(z_l - \tilde{y}_l)$
- 4 **fin-desde**

Luego podemos agrupar los valores del estado, del estado adjunto, los valores deseados y las funciones de control definiendo el vector de estado discreto $\mathbf{z} \in \mathbb{R}^{\hat{q} \times \hat{l}}$, el vector de estado adjunto discreto $\mathbf{p} \in \mathbb{R}^{\hat{q} \times \hat{l}}$, el vector de valores deseados $\tilde{\mathbf{y}} \in \mathbb{R}^{\hat{q} \times \hat{l}}$ y el vector de control discreto $\mathbf{v} \in \mathbb{R}^{\hat{p} \times \hat{l}}$

$$\begin{aligned} \mathbf{z} &= [z_1^T \ z_2^T \ \cdots \ z_{\hat{l}}^T]^T, & \mathbf{v} &= [v_1^T \ v_2^T \ \cdots \ v_{\hat{l}}^T]^T, \\ \mathbf{p} &= [p_0^T \ p_1^T \ \cdots \ p_{\hat{l}-1}^T]^T, & \tilde{\mathbf{y}} &= [\tilde{y}_1^T \ \tilde{y}_2^T \ \cdots \ \tilde{y}_{\hat{l}}^T]^T, \end{aligned} \quad (2.39)$$

Gradiente del funcional discreto

Consideramos ahora la discretización temporal de la expresión (2.33), la cual se transforma en el siguiente esquema iterativo:

$$\begin{cases} z_0 = z_0^h, F_1 z_l = F_0 z_{l-1} + \tau B_h v_l + \tau c_h, \forall l = 1, \dots, \hat{l}, \\ p_{\hat{l}} = -s(z_{\hat{l}} - \tilde{y}_{\hat{l}}), F_1 p_l = F_0 p_{l+1} - \tau q M_h(z_l - \tilde{y}_l), \forall l = \hat{l} - 1, \dots, 0, \\ \nabla \mathcal{J}_{h_l}^\tau = \underline{g}^i = r R_h \underline{v}_l^i - B_h \underline{p}_{l-1}^i, \forall l = 1, \dots, \hat{l}. \end{cases} \quad (2.40)$$

donde $\nabla \mathcal{J}_{h_l}^\tau \in \mathbb{R}^{\hat{p}}$. Luego si agrupamos los gradientes de cada tiempo en un vector $\mathbf{g} = \nabla \mathcal{J}_h^\tau \in \mathbb{R}^{\hat{p}\hat{l}}$ como sigue:

$$\mathbf{g} = \nabla \mathcal{J}_h^\tau = \left[(\nabla \mathcal{J}_{h_1}^\tau)^T (\nabla \mathcal{J}_{h_2}^\tau)^T \dots (\nabla \mathcal{J}_{h_{\hat{l}}}^\tau)^T \right]^T, \quad (2.41)$$

y si definimos la matrices $\mathbf{G} \in \mathbb{R}^{\hat{p}\hat{l} \times \hat{p}\hat{l}}$ y $\mathbf{N} \in \mathbb{R}^{\hat{q}\hat{l} \times \hat{p}\hat{l}}$:

$$\mathbf{G} = \begin{bmatrix} r R_h & & & \\ & r R_h & & \\ & & \dots & \\ & & & r R_h \end{bmatrix}, \quad \mathbf{N} = \begin{bmatrix} B_h & & & \\ & B_h & & \\ & & \dots & \\ & & & B_h \end{bmatrix}, \quad (2.42)$$

luego, el esquema (2.40) toma la forma:

$$\begin{cases} z_0 = z_0^h, F_1 z_l = F_0 z_{l-1} + \tau B_h v_l + \tau c_h, \forall l = 1, \dots, \hat{l}, \\ p_{\hat{l}} = -s(z_{\hat{l}} - \tilde{y}_{\hat{l}}), F_1 p_l = F_0 p_{l+1} - \tau q M_h(z_l - \tilde{y}_l), \forall l = \hat{l} - 1, \dots, 0, \\ \mathbf{g}^i = \mathbf{G} \mathbf{v}^i - \mathbf{N}^T \mathbf{p}^i \end{cases} \quad (2.43)$$

Producto de la Matriz Hessiana del funcional discreto y la función de control

Con un procedimiento análogo al utilizado para obtener la expresión (2.43) obtenemos la discretización temporal de la expresión (2.34) toma la forma:

$$\begin{cases} \psi_0 = 0, F_1 z_l = F_0 \psi_{l-1} + \tau B_h d_l, \forall l = 1, \dots, \hat{l}, \\ \psi_{p_i} = -s\psi_{\hat{l}}, F_1 \psi_{p_l} = F_0 \psi_{p_{l+1}} - \tau q M_h \psi_l, \forall l = \hat{l} - 1, \dots, 0, \\ \mathbf{H}d^i = \mathbf{G}d^i - \mathbf{N}^T \psi_p^i \end{cases} \quad (2.44)$$

donde los vectores $\mathbf{d}^i \in \mathbb{R}^{\hat{p}l}$ y $\psi_p^i \in \mathbb{R}^{\hat{q}l}$ corresponden a las discretizaciones temporales de \underline{d}^i y $\underline{\psi}_p^i$ respectivamente. Finalmente el Método Gradiente Conjugado en forma discreta puede expresarse como sigue:

Algoritmo GC discreto (2.4.1)

- 1 Sea $\underline{v}^0 \in \mathbb{R}^{\hat{p}l}$ arbitraria.
- 2 $i = 0$, Calcular el gradiente del funcional $\mathbf{g}^{(0)}$
- 3 mientras $\frac{\|\mathbf{g}^i\|}{\|\mathbf{g}^0\|} \leq tol$
- 4 **if** $i = 0$ **then**,
- 5 $d^i = -\mathbf{g}^i$
- 6 **else**
- 7 $d^{i+1} = -\mathbf{g}^i + \frac{\mathbf{g}^{i,T} H d^i}{d^{i,T} H d^i} d^i$
- 8 **end**
- 9 $\underline{v}^{i+1} = \underline{v}^i + \frac{\mathbf{g}^{i,T} \mathbf{g}^i}{d^{i,T} H d^i} d^i$
- 10 $\mathbf{g}^{i+1} = \mathbf{g}^i + \alpha^i H d^i$
- 11 $i = i + 1$
- 12 **end**

CAPÍTULO 3

ALGORITMO PARAREAL

En el capítulo anterior presentamos el problema de control óptimo (2.6), y resolvimos el mismo mediante el método del Gradiente Conjugado. En éste capítulo se realiza una paralelización temporal del problema (2.6) con el objetivo de resolverlo en plataforma distribuida, reduciendo así los costos computacionales. La resolución se realiza nuevamente con el método del Gradiente Conjugado, y para acelerar la convergencia del mismo utilizamos el algoritmo Parareal propuesto en [1].

3.1. Formulación Matemática

Por comodidad, re-escribimos la ecuación (2.1)

$$\begin{aligned}\dot{\underline{z}} &= f(\underline{z}, \underline{v}, t), \quad \text{para } t_o \leq t \leq t_f \\ \underline{z}(t_o) &= \underline{z}_o^h,\end{aligned}\tag{3.1}$$

donde, $\underline{v} \in \mathbb{R}^{\hat{p}} \times [t_o, t_f]$, $\underline{z} \in \mathbb{R}^{\hat{q}} \times [t_o, t_f]$ y $f : \mathbb{R}^{\hat{q}} \times \mathbb{R}^{\hat{p}} \times [t_o, t_f] \rightarrow \mathbb{R}^{\hat{q}} \times [t_o, t_f]$.

Ahora, dividimos el dominio temporal $[t_o, t_f]$ en sub-intervalos de tamaño

$$\Delta T = \frac{t_f - t_o}{\hat{k}}:$$

$$t_o = T_o < T_1 < \dots < T_k = k \Delta T < T_{k+1} < \dots < T_{\hat{k}} = t_f$$

Entonces definimos los estados $\{\underline{z}_0, \underline{z}_1, \dots, \underline{z}_k, \dots, \underline{z}_{\hat{k}-1}\}$ de modo que

$\underline{z}_k \in \mathbb{R}^q \times [T_k, T_{k+1}]$ sea la solución de:

$$\left. \begin{aligned} \dot{\underline{z}}_k &= f(\underline{z}_k, \underline{v}_k, t) & t \in [T_k, T_{k+1}], \\ \underline{z}_k(T_k^+) &= Z_k, \end{aligned} \right\} k = 0, 1, \dots, \hat{k} - 1 \quad (3.2)$$

La solución de (3.2) será también solución de (3.1) si y sólo si, para cualquier $k = 0, 1, \dots, \hat{k} - 1$ se cumple que:

$$\begin{aligned} \underline{z}_k &= \underline{z} \quad \forall t \in [T_k, T_{k+1}], \\ \underline{v}_k &= \underline{v} \quad \forall t \in [T_k, T_{k+1}], \text{ y} \\ Z_k &= \underline{z}(T_k). \end{aligned} \quad (3.3)$$

La condición impuesta en la ecuación (3.3) establece que el estado \underline{z}_k definido en el sub-intervalo $[T_k, T_{k+1}]$, coincida en el tiempo final con la condición inicial del estado \underline{z}_{k+1} , dicho de otra manera, satisfacemos (3.3) si y sólo si:

$$\begin{aligned} Z_0 &= \underline{z}_{-1}(T_0) = \underline{z}(t_0), & \text{si } k = 0; \\ Z_k &= \underline{z}_{k-1}(T_k^-), & \text{si } k = 1, 2, \dots, \hat{k} - 1. \end{aligned} \quad (3.4)$$

Notemos que con la división del dominio temporal en \hat{k} sub-intervalos, estamos reduciendo el tamaño del problema (3.1), en \hat{k} problemas más pequeños (3.2), este proceso introduce al problema de control una restricción adicional

PROBLEMA P2, Problema de minimización con restricciones

$$\left\{ \begin{array}{l} \text{minimizar } \mathcal{J}(\underline{z}_k, \underline{v}_k, Z_k) = \frac{1}{2} \sum_{k=0}^{\hat{k}-1} \int_{T_k}^{T_{k+1}} (r \|\underline{v}_k\|_{R_h}^2 + q \|\underline{z}_k - \tilde{y}_k\|_{M_h}^2) dt \\ \quad + \frac{s}{2} \|z_{\hat{k}-1}(T_{\hat{k}}) - \tilde{y}_{\hat{k}-1}(T_{\hat{k}})\|_{M_h}^2 \\ \quad + \frac{1}{2\varepsilon \Delta T} \sum_{k=1}^{\hat{k}-1} \|\underline{z}_{k-1}(T_k^-) - Z_k\|_{M_h}^2, \\ \text{sujeto a } \left\{ \begin{array}{l} \dot{\underline{z}}_k = f(\underline{z}_k, \underline{v}_k, t), \quad t \in [T_k, T_{k+1}], \\ \underline{z}_k(T_k^+) = Z_k \end{array} \right. \quad k = 0, 1, \dots, \hat{k} - 1. \end{array} \right. \quad (3.6)$$

3.2. Condiciones de Optimalidad

Para convertir el problema **P2** (3.6), en un problema de minimización sin restricciones, definimos el funcional Lagrangiano (\mathcal{L}), tomando la misma consideración del capítulo anterior, respecto a las ecuaciones de estado (2.13):

$$\begin{aligned} \min \mathcal{L}(\underline{z}, \underline{v}, Z, \underline{p}, \underline{\eta}) &= \mathcal{J}(\underline{z}, \underline{v}, Z) + \sum_{k=0}^{\hat{k}-1} \int_{T_k}^{T_{k+1}} \underline{p}_k^T (M_h \dot{\underline{z}}_k - A_h \underline{z}_k - B_h \underline{v}_k - c_h) dt \\ &\quad + \sum_{k=0}^{\hat{k}-1} \underline{\eta}_k^T (z_k(T_k^+) - Z_k). \end{aligned} \quad (3.7)$$

Minimizar (3.7), es equivalente a minimizar (3.6). La solución que minimiza éstas, es la que minimiza (\mathcal{L}) en cada sub-intervalo $[T_k, T_{k+1}]$, $\forall k = 0, 1, \dots, \hat{k}-1$ y la denominamos $\underline{z}_k^*(t)$, $\underline{v}_k^*(t)$ y Z_k^* . En las ecuaciones siguientes escribiremos \underline{z}_k^* y \underline{v}_k^* para referirnos a $\underline{z}_k^*(t)$ y $\underline{v}_k^*(t)$ respectivamente.

La variación total de primer orden $\delta\mathcal{L}$ tiene la forma:

$$\delta\mathcal{L} = \sum_{k=0}^{\hat{k}-1} \frac{\partial\mathcal{L}}{\partial z_k} \delta z_k + \sum_{k=0}^{\hat{k}-1} \frac{\partial\mathcal{L}}{\partial v_k} \delta v_k + \sum_{k=0}^{\hat{k}-1} \frac{\partial\mathcal{L}}{\partial Z_k} \delta Z_k + \sum_{k=0}^{\hat{k}-1} \frac{\partial\mathcal{L}}{\partial p_k} \delta p_k + \sum_{k=0}^{\hat{k}-1} \frac{\partial\mathcal{L}}{\partial \eta_k} \delta \eta_k.$$

Las condiciones de optimalidad de Karush-Kuhn-Tucker (KKT) son obtenidas anulando la variación total de primer orden $\delta\mathcal{L}$ al ser evaluada en los valores z_k^* , v_k^* y Z_k^* , (véase [22], [23]). Por simplicidad, la evaluación en la solución óptima $|_{z_k^*, v_k^*, Z_k^*}$ representaremos por $|_*$ en las ecuaciones siguientes.

$$\begin{aligned} \delta\mathcal{L} = & \sum_{k=0}^{\hat{k}-1} \frac{\partial\mathcal{L}}{\partial z_k} \Big|_* \delta z_k + \sum_{k=0}^{\hat{k}-1} \frac{\partial\mathcal{L}}{\partial v_k} \Big|_* \delta v_k + \sum_{k=0}^{\hat{k}-1} \frac{\partial\mathcal{L}}{\partial Z_k} \Big|_* \delta Z_k \\ & + \sum_{k=0}^{\hat{k}-1} \frac{\partial\mathcal{L}}{\partial p_k} \Big|_* \delta p_k + \sum_{k=0}^{\hat{k}-1} \frac{\partial\mathcal{L}}{\partial \eta_k} \Big|_* \delta \eta_k = 0. \end{aligned} \quad (3.8)$$

Las distintas derivadas parciales del funcional \mathcal{L} son calculadas utilizando la definición (2.1). De la ecuación (3.8) obtenemos las siguientes condiciones de optimalidad:

Ecuaciones de Estado

$$\begin{aligned} \frac{\partial\mathcal{L}}{\partial p_k} \Big|_* &= (M_h \dot{z}_k - A_h z_k - B_h v_k - c_h) \Big|_* = 0 \quad t \in [T_k, T_{k+1}], \\ \frac{\partial\mathcal{L}}{\partial \eta_k} \Big|_* &= (z_k(T_k^+) - Z_k) \Big|_* = 0, \end{aligned} \quad (3.9)$$

donde $k \in \{0, 1, \dots, \hat{k} - 1\}$.

Estado Adjunto

$$\left. \frac{\partial \mathcal{L}}{\partial \underline{z}_k} \right|_* = \begin{cases} (M_h \dot{\underline{p}}_k + A_h \underline{p}_k - q M_h (z_k - \tilde{y}_k)) \Big|_* = 0 & t \in [T_k, T_{k+1}], \\ (\underline{p}_k(T_{k+1}^-) + \frac{1}{\varepsilon \Delta T} (z_k(T_{k+1}^-) - Z_{k+1})) \Big|_* = 0, \\ \forall k = 0, 1, \dots, \hat{k} - 2. \end{cases} \quad (3.10)$$

$$\left. \frac{\partial \mathcal{L}}{\partial \underline{z}_{\hat{k}-1}} \right|_* = \begin{cases} (M_h \dot{\underline{p}}_{\hat{k}-1} + A_h \underline{p}_{\hat{k}-1} - q M_h (z_{\hat{k}-1} - \tilde{y}_{\hat{k}-1})) \Big|_* = 0 & t \in [T_{\hat{k}-1}, T_{\hat{k}}], \\ (\underline{p}_{\hat{k}-1}(T_{\hat{k}}) + s (z_{\hat{k}-1}(T_{\hat{k}}) - \tilde{y}_{\hat{k}-1}(T_{\hat{k}}))) \Big|_* = 0. \end{cases} \quad (3.11)$$

Condición de Euler

$$\left. \frac{\partial \mathcal{L}}{\partial \underline{v}_k} \right|_* = (r R_h \underline{v}_k - B_h^T \underline{p}_k) \Big|_* = 0 \quad t \in [T_k, T_{k+1}] \quad \left. \vphantom{\frac{\partial \mathcal{L}}{\partial \underline{v}_k}} \right\} k = 0, 1, \dots, \hat{k} - 1. \quad (3.12)$$

$$\left. \frac{\partial \mathcal{L}}{\partial Z_k} \right|_* = (\underline{p}_k(T_k^+) + \frac{1}{\varepsilon \Delta T} (z_{k-1}(T_k^-) - Z_k)) \Big|_* = 0$$

La deducción de las condiciones de optimalidad (3.9), (3.10), (3.11) y (3.12) se encuentran en el apéndice (B.1).

3.3. Método de Solución del Problema de Control Paralelo

En la sección 3.1 planteamos el problema de control óptimo **P2** (3.6) de modo que pueda ser resuelto en varios procesos computacionales, o dicho de otra manera para ser resuelto en forma paralela. Como podemos ver el conjunto de ecuaciones (3.9), (3.10), (3.11) y (3.12), son equivalentes a (3.6).

En esta sección presentaremos un método para hallar la solución de (3.6) utilizando nuevamente el método del Gradiente Conjugado. Corrigiendo ahora en forma iterativa dos variables arbitrarias, la función de control \underline{v}_k^i y las condiciones iniciales Z_k^i introducida en la sección 3.1. Una vez arbitrado ambos valo-

res, obtenemos $\underline{z}_k^i(v_k^i, Z_k^i)$ de la ecuación (3.9) y $\underline{p}_k^i(v_k^i, Z_k^i)$ de las ecuaciones (3.10) y (3.11). En consecuencia la variación de primer orden del Lagrangiano \mathcal{L} (3.8) es igual al gradiente del funcional \mathcal{J} (3.6) al ser evaluados en $\underline{z}_k^i, \underline{v}_k^i, Z_k^i$ y \underline{p}_k^i , a los que representaremos por $|_i$ (véase Apéndice (B.2)).

$$\begin{aligned} \delta \mathcal{L} &= \sum_{k=0}^{\hat{k}-1} \left. \frac{\partial \mathcal{L}}{\partial v_k} \right|_i \delta v_k + \sum_{k=0}^{\hat{k}-1} \left. \frac{\partial \mathcal{L}}{\partial Z_k} \right|_i \delta Z_k \\ &= \sum_{k=0}^{\hat{k}-1} \left. \nabla_{v_k} \mathcal{J} \right|_i \delta v_k + \sum_{k=0}^{\hat{k}-1} \left. \nabla_{Z_k} \mathcal{J} \right|_i \delta Z_k, \end{aligned} \quad (3.13)$$

tomando cada sub-intervalo por separado y expresando la ecuación (3.13) en forma matricial tenemos:

$$\begin{bmatrix} \left. \frac{\partial \mathcal{L}}{\partial v_k} \right|_i \\ \left. \frac{\partial \mathcal{L}}{\partial Z_k} \right|_i \end{bmatrix}^T \begin{bmatrix} \delta v_k \\ \delta Z_k \end{bmatrix} = \begin{bmatrix} \left. \nabla_{v_k} \mathcal{J} \right|_i \\ \left. \nabla_{Z_k} \mathcal{J} \right|_i \end{bmatrix}^T \begin{bmatrix} \delta v_k \\ \delta Z_k \end{bmatrix}; \quad \forall k = 0, 1, \dots, \hat{k} - 1, \quad (3.14)$$

la ecuación (3.14) nos indica que minimizar \mathcal{L} (3.8) es equivalente a minimizar \mathcal{J} (3.6), por lo que en adelante trabajaremos únicamente con el funcional \mathcal{J} (3.6) para obtener nuestro algoritmo.

Notemos que el gradiente del funcional $\nabla \mathcal{J}$ del problema (3.6) posee dos componentes $\nabla \mathcal{J} = \left[\nabla_{v_k} \mathcal{J}^T \quad \nabla_{Z_k} \mathcal{J}^T \right]^T$, que son obtenidas a partir de las derivadas totales de \mathcal{J} respecto a las variables independientes v_k y Z_k respectivamente (véase Apéndice (B.2)). Y mientras las variables \underline{v}_k^i, Z_k^i arbitradas no sean la solución óptima, el gradiente del funcional de la ecuación (3.6) no será nulo, entonces podemos encontrar un nuevo valor $\underline{v}_k^{i+1}, Z_k^{i+1}$ tal que:

$$\nabla \mathcal{J}(\underline{z}_k^{i+1}(v_k^{i+1}, Z_k^{i+1}), v_k^{i+1}, Z_k^{i+1}) < \nabla \mathcal{J}(\underline{z}_k^i(v_k^i, Z_k^i), v_k^i, Z_k^i), \quad (3.15)$$

en el cual el esquema de corrección es el siguiente:

$$\begin{aligned} \underline{v}_k^{i+1} &= \underline{v}_k^i + \alpha^i dv_k^i & \forall k = 0, 1, \dots, \hat{k} - 1, \\ Z_k^{i+1} &= Z_k^i + \alpha^i dz_k^i & \forall k = 1, 2, \dots, \hat{k} - 1, \end{aligned} \quad (3.16)$$

donde α^i es un escalar obtenido del mismo modo que se obtuvo en (2.20); dv_k y dz_k son las direcciones de búsqueda para \underline{v}_k y Z_k respectivamente.

Al gradiente del funcional $\nabla \mathcal{J}$ en la iteración i lo denotaremos por:

$$\begin{aligned} g_k^i &= \nabla \mathcal{J} \Big|_{\underline{z}_k^i, \underline{v}_k^i, Z_k^i, p_k^i}, \\ &= \begin{bmatrix} gv_k^i \\ gz_k^i \end{bmatrix} \end{aligned} \quad (3.17)$$

entonces el valor de α^i es:

$$\alpha^i = - \frac{\sum_{k=0}^{\hat{k}-1} (g_k^i, d_k^i)}{\sum_{k=0}^{\hat{k}-1} (d_k^i, H d_k^i)} \quad (3.18)$$

donde $d_k = [dv_k^T \ dz_k^T]^T$, y $H := \nabla^2 \mathcal{J} \Big|_{\underline{z}_k^i, \underline{v}_k^i, Z_k^i, p_k^i}$ es la matriz Hessiana del funcional (3.6).

A continuación construimos las direcciones de búsqueda de (3.16) de modo que sean ortogonales respecto de la matriz Hessiana, que es como utiliza el algoritmo del gradiente conjugado.

La primera dirección, tomamos la dirección contraria del gradiente del fun-

cional, es decir:

$$\begin{aligned} d_k^0 &= -g_k^0 \\ \begin{bmatrix} dv_k^0 \\ dz_k^0 \end{bmatrix} &= - \begin{bmatrix} gv_k^0 \\ gz_k^0 \end{bmatrix} \end{aligned} \quad (3.19)$$

y las siguientes direcciones con el siguiente esquema:

$$\begin{aligned} dv_k^{i+1} &= -gv_k^{i+1} + \beta^i dv_k^i \quad \forall k = 0, 1, \dots, \hat{k} - 1, \\ dz_k^{i+1} &= -gz_k^{i+1} + \beta^i dz_k^i \quad \forall k = 1, 2, \dots, \hat{k} - 1, \end{aligned} \quad (3.20)$$

donde

$$\beta^i = \frac{\sum_{k=0}^{\hat{k}-1} (g_k^{i+1}, H d_k^i)}{\sum_{k=0}^{\hat{k}-1} (d_k^i, H d_k^i)}. \quad (3.21)$$

El algoritmo corrige las funciones de control y las condiciones iniciales (3.16) en forma iterativa, la corrección se realiza hasta cumplirse algún criterio de parada, en el contexto de este trabajo el criterio de parada que utilizamos es el siguiente:

$$\frac{\|g^i\|}{\|g^0\|} \leq tol, \quad (3.22)$$

donde $g := [g_0^T \quad g_1^T \quad \dots \quad g_{\hat{k}-1}^T]^T$.

Como podemos ver, utilizamos el mismo esquema del algoritmo (2.4.1), con algunas modificaciones que detallamos a continuación:

Algoritmo GC paralelo (3.3)

- 1 Sean $\underline{v}^0 \in \mathbb{R}^{\hat{p}} \times [t_0, t^f]$ y $Z^0 \in \mathbb{R}^{\hat{k}(\hat{k}-1)}$ valores arbitrarias.
- 2 Obtener \underline{z}_k^0 y \underline{p}_k^0
- 3 $i = 0$, Calcular el gradiente del funcional $g^{(0)}$
- 4 **mientras** $\frac{\|g^i\|}{\|g^0\|} \leq tol$
- 5 **if** $i = 0$ **then**,
- 6 $d_k^i = -g_k^i$
- 7 **else**
- 8 $dv_k^i = -gv_k^i + \beta^i dv_k^{i-1}$
- 9 $dz_k^i = -gz_k^i + \beta^i dz_k^{i-1}$
- 10 **end**
- 11 $\underline{v}^{i+1} = \underline{v}^i + \alpha^i d_v^i$
- 12 $Z^{i+1} = Z^i + \alpha^i d_Z^i$
- 13 $g^{i+1} = g^i + \alpha^i H d^i$
- 14 $i = i + 1$
- 15 **end**

Notemos que en el algoritmo (3.3), los valores de \underline{v}_k , \underline{p}_k , \underline{v}_k y Z_k , así como las direcciones de búsqueda respectivas son calculados $\forall k = 0, 1, \dots, \hat{k} - 1$ en \hat{k} procesos computacionales en forma paralela. En otras palabras los cálculos para cada valor de k se realizan en forma simultanea en un proceso computacional distinto.

DESARROLLO DE UNA APLICACIÓN PARALELA EN EL TIEMPO PARA LA SIMULACIÓN Y CONTROL DE ECUACIONES DIFERENCIALES PARCIALES PARABÓLICAS

Como podemos observar en el algoritmo (2.4.1), luego de estimar un valor inicial para la función de control \underline{v}_k , y para las condiciones iniciales Z_k , procedemos a resolver la ecuación de estados (3.9) y los estados adjuntos (3.10) y (3.11) en cada sub-dominio $[T_k, T_{k+1}]$. Para ello en cada sub-dominio utilizamos la misma discretización temporal hecha en la sección (2.4.1), entonces \underline{z}_k , \underline{v}_k , $\underline{\tilde{y}}_k$ y \underline{p}_k discretos tienen la siguiente forma:

$$\begin{aligned} \mathbf{z}_k &= \begin{bmatrix} z_{k,1}^T & z_{k,2}^T & \cdots & z_{k,\hat{l}}^T \end{bmatrix}^T, & \mathbf{v}_k &= \begin{bmatrix} v_{k,1}^T & v_{k,2}^T & \cdots & v_{k,\hat{l}}^T \end{bmatrix}^T, \\ \mathbf{p}_k &= \begin{bmatrix} p_{k,1}^T & p_{k,2}^T & \cdots & p_{k,\hat{l}}^T \end{bmatrix}^T, & \tilde{\mathbf{y}}_k &= \begin{bmatrix} \tilde{y}_{k,1}^T & \tilde{y}_{k,2}^T & \cdots & \tilde{y}_{k,\hat{l}}^T \end{bmatrix}^T, \end{aligned} \quad (3.23)$$

La ecuación (3.23) indica que cada sub-dominio $[T_k, T_{k+1}]$ dividimos en \hat{l} intervalos iguales, donde el paso entre intervalos es $\tau = (T_{k+1} - T_k)/\hat{l}$, y cada nodo asociado a estos intervalos de tiempo son denotados como $t_{k,l} = T_k + l\tau$. Entonces la ecuación de estados (3.9) discreta, tiene la misma forma que la ecuación (2.36), es decir:

$$\begin{aligned} M_h \left(\frac{z_{k,l} - z_{k,l-1}}{\tau} \right) &= A_h z_{k,l} + B_h v_{k,l} + c_h, \quad l = 0, 1, \dots, \hat{l} - 1; \\ z_{k,0} &= Z_k \end{aligned} \quad (3.24)$$

Utilizando las matrices $F_0 := M_h$ y $F_1 := M_h - \tau A_h$, podemos obtener los estados $z_{k,l}$ mediante el siguiente esquema iterativo:

$$\begin{aligned} z_{k,0} &= Z_k \\ F_1 z_{k,l} &= F_0 z_{k,l-1} + \tau B_h v_{k,l} + \tau c_h; \quad \text{para } l = 1 \text{ hasta } \hat{l} \end{aligned} \quad (3.25)$$

Del mismo modo, los estados adjuntos (3.10) y (3.11) discretos tienen la siguiente forma:

$$\begin{aligned}
 M_h \left(\frac{p_{k,l+1} - p_{k,l}}{\tau} \right) &= -A_h p_{k,l} + q M_h(z_{k,l} - \tilde{y}_{k,l}), \quad l = 0, 1, \dots, \hat{l} - 1; \\
 p_{k,\hat{l}} &= -\frac{1}{\varepsilon \Delta T} (z_{k,\hat{l}} - Z_{k+1}) \quad \forall k = 0, 1, \dots, \hat{k} - 2, \quad (3.26) \\
 p_{\hat{k}-1,\hat{l}} &= -s(z_{\hat{k}-1,\hat{l}} - \tilde{y}_{\hat{k}-1,\hat{l}}).
 \end{aligned}$$

Entonces podemos obtener el estado adjunto $p_{k,l}$ mediante el siguiente esquema iterativo:

$$\begin{aligned}
 p_{k,\hat{l}} &= -\frac{1}{\varepsilon \Delta T} (z_{k,\hat{l}} - Z_{k+1}) \quad \forall k = 0, 1, \dots, \hat{k} - 2, \\
 p_{\hat{k}-1,\hat{l}} &= -s(z_{\hat{k}-1,\hat{l}} - \tilde{y}_{\hat{k}-1,\hat{l}}), \\
 F_1 p_{k,l} &= F_0 p_{k,l+1} - \tau q M_h(z_{k,l} - \tilde{y}_{k,l}); \quad \text{para } l = \hat{l} - 1 \text{ hasta } 0. \quad (3.27)
 \end{aligned}$$

Las ecuaciones (3.25) y (3.27) son resueltas en forma simultánea en \hat{k} procesos, donde cada proceso resuelve el sub-dominio $[T_k, T_{k+1}]$ para $k = 0, 1, \dots, \hat{k} - 1$ respectivamente, de esta manera hallamos el estado z en todo el dominio $[t_o, t_f]$.

El siguiente paso del algoritmo (2.4.1) es calcular el gradiente del funcional $\nabla \mathcal{J}$ evaluando en los valores \underline{z}_k^i , \underline{v}_k , \underline{p}_k^i y Z_k^i , el cual es denotado por g_k^i . El gradiente del funcional tiene dos componentes es decir, $\nabla \mathcal{J} = \left[\nabla_{\underline{v}_k} \mathcal{J}^T \quad \nabla_{Z_k} \mathcal{J}^T \right]^T$, los cuales calculamos por separado. Primero calculamos $\frac{D\mathcal{J}}{D\underline{v}_k}$:

$$\left. \begin{aligned}
 \frac{D\mathcal{J}}{D\underline{v}_k} &= \frac{D\mathcal{J}}{D\underline{v}_k} + \frac{D\mathcal{J}}{D\underline{z}_k} \frac{\partial \underline{z}_k}{\partial \underline{v}_k}, \\
 \nabla_{\underline{v}_k} \mathcal{J} &= r R_h \underline{v}_k - B_h^T \underline{p}_k
 \end{aligned} \right\} k = 0, 1, \dots, \hat{k} - 1. \quad (3.28)$$

entonces:

$$gv_k^i = r R_h v_k^i - B_h^T p_k^i, \quad (3.29)$$

Discretizando la ecuación (3.29), utilizando los vectores v_k y p_k obtenemos gv_k^i mediante el siguiente esquema iterativo:

$$gv_{k,l}^i = r R_h v_{k,l}^i - B_h^T p_{k,l}^i; \quad \text{para } l = 1 \text{ hasta } \hat{l}. \quad (3.30)$$

Ahora calculamos $\frac{D\mathcal{J}}{DZ_k}$:

$$\left. \begin{aligned} \frac{D\mathcal{J}}{DZ_k} &= \frac{D\mathcal{J}}{DZ_k} + \frac{D\mathcal{J}}{Dz_k} \frac{\partial z_k}{\partial Z_k}, \\ \nabla_{Z_k} \mathcal{J} &= -M_h p_k(T_k^+) - \frac{1}{\varepsilon \Delta T} M_h \left(z_{k-1}(T_k^-) - Z_k \right) \end{aligned} \right\} k = 0, 1, \dots, \hat{k} - 1. \quad (3.31)$$

entonces:

$$gz_k^i = -M_h p_k^i(T_k^+) - \frac{1}{\varepsilon \Delta T} M_h \left(z_{k-1}^i(T_k^-) - Z_k^i \right) \quad (3.32)$$

Al considerar el vector discreto p_k y de la ecuación (3.26) podemos ver que:

$$\begin{aligned} p_{k,0} &= p_k(T_k^+), \\ p_{k-1,\hat{l}} &= -\frac{1}{\varepsilon \Delta T} \left(z_{k-1}(T_k^-) - Z_k \right), \end{aligned} \quad (3.33)$$

Entonces la ecuación (3.32) discreta queda como sigue:

$$gz_k^i = M_h \left(-p_{k,0}^i + p_{k-1,\hat{l}}^i \right). \quad (3.34)$$

La deducción de las ecuaciones (3.28) y (3.31) se encuentran en el apéndice (B.2).

Para calcular α^i y β^{i-1} , necesitamos calcular $H d_k^i$, el cual calculamos resolviendo las siguientes ecuaciones (véase apéndice (B.3)):

$$\left\{ \begin{array}{l} \left\{ \begin{array}{l} \underline{\psi}_k(T_k^+) = dz_k^i \\ M_h \underline{\dot{\psi}}_k = A_h \underline{\psi}_k + B_h d\underline{v}_k^i \end{array} \right. \quad \forall k = 0, 1, \dots, \hat{k} - 1. \\ \left\{ \begin{array}{l} \underline{\psi}_{p_k}(T_{k+1}^-) = -\frac{1}{\varepsilon \Delta T} (\underline{\psi}_k(T_{k+1}^-) - dz_{k+1}^i) \quad \forall k = 0, 1, \dots, \hat{k} - 2. \\ \underline{\psi}_{p_{\hat{k}-1}}(T_{\hat{k}}) = -s \left(\underline{\psi}_{\hat{k}-1}(T_{\hat{k}}) \right) \\ M_h \underline{\dot{\psi}}_{p_k} = -A_h \underline{\psi}_{p_k} + q M_h \underline{\psi}_k \end{array} \right. \end{array} \right. \quad (3.35)$$

donde $\underline{\psi}_k$ y $\underline{\psi}_{p_k}$ son variables auxiliares. Luego el producto $H d_k^i$ es:

$$H \begin{bmatrix} d\underline{v}_k^i \\ dz_k^i \end{bmatrix} = \begin{bmatrix} r R_h d\underline{v}_k^i - B_h^T \underline{\psi}_{p_k} \\ -M_h \underline{\psi}_{p_k}(T_k^+) - \frac{1}{\varepsilon \Delta T} M_h \left(\underline{\psi}_{k-1}(T_k^-) - dz_k^i \right) \end{bmatrix} \quad (3.36)$$

Discretizando (3.35) del mismo modo realizado en (3.24) y (3.26) podemos hallar

$H d_k^i$ en forma iterativa como sigue:

$$\left\{ \begin{array}{l} \left\{ \begin{array}{l} \psi_{k,1} = dz_k^i \\ F_1 \psi_{k,l} = F_0 \psi_{k,l-1} + \tau B_h d\underline{v}_{k,l}^i; \quad \text{para } l = 1 \text{ hasta } \hat{l} \end{array} \right. \\ \left\{ \begin{array}{l} \psi_{p_{k,\hat{l}}} = -\frac{1}{\varepsilon \Delta T} (\psi_{k,\hat{l}} - dz_{k+1}^i) \quad \forall k = 0, 1, \dots, \hat{k} - 2. \\ \psi_{p_{\hat{k}-1,\hat{l}}} = -s \psi_{\hat{k}-1,\hat{l}} \\ F_1 \psi_{p_{k,l}} = F_0 \psi_{p_{k,l+1}} - \tau q M_h \psi_{k,l}; \quad \text{para } l = \hat{l} - 1 \text{ hasta } 0 \end{array} \right. \end{array} \right. \quad (3.37)$$

luego $H d_k^i$ discreto es obtenido mediante el siguiente esquema iterativo:

$$\left\{ \begin{array}{l} H_v d\underline{v}_{k,l}^i = r R_h d\underline{v}_{k,l}^i - B_h^T \psi_{p_{k,l}} \\ H_z dz_k^i = M_h (-\psi_{p_{k,0}} + \psi_{p_{k-1,\hat{l}}}) \end{array} \right. \quad \text{para } l = 1 \text{ hasta } \hat{l} \quad (3.38)$$

Notemos que para obtener α y β en el algoritmo (2.4.1), los \hat{k} procesos computacionales calculan los valores (g_k^i, Hd_k^i) , (d_k^i, Hd_k^i) y (g_k^i, d_k^i) para $k = 0, 1, \dots, \hat{k} - 1$ respectivamente, luego envía los valores calculados a los $\hat{k} - 1$ procesos computacionales restantes, y de este modo cada uno de ellos calculan α^i y β^{i-1} . Para luego actualizar las funciones de control v_k^{i+1} y las condiciones iniciales Z_k^{i+1} .

El algoritmo realiza el proceso ya descrito en forma iterativa, hasta que se cumpla con el criterio de parada. Luego calcula el estado z_k con los valores de v_k y Z_k ya obtenidos, encontrando de este modo el estado óptimo z_k^* buscado.

3.4. Gradiente Conjugado Precondicionado

En la sección anterior se presentó el algoritmo del gradiente conjugado como método de solución del problema de control paralelo **P2** (3.6), este método presenta un bajo costo computacional, además posee buena estabilidad numérica.

La velocidad de convergencia del algoritmo depende del condicionamiento de la matriz Hessiana H . La matriz H del problema (3.6) no posee un buen condicionamiento, lo que genera que el algoritmo (2.4.1) converja lentamente a la solución óptima.

Para mejorar el condicionamiento de la matriz H utilizamos la técnica denominada precondicionamiento y así acelerar la velocidad de convergencia del algoritmo gradiente conjugado. La matriz P llamada matriz de precondicionamiento,

DESARROLLO DE UNA APLICACIÓN PARALELA EN EL TIEMPO PARA LA SIMULACIÓN Y CONTROL DE ECUACIONES DIFERENCIALES PARCIALES PARABÓLICAS

que presentamos en esta sección fue propuesto en [1] y se denomina algoritmo parareal.

A continuación deduciremos la matriz P del algoritmo parareal, para ello expresamos la diferencia entre el estado final de un sub-dominio y el estado inicial del sub-dominio siguiente, al cual llamamos saltos.

Ésta diferencia es obtenida marchando la malla fina, como puede observarse en la siguiente figura.

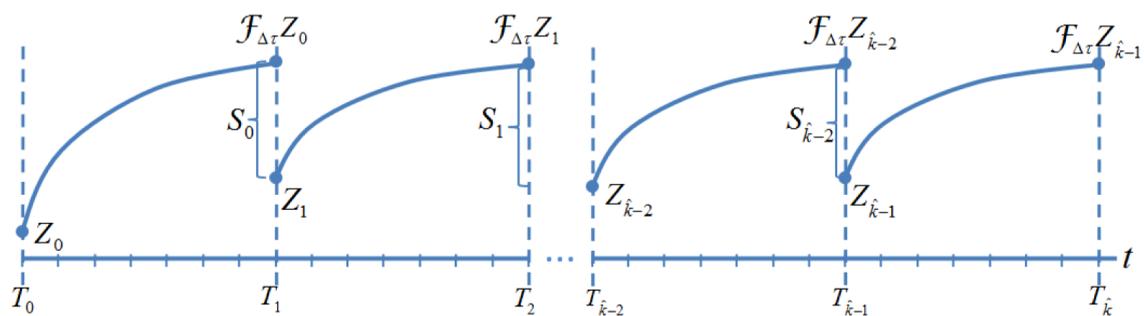


Figura 3.1: Temperatura calculada en forma Paralela.

Fuente: Elaboración Propia.

Para que la solución de la ecuación de estados calculada en cada sub-dominio sea equivalente a la solución de la ecuación de estado calculada de forma secuencial, se debe cumplir que los saltos S_0, S_1, \dots, S_{k-2} , que se ilustran en la figura 3.1 deben ser iguales a cero.

Entonces podemos expresar éstos saltos mediante la siguiente ecuación ma-

tricial (véase apéndice B.4):

$$\begin{bmatrix} I & & & & \\ -\mathcal{F}_{\Delta\tau} & I & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & -\mathcal{F}_{\Delta\tau} & I \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_{\hat{k}-1} \end{bmatrix} = \begin{bmatrix} S_0 + \mathcal{F}_{\Delta\tau} Z_0 \\ S_1 \\ \vdots \\ S_{\hat{k}-2} \end{bmatrix} \quad (3.39)$$

$CZ = S,$

donde $\mathcal{F}_{\Delta\tau} = (F_1^{-1}F_0)^{\hat{l}}$ y $S_k = \tau \sum_{i=1}^{\hat{l}} (F_1^{-1}F_0)^{(\hat{l}-i)} F_1^{-1} (B_h v_{k,i} + c_h).$

El valor de $\mathcal{F}_{\Delta\tau}$ se obtiene calculando el estado en cada punto de la malla fina, el cual consiste en la discretización de la malla gruesa, como se puede ver en la figura 3.1.

Del mismo modo podemos utilizar $\mathcal{F}_{\Delta\tau}$ para hallar el estado adjunto como se ilustra en la siguiente figura:

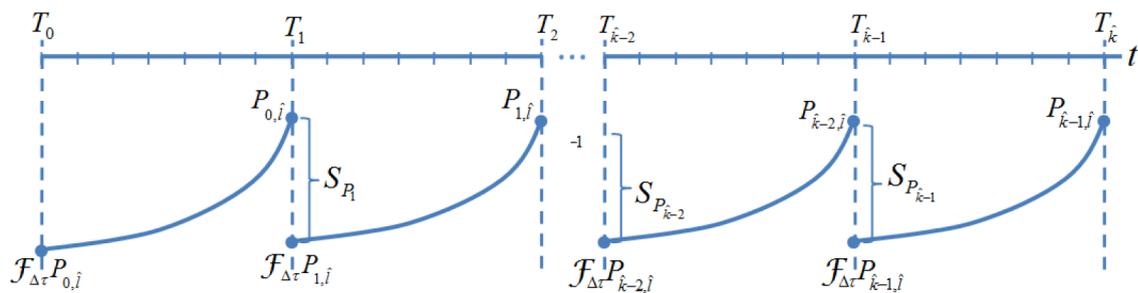


Figura 3.2: Estado Adjunto en Paralelo. Fuente: Elaboración Propia.

El gradiente de las condiciones iniciales $\nabla_Z \mathcal{J}$ dado por la ecuación (3.34), podemos expresar como una relación entre el valor final del adjunto

$p_{k,\hat{l}} = \frac{1}{\varepsilon\Delta T}(z_{k,\hat{l}} - Z_{k+1})$ y los saltos del estado adjunto S_{p_k} como se ilustra en la figura 3.2, es decir:

$$\nabla_Z \mathcal{J} = \mathbf{C}^T \mathbf{K} \begin{bmatrix} z_{0,\hat{l}} - Z_1 \\ z_{1,\hat{l}} - Z_2 \\ \vdots \\ z_{\hat{k}-2,\hat{l}} - Z_{\hat{k}-1} \end{bmatrix} - \begin{bmatrix} F_0 S_{p_1} \\ F_0 S_{p_2} \\ \vdots \\ F_0 (S_{p_{\hat{k}-1}} + \mathcal{F}_{\Delta T}^T p_{\hat{k}-1,\hat{l}}) \end{bmatrix} \quad (3.40)$$

donde $S_{p_k} = \tau q \sum_{i=1}^{\hat{l}} (F_1^{-1} F_0)^{(\hat{l}-i)} F_1^{-1} M_h(z_{k,\hat{l}-i} + \tilde{y}_{k,\hat{l}-i})$,

$$\mathbf{K} = \begin{bmatrix} \frac{1}{\varepsilon\Delta T} F_0 & & & \\ & \frac{1}{\varepsilon\Delta T} F_0 & & \\ & & \ddots & \\ & & & \frac{1}{\varepsilon\Delta T} F_0 \end{bmatrix}$$

utilizando la ecuación (3.39) en (3.40) tenemos:

$$\nabla_Z \mathcal{J} = \mathbf{C}^T \mathbf{K} \mathbf{C} \mathbf{Z} - \mathbf{F}, \quad (3.41)$$

donde $\mathbf{F} = \mathbf{C}^T \mathbf{K} \mathbf{S} + \mathbf{S}_p$.

Notemos que para hallar las condiciones iniciales Z_k^* del problema (3.6), se debe cumplir que $\nabla_Z \mathcal{J}$ sea igual a cero. Entonces para hallar Z_k^* tenemos que resolver la siguiente ecuación matricial:

$$\mathbf{C}^T \mathbf{K} \mathbf{C} \mathbf{Z} = \mathbf{F}, \quad (3.42)$$

para resolver esta ecuación podemos invertir el producto de las matrices $\mathbf{C}^T \mathbf{K} \mathbf{C}$, lo cual presenta un elevado costo computacional, pues necesitamos obtener $\mathcal{F}_{\Delta T}$, y éste se obtiene marchando la malla fina.

Una forma aproximada de obtener Z es por medio de la discretización de la malla gruesa por el método de Euler Implícito (Backward Euler), como se ilustra en la figura:

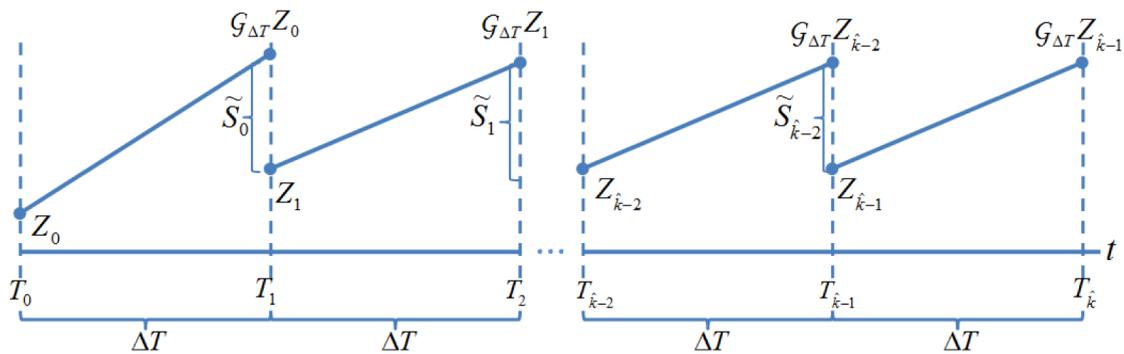


Figura 3.3: Estados malla gruesa. Fuente: Elaboración Propia.

La discretización se realiza utilizando pasos del tiempo de tamaño ΔT , y expresamos como sigue:

$$M_h \frac{Z_{k+1} - Z_k}{\Delta T} = A_h Z_{k+1} + B_h v_{k,0} + c_h,$$

$$Z_{k+1} = (M_h - \Delta T A_h)^{-1} M_h Z_k + (M_h - \Delta T A_h)^{-1} \Delta T (B_h v_{k,0} + c_h).$$

Entonces la aproximación de Z podemos expresar mediante la siguiente ecuación matricial:

$$\tilde{C} Z = \tilde{S},$$

donde:

$$\tilde{\mathbf{C}} = \begin{bmatrix} I & & & & \\ -\mathcal{G}_{\Delta T} & I & & & \\ & & \ddots & \ddots & \\ & & & & -\mathcal{G}_{\Delta T} & I \end{bmatrix}; \quad \mathcal{G}_{\Delta T} = (M_h - \Delta T A_h)^{-1} M_h. \quad (3.43)$$

Notemos que $\tilde{\mathbf{C}}$ es una aproximación de \mathbf{C} , además invertir $\tilde{\mathbf{C}}$ presenta un bajo costo computacional, debido a que $\mathcal{G}_{\Delta T}$ puede obtenerse inmediatamente. Entonces para acelerar la convergencia del algoritmo gradiente conjugado utilizamos la siguiente matriz como preconditionador:

$$\mathbf{P} = \tilde{\mathbf{C}}^T \mathbf{K} \tilde{\mathbf{C}}. \quad (3.44)$$

Luego modificamos el algoritmo 3.3 para que utilice el preconditionador \mathbf{P} de modo a acelerar la convergencia del Gradiente Conjugado, este algoritmo es llamado Gradiente Conjugado Precondicionado y se detalla a continuación.

Algoritmo GC paralelo preconditionado (3.4)

- 1 Sean $\underline{v}^0 \in \mathbb{R}^{\hat{p}\hat{l}\hat{k}}$ y $Z^0 \in \mathbb{R}^{\hat{q}(\hat{k}-1)}$ arbitrarias.
- 2 $i = 0$, Calcular el gradiente del funcional \mathbf{g}^0
- 3 **mientras** $\frac{\|\mathbf{g}^i\|}{\|\mathbf{g}^0\|} \leq tol$
- 4 $\mathbf{g}_p^{(i)} = [\mathbf{g}_v^i \ P^{-1} \mathbf{g}_Z^i]^T$
- 5 **if** $i = 0$ **then**,
- 6 $d^i = -\mathbf{g}_p^i$

DESARROLLO DE UNA APLICACIÓN PARALELA EN EL TIEMPO PARA LA SIMULACIÓN Y CONTROL DE ECUACIONES DIFERENCIALES PARCIALES PARABÓLICAS

7 else

8
$$d^{i+1} = -\mathbf{g}_p^i + \frac{\mathbf{g}_p^{i,T} H d^i}{d^{i,T} H d^i} d^i$$

9 end

10

$$\begin{bmatrix} \underline{v}^{i+1} \\ Z^{i+1} \end{bmatrix} = \begin{bmatrix} \underline{v}^{(i)} \\ Z^i \end{bmatrix} + \frac{g^{i,T} g_p^i}{d^{i,T} H d^i} d^i$$

11
$$\mathbf{g}^{i+1} = \mathbf{g}^i + \alpha^i H d^i$$

12
$$i = i + 1$$

13 end

CAPÍTULO 4

RESULTADOS NUMÉRICOS

4.1. Control óptimo del transitorio térmico de una placa de circuitos electrónicos mediante un sistema de enfriamiento en los bordes

El control óptimo del transitorio térmico de una placa de circuitos electrónicos con un sistema de enfriamiento en los bordes es modelado por la ecuación de conducción de calor con condiciones de frontera de Neumann [20]. Consecuentemente controlar éste transitorio involucra la solución de una ecuación diferencial parcial parabólica.

4.1.1. Ecuación de conducción de calor con condiciones de frontera de Neumann

Para enunciar correctamente el problema introducimos la siguiente definición:

Definición 4.1 (Espacio de Lebesgue $L^2(\Omega)$). *Se define como el espacio de funciones cuadrado integrables sobre Ω .*

$$L^2(\Omega) := \left\{ f : \Omega \rightarrow R \text{ y } \int_{\Omega} |f|^2 dx < \infty \right\} \quad (4.1)$$

Para modelar el transitorio térmico de un circuito térmico consideramos la ecuación de conducción de calor obtenida a partir de la ley de conducción de

calor de Fourier y el balance energético.

La conducción, la convección y la radiación son los tres modos de transferencia de energía térmica en los cuerpos materiales, debido a la diferencia de temperatura entre ellos.

La conducción de calor es un modo de transferencia de energía térmica por contacto directo entre moléculas que están a diferentes temperaturas y es debida al movimiento de los electrones libres presentes en las moléculas.

La transferencia de energía térmica se rige por la ley de conducción de calor de Fourier, que establece que el flujo de calor q es proporcional al gradiente de temperatura ∇z [19].

Expresado en:

$$\vec{q} = -\lambda \nabla z. \quad (4.2)$$

donde z es la temperatura, y λ la constante de proporcionalidad, es la conductividad térmica del material.

La ecuación (4.2) expresa que la energía térmica fluye desde las regiones con mayor temperatura hacia las de menor temperatura.

Este flujo de calor depende fuertemente de las propiedades físicas del medio, específicamente de la conductividad térmica λ de los materiales, dándose lugar en sólidos, líquidos y gases, siempre que exista diferencia de temperatura.

Consideremos un cuerpo determinado por un dominio $\Omega \subset R^2$, cuya frontera es una superficie $\Gamma = \partial\Omega$ como se ilustra en la siguiente figura (4.1).

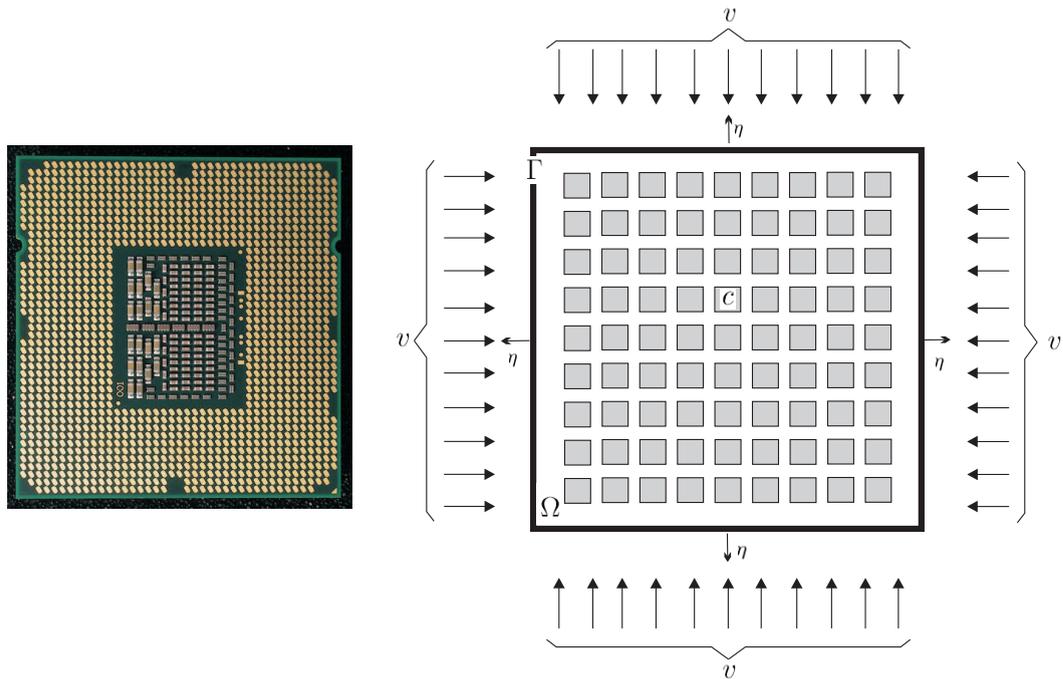


Figura 4.1: Ejemplo de aplicación: Control de temperatura de circuitos electrónicos. Fuente: Elaboración Propia.

La energía térmica en un dominio la expresaremos en función a la densidad de masa ρ y a la densidad de energía e del material que ocupa el dominio.

$$E_i = \int_{\Omega} \rho e \, d\Omega. \quad (4.3)$$

La variación de energía en el dominio con respecto al tiempo es:

$$\frac{dE_i}{dt} = \frac{d}{dt} \int_{\Omega} \rho e \, d\Omega. \quad (4.4)$$

La densidad de energía e , puede ser expresada como el producto entre el calor específico del cuerpo c_p a presión constante y la temperatura z , es decir:

$$e(x, y, t) = c_p(x, y) z(x, y, t), \quad (4.5)$$

donde c_p el calor específico es una propiedad física de los materiales que determina la capacidad de una sustancia de transferir la energía cinética de sus moléculas a otras moléculas adyacentes o a sustancias con las que está en contacto.

Luego teniendo en cuenta (4.5) las expresiones (4.3) y (4.4), quedan como:

$$E_i = \int_{\Omega} \rho c_p z d\Omega. \quad \text{y} \quad \frac{dE_i}{dt} = \int_{\Omega} \rho c_p \frac{dz}{dt} d\Omega. \quad (4.6)$$

La energía entrante al cuerpo se obtiene mediante la siguiente integral de superficie, entre el flujo de calor y un versor diferencial de la frontera $d\vec{\Gamma} = \vec{n} d\Gamma$, donde \vec{n} es el vector normal al diferencial de la frontera, consecuentemente la energía entrante es igual a:

$$E_e = \int_{\partial\Omega} \vec{q} \cdot d\vec{\Gamma} \quad (4.7)$$

La energía térmica total generada E_g en el dominio se obtiene integrando la función c que representa la densidad de energía térmica generada por las fuentes presentes en la región de estudio.

$$E_g = \int_{\Omega} c d\Omega \quad (4.8)$$

La ecuación de balance de energía térmica en un cuerpo deriva de la ley de la conservación de energía y establece que la variación de energía con respecto al tiempo $\frac{dE_i}{dt}$ en el dominio, es debida a la energía que entra E_e en el dominio y a la energía total generada E_g en el dominio [19]. Expresada en:

$$\frac{d}{dt} E_i = E_e + E_g, \quad (4.9)$$

y utilizando las expresiones (4.4),(4.7) y (4.8), la ecuación anterior, queda como:

$$\int_{\Omega} \rho c_p \frac{dz}{dt} d\Omega. = - \int_{\partial\Omega} \vec{q} \cdot d\vec{\Gamma} + \int_{\Omega} c d\Omega. \quad (4.10)$$

Notemos que si utilizamos el Sistema Métrico Internacional tenemos: ρ es la densidad de masa del cuerpo [kg/m^3], c_p es la densidad de energía interna del cuerpo [W/Kkg] y f es la densidad de calor generado por las fuentes [W/m^3].

De forma a simplificar la expresión (4.10) utilizaremos el teorema de la divergencia que expresa lo siguiente:

$$\int_{\partial\Omega} \vec{F} \cdot \vec{n} d\Gamma = \int_{\partial\Omega} \vec{F} \cdot d\vec{\Gamma} = \int_{\Omega} \nabla \cdot \vec{F} d\Omega. \quad (4.11)$$

donde Γ es una superficie suave y Ω es un área cerrada que es limitada por Γ .

Entonces aplicamos el teorema de la divergencia a (4.7) tenemos:

$$\int_{\partial\Omega} \vec{q} \cdot d\vec{\Gamma} = \int_{\Omega} \nabla \cdot \vec{q} d\Omega. \quad (4.12)$$

Reemplazando la expresión (4.12) en la ecuación de balance de energía (4.10) se tiene:

$$\int_{\Omega} \left(\rho c_p \frac{\partial z}{\partial t} + \nabla \cdot \vec{q} - c \right) d\Omega = 0. \quad (4.13)$$

Luego al utilizar la ley de Fourier (4.2) en la ecuación anterior tenemos:

$$\int_{\Omega} \left(\rho c_p \frac{\partial z}{\partial t} - \nabla \cdot \lambda \nabla z - c \right) d\Omega = 0. \quad (4.14)$$

La ecuación (4.14) sugiere que la siguiente ecuación se cumpla para todo $P \{x, y\} \in \Omega$.

$$\rho c_p \partial_t z = \nabla \cdot (\lambda \nabla z) + c. \quad (4.15)$$

En el intervalo de tiempo $t \in [t_o, t_f]$, asumiendo condiciones de frontera de Neumann. Las condiciones de frontera de Neumann consisten en considerar que existe un flujo de calor v saliente del dominio considerado, es decir:

$$\frac{\partial z}{\partial \eta} = v \quad \text{en } \Gamma \times [t_o, t_f], \quad (4.16)$$

donde η es la dirección normal a la frontera. La función v representa el mecanismo de enfriamiento en los bordes de la placa de circuito electrónico. El cuadrado de ésta función debe ser integrable, es decir $v \in L^2(\Gamma)$, de manera que más adelante podamos medir el esfuerzo para controlar el sistema.

Las condiciones iniciales del problema están definidas mediante la función z_o sobre todo el dominio.

La variable z es la función de temperatura en el punto $P\{x, y\} \in \Omega$ y en el tiempo t y debe ser por lo menos dos veces derivable con respecto al espacio, es decir $z \in C^2(\Omega)$, para que verifique (4.15).

La función c representa el flujo de calor generado por las fuentes presentes en el dominio considerado. En el contexto de este trabajo el flujo de calor generado será considerado constante por partes en el espacio y constante en todo el dominio temporal.

En el dominio Ω asumimos condiciones de continuidad, homogeneidad e isotropía de la región de estudio. Entonces la temperatura en el circuito elec-

trónico será modelada por la siguiente ecuación:

$$\begin{cases} \rho c_p \frac{\partial z}{\partial t} = \lambda \Delta z + c, & \text{en } \Omega \times [t_o, t_f], \\ \frac{\partial z}{\partial \eta} = v & \text{en } \Gamma \times [t_o, t_f], \\ z(t_o, \Omega) = z_o, \end{cases} \quad (4.17)$$

donde $z \in C^2(\Omega) \times [t_o, t_f]$, $z_o \in C^2(\Omega)$, $v \in L^2(\Gamma) \times [t_o, t_f]$ y Δ es el laplaciano.

4.1.2. Problema de control óptimo

El objetivo del problema de control es hallar una solución óptima (v^* y z^*) que minimice el gasto de energía del sistema de enfriamiento y el error cuadrático medio con respecto a una temperatura deseada. Y para formular el problema de control definimos el siguiente funcional de costo:

$$\mathcal{J}(z, v) = \frac{1}{2} \int_{t_o}^{t_f} \left(r \|v\|_{L^2(\Gamma)}^2 + q \|(z(t) - \tilde{y}(t))\|_{L^2(\Omega)}^2 \right) dt + \frac{1}{2} s \|z(t_f) - \tilde{y}(t_f)\|_{L^2(\Omega)}^2, \quad (4.18)$$

donde la norma $L^2(\Gamma)$, utilizada para definir el funcional de costo, se obtiene como sigue:

$$\|v\|_{L^2(\Gamma)}^2 = \int_{\Gamma} v^2 d\Gamma, \quad (4.19)$$

y la norma $L^2(\Omega)$ es:

$$\|z(t) - \tilde{y}(t)\|_{L^2(\Omega)}^2 = \int_{\Omega} (z - \tilde{y})^2 d\Omega. \quad (4.20)$$

donde, $M_h = M_h^T$, $A_h = A_h^T \in \mathbb{R}^{\hat{q} \times \hat{q}}$; $B_h \in \mathbb{R}^{\hat{q} \times \hat{p}}$; $R_h \in \mathbb{R}^{\hat{q} \times \hat{p}}$; $c_h \in \mathbb{R}^{\hat{q}}$. Los elementos de éstas matrices y el vector c_h se obtienen a partir de las funciones bases en el apéndice (C).

4.2. Método Gradiente Conjugado

En ésta sección se presentan los resultados numéricos relativos a algunos casos del control de temperatura para visualizar la influencia de los parámetros del controlador en la cantidad de iteraciones del algoritmo (2.4.1) del capítulo (2).

Consideramos una placa cuadrada de $5[cm]$ de lado (la conductividad térmica $\lambda = 1[\frac{W}{mK}]$ y el producto de la densidad por el calor específico de la placa $\rho c_p = 10^4[J/m^2K]$), un componente generador de calor con $2cm$ de lado, una densidad de calor generado de $4 \times 10^5[\frac{W}{m^2}]$, una temperatura inicial de $320[^\circ K]$ y el objetivo de enfriar la placa hasta los $300[^\circ K]$ en un tiempo de $10[s]$ (véase [20]). La discretización espacial consta de \hat{q} elementos triangulares uniformemente distribuidos en el dominio espacial y \hat{p} elementos de control en la frontera. Los valores del estado y del estado adjunto en cada tiempo del dominio temporal discreto se obtienen mediante el esquema Backward Euler y Forward Euler respectivamente. El dominio de tiempo es particionado con un paso de tiempo $\tau = (t_f - t_o)/(\hat{l})$.

El criterio de parada se basa en el decrecimiento de la norma del gradiente del funcional con respecto al primer gradiente del funcional. El algoritmo iterativo

corrige las funciones de control hasta que la relación $\frac{|g^i|}{|g^1|}$ sea menor que una tolerancia.

4.2.1. Ajuste del controlador

En ésta sección se muestran los resultados del ajuste de los parámetros r , q y s del funcional de costo para controlar el sistema, mantenemos constantes la cantidad de elementos del dominio espacial ($\hat{q} = 81$ y $\hat{p} = 32$), temporal ($\hat{l} = 256$) y establecemos la tolerancia igual a ($tol = 10^{-9}$). A continuación presentamos una tabla que relaciona el parámetro s con la cantidad de iteraciones necesarias para la convergencia del Método Gradiente Conjugado.

Tabla 4.1: Número de iteraciones del Método Gradiente Conjugado para un valor fijo de $r = 1$, $q = 0$.

s	1	10^2	10^3	$5 \cdot 10^3$	10^4	$5 \cdot 10^4$	10^5
Iter.	3	4	5	6	7	10	12
$ g^i $	$1,1 \cdot 10^{-11}$	$3,0 \cdot 10^{-9}$	$7,6 \cdot 10^{-8}$	$2,5 \cdot 10^{-6}$	$1,5 \cdot 10^{-6}$	$1,3 \cdot 10^{-5}$	$8,7 \cdot 10^{-5}$

Fuente: Elaboración Propia.

Notemos que existe una relación de compromiso entre la cantidad de iteraciones del algoritmo, gasto de energía en el control y resultados del control.

DESARROLLO DE UNA APLICACIÓN PARALELA EN EL TIEMPO PARA LA SIMULACIÓN Y CONTROL DE ECUACIONES DIFERENCIALES PARCIALES PARABÓLICAS

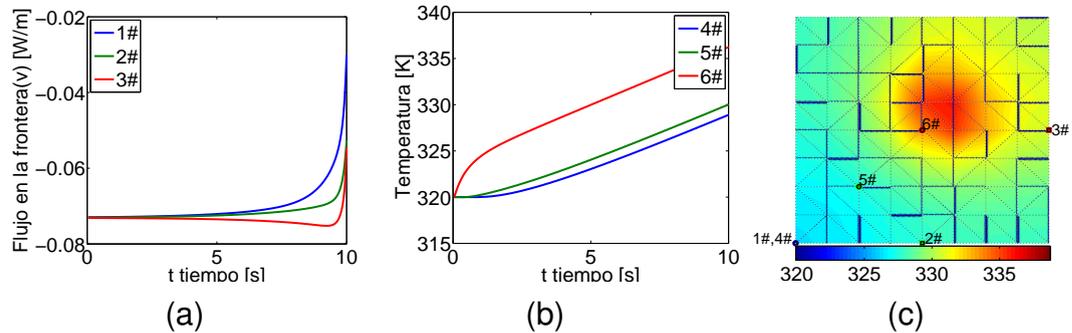


Figura 4.2: Caso: $s = 1$, (a) Función de control en 1, 2, 3, (b) Temperatura en 4, 5, 6, (c) Temperatura [K] p/t=10[s]. Fuente: Elaboración Propia.

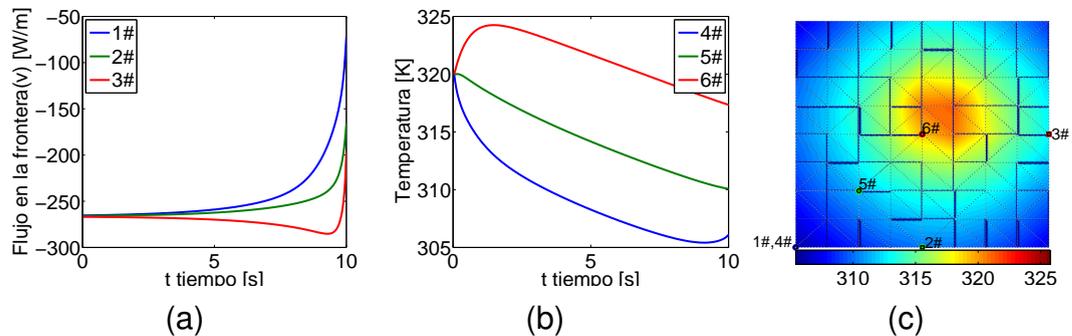


Figura 4.3: Caso: $s = 10000$, (a) Función de control en 1, 2, 3, (b) Temperatura en 4, 5, 6, (c) Temperatura [K] p/t=10[s]. Fuente: Elaboración Propia.

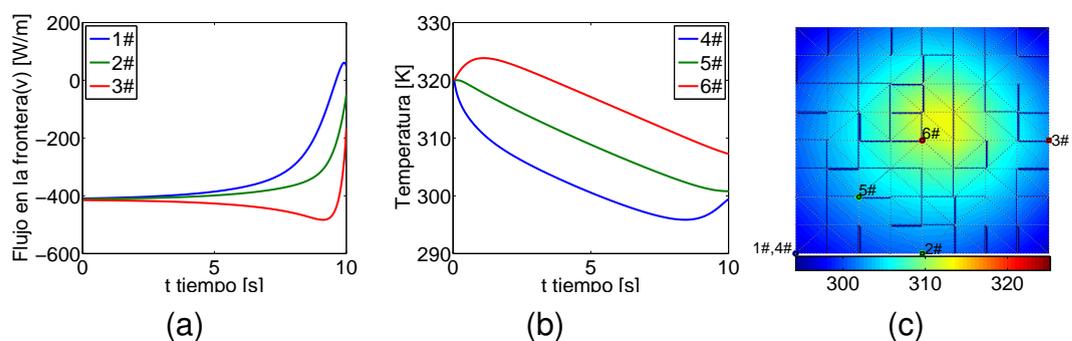


Figura 4.4: Caso: $s = 100000$, (a) Función de control en 1, 2, 3, (b) Temperatura en 4, 5, 6, (c) Temperatura [K] p/t=10[s]. Fuente: Elaboración Propia.

Seleccionamos el valor de $s = 10000$ porque es el valor que controla adecuadamente la temperatura sin elevar notablemente el gasto del control (ver

DESARROLLO DE UNA APLICACIÓN PARALELA EN EL TIEMPO PARA LA SIMULACIÓN Y CONTROL DE ECUACIONES DIFERENCIALES PARCIALES PARABÓLICAS

gráfico). Entonces a continuación ajustamos el escalado del funcional dividiendo todos los parámetros r y s por una constante para reducir la norma del gradiente del funcional. En la siguiente tabla se muestran los resultados de este ajuste.

Tabla 4.2: Norma del gradiente del funcional al modificar las relaciones $r/const$ y $s/const$.

$const$	1	10	10^{-2}	10^{-3}
Iter.	7	7	7	7
$ g^i $	$1,5 \cdot 10^{-6}$	$1,5 \cdot 10^{-7}$	$1,5 \cdot 10^{-8}$	$1,5 \cdot 10^{-9}$

Fuente: Elaboración Propia.

Seleccionamos el valor de $const = 10^{-3}$ porque reduce el valor del gradiente (esto permite mayor precisión), consecuentemente los parámetros quedan como sigue $r = 10^{-3}$ y $s = 10$. A continuación presentamos una tabla que relaciona el parámetro q con la cantidad de iteraciones necesarias para la convergencia.

Tabla 4.3: Ajuste de q .

q	10^{-2}	10^{-1}	1	10
Iter.	7	7	8	16
$ g^i $	$1,6 \cdot 10^{-9}$	$4,7 \cdot 10^{-9}$	$3,77 \cdot 10^{-8}$	$1,19 \cdot 10^{-7}$

Fuente: Elaboración Propia.

Seleccionamos el valor de $q = 1$ porque es el valor que controla adecuadamente la temperatura (sin sobrepicos).

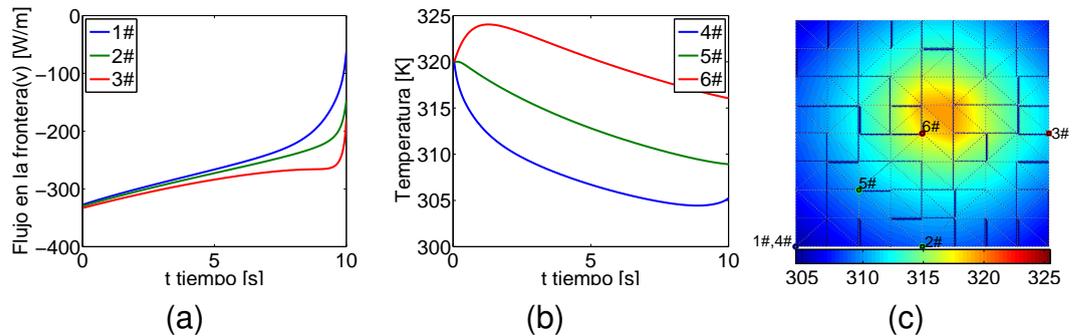


Figura 4.5: Caso: $q = 0,1$, (a) Función de control en 1, 2, 3, (b) Temperatura en 4, 5, 6, (c) Temperatura [K] p/t=10[s]. Fuente: Elaboración Propia.

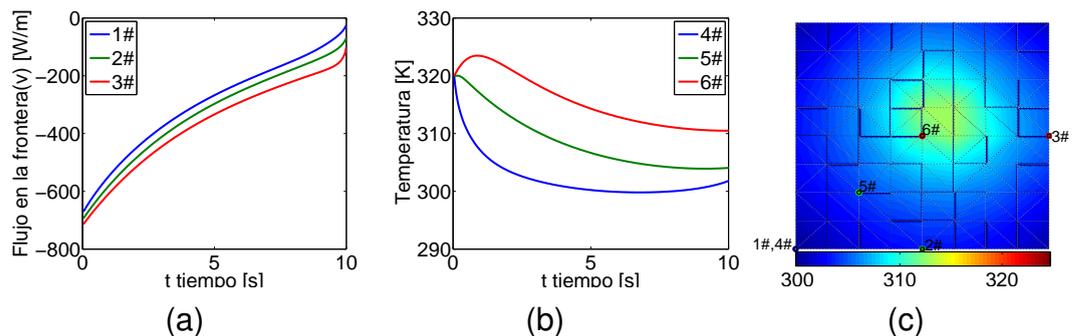


Figura 4.6: Caso: $q = 1$, (a) Función de control en 1, 2, 3, (b) Temperatura en 4, 5, 6, (c) Temperatura [K] p/t=10[s]. Fuente: Elaboración Propia.

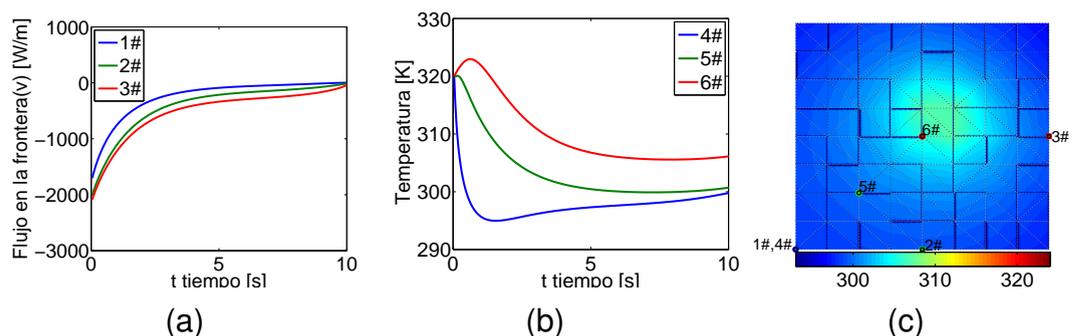


Figura 4.7: Caso: $q = 10$, (a) Función de control en 1, 2, 3, (b) Temperatura en 4, 5, 6, (c) Temperatura [K] p/t=10[s]. Fuente: Elaboración Propia.

4.2.2. Cantidad de Iteraciones vs tolerancia

En ésta sección mostramos la relación entre la cantidad de iteraciones del algoritmo y la tolerancia utilizada para definir el criterio de parada. Consideramos los parámetros $r = 10^{-3}$, $q = 1$ y $s = 10$ constantes, así también como la cantidad de elementos del dominio espacial ($\hat{q} = 81$ y $\hat{p} = 32$), temporal $\hat{l} = 256$).

Tabla 4.4: Cantidad de Iteraciones vs tolerancia

<i>tol</i>	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}
Iter.	6	7	8	8	9
$ g^i $	$1,1 \cdot 10^{-5}$	$7,0 \cdot 10^{-7}$	$3,8 \cdot 10^{-8}$	$3,8 \cdot 10^{-8}$	$9,9 \cdot 10^{-10}$

Fuente: Elaboración Propia.

4.2.3. Cantidad de Iteraciones vs cantidad de elementos del dominio espacial

En ésta sección mostramos la relación entre la cantidad de iteraciones del algoritmo y la cantidad de elementos del dominio espacial. Consideramos los parámetros $r = 10^{-3}$, $q = 1$ y $s = 10$ constantes, así también como la cantidad de elementos del dominio temporal $\hat{l} = 256$ y la tolerancia utilizada para definir el criterio de parada $tol = 10^{-9}$.

Tabla 4.5: Influencia de la cantidad elementos del dominio espacial

\hat{q}	25	81	289	1089
Iter.	12	8	7	6
$ g^i $	$1,2 \cdot 10^{-7}$	$3,8 \cdot 10^{-8}$	$1,9 \cdot 10^{-9}$	$3,6 \cdot 10^{-9}$

Fuente: Elaboración Propia.

4.3. Método Gradiente Conjugado Precondicionado paralelo

En ésta sección se presentan los resultados numéricos relativos a algunos casos del control de temperatura para visualizar la influencia del parámetro ε del controlador en la cantidad de iteraciones del algoritmo paralelo, evaluar la escalabilidad y el speed-up del algoritmo.

4.3.1. Ajuste de la penalización ε

En ésta sección ajustamos el parámetro de penalización ε del funcional, mantenemos constante la cantidad de elementos del dominio espacial ($\hat{q} = 169$ y $\hat{p} = 48$), temporal ($\hat{l} = 12$ y $\hat{k} = 12$) y establecemos la tolerancia igual a ($tol = 10^{-9}$). A continuación presentamos una tabla que relaciona el parámetro ε con la cantidad de iteraciones necesarias para la convergencia del algoritmo paralelo

Notemos que existe una relación de compromiso entre la cantidad de iteraciones del algoritmo, el parámetro de penalización y los resultados del control.

Tabla 4.6: Ajuste de ε

ε	1	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}
Iter.	83	55	43	39	40	49
$ g^i $	$3,8 \cdot 10^{-5}$	$7,9 \cdot 10^{-5}$	$6,4 \cdot 10^{-4}$	$3,9 \cdot 10^{-3}$	$3,7 \cdot 10^{-2}$	$6,55 \cdot 10^{-1}$

Fuente: Elaboración Propia.

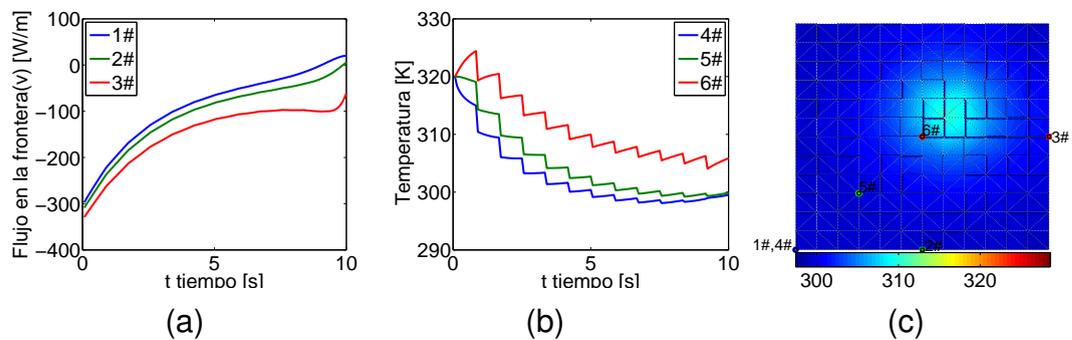


Figura 4.8: Caso: $\varepsilon = 10^{-1}$, (a) Función de control en 1, 2, 3, (b) Temperatura en 4, 5, 6, (c) Temperatura [K] p/t=10[s]. Fuente: Elaboración Propia.

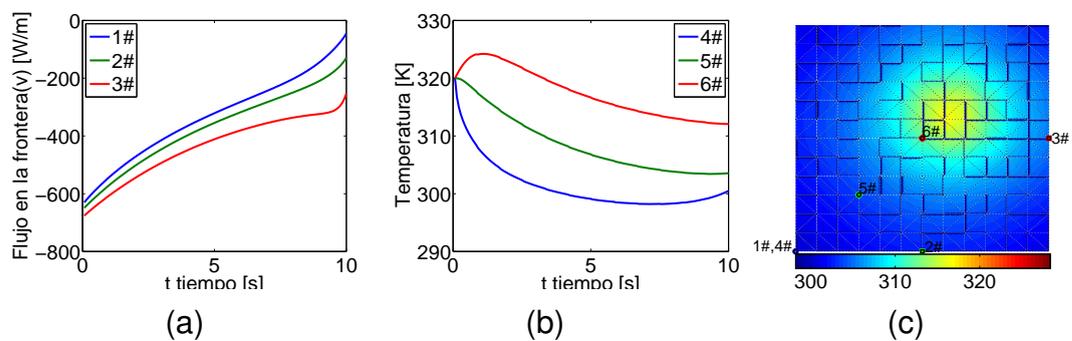


Figura 4.9: Caso: $\varepsilon = 10^{-3}$, (a) Función de control en 1, 2, 3, (b) Temperatura en 4, 5, 6, (c) Temperatura [K] p/t=10[s]. Fuente: Elaboración Propia.

DESARROLLO DE UNA APLICACIÓN PARALELA EN EL TIEMPO PARA LA SIMULACIÓN Y CONTROL DE ECUACIONES DIFERENCIALES PARCIALES PARABÓLICAS

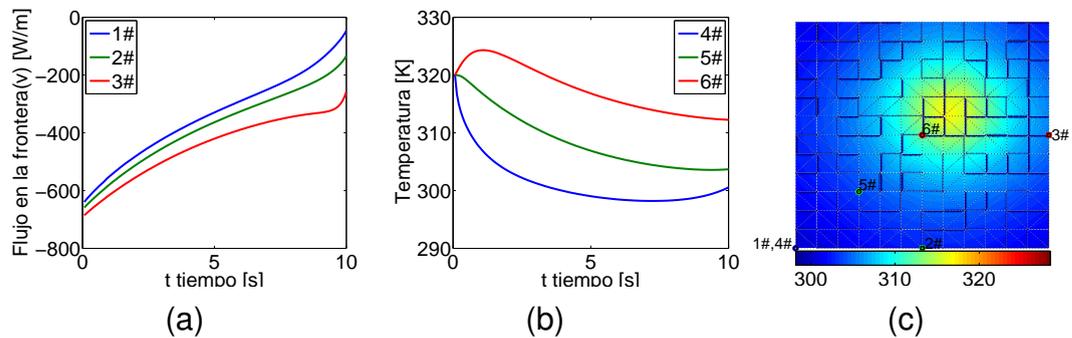


Figura 4.10: Caso: $\varepsilon = 10^{-4}$, (a) Función de control en 1, 2, 3, (b) Temperatura en 4, 5, 6, (c) Temperatura [K] p/t=10[s]. Fuente: Elaboración Propia.

Seleccionamos el valor de $\varepsilon = 10^{-3}$ porque es el valor que permite obtener una solución similar a la solución del algoritmo secuencial y sin elevar la cantidad de interacciones (ver gráfico).

4.3.2. Cantidad de Iteraciones vs tolerancia

En esta sección mostramos la relación entre la cantidad de iteraciones del algoritmo y la tolerancia utilizada para definir el criterio de parada. Consideramos los parámetros $r = 10^{-3}$, $q = 1$, $s = 10$ y $\varepsilon = 10^{-3}$ constantes, como así también la cantidad de elementos del dominio espacial ($\hat{q} = 169$ y $\hat{p} = 48$), temporal ($\hat{l} = 16$ y $\hat{k} = 16$).

Tabla 4.7: Cantidad de Iteraciones vs tolerancia

<i>tol</i>	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}
Iter.	4	11	19	27	42
$ g^i $	84,1	3,9	$6,0 \cdot 10^{-1}$	$6,9 \cdot 10^{-2}$	$3,1 \cdot 10^{-3}$

Fuente: Elaboración Propia.

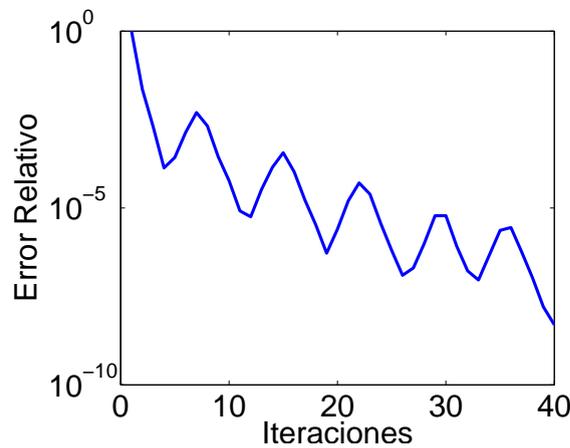


Figura 4.11: Norma del gradiente del funcional en función a la cantidad de iteraciones. Fuente: Elaboración Propia.

4.3.3. Cantidad de iteraciones vs cantidad de elementos del dominio espacial

En ésta sección mostramos la relación entre la cantidad de iteraciones del algoritmo y la cantidad de elementos del dominio espacial. Consideramos los parámetros $r = 10^{-3}$, $q = 1$, $s = 10$ y $\varepsilon = 10^{-3}$ constantes, como así también

la cantidad de elementos del dominio temporal ($\hat{l} = 16$ y $\hat{k} = 16$) y la tolerancia utilizada para definir el criterio de parada $tol = 10^{-9}$.

Tabla 4.8: Cantidad de iteraciones vs cantidad de elementos del dominio espacial

\hat{q}	25	81	289
Iter.	53	48	42
$ g^i $	$1,5 \cdot 10^{-2}$	$4,7 \cdot 10^{-3}$	$4,4 \cdot 10^{-3}$

Fuente: Elaboración Propia.

4.3.4. Escalabilidad débil

En ésta sección mostramos los resultados numéricos correspondientes a la escalabilidad del algoritmo paralelo. Un algoritmo paralelo es escalable cuando al aumentar la cantidad de procesadores \hat{k} la cantidad de iteraciones no varía significativamente.

Tabla 4.9: Escalabilidad débil $\hat{l} = 4$.

\hat{k}	4	8	16	32	64
Iter.	35	37	42	44	39

Fuente: Elaboración Propia.

4.3.5. Escalabilidad Fuerte

En ésta sección mostramos los resultados numéricos correspondientes a la evaluación de la escalabilidad fuerte del algoritmo paralelo.

Tabla 4.10: Escalabilidad Fuerte.

\hat{k}	4	8	16	32	64
\hat{l}	64	32	16	8	4
Iter.	39	39	42	41	38

Fuente: Elaboración Propia.

CAPÍTULO 5

CONCLUSIONES Y TRABAJOS FUTUROS

5.1. Conclusión

En el presente trabajo se implementan dos algoritmos para resolver problemas de control óptimo que involucran ecuaciones diferenciales parciales parabólicas. El primero se basa en el Método Gradiente Conjugado y el segundo en el Gradiente Conjugado Precondicionado, éste último posibilita resolver el problema en forma paralela.

Los resultados numéricos presentados demuestran que el primer algoritmo reduce el costo computacional de la resolución del problema de control, con respecto a los métodos directos.

Los parámetros r , q y s modifican la respuesta del sistema:

- Una relación q/r grande hace que el controlador lleve el estado, al estado deseado más rápidamente, sin importar el esfuerzo de control. La cantidad de iteraciones de los algoritmos varían de manera directamente proporcional a ésta relación.
- La relación s/r modifica la diferencia entre la temperatura en el tiempo final y la temperatura de referencia. Ésta relación también influye en la cantidad

DESARROLLO DE UNA APLICACIÓN PARALELA EN EL TIEMPO PARA LA SIMULACIÓN Y CONTROL DE ECUACIONES DIFERENCIALES PARCIALES PARABÓLICAS

de iteraciones necesarias para llegar a la convergencia, de ambos algoritmos

- Si las relaciones entre s/r o q/r son muy elevados el algoritmo puede presentar problemas de convergencia.

En el segundo algoritmo:

- Al incrementar la cantidad de procesadores pero manteniendo constante el tamaño de la malla fina, el número de iteraciones se mantiene aproximadamente constante. Esto significa que estamos resolviendo con mayor precisión el problema.
- Al incrementar la cantidad de procesadores y en la misma razón disminuir el tamaño de la malla fina, el número de iteraciones también se mantiene aproximadamente constante. Esto significa que podemos disminuir el tiempo de cálculo incrementando la cantidad de procesadores.
- Los experimentos numéricos permiten detectar que el factor de penalización modifica el condicionamiento del problema, consecuentemente se debe realizar un ajuste cuidadoso para minimizar la cantidad de iteraciones y tener una solución equivalente a la del primer algoritmo.

El incremento de la cantidad de elementos del dominio espacial disminuye la cantidad de iteraciones de los algoritmos mencionados.

5.1.1. Trabajos Futuros

- Aplicar el algoritmo a otros problemas de la ciencias e ingeniería.
- Utilizar el algoritmo Parareal como preconditionador para métodos iterativos aplicados a sistemas no lineales.
- Analizar las causas de las oscilaciones del algoritmo paralelo.
- Implementar el algoritmo en una plataforma computacional distribuida.

APÉNDICE A

APÉNDICE: PROBLEMA DE CONTROL

A.1. Condiciones de optimalidad KKT del Lagrangiano

La ecuación (2.8) es la condición de optimalidad de KKT:

$$\delta \mathcal{L} = \left. \frac{\partial \mathcal{L}}{\partial \underline{v}} \right|_{(\underline{z}^*, \underline{v}^*)} \delta \underline{v} + \left. \frac{\partial \mathcal{L}}{\partial \underline{p}} \right|_{(\underline{z}^*, \underline{v}^*)} \delta \underline{p} + \left. \frac{\partial \mathcal{L}}{\partial \underline{z}} \right|_{(\underline{z}^*, \underline{v}^*)} \delta \underline{z} + \left. \frac{\partial \mathcal{L}}{\partial \underline{\eta}_o} \right|_{(\underline{z}^*, \underline{v}^*)} \delta \underline{\eta}_o = 0.$$

Como las variaciones respecto a cada variable no son nulas, significa que para hacer $\delta \mathcal{L} = 0$, se debe cumplir que:

$$\left. \frac{\partial \mathcal{L}}{\partial \underline{p}} \right|_{(\underline{z}^*, \underline{v}^*)} = 0, \quad \left. \frac{\partial \mathcal{L}}{\partial \underline{z}} \right|_{(\underline{z}^*, \underline{v}^*)} = 0, \quad \left. \frac{\partial \mathcal{L}}{\partial \underline{v}} \right|_{(\underline{z}^*, \underline{v}^*)} = 0, \quad \left. \frac{\partial \mathcal{L}}{\partial \underline{\eta}_o} \right|_{(\underline{z}^*, \underline{v}^*)} = 0. \quad (\text{A.1})$$

Primeramente re-escribimos el Lagrangiano definido en (2.7):

$$\mathcal{L}(\underline{z}, \underline{v}, \underline{p}, \underline{\eta}_o) = \mathcal{J}_h(\underline{z}, \underline{v}) + \int_{t_o}^{t_f} \underline{p}^T (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) dt + \underline{\eta}_o^T (\underline{z}(t_o) - \underline{z}_o^h), \quad (\text{A.2})$$

A continuación calculamos cada una de las derivadas de las expresiones (A.1), mediante la derivada de Gâteaux (definición 2.1).

Cálculo de $\frac{\partial \mathcal{L}}{\partial \underline{p}}$:

$$\frac{D\mathcal{L}}{D\underline{p}} = \lim_{\epsilon \rightarrow 0} \left(\frac{\mathcal{L}(\underline{p} + \epsilon \phi, \underline{z}, \underline{v}, \underline{\eta}_o) - \mathcal{L}(\underline{p}, \underline{z}, \underline{v}, \underline{\eta}_o)}{\epsilon} \right) = \left. \frac{d\mathcal{L}(\underline{p} + \epsilon \phi, \underline{z}, \underline{v}, \underline{\eta}_o)}{d\epsilon} \right|_{\epsilon=0} \quad (\text{A.3})$$

Utilizando la definición del Lagrangiano (A.2), (A.3) toma la siguiente forma:

$$\begin{aligned} \frac{D\mathcal{L}}{D\underline{p}} &= \left. \frac{d\mathcal{J}_h(\underline{z}, \underline{v}) + \int_{t_o}^{t_f} (\underline{p} + \epsilon \phi)^T (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) dt + \eta_0^T (\underline{z}(t_o) - \underline{z}_o^h)}{d\epsilon} \right|_{\epsilon=0}, \\ &= \left. \frac{d\mathcal{J}_h(\underline{z}, \underline{v})}{d\epsilon} \right|_{\epsilon=0} + \left. \frac{d \int_{t_o}^{t_f} (\underline{p} + \epsilon \phi)^T (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) dt}{d\epsilon} \right|_{\epsilon=0} \\ &\quad + \left. \frac{d\eta_0^T (\underline{z}(t_o) - \underline{z}_o^h)}{d\epsilon} \right|_{\epsilon=0}, \end{aligned} \quad (\text{A.4})$$

luego $\left. \frac{d\mathcal{J}_h(\underline{z}, \underline{v})}{d\epsilon} \right|_{\epsilon=0}$ y $\left. \frac{d\eta_0^T (\underline{z}(t_o) - \underline{z}_o^h)}{d\epsilon} \right|_{\epsilon=0}$ son nulos porque no dependen de ϵ . Consecuentemente (A.4) queda:

$$\begin{aligned} \frac{D\mathcal{L}}{D\underline{p}} &= \left. \frac{d \int_{t_o}^{t_f} (\underline{p} + \epsilon \phi)^T (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) dt}{d\epsilon} \right|_{\epsilon=0}, \\ &= \left. \frac{d \int_{t_o}^{t_f} \underline{p}^T (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) dt}{d\epsilon} \right|_{\epsilon=0} + \left. \frac{d \int_{t_o}^{t_f} \epsilon (\phi)^T (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) dt}{d\epsilon} \right|_{\epsilon=0} \end{aligned} \quad (\text{A.5})$$

donde la expresión $\left. \frac{d \int_{t_o}^{t_f} (\underline{p})^T (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) dt}{d\epsilon} \right|_{\epsilon=0}$ también es nula porque no depende de ϵ , es decir (A.5) queda:

$$\begin{aligned} \frac{D\mathcal{L}}{D\underline{p}} &= \left. \int_{t_o}^{t_f} \phi^T (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) dt \right|_{\epsilon=0}, \\ &= \int_{t_o}^{t_f} \phi^T (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) dt, \\ &= \int_{t_o}^{t_f} (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t))^T \phi dt, \\ &= (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t), \phi), \end{aligned} \quad (\text{A.6})$$

luego, aplicamos la condición de optimalidad $\left. \frac{\partial \mathcal{L}}{\partial \underline{p}} \right|_{(\underline{z}^*, \underline{v}^*)} = 0$ a la ecuación (A.6), y tenemos:

$$\left. \frac{\partial \mathcal{L}}{\partial \underline{p}} \right|_{(\underline{z}^*, \underline{v}^*)} = \dot{\underline{z}} - f(\underline{z}, \underline{v}, t)|_{(\underline{z}^*, \underline{v}^*)} = 0, \quad \forall \phi \neq 0. \quad (\text{A.7})$$

Cálculo de $\frac{\partial \mathcal{L}}{\partial \underline{z}}$:

$$\frac{D\mathcal{L}}{D\underline{z}} = \lim_{\epsilon \rightarrow 0} \left(\frac{\mathcal{L}(\underline{p}, \underline{z} + \epsilon \phi, \underline{v}, \underline{\eta}_o) - \mathcal{L}(\underline{p}, \underline{z}, \underline{v}, \underline{\eta}_o)}{\epsilon} \right) = \left. \frac{d\mathcal{L}(\underline{p}, \underline{z} + \epsilon \phi, \underline{v}, \underline{\eta}_o)}{d\epsilon} \right|_{\epsilon=0} \quad (\text{A.8})$$

Utilizando la definición del Lagrangiano (A.2) y procediendo de la misma manera como hicimos al calcular $\frac{\partial \mathcal{L}}{\partial \underline{p}}$, (A.8) toma la forma siguiente:

$$\begin{aligned} \frac{D\mathcal{L}}{D\underline{z}} &= \left. \frac{d \left(\mathcal{J}_h(\underline{z} + \epsilon \phi, \underline{v}) + \int_{t_o}^{t_f} \underline{p}^T \left(\frac{d(\underline{z} + \epsilon \phi)}{dt} - f(\underline{z} + \epsilon \phi, \underline{v}, t) \right) dt \right)}{d\epsilon} \right|_{\epsilon=0} \\ &\quad + \left. \frac{d \left(\underline{\eta}_0^T (\underline{z}(t_o) - \underline{z}_o^h) \right)}{d\epsilon} \right|_{\epsilon=0}, \\ \frac{D\mathcal{L}}{D\underline{z}} &= \left. \frac{d\mathcal{J}_h(\underline{z} + \epsilon \phi, \underline{v})}{d\epsilon} \right|_{\epsilon=0} + \left. \frac{d \int_{t_o}^{t_f} \underline{p}^T \left(\frac{d(\underline{z} + \epsilon \phi)}{dt} - f(\underline{z} + \epsilon \phi, \underline{v}, t) \right) dt}{d\epsilon} \right|_{\epsilon=0}. \quad (\text{A.9}) \end{aligned}$$

Luego, si expandimos en series de Taylor los términos de (A.9), el primer término es igual a:

$$\mathcal{J}_h(\underline{z} + \epsilon \phi, \underline{v}) = \mathcal{J}_h(\underline{z}, \underline{v}) + \left(\left. \frac{\partial \mathcal{J}_h}{\partial \underline{z}} \right|_{(\underline{z}, \underline{v})} \right)^T \epsilon \phi + O(\epsilon^2), \quad (\text{A.10})$$

derivando tenemos:

$$\left. \frac{d\mathcal{J}_h(\underline{z} + \epsilon \phi, \underline{v})}{d\epsilon} \right|_{\epsilon=0} = \left(\left. \frac{\partial \mathcal{J}_h}{\partial \underline{z}} \right|_{(\underline{z}, \underline{v})} \right)^T \phi. \quad (\text{A.11})$$

El segundo término de (A.9) es igual a:

$$\begin{aligned} \int_{t_o}^{t_f} \underline{p}^T \left(\frac{d(\underline{z} + \epsilon \phi)}{dt} - f(\underline{z} + \epsilon \phi, \underline{v}, t) \right) dt &= \int_{t_o}^{t_f} \underline{p}^T (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) dt \\ &\quad + \left(\left. \frac{\partial \left(\int_{t_o}^{t_f} \underline{p}^T (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) dt \right)}{\partial \underline{z}} \right|_{(\underline{z}, \underline{v})} \right)^T \epsilon \phi + O(\epsilon^2), \quad (\text{A.12}) \end{aligned}$$

derivando tenemos:

$$\begin{aligned} \frac{d \int_{t_o}^{t_f} \underline{p}^T \left(\frac{d(\underline{z} + \epsilon \phi) - f(\underline{z} + \epsilon \phi, \underline{v}, t)}{dt} \right) dt}{d\epsilon} \Big|_{\epsilon=0} &= \left(\frac{\partial \left(\int_{t_o}^{t_f} \underline{p}^T \dot{\underline{z}} dt \right)}{\partial \underline{z}} \Big|_{(\underline{z}, \underline{v})} \right)^T \phi \\ &- \left(\frac{\partial \left(\int_{t_o}^{t_f} \underline{p}^T f(\underline{z}, \underline{v}, t) dt \right)}{\partial \underline{z}} \Big|_{(\underline{z}, \underline{v})} \right)^T \phi, \end{aligned} \quad (\text{A.13})$$

Luego si reemplazamos (A.11) y (A.13) en (A.9) tenemos:

$$\frac{D\mathcal{L}}{D\underline{z}} = \left(\frac{\partial \mathcal{J}_h}{\partial \underline{z}} \Big|_{(\underline{z}, \underline{v})} \right)^T \phi + \left(\frac{\partial \left(\int_{t_o}^{t_f} \underline{p}^T (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) dt \right)}{\partial \underline{z}} \Big|_{(\underline{z}, \underline{v})} \right)^T \phi, \quad (\text{A.14})$$

y aplicando la condición de optimalidad $\frac{\partial \mathcal{L}}{\partial \underline{z}} \Big|_{(\underline{z}^*, \underline{u}^*)} = 0$ a la ecuación (A.14), tenemos:

$$\frac{\partial \mathcal{L}}{\partial \underline{z}} \Big|_{(\underline{z}^*, \underline{u}^*)} = \left(\frac{\partial \mathcal{J}_h}{\partial \underline{z}} \Big|_{(\underline{z}^*, \underline{v}^*)} \right) + \left(\frac{\partial \left(\int_{t_o}^{t_f} \underline{p}^T (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) dt \right)}{\partial \underline{z}} \Big|_{(\underline{z}^*, \underline{v}^*)} \right) \Big|_{(\underline{z}^*, \underline{u}^*)} = 0. \quad (\text{A.15})$$

A la ecuación (A.15) se le denomina problema adjunto.

Cálculo de $\frac{\partial \mathcal{L}}{\partial \underline{v}}$:

$$\begin{aligned} \frac{D\mathcal{L}}{D\underline{v}} &= \lim_{\epsilon \rightarrow 0} \left(\frac{\mathcal{L}(\underline{p}, \underline{z}, \underline{v} + \epsilon \phi, \underline{\eta}_o) - \mathcal{L}(\underline{p}, \underline{z}, \underline{v}, \underline{\eta}_o)}{\epsilon} \right) \\ &= \frac{d \mathcal{L}(\underline{p}, \underline{z}, \underline{v} + \epsilon \phi, \underline{\eta}_o)}{d\epsilon} \Big|_{\epsilon=0}. \end{aligned} \quad (\text{A.16})$$

Utilizando la definición del Lagrangiano (A.2), expandiendo en series de Taylor los términos de (A.16) y procediendo de la misma manera como hicimos al calcular

$\frac{\partial \mathcal{L}}{\partial \underline{z}}$, (A.16) toma la siguiente forma:

$$\begin{aligned} \frac{D\mathcal{L}}{D\underline{v}} &= \frac{d \left(\mathcal{J}_h(\underline{z}, \underline{v} + \epsilon \phi) + \int_{t_o}^{t_f} \underline{p}^T (\dot{\underline{z}} - f(\underline{z}, \underline{v} + \epsilon \phi, t)) dt + \underline{\eta}_0^T (\underline{z}(t_o) - z_o^h) \right)}{d\epsilon} \Bigg|_{\epsilon=0}, \\ &= \frac{d\mathcal{J}_h(\underline{z}, \underline{v} + \epsilon \phi)}{d\epsilon} \Bigg|_{\epsilon=0} + \frac{d \int_{t_o}^{t_f} \underline{p}^T (\dot{\underline{z}} - f(\underline{z}, \underline{v} + \epsilon \phi, t)) dt}{d\epsilon} \Bigg|_{\epsilon=0}. \\ &= \left(\frac{\partial \mathcal{J}_h}{\partial \underline{v}} \Bigg|_{(\underline{z}, \underline{v})} \right)^T \phi + \left(\frac{\partial \left(\int_{t_o}^{t_f} \underline{p}^T (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) dt \right)}{\partial \underline{v}} \Bigg|_{(\underline{z}, \underline{v})} \right)^T \phi. \end{aligned} \quad (\text{A.17})$$

La ecuación (A.16) se obtiene con el mismo razonamiento realizado para obtener el problema adjunto, y aplicando la condición de optimalidad $\frac{\partial \mathcal{L}}{\partial \underline{v}} \Bigg|_{(\underline{z}^*, \underline{v}^*)} = 0$ a la ecuación (A.16), tenemos:

$$\frac{\partial \mathcal{L}}{\partial \underline{v}} \Bigg|_{(\underline{z}^*, \underline{v}^*)} = \left(\frac{\partial \mathcal{J}_h}{\partial \underline{v}} \Bigg|_{(\underline{z}^*, \underline{v}^*)} \right) + \left(\frac{\partial \left(\int_{t_o}^{t_f} \underline{p}^T (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) dt \right)}{\partial \underline{v}} \Bigg|_{(\underline{z}^*, \underline{v}^*)} \right) = 0. \quad (\text{A.18})$$

A la ecuación (A.18) se le denomina condición de Euler.

Cálculo de $\frac{\partial \mathcal{L}}{\partial \underline{\eta}_0}$:

$$\frac{D\mathcal{L}}{D\underline{\eta}_0} = \lim_{\epsilon \rightarrow 0} \left(\frac{\mathcal{L}(\underline{p}, \underline{z}, \underline{v}, \underline{\eta}_0 + \epsilon \phi) - \mathcal{L}(\underline{p}, \underline{z}, \underline{v}, \underline{\eta}_0)}{\epsilon} \right) = \frac{d\mathcal{L}(\underline{p}, \underline{z}, \underline{v}, \underline{\eta}_0 + \epsilon \phi)}{d\epsilon} \Bigg|_{\epsilon=0} \quad (\text{A.19})$$

Utilizando la definición del Lagrangiano (A.2), (A.19) toma la siguiente forma:

$$\begin{aligned} \frac{D\mathcal{L}}{D\underline{\eta}_0} &= \frac{d \mathcal{J}_h(\underline{z}, \underline{v}) + \left(\int_{t_o}^{t_f} (\underline{p} + \epsilon \phi)^T (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) dt \right)}{d\epsilon} \Bigg|_{\epsilon=0} \\ &\quad + \frac{d \left((\underline{\eta}_0 + \epsilon \phi)^T (\underline{z}(t_o) - z_o^h) \right)}{d\epsilon} \Bigg|_{\epsilon=0}. \\ \frac{D\mathcal{L}}{D\underline{\eta}_0} &= \phi^T (\underline{z}(t_o) - z_o^h) = (\underline{z}(t_o) - z_o^h)^T \phi. \end{aligned} \quad (\text{A.20})$$

Luego aplicando la condición de optimalidad $\frac{\partial \mathcal{L}}{\partial \underline{\eta}_0} \Bigg|_{(\underline{z}^*, \underline{v}^*)} = (\underline{z}(t_o) - z_o^h) = 0 \quad \forall \phi$,

Utilizando la definición (A.22) y procediendo de manera análoga como hicimos para obtener (A.5), tenemos:

$$\begin{aligned}
 \frac{D\mathcal{L}}{D\underline{p}} &= \left. \frac{d\mathcal{J}_h(\underline{z}, \underline{v}) + \int_{t_o}^{t_f} (\underline{p} + \epsilon \phi)^T (M_h \dot{\underline{z}} - A_h \underline{z} - B_h \underline{v} - c_h) dt}{d\epsilon} \right|_{\epsilon=0} \\
 &\quad + \left. \frac{d\left(\underline{\eta}_0^T (\underline{z}(t_o) - \underline{z}_o^h)\right)}{d\epsilon} \right|_{\epsilon=0}, \\
 &= \int_{t_o}^{t_f} \phi^T (M_h \dot{\underline{z}} - A_h \underline{z} - B_h \underline{v} - c_h) dt, \\
 \frac{D\mathcal{L}}{D\underline{p}} &= \int_{t_o}^{t_f} (M_h \dot{\underline{z}} - A_h \underline{z} - B_h \underline{v} - c_h)^T \phi dt, \tag{A.24}
 \end{aligned}$$

luego, aplicando la condición de optimalidad $\left. \frac{\partial \mathcal{L}}{\partial \underline{p}} \right|_{(\underline{z}^*, \underline{v}^*)} = 0$ a la ecuación (A.24)

tenemos:

$$\begin{aligned}
 \left. \frac{\partial \mathcal{L}}{\partial \underline{p}} \right|_{(\underline{z}^*, \underline{v}^*)} &= 0 \\
 \Rightarrow (M_h \dot{\underline{z}} - A_h \underline{z} - B_h \underline{v} - c_h)|_{(\underline{z}^*, \underline{v}^*)} &= 0, \text{ para } t_o \leq t \leq t_f. \tag{A.25}
 \end{aligned}$$

Problema adjunto:

Primeramente calculamos de $\frac{\partial \mathcal{L}}{\partial \underline{z}}$:

$$\begin{aligned}
 \frac{D\mathcal{L}}{D\underline{z}} &= \lim_{\epsilon \rightarrow 0} \left(\frac{\mathcal{L}(\underline{p}, \underline{z} + \epsilon \phi, \underline{v}, \underline{\eta}_o) - \mathcal{L}(\underline{p}, \underline{z}, \underline{v}, \underline{\eta}_o)}{\epsilon} \right), \\
 &= \left. \frac{d\mathcal{L}(\underline{p}, \underline{z} + \epsilon \phi, \underline{v}, \underline{\eta}_o)}{d\epsilon} \right|_{\epsilon=0}. \tag{A.26}
 \end{aligned}$$

Utilizando la definición del Lagrangiano (A.22), (A.26) toma la siguiente forma:

$$\begin{aligned}
 \frac{D\mathcal{L}}{D\underline{z}} &= \left. \frac{d\mathcal{J}_h(\underline{z} + \epsilon \phi, \underline{v})}{d\epsilon} \right|_{\epsilon=0} + \left. \frac{d \int_{t_o}^{t_f} \underline{p}^T \left(M_h \frac{d(\underline{z} + \epsilon \phi)}{dt} - A_h (\underline{z} + \epsilon \phi) \right) dt}{d\epsilon} \right|_{\epsilon=0} \\
 &\quad + \left. \frac{d \int_{t_o}^{t_f} \underline{p}^T (-B_h \underline{v} - c_h) dt}{d\epsilon} \right|_{\epsilon=0}. \tag{A.27}
 \end{aligned}$$

Utilizando la definición del funcional (A.22) el primer término de la expresión anterior (A.27) es equivalente a:

$$\left. \frac{d\mathcal{J}_h(\underline{z} + \epsilon \phi, \underline{v})}{d\epsilon} \right|_{\epsilon=0} = \left. \frac{d \left(\frac{1}{2} \int_{t_o}^{t_f} (r \|\underline{v}(t)\|_{R_h}^2 + q \|\underline{z} + \epsilon \phi - \tilde{\underline{y}}(t)\|_{M_h}^2) dt \right)}{d\epsilon} \right|_{\epsilon=0} + \left. \frac{d \left(\frac{1}{2} s \|\underline{z}(t_f) + \epsilon \phi(t_f) - \tilde{\underline{y}}(t_f)\|_{M_h}^2 \right)}{d\epsilon} \right|_{\epsilon=0}, \quad (\text{A.28})$$

donde $\left. \frac{d \left(\frac{1}{2} \int_{t_o}^{t_f} r \|\underline{v}(t)\|_{R_h}^2 dt \right)}{d\epsilon} \right|_{\epsilon=0}$ es nulo porque no depende de ϵ , entonces A.28 toma la siguiente forma:

$$\left. \frac{d\mathcal{J}_h(\underline{z} + \epsilon \phi, \underline{v})}{d\epsilon} \right|_{\epsilon=0} = \frac{1}{2} \int_{t_o}^{t_f} q \phi^T M_h (\underline{z} - \tilde{\underline{y}}) + q (\underline{z} - \tilde{\underline{y}})^T M_h \phi dt + \frac{1}{2} s (\phi(t_f)^T M_h (\underline{z}(t_f) - \tilde{\underline{y}}(t_f))) + \frac{1}{2} s \left((\underline{z}(t_f) - \tilde{\underline{y}}(t_f))^T M_h \phi(t_f) \right), \quad (\text{A.29})$$

$$(\text{A.30})$$

además teniendo en cuenta que $\phi^T M_h (\underline{z} - \tilde{\underline{y}}) = (\underline{z} - \tilde{\underline{y}})^T M_h \phi$ y

$\phi(t_f)^T M_h (\underline{z}(t_f) - \tilde{\underline{y}}(t_f)) = (\underline{z}(t_f) - \tilde{\underline{y}}(t_f))^T M_h \phi(t_f)$, (A.29) se puede expresar como

sigue:

$$\left. \frac{d\mathcal{J}_h(\underline{z} + \epsilon \phi, \underline{v})}{d\epsilon} \right|_{\epsilon=0} = \int_{t_o}^{t_f} q \phi^T M_h (\underline{z} - \tilde{\underline{y}}) dt + s (\phi(t_f)^T M_h (\underline{z}(t_f) - \tilde{\underline{y}}(t_f))), \quad (\text{A.31})$$

El segundo término y el tercer término de (A.27) pueden ser simplificados teniendo en cuenta que los términos que no dependen de ϵ son nulos, entonces

tenemos:

$$\frac{d \int_{t_o}^{t_f} \underline{p}^T \left(M_h \frac{d(\underline{z} + \epsilon \phi)}{dt} - A_h (\underline{z} + \epsilon \phi) - B_h \underline{v} - c_h \right) dt}{d\epsilon} \Bigg|_{\epsilon=0} = \frac{d \int_{t_o}^{t_f} \underline{p}^T \left(M_h \frac{d(\underline{z} + \epsilon \phi)}{dt} - A_h (\underline{z} + \epsilon \phi) \right) dt}{d\epsilon} \Bigg|_{\epsilon=0}, \quad (\text{A.32})$$

donde el primer término de ésta expresión puede ser simplificado integrando por partes:

$$\int_{t_o}^{t_f} \underline{p}^T M_h \frac{d(\underline{z} + \epsilon \phi)}{dt} dt = - \int_{t_o}^{t_f} \dot{\underline{p}}^T M_h (\underline{z} + \epsilon \phi) dt + \underline{p}^T(t_f) M_h (\underline{z} + \epsilon \phi)(t_f) - \underline{p}^T(t_o) M_h (\underline{z} + \epsilon \phi)(t_o), \quad (\text{A.33})$$

derivando (A.33) con respecto a ϵ y luego evaluando en $\epsilon = 0$ tenemos:

$$\begin{aligned} \frac{d \int_{t_o}^{t_f} \underline{p}^T M_h \frac{d(\underline{z} + \epsilon \phi)}{dt} dt}{d\epsilon} \Bigg|_{\epsilon=0} &= - \frac{d \int_{t_o}^{t_f} \dot{\underline{p}}^T M_h (\underline{z} + \epsilon \phi) dt}{d\epsilon} \Bigg|_{\epsilon=0} \\ &+ \frac{d \underline{p}^T(t_f) M_h (\underline{z} + \epsilon \phi)(t_f)}{d\epsilon} \Bigg|_{\epsilon=0} \\ &- \frac{d \underline{p}^T(t_o) M_h (\underline{z} + \epsilon \phi)(t_o)}{d\epsilon} \Bigg|_{\epsilon=0}, \\ \frac{d \int_{t_o}^{t_f} \underline{p}^T M_h \frac{d(\underline{z} + \epsilon \phi)}{dt} dt}{d\epsilon} \Bigg|_{\epsilon=0} &= - \int_{t_o}^{t_f} \dot{\underline{p}}^T M_h \phi dt + \underline{p}^T(t_f) M_h \phi(t_f) \\ &- \underline{p}^T(t_o) M_h \phi(t_o), \end{aligned} \quad (\text{A.34})$$

además como $\phi(t_o) = 0$,

$$\frac{d \int_{t_o}^{t_f} \underline{p}^T M_h \frac{d(\underline{z} + \epsilon \phi)}{dt} dt}{d\epsilon} \Bigg|_{\epsilon=0} = - \int_{t_o}^{t_f} \dot{\underline{p}}^T M_h \phi dt + \underline{p}^T(t_f) M_h \phi(t_f), \quad (\text{A.35})$$

y reemplazando (A.35) en (A.32) tenemos:

$$\frac{d \int_{t_o}^{t_f} \underline{p}^T \left(M_h \frac{d(\underline{z} + \epsilon \phi)}{dt} - A_h (\underline{z} + \epsilon \phi) - B_h \underline{v} - c_h \right) dt}{d\epsilon} \Bigg|_{\epsilon=0} = - \int_{t_o}^{t_f} (\underline{\dot{p}}^T M_h + \underline{p}^T A_h) \phi dt + \underline{p}^T(t_f) M_h \phi(t_f). \quad (\text{A.36})$$

Teniendo en cuenta que $(\underline{\dot{p}}^T M_h + \underline{p}^T A_h) \phi = (\phi)^T (M_h \underline{\dot{p}} + A_h \underline{p})$ y

$\underline{p}^T(t_f) M_h \phi(t_f) = (\phi(t_f))^T M_h \underline{p}(t_f)$, (A.36) se puede expresar como sigue:

$$\frac{d \int_{t_o}^{t_f} \underline{p}^T \left(M_h \frac{d(\underline{z} + \epsilon \phi)}{dt} - A_h (\underline{z} + \epsilon \phi) - B_h \underline{v} - c_h \right) dt}{d\epsilon} \Bigg|_{\epsilon=0} = - \int_{t_o}^{t_f} (\phi)^T (M_h \underline{\dot{p}} + A_h \underline{p}) dt + (\phi(t_f))^T M_h \underline{p}(t_f). \quad (\text{A.37})$$

Finalmente reemplazamos (A.31) y (A.37) en (A.27):

$$\frac{D\mathcal{L}}{D\underline{z}} = \int_{t_o}^{t_f} (\phi)^T (-M_h \underline{\dot{p}} - A_h \underline{p} + q M_h (\underline{z} - \underline{\tilde{y}})) dt + (\phi(t_f))^T (M_h \underline{p}(t_f) + M_h s (\underline{z}(t_f) - \underline{\tilde{y}}(t_f))), \quad (\text{A.38})$$

luego aplicamos la condición de optimalidad $\frac{\partial \mathcal{L}}{\partial \underline{z}} \Big|_{(\underline{z}^*, \underline{u}^*)} = 0$ a la ecuación (A.38), y

tenemos:

$$M_h^T \underline{\dot{p}}^* = -A_h^T \underline{p}^* + q M_h (\underline{z}^* - \underline{\tilde{y}}), \quad \text{para } t_o \leq t \leq t_f, \\ M_h^T \underline{p}^*(t_f) = -s M_h (\underline{z}^*(t_f) - \underline{\tilde{y}}(t_f)). \quad (\text{A.39})$$

Nótese que si M_h y A_h son simétricas el problema adjunto se resume a:

$$M_h \underline{\dot{p}}^* = -A_h \underline{p}^* + q M_h (\underline{z}^* - \underline{\tilde{y}}), \quad \text{para } t_o \leq t \leq t_f, \\ \underline{p}^*(t_f) = -s (\underline{z}^*(t_f) - \underline{\tilde{y}}(t_f)). \quad (\text{A.40})$$

Condición de Euler

A continuación obtenemos la condición de optimalidad (A.18) que se denomina condición de Euler para la planta lineal (2.13). Entonces primeramente calculamos $\frac{\partial \mathcal{L}}{\partial \underline{v}}$:

$$\begin{aligned} \frac{D\mathcal{L}}{D\underline{v}} &= \lim_{\epsilon \rightarrow 0} \left(\frac{\mathcal{L}(\underline{p}, \underline{z}, \underline{v} + \epsilon \phi, \underline{\eta}_o) - \mathcal{L}(\underline{p}, \underline{z}, \underline{v}, \underline{\eta}_o)}{\epsilon} \right), \\ &= \left. \frac{d\mathcal{L}(\underline{p}, \underline{z}, \underline{v} + \epsilon \phi, \underline{\eta}_o)}{d\epsilon} \right|_{\epsilon=0}. \end{aligned} \quad (\text{A.41})$$

Utilizando la definición del Lagrangiano (A.22), (A.41) toma la siguiente forma:

$$\begin{aligned} \frac{D\mathcal{L}}{D\underline{v}} &= \left. \frac{d \left(\mathcal{J}_h(\underline{z}, \underline{v} + \epsilon \phi) + \int_{t_o}^{t_f} \underline{p}^T (M_h \dot{\underline{z}} - A_h \underline{z} - B_h (\underline{v} + \epsilon \phi) - c_h) dt \right)}{d\epsilon} \right|_{\epsilon=0} \\ &+ \left. \frac{d \left(\underline{\eta}_0^T (\underline{z}(t_o) - \underline{z}^h) \right)}{d\epsilon} \right|_{\epsilon=0}, \end{aligned} \quad (\text{A.42})$$

y eliminando los términos que no dependen de ϵ tenemos:

$$\frac{D\mathcal{L}}{D\underline{v}} = \left. \frac{d\mathcal{J}_h(\underline{z}, (\underline{v} + \epsilon \phi))}{d\epsilon} \right|_{\epsilon=0} - \int_{t_o}^{t_f} \underline{p}^T B_h \phi dt, \quad (\text{A.43})$$

además teniendo en cuenta que $\underline{p}^T B_h \phi = \phi^T B_h \underline{p}$, (A.43) se puede expresar como sigue:

$$\frac{D\mathcal{L}}{D\underline{v}} = \left. \frac{d\mathcal{J}_h(\underline{z}, (\underline{v} + \epsilon \phi))}{d\epsilon} \right|_{\epsilon=0} - \int_{t_o}^{t_f} \phi^T B_h \underline{p} dt. \quad (\text{A.44})$$

Utilizando la definición del funcional (A.22) el primer término de ésta expresión

puede ser simplificado como sigue:

$$\begin{aligned} \left. \frac{d\mathcal{J}_h(\underline{z}, (\underline{v} + \epsilon \phi))}{d\epsilon} \right|_{\epsilon=0} &= \left. \frac{d\frac{1}{2} \int_{t_o}^{t_f} r \|\underline{v} + \epsilon \phi\|_{R_h}^2 + q \|(z - \tilde{y})(t)\|_{M_h}^2 dt}{d\epsilon} \right|_{\epsilon=0} + \\ &+ \left. \frac{d\frac{1}{2} s \|\underline{z}(t_f) - \tilde{y}(t_f)\|_{M_h}^2}{d\epsilon} \right|_{\epsilon=0}, \\ \left. \frac{d\mathcal{J}_h(\underline{z}, (\underline{v} + \epsilon \phi))}{d\epsilon} \right|_{\epsilon=0} &= \frac{1}{2} \int_{t_o}^{t_f} (r \phi^T R_h \underline{v} + r \underline{v}^T R_h \phi) dt \end{aligned} \quad (\text{A.45})$$

donde los términos $\left. \frac{d\frac{1}{2} \int_{t_o}^{t_f} q \|(z - \tilde{y})(t)\|_{M_h}^2 dt}{d\epsilon} \right|_{\epsilon=0}$ y $\left. \frac{d\frac{1}{2} s \|\underline{z}(t_f) - \tilde{y}(t_f)\|_{M_h}^2}{d\epsilon} \right|_{\epsilon=0}$ son nulos porque no dependen de ϵ . Además teniendo en cuenta que $\phi^T R_h \underline{v} = \underline{v}^T R_h \phi$, (A.45) se puede expresar como sigue:

$$\left. \frac{d\mathcal{J}_h(\underline{z}, \underline{v} + \epsilon \phi)}{d\epsilon} \right|_{\epsilon=0} = \int_{t_o}^{t_f} r \phi^T R_h \underline{v} dt. \quad (\text{A.46})$$

Finalmente reemplazando (A.46) en (A.44):

$$\frac{D\mathcal{L}}{D\underline{v}} = \int_{t_o}^{t_f} \phi^T (r R_h \underline{v} - B_h^T \underline{p}) dt, \quad (\text{A.47})$$

Luego aplicamos la condición de optimalidad $\left. \frac{\partial \mathcal{L}}{\partial \underline{v}} \right|_{(\underline{z}^*, \underline{v}^*)} = 0$ a la ecuación (A.47), tenemos:

$$\left. \frac{\partial \mathcal{L}}{\partial \underline{v}} \right|_{(\underline{z}^*, \underline{v}^*)} = (r R_h \underline{v} - B_h^T \underline{p}) \Big|_{(\underline{z}^*, \underline{v}^*)} = 0 \quad (\text{A.48})$$

La condición de optimalidad $\left. \frac{\partial \mathcal{L}}{\partial \underline{\eta}_0} \right|_{(\underline{z}^*, \underline{v}^*)} = (\underline{z}(t_o) - z_o^h) = 0 \quad \forall \phi$, es equivalente a la obtenida en la sección anterior en (A.21):

$$\underline{z}(t_o) = z_o^h, \quad (\text{A.49})$$

A.2. Apéndice: Método de solución del problema de control óptimo

A.2.1. Gradiente conjugado para un problema cuadrático

La idea de conjugar las direcciones de búsqueda es motivada por el deseo de acelerar la convergencia del método de máximo descenso (las direcciones contrarias al gradiente del funcional) y evitando el alto costo computacional asociado al método de Newton (la dirección es la inversa del Hessiano por el gradiente del funcional). El método Gradiente conjugado es utilizado para minimizar funcionales de tipo cuadráticos, los cuales pueden ser escritos como sigue:

$$\text{minimizar } \mathcal{J} = \frac{1}{2}x^T Hx - b^T x, \quad (\text{A.50})$$

donde $H \in \mathbb{R}^{n \times n}$ (Definida positiva), $x, b \in \mathbb{R}^n$. Nótese que el problema de minimización tiene solución única y equivalente a la solución de la ecuación lineal:

$$Hx = b. \quad (\text{A.51})$$

Dado x^0 un punto inicial arbitrario, podemos obtener un mejor valor de x^{i+1} , es decir: $\mathcal{J}(x^{i+1}) < \mathcal{J}(x^i) < \mathcal{J}(x^0)$, lo cual logramos con el siguiente esquema de corrección:

$$x^{i+1} = x^i + \alpha^i d^i, \quad (\text{A.52})$$

donde α^i se determina de tal manera que tengamos un máximo descenso en la dirección d^i , es decir:

$$\begin{aligned} \frac{d\mathcal{J}(x^{i+1})}{d\alpha^i} &= 0, \\ &= \frac{d\mathcal{J}(x^i + \alpha^i d^i)}{d\alpha^i}, \\ &= \nabla \mathcal{J}^T(x^{i+1})d^i \end{aligned} \quad (\text{A.53})$$

donde $\nabla \mathcal{J}(x^{i+1}) = Hx^{i+1} - b$, la condición (A.53) puede expresarse como:

$$\begin{aligned} \nabla \mathcal{J}^T(x^{i+1})d^i &= (H(x^i + \alpha^i d^i) - b)^T d^i = 0, \\ (Hx^i - b)^T d^i + \alpha^i (Hd^i)^T d^i &= 0, \\ \nabla \mathcal{J}^T(x^i)d^i + \alpha^i (Hd^i)^T d^i &= 0, \end{aligned} \quad (\text{A.54})$$

luego:

$$\alpha^i = -\frac{\nabla \mathcal{J}^T(x^i)d^i}{(d^i)^T H d^i} \quad (\text{A.55})$$

A continuación presentamos un teorema que es necesario para construir el conjunto de direcciones de búsqueda:

Teorema A.1 (Vectores conjugados respecto a A). *Si $A \in \mathbb{R}^{n \times n}$ es definida positiva y tenemos un conjunto de n vectores conjugados respecto a A , entonces éstos vectores son linealmente independientes.*

Demostración. Supongamos que los vectores \underline{d}^i , $i = 0, 1, \dots, n$ son linealmente dependientes, entonces existen a^i , $i = 0, 1, \dots, n$ tales que:

$$a^0 \underline{d}^0 + a^1 \underline{d}^1 + \dots + a^n \underline{d}^n = 0$$

Multiplicando por A y tomando el producto interno con cada \underline{d}^i , tenemos:

$$a^i (\underline{d}^i, A \underline{d}^i) = 0, \quad i = 0, 1, \dots, n$$

Como la matriz A es definida positiva, $(\underline{d}^i, A \underline{d}^i) > 0$, tenemos

$$a^i = 0 \quad \forall i = 0, 1, \dots, n.$$

Por lo tanto los n vectores conjugados son linealmente independientes. \square

Utilizando la matriz H del problema (A.50) podemos definir n vectores conjugados respecto a H . Nótese que $\nabla^2 \mathcal{J} = H$, es decir H es la matriz Hessiana del funcional (A.50). El teorema A.1 permite afirmar que éstos n vectores son linealmente independientes, este hecho hace que cada x^{i+1} pueda expresarse como la combinación lineal de las direcciones de búsqueda y el vector inicial arbitrario x^0 , es decir:

$$x^{i+1} = x^0 + \sum_{j=0}^i \alpha^j \underline{d}^j. \quad (\text{A.56})$$

Como cada x^{i+1} se obtiene aplicando la condición de máximo descenso (A.53) tenemos:

$$\frac{d\mathcal{J}(x^{i+1})}{d\alpha^i} = \nabla \mathcal{J}^T(x^{i+1}) \underline{d}^i = 0, \quad \forall i = 0, 1, \dots, n, \quad (\text{A.57})$$

por otro lado:

$$\nabla \mathcal{J}^T(x^{n+1}) \underline{d}^i = (H x^{n+1} - b)^T \underline{d}^i, \quad \forall i = 0, 1, \dots, n, \quad (\text{A.58})$$

además de (A.56) tenemos:

$$x^{n+1} = x^{i+1} + \sum_{j=i+1}^n \alpha^j d^j, \quad \forall i = 0, 1, \dots, n-1, \quad (\text{A.59})$$

reemplazando (A.59) en (A.58) tenemos:

$$\nabla \mathcal{J}^T(x^{n+1})d^i = (x^{i+1} + \sum_{j=i+1}^n \alpha^j d^j)^T H d^i - b^T d^i, \quad \forall i = 0, 1, \dots, n-1, \quad (\text{A.60})$$

además como los vectores d^i son H -ortogonales y utilizando la ecuación (A.57)

tenemos:

$$\nabla \mathcal{J}^T(x^{n+1})d^i = \nabla \mathcal{J}^T(x^{i+1})d^i = 0, \quad \forall i = 0, 1, \dots, n, \quad (\text{A.61})$$

con esta expresión podemos concluir que **el gradiente del funcional en la iteración $i+1$ es ortogonal al espacio generado por las i direcciones anteriores.**

Además (A.61) es equivalente a:

$$\frac{d\mathcal{J}(x^0 + \sum_{i=0}^n \alpha^i \underline{d}^i)}{d\alpha^i} = 0, \quad \forall i = 0, 1, \dots, n. \quad (\text{A.62})$$

Esto significa que x^{n+1} minimiza el funcional (A.50) en el espacio vectorial generado por las d^0, d^2, \dots, d^n y x^0 . Nótese que el espacio generado por las n direcciones de búsqueda es \mathbb{R}^n , consecuentemente $\mathcal{J}(x^{n+1})$ es mínimo en \mathbb{R}^n , lo cual implica que x^{n+1} es la solución óptima.

Entonces el esquema de corrección de los valores x^i para el Método del gradiente conjugado se resume en el siguiente esquema iterativo:

Algorithm 1: Corrección de los x^i para el CG

Dado : $x^i, \nabla \mathcal{J}(x^i) \neq 0$

1 Calcular $\alpha^i = -\frac{\nabla \mathcal{J}^T(x^i)d^i}{(d^i)^T H d^i}$;

2 $x^{i+1} = x^i + \alpha^i d^i$;

En la siguiente sección veremos como aplicar el procedimiento de ortogonalización de Gram-Schmidt para generar las direcciones de búsqueda conjugadas.

A.2.2. Generación de direcciones conjugadas

Dado un conjunto de n vectores linealmente independientes $\{\xi^0, \xi^1, \dots, \xi^n\}$, y una matriz $A \in \mathbb{R}^{n \times n}$ definida positiva podemos construir un conjunto de direcciones conjugadas $\{d^0, d^1, \dots, d^n\}$, respecto a A usando un procedimiento de ortogonalización de Gram-Schmidt. Éste procedimiento garantiza que el espacio generado por los vectores ξ^i sea el mismo espacio generado por los vectores d^i para $i = 0, 1, \dots, n$ [27].

El primer paso del proceso de ortogonalización consiste en seleccionar la dirección inicial:

$$d^0 = \xi^0, \tag{A.63}$$

luego, supongamos ya hemos seleccionado $i < n$ vectores A -ortogonales,

$$\{d^0, d^1, \dots, d^i\}.$$

El esquema de selección de la siguiente dirección de búsqueda d^{i+1} es de la

forma:

$$d^{i+1} = \xi^{i+1} + \sum_{m=0}^i c^{(i+1)m} d^m, \quad (\text{A.64})$$

donde $c^{(i+1)m}$ son escalares que serán obtenidos a continuación. Para que la dirección d^{i+1} sea A -ortogonal a las direcciones anteriores, tenemos que:

$$(d^{i+1})^T A d^j = (\xi^{i+1})^T A d^j + \left(\sum_{m=0}^i c^{(i+1)m} d^m \right)^T A d^j = 0, \quad \forall j = 0, 1, \dots, i. \quad (\text{A.65})$$

Como los vectores $\{d^0, d^1, \dots, d^i\}$ ya son A -ortogonales, es decir:

$$(d^m)^T A d^j = 0 \quad \text{si } m \neq j, \quad (\text{A.66})$$

y reemplazando las ecuaciones (A.66) en (A.65) tenemos los escalares:

$$c^{(i+1)j} = -\frac{\xi^{i+1} A d^j}{(d^j)^T A d^j}, \quad \forall j = 0, 1, \dots, i, \quad (\text{A.67})$$

notemos que $(d^j)^T A d^j \neq 0 \quad \forall j = 0, 1, \dots, i$. Además la dirección $d^{i+1} \neq 0$ porque la ecuación (A.64) sería una combinación lineal de los vectores

$$\{\xi^0, \xi^1, \dots, \xi^n\},$$

lo cual no es posible porque los vectores ξ^i son linealmente independientes.

Entonces, el procedimiento de ortogonalización de Gram-Schmidt se resume en el siguiente algoritmo:

Algorithm 2: Proceso de Ortogonalización de Gram-Schmidt

Punto de partida: $d^0 = \xi^0$

1 $i = 0$;

2 **repeat**

3
$$d^{i+1} = \xi^{i+1} + \sum_{m=0}^i -\frac{\xi^{i+1} A d^m}{(d^m)^T A d^m} d^m ;$$

4 $i=i+1$;

5 **until** $i = n$;

Es importante notar que si los vectores $\{\xi^0, \xi^1 \dots, \xi^i\}$ son linealmente independientes y el vector ξ^{i+1} es linealmente dependiente a éstos, el algoritmo (2) aún es valido, pero la nueva dirección d^{i+1} dada será cero. Ésta propiedad podemos utilizar para construir el conjunto de direcciones conjugadas, sin importar que los vectores $\{\xi^0, \xi^1 \dots, \xi^k\}$ sean linealmente independientes. Ésta construcción realizamos mediante una leve modificación del algoritmo (2), donde cada d^{i+1} es obtenida en el paso 3. Si ésta nueva dirección es cero, no es adherida al conjunto $\{d^0, d^1, \dots, d^i\}$ de direcciones conjugadas previamente obtenidas [27].

Para obtener las direcciones de búsqueda del método del Gradiente Conjugado, utilizamos en procedimiento de Gram-Schmidt, donde construimos las direcciones de búsqueda conjugadas a partir de la dirección contraria al gradiente del funcional en cada iteración $g^i = \nabla \mathcal{J}(x^i)$, es decir hacemos:

$$\xi^i = -g^i, \quad \forall i = 0, 1, \dots, n, \quad (\text{A.68})$$

note que en el caso del funcional (A.50) $-g^i = b - Hx^i$.

Entonces las direcciones son dadas por:

$$\begin{aligned} d^0 &= -g^0, \\ d^{i+1} &= -g^{i+1} + \sum_{m=0}^i \frac{(g^{i+1})^T H d^m}{(d^m)^T H d^m} d^m. \end{aligned} \quad (\text{A.69})$$

Podemos ver que:

$$\begin{aligned} g^{m+1} &= Hx^{m+1} - b, \\ &= H(x^m + \alpha^m d^m) - b, \\ &= Hx^m - b + \alpha^m H d^m, \\ &= g^m + \alpha^m H d^m, \end{aligned} \quad (\text{A.70})$$

multiplicando por g^{i+1} tenemos:

$$\begin{aligned} (g^{i+1})^T g^{m+1} &= (g^{i+1})^T (g^m + \alpha^m H d^m), \\ (g^{i+1})^T H d^m &= \frac{(g^{i+1})^T g^{m+1} - (g^{i+1})^T g^m}{\alpha^m}. \end{aligned} \quad (\text{A.71})$$

Como en la ecuación (A.61) se concluyó que g^{i+1} es ortogonal al espacio generado por $\{d^1, d^2, \dots, d^i\}$. Además notemos que, debido al esquema de obtención de las direcciones de búsqueda (A.69), g^i pertenece al espacio generado por $\{d^1, d^2, \dots, d^i\}$. En consecuencia g^{i+1} es ortogonal a g^i y la ecuación (A.71) queda:

$$(g^{i+1})^T H d^m = \begin{cases} 0 & \forall m = 0, 1, \dots, i-1; \\ \frac{(g^{i+1})^T g^{i+1}}{\alpha^i} & m = i; \end{cases} \quad (\text{A.72})$$

Entonces los términos de la sumatoria de (A.69) son iguales a cero, excepto para $m = i$, por lo tanto:

$$\begin{aligned}d^0 &= -g^0, \\d^{i+1} &= -g^{i+1} + \beta^i d^i,\end{aligned}\tag{A.73}$$

donde

$$\beta^i = \frac{(g^{i+1})^T H d^i}{(d^i)^T H d^i}.\tag{A.74}$$

Nótese que para generar la nueva dirección de búsqueda, ya no se necesitan todas las direcciones anteriores, como en el caso del algoritmo (2). Lo cual implica menos costo computacional.

Una forma alternativa de calcular el valor del escalar β^i será obtenida a continuación. Primeramente reemplazamos (A.72) en (A.74):

$$\beta^i = \frac{(g^{i+1})^T g^{i+1}}{\alpha^i (d^i)^T H d^i},\tag{A.75}$$

Por otro lado si multiplicamos la ecuación (A.70) por d^i tenemos:

$$\begin{aligned}(d^i)^T g^{i+1} &= (d^i)^T (g^i + \alpha^i H d^i), \\ \alpha^i (d^i)^T H d^i &= (d^i)^T (g^{i+1} - g^i), \\ \alpha^i (d^i)^T H d^i &= -(d^i)^T (g^i),\end{aligned}\tag{A.76}$$

donde hemos utilizado la ecuación (A.61) para simplificar la expresión. Además

si reemplazamos la ecuación (A.73) en (A.76) tenemos:

$$\begin{aligned}\alpha^i (d^i)^T H d^i &= -(-g^i + \beta^{i-1} d^{i-1})^T (g^i), \\ &= (g^i)^T (g^i).\end{aligned}\tag{A.77}$$

Finalmente reemplazando (A.77) en (A.75) obtenemos la formula conocida como de Fletcher Reeves:

$$\beta^i = \frac{(g^{i+1})^T g^{i+1}}{(g^i)^T g^i}.\tag{A.78}$$

Entonces el algoritmo utilizado por el Método Gradiente Conjugado para obtener las direcciones de búsqueda conjugadas se resume a continuación:

Algorithm 3: Generación de direcciones conjugadas para el CG

Punto de partida: $d^0 = g^0$

1 $i = 0$;

2 **repeat**

3 $d^{i+1} = g^{i+1} + \frac{(g^{i+1})^T g^{i+1}}{(g^i)^T g^i} d^i$;

4 $i=i+1$;

5 **until** $i = n$;

A.2.3. Cálculo del incremento para la función de control

El valor $\underline{\alpha}^i$ debe estar dado de manera que el gradiente del funcional en $k + 1$ sea mínimo, para lo cual se debe cumplir que $\frac{d(\mathcal{J}_h|_{\underline{z}^{i+1}, \underline{v}^{i+1}})}{d\underline{\alpha}^i} = 0$. A continuación para obtener el valor de $\underline{\alpha}^i$, utilizamos la solución analítica de la ecuación de estado (2.13), que es:

$$\begin{aligned}\underline{z} &= e^{-M_h^{-1}A_h(t-t_o)} z_0^h + \int_{t_o}^t e^{-M_h^{-1}A_h(t-\tau)} M_h^{-1} (B_h \underline{v} + c_h) d\tau, \\ \underline{z} &= \mathcal{Z}(z_0^h, \underline{v}^{i+1}, c_h)\end{aligned}\tag{A.79}$$

donde $e^{-M_h^{-1}A_h t} \in \mathbb{R}^{\hat{q} \times \hat{q}} \times [t_o, t_f]$ es la solución homogénea del sistema de ecuaciones diferenciales ordinarias. En las siguientes ecuaciones utilizaremos $\underline{z}(z_0, \underline{v}, c_h)$ para referirnos a (A.79). Luego, si:

$$\underline{z}^{i+1} = \mathcal{Z}(z_0^h, \underline{v}^{i+1}, c_h),\tag{A.80}$$

donde $\underline{v}^{i+1} = \underline{v}^i + \underline{\alpha}^i \underline{d}^i$

$$\underline{z}^{i+1} = \mathcal{Z}(z_0^h, \underline{v}^i + \underline{\alpha}^i \underline{d}^i, c_h)\tag{A.81}$$

Primeramente evaluamos el funcional en \underline{v}^{i+1} y \underline{z}^{i+1}

$$\begin{aligned}\mathcal{J}_h|_{\underline{z}^{i+1}, \underline{v}^{i+1}} &= \frac{1}{2} \int_{t_o}^{t_f} (r \|\underline{v}^{i+1}\|_{R_h}^2 + q \|\underline{z}^{i+1} - \tilde{y}(t)\|_{M_h}^2) dt \\ &+ \frac{s}{2} \|\underline{z}^{i+1}(t_f) - \tilde{y}(t_f)\|_{M_h}^2,\end{aligned}\tag{A.82}$$

luego reemplazamos (A.81) en (A.82) tenemos $\mathcal{J}_h|_{\underline{z}^{i+1}, \underline{v}^{i+1}}$

$$\begin{aligned}\mathcal{J}_h|_{\underline{z}^{i+1}, \underline{v}^{i+1}} &= \frac{1}{2} \int_{t_o}^{t_f} (r \|\underline{v}^i + \underline{\alpha}^i \underline{d}^i\|_{R_h}^2 + q \|\mathcal{Z}(z_0^h, \underline{v}^i + \underline{\alpha}^i \underline{d}^i, c_h) - \tilde{y}\|_{M_h}^2) dt \\ &+ \frac{s}{2} \|\mathcal{Z}(z_0^h, \underline{v}^i + \underline{\alpha}^i \underline{d}^i, c_h)(t_f) - \tilde{y}(t_f)\|_{M_h}^2\end{aligned}\tag{A.83}$$

Derivando (A.83) respecto de $\underline{\alpha}^i$, tenemos:

$$\begin{aligned}
 \frac{d \mathcal{J}_h|_{\underline{z}^{i+1}, \underline{v}^{i+1}}}{d \underline{\alpha}^i} &= \int_{t_o}^{t_f} \frac{1}{2} r (\underline{v}^i + \underline{\alpha}^i \underline{d}^i)^T R_h \underline{d}^i + \frac{1}{2} r (\underline{d}^i)^T R_h (\underline{v}^i + \underline{\alpha}^i \underline{d}^i) dt \\
 &+ \int_{t_o}^{t_f} \frac{1}{2} q (\mathcal{Z}(z_0^h, \underline{v}^i + \underline{\alpha}^i \underline{d}^i, c_h) - \underline{\tilde{y}})^T M_h (\mathcal{Z}(0, \underline{d}^i, 0)) dt \\
 &+ \int_{t_o}^{t_f} \frac{1}{2} q (\mathcal{Z}(0, \underline{d}^i, 0))^T M_h (\mathcal{Z}(z_0^h, \underline{v}^i + \underline{\alpha}^i \underline{d}^i, c_h) - \underline{\tilde{y}}) dt \\
 &+ \frac{1}{2} s (\mathcal{Z}(z_0^h, \underline{v}^i + \underline{\alpha}^i \underline{d}^i, c_h)(t_f) - \underline{\tilde{y}}(t_f))^T M_h (\mathcal{Z}(0, \underline{d}^i, 0)(t_f)) \\
 &+ \frac{1}{2} s (\mathcal{Z}(0, \underline{d}^i, 0)(t_f))^T M_h (\mathcal{Z}(z_0^h, \underline{v}^i + \underline{\alpha}^i \underline{d}^i, c_h)(t_f) - \underline{\tilde{y}}(t_f)),
 \end{aligned}
 \tag{A.84}$$

luego simplificamos la expresión (A.84) utilizando la siguiente propiedad del producto de matrices y vectores $w^T H u = u^T H w$ donde $u, w \in \mathbb{R}^n$ y

$$H = H^T \in \mathbb{R}^n \times \mathbb{R}^n$$

$$\begin{aligned}
 \frac{d \mathcal{J}_h|_{\underline{z}^{i+1}, \underline{v}^{i+1}}}{d \underline{\alpha}^i} &= \int_{t_o}^{t_f} r (\underline{v}^i + \underline{\alpha}^i \underline{d}^i)^T R_h \underline{d}^i dt \\
 &+ \int_{t_o}^{t_f} q (\mathcal{Z}(z_0^h, \underline{v}^i + \underline{\alpha}^i \underline{d}^i, c_h) - \underline{\tilde{y}})^T M_h (\mathcal{Z}(0, \underline{d}^i, 0)) dt \\
 &+ s (\mathcal{Z}(z_0^h, \underline{v}^i + \underline{\alpha}^i \underline{d}^i, c_h)(t_f) - \underline{\tilde{y}}(t_f))^T M_h \mathcal{Z}(0, \underline{d}^i, 0)(t_f),
 \end{aligned}
 \tag{A.85}$$

y aplicando la propiedad distributiva del producto escalar tenemos:

$$\begin{aligned}
 \frac{d \mathcal{J}_h|_{\underline{z}^{i+1}, \underline{v}^{i+1}}}{d \underline{\alpha}^i} &= \int_{t_o}^{t_f} r (\underline{v}^i)^T R_h \underline{d}^i + (\underline{\alpha}^i \underline{d}^i)^T R_h \underline{d}^i dt \\
 &+ \int_{t_o}^{t_f} \underline{\alpha}^i q (\mathcal{Z}(0, \underline{d}^i, 0))^T M_h (\mathcal{Z}(0, \underline{d}^i, 0)) dt \\
 &+ q \int_{t_o}^{t_f} (\mathcal{Z}(z_0^h, \underline{v}^i, c_h) - \underline{\tilde{y}})^T M_h (\mathcal{Z}(0, \underline{d}^i, 0)) dt \\
 &+ s (\mathcal{Z}(z_0^h, \underline{v}^i, c_h)(t_f) - \underline{\tilde{y}}(t_f))^T M_h (\mathcal{Z}(0, \underline{d}^i, 0)(t_f)) \\
 &+ \underline{\alpha}^i s (\mathcal{Z}(0, \underline{d}^i, 0)(t_f))^T M_h (\mathcal{Z}(0, \underline{d}^i, 0)(t_f)). \quad (\text{A.86})
 \end{aligned}$$

Finalmente aplicamos la condición de máximo descenso (2.20) y tenemos:

$$\begin{aligned}
 \frac{d \mathcal{J}_h|_{\underline{z}^{i+1}, \underline{v}^{i+1}}}{d \underline{\alpha}^i} &= \underline{\alpha}^i \left(\int_{t_o}^{t_f} r \|\underline{d}^i\|_{R_h}^2 + q \|\mathcal{Z}(0, \underline{d}^i, 0)\|_{M_h}^2 dt + s \|\mathcal{Z}(0, \underline{d}^i, 0)(t_f)\|_{M_h}^2 \right) \\
 &+ \int_{t_o}^{t_f} r (\underline{d}^i)^T R_h \underline{v}^i + q (\mathcal{Z}(z_0^h, \underline{v}^i, c_h) - \underline{\tilde{y}})^T M_h \mathcal{Z}(0, \underline{d}^i, 0) dt \\
 &+ (\mathcal{Z}(z_0^h, \underline{v}^i, c_h)(t_f) - \underline{\tilde{y}}(t_f))^T M_h \mathcal{Z}(0, \underline{d}^i, 0)(t_f) = 0, \quad (\text{A.87})
 \end{aligned}$$

de ésta expresión podemos obtener el valor $\underline{\alpha}^i$ como sigue:

$$\begin{aligned}
 \underline{\alpha}^i &= \frac{\int_{t_o}^{t_f} r (\underline{d}^i)^T R_h \underline{v}^i + q (\mathcal{Z}(z_0^h, \underline{v}^i, c_h) - \underline{\tilde{y}})^T M_h \mathcal{Z}(0, \underline{d}^i, 0) dt}{\int_{t_o}^{t_f} r \|\underline{d}^k\|_{R_h}^2 + q \|\mathcal{Z}(0, \underline{d}^i, 0)\|_{M_h}^2 dt + s \|\mathcal{Z}(0, \underline{d}^i, 0)\|_{M_h}^2} \\
 &+ \frac{s (\mathcal{Z}(z_0^h, \underline{v}^i, c_h) - \underline{\tilde{y}})^T M_h (\mathcal{Z}(0, \underline{d}^i, 0)) \Big|_{(t_f)}}{\int_{t_o}^{t_f} r \|\underline{d}^k\|_{R_h}^2 + q \|\mathcal{Z}(0, \underline{d}^i, 0)\|_{M_h}^2 dt + s \|\mathcal{Z}(0, \underline{d}^i, 0)\|_{M_h}^2} \quad (\text{A.88})
 \end{aligned}$$

A.2.4. Gradiente del Lagrangiano y del funcional costo

Asumiendo que tenemos una función de control arbitraria \underline{v}^i , podemos obtener el estado de la planta \underline{z}^i resolviendo las ecuaciones de estado (A.25) y con

las condiciones iniciales (A.49):

$$\begin{aligned} M_h \dot{\underline{z}} &= A_h \underline{z} + B_h \underline{v} + c_h, \quad \text{para } t_o \leq t \leq t_f, \\ \underline{z}(t_o) &= z_o^h. \end{aligned} \quad (\text{A.89})$$

Luego podemos obtener el estado adjunto \underline{p}^i resolviendo la ecuación adjunta (A.39):

$$\begin{aligned} M_h^T \dot{\underline{p}} &= -A_h^T \underline{p} + q M_h (\underline{z} - \tilde{\underline{y}}), \quad \text{para } t_o \leq t \leq t_f, \\ M_h^T \underline{p}(t_f) &= -s M_h (\underline{z}(t_f) - \tilde{\underline{y}}(t_f)). \end{aligned} \quad (\text{A.90})$$

Consecuentemente las variaciones de primer orden del Lagrangiano (2.7), $\frac{\partial \mathcal{L}_h}{\partial \underline{p}}$, $\frac{\partial \mathcal{L}_h}{\partial \underline{z}}$ y $\frac{\partial \mathcal{L}_h}{\partial \underline{n}_0}$ son nulas como puede verse en (2.10) y (2.11). Entonces la variación total del Lagrangiano es:

$$\delta \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \underline{v}} \delta \underline{v} \quad (\text{A.91})$$

A continuación calculamos la derivada total del funcional (2.6) con respecto a la función de control a partir de \underline{v}^i , \underline{z}^i y \underline{p}^i .

$$\frac{D \mathcal{J}_h}{D \underline{v}} \Big|_{\underline{z}^i, \underline{v}^i} = \frac{D \mathcal{J}_h}{D \underline{v}} \Big|_{\underline{z}^i, \underline{v}^i} + \frac{D \mathcal{J}_h}{D \underline{z}} \Big|_{\underline{z}^i, \underline{v}^i} \frac{\partial \underline{z}}{\partial \underline{v}} \Big|_{\underline{z}^i, \underline{v}^i}, \quad (\text{A.92})$$

Como utilizamos el estado adjunto \underline{p}^i , obtenido resolviendo (A.90), esto es equivalente a decir que cumple (A.15):

$$\left(\frac{\partial \mathcal{J}_h}{\partial \underline{z}} \Big|_{(\underline{z}^*, \underline{v}^*)} \right) + \left(\frac{\partial \left(\int_{t_o}^{t_f} \underline{p}^T (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) dt \right)}{\partial \underline{z}} \Big|_{\underline{z}^*, \underline{v}^*} \right) = 0, \quad (\text{A.93})$$

entonces reemplazando (A.93) en (A.92) tenemos

$$\frac{\mathcal{D}\mathcal{J}_h}{\mathcal{D}\underline{v}} \Big|_{\underline{z}^i, \underline{v}^i} = \left(\frac{\partial \mathcal{J}_h}{\partial \underline{v}} \Big|_{(\underline{z}^*, \underline{v}^*)} \right)^T \phi - \left(\frac{\partial \left(\int_{t_o}^{t_f} \underline{p}^T (\dot{\underline{z}} - f(\underline{z}, \underline{v}, t)) dt \right)}{\partial \underline{z}} \Big|_{(\underline{z}^*, \underline{v}^*)} \right)^T \phi \frac{\partial \underline{z}}{\partial \underline{v}} \Big|_{\underline{z}^i, \underline{v}^i}, \quad (\text{A.94})$$

luego, asumimos que las ecuaciones de estado son lineales (A.89), entonces (A.94) toma la forma:

$$\frac{\mathcal{D}\mathcal{J}_h}{\mathcal{D}\underline{v}} \Big|_{\underline{z}^i, \underline{v}^i} = \left(\frac{\partial \mathcal{J}_h}{\partial \underline{z}} \Big|_{(\underline{z}^*, \underline{v}^*)} \right)^T \phi - \left(\frac{\partial \left(\int_{t_o}^{t_f} \underline{p}^T (M_h \dot{\underline{z}} - A_h \underline{z} - B_h \underline{v} - c_h) dt \right)}{\partial \underline{z}} \frac{\partial \underline{z}}{\partial \underline{v}} \Big|_{\underline{z}^i, \underline{v}^i} \right)^T \phi \quad (\text{A.95})$$

Por otro lado, hallando el producto interno del estado adjunto p por las ecuaciones de estado tenemos:

$$\int_{t_o}^{t_f} \underline{p}^T (M_h \dot{\underline{z}} - A_h \underline{z} - B_h \underline{v} - c_h) dt = 0, \quad (\text{A.96})$$

derivando en forma implícita con respecto a \underline{v} la expresión (A.96) queda como sigue:

$$\begin{aligned} & \frac{\mathcal{D} \left(\int_{t_o}^{t_f} \underline{p}^T (M_h \dot{\underline{z}} - A_h \underline{z} - B_h \underline{v} - c_h) dt \right)}{\mathcal{D}\underline{v}} = \\ & \frac{\partial \left(\int_{t_o}^{t_f} \underline{p}^T (M_h \dot{\underline{z}} - A_h \underline{z} - B_h \underline{v} - c_h) dt \right)^T}{\partial \underline{v}} \phi \\ & + \frac{\partial \left(\int_{t_o}^{t_f} \underline{p}^T (M_h \dot{\underline{z}} - A_h \underline{z} - B_h \underline{v} - c_h) dt \right)^T}{\partial \underline{z}} \frac{\partial \underline{z}}{\partial \underline{v}} \phi = 0, \end{aligned} \quad (\text{A.97})$$

entonces:

$$\frac{\partial \left(\int_{t_o}^{t_f} \underline{p}^T (M_h \dot{\underline{z}} - A_h \underline{z} - B_h \underline{v} - c_h) dt \right)^T}{\partial \underline{z}} \frac{\partial \underline{z}}{\partial \underline{v}} \phi = \frac{\partial \left(\int_{t_o}^{t_f} \underline{p}^T (M_h \dot{\underline{z}} - A_h \underline{z} - B_h \underline{v} - c_h) dt \right)^T}{\partial \underline{v}} \phi, \quad (\text{A.98})$$

luego reemplazamos (A.98) en (A.95)

$$\frac{D \mathcal{J}_h}{D \underline{v}} \Big|_{\underline{z}^i, \underline{v}^i} = \left(\frac{\partial \mathcal{J}_h}{\partial \underline{v}} \Big|_{(\underline{z}^i, \underline{v}^i)} \right)^T \phi + \left(\frac{\partial \left(\int_{t_o}^{t_f} \underline{p}^T (M_h \dot{\underline{z}} - A_h \underline{z} - B_h \underline{v} - c_h) dt \right)}{\partial \underline{v}} \Big|_{\underline{z}^i, \underline{v}^i} \right)^T \phi. \quad (\text{A.99})$$

Luego podemos observar que la relación (A.99) es equivalente a (A.17) entonces podemos usar (A.47) para obtener el gradiente del funcional para el sistema lineal:

$$\frac{D \mathcal{J}_h}{D \underline{v}} \Big|_{\underline{z}^i, \underline{v}^i} = \left(\int_{t_o}^{t_f} \phi^T (r R_h \underline{v} - B_h \underline{p}) dt \Big|_{\underline{z}^i, \underline{v}^i} \right), \quad (\text{A.100})$$

consecuentemente podemos afirmar que $\frac{D \mathcal{J}_h}{D \underline{v}} \Big|_{\underline{z}^i, \underline{v}^i} = \frac{D \mathcal{L}}{D \underline{v}} \Big|_{\underline{z}^i, \underline{v}^i}$.

Finalmente la derivada del funcional (en el sentido de Gâteaux) con respecto a la función control puede ser escrita también en términos del producto interno del gradiente del funcional con la dirección de derivación ϕ .

$$\frac{D \mathcal{J}_h}{D \underline{v}} \Big|_{\underline{z}^i, \underline{v}^i} = \int_{t_o}^{t_f} \nabla \mathcal{J}_{h_v}^T \phi dt \Big|_{\underline{z}^i, \underline{v}^i} = \int_{t_o}^{t_f} \phi^T \nabla \mathcal{J}_{h_v} dt \Big|_{\underline{z}^i, \underline{v}^i}, \quad (\text{A.101})$$

luego, comparando (A.101) con (A.100), obtenemos el gradiente del funcional:

$$\nabla \mathcal{J}_{h_v} \Big|_{\underline{z}^i, \underline{v}^i} = r R_h \underline{v} - B_h \underline{p} \Big|_{\underline{z}^i, \underline{v}^i} \quad \text{para } t_o \leq t \leq t_f, \quad (\text{A.102})$$

$$\text{y } \frac{\partial \mathcal{L}}{\partial \underline{v}} \Big|_{\underline{z}^i, \underline{v}^i} = \nabla \mathcal{J}_{h_v} \Big|_{\underline{z}^i, \underline{v}^i}$$

En conclusión, el gradiente del funcional lo podemos obtener resolviendo el siguiente sistema de ecuaciones diferenciales:

$$\left\{ \begin{array}{l} \left\{ \begin{array}{l} M_h \dot{z}^i = A_h z^i + B_h v^i + c_h, \quad \text{para } t_o \leq t \leq t_f, \\ z(t_o) = z_o^h, \end{array} \right. \\ \left\{ \begin{array}{l} M_h \dot{p}^i = -A_h p^i + q M_h (z^i - \tilde{y}), \quad \text{para } t_o \leq t \leq t_f, \\ M_h^T p(t_f) = -s M_h (z(t_f) - \tilde{y}(t_f)), \end{array} \right. \\ \nabla \mathcal{J}_{h_v} |_{z^i, v^i} = r R_h v^i - B_h p^i \quad \text{para } t_o \leq t \leq t_f, \end{array} \right. \quad (\text{A.103})$$

A.2.5. Producto de la matriz Hessiana por un vector

Para calcular el incremento de las funciones de control definido por la ecuación (2.23) necesitamos el producto de la matriz Hessiana del Lagrangiano (2.7) por la dirección de búsqueda \underline{d}^i . Entonces aplicando la definición 2.8 al Lagrangiano (2.7), tenemos que la matriz Hessiana es igual a:

$$\mathbf{H} := \nabla^2 \mathcal{L} = \nabla(\nabla \mathcal{L}(z, \underline{u}, \underline{p}, \eta_o)), \quad (\text{A.104})$$

pero como en la sección anterior mostramos que:

$$\nabla \mathcal{L}(z^i, \underline{u}^i, \underline{p}^i) = \nabla \mathcal{J}(z^i, \underline{u}^i, \underline{p}^i), \quad (\text{A.105})$$

donde el gradiente del funcional (2.5) esta dado por la solución de la ecuación (A.103). La matriz Hessiana del funcional de costo (2.5) es igual a la matriz Hessiana del Lagrangiano (2.7).

Primeramente consideramos la solución analítica de las ecuaciones de esta-

do y las ecuaciones adjuntas (como en (A.79)):

$$\begin{aligned}\underline{z} &= e^{-M_h^{-1}A_h(t-t_o)} z_0^h + \int_{t_o}^t e^{-M_h^{-1}A_h(t-\tau)} M_h^{-1} (B_h \underline{v} + c_h) d\tau, \\ \underline{z} &= \mathcal{Z}(z_0^h, \underline{v}^i, c_h),\end{aligned}\tag{A.106}$$

y

$$\begin{aligned}\underline{p} &= e^{M_h^{-1}A_h(t-t_f)} - s (\underline{z}(t_f) - \tilde{y}(t_f)) + \int_{t_f}^t e^{M_h^{-1}A_h(t-\tau)} M_h^{-1} (q M_h (\underline{z} - \tilde{y})) d\tau, \\ \underline{p} &= \mathcal{P}(\underline{z}^i, \tilde{y}),\end{aligned}\tag{A.107}$$

donde \mathcal{Z} y \mathcal{P} son operadores lineales. Además reemplazando éstas ecuaciones en (A.103), tenemos:

$$\begin{aligned}\nabla \mathcal{J}_{h_v} \Big|_{\underline{z}^i, \underline{v}^i} &= r R_h \underline{v}^i - B_h \mathcal{P}(\underline{z}^i, \tilde{y}), \\ \nabla \mathcal{J}_{h_v} \Big|_{\underline{z}^i, \underline{v}^i} &= r R_h \underline{v}^i - B_h \mathcal{P}(\mathcal{Z}(z_0^h, \underline{v}^i, c_h), \tilde{y}).\end{aligned}\tag{A.108}$$

A continuación calculamos la derivada direccional del gradiente del funcional para obtener el producto del Hessiano del Lagrangiano por la dirección de búsqueda:

$$\begin{aligned}\frac{D \nabla \mathcal{J}_h}{D \underline{v}^i} &= \frac{D (r R_h \underline{v}^i - B_h \mathcal{P}(\mathcal{Z}(z_0^h, \underline{v}^i, c_h), \tilde{y}))}{D \underline{v}^i} = \nabla^2 \mathcal{J}_h \phi = \mathbf{H} \phi \\ &= \left. \frac{d r R_h (\underline{v}^i + \epsilon \phi)}{d \epsilon} \right|_{\epsilon=0} - \left. \frac{d (B_h \mathcal{P}(\mathcal{Z}(z_0^h, \underline{v}^i + \epsilon \phi, c_h), \tilde{y}))}{d \epsilon} \right|_{\epsilon=0} \\ &= (r R_h \phi) \Big|_{\epsilon=0} - (B_h \mathcal{P}(\mathcal{Z}(0, \phi, 0), 0)) \Big|_{\epsilon=0} \\ &= r R_h \phi - B_h \mathcal{P}(\mathcal{Z}(0, \phi, 0), 0),\end{aligned}\tag{A.109}$$

si hacemos $\phi = \underline{d}^i$, tenemos lo que buscamos, es decir:

$$\mathbf{H} \underline{d}^i = r R_h \underline{d}^i - B_h \mathcal{P}(\mathcal{Z}(0, \underline{d}^i, 0), 0).\tag{A.110}$$

DESARROLLO DE UNA APLICACIÓN PARALELA EN EL TIEMPO PARA LA SIMULACIÓN Y CONTROL DE ECUACIONES DIFERENCIALES PARCIALES PARABÓLICAS

Luego introducimos la siguientes definiciones $\underline{\psi}^i = \mathcal{Z}(0, \underline{d}^i, 0)$ y

$\underline{\psi}_p^i = \mathcal{P}(\mathcal{Z}(0, \underline{d}^i, 0))$ para reemplazar los operadores \mathcal{P} y \mathcal{Z} por sus respectivas ecuaciones diferenciales nuevamente, es decir:

$$\left\{ \begin{array}{l} \left\{ \begin{array}{l} M_h \dot{\underline{\psi}}^i = A_h \underline{\psi}^i + B_h \underline{d}^i, \quad \text{para } t_o \leq t \leq t_f, \\ \underline{\psi}^i(t_o) = 0, \end{array} \right. \\ \left\{ \begin{array}{l} M_h \dot{\underline{\psi}}_p^i = -A_h \underline{\psi}_p^i + q M_h \underline{\psi}^i, \quad \text{para } t_o \leq t \leq t_f, \\ \underline{\psi}_p^i(t_f) = -s \underline{\psi}^i(t_f), \end{array} \right. \\ \mathbf{H} \underline{d}^i = r R_h \underline{d}^i - B_h \underline{\psi}_p^i \quad \text{para } t_o \leq t \leq t_f, \end{array} \right. \quad (\text{A.111})$$

APÉNDICE B

APÉNDICE: ALGORITMO PARAREAL

B.1. Condiciones de KKT del Lagrangiano

De la ecuación (3.8) obtenemos las condiciones de optimalidad, como las variaciones respecto a cada una de las variables de $\delta\mathcal{L}$ en cada sub-intervalo $k = \{0, 1, \dots, \hat{k} - 1\}$ no son nulas, se debe cumplir que:

$$\begin{aligned} \left. \frac{\partial \mathcal{L}}{\partial \underline{z}_k} \right|_{\underline{z}_k^*, \underline{v}_k^*, Z_k^*} &= 0, & \left. \frac{\partial \mathcal{L}}{\partial \underline{v}_k} \right|_{\underline{z}_k^*, \underline{v}_k^*, Z_k^*} &= 0, & \left. \frac{\partial \mathcal{L}}{\partial Z_k} \right|_{\underline{z}_k^*, \underline{v}_k^*, Z_k^*} &= 0, \\ \left. \frac{\partial \mathcal{L}}{\partial \underline{p}_k} \right|_{\underline{z}_k^*, \underline{v}_k^*, Z_k^*} &= 0, & \left. \frac{\partial \mathcal{L}}{\partial \eta_k} \right|_{\underline{z}_k^*, \underline{v}_k^*, Z_k^*} &= 0. \end{aligned}$$

para $k = \{0, 1, \dots, \hat{k} - 1\}$.

A continuación calculamos cada una de éstas derivadas, utilizando la definición (2.1). En las ecuaciones siguientes omitiremos la indicación $k = \{0, 1, \dots, \hat{k} - 1\}$, quedando sub-entendido que las mismas se cumplen cada en uno de estos sub-intervalos, y la evaluación en la solución óptima $|_{\underline{z}_k^*, \underline{v}_k^*, Z_k^*}$ representaremos por $|_*$.

Cálculo de $\frac{\partial \mathcal{L}}{\partial \underline{z}_k}$:

Aplicando la definición (2.1) a la ecuación (3.7) tenemos:

$$\left. \frac{D\mathcal{L}}{D\underline{p}_k} = \frac{D\mathcal{J}(\underline{z}_k, \underline{v}_k, Z_k)}{D\underline{p}_k} + \frac{d \left(\int_{T_k}^{T_{k+1}} (\underline{p}_k + \lambda \phi_k)^T (M_h \dot{\underline{z}}_k - A_h \underline{z}_k) dt \right)}{d\lambda} \right|_{\lambda=0}$$

$$\begin{aligned}
 & - \left. \frac{d \left(\int_{T_k}^{T_{k+1}} (\underline{p}_k + \lambda \phi_k)^T (B_h \underline{v}_k + c_h) dt - \eta_k^T (\underline{z}_k(T_k^+) - Z_k) \right)}{d\lambda} \right|_{\lambda=0}, \\
 & = \int_{T_k}^{T_{k+1}} \phi_k^T (M_h \dot{\underline{z}}_k - A_h \underline{z}_k - B_h \underline{v}_k - c_h) dt, \tag{B.1}
 \end{aligned}$$

el cual podemos expresar como:

$$\frac{D\mathcal{L}}{D\underline{p}_k} = \left(\phi_k, \frac{\partial \mathcal{L}}{\partial \underline{p}_k} \right).$$

entonces, aplicando la condición de optimalidad $\left. \frac{\partial \mathcal{L}}{\partial \underline{p}_k} \right|_* = 0$ tenemos:

$$\left. \frac{\partial \mathcal{L}}{\partial \underline{p}_k} \right|_* = (M_h \dot{\underline{z}}_k - A_h \underline{z}_k - B_h \underline{v}_k - c_h)|_* = 0 \quad t \in [T_k, T_{k+1}]. \tag{B.2}$$

Cálculo de $\frac{\partial \mathcal{L}}{\partial \eta_k}$:

$$\begin{aligned}
 \frac{D\mathcal{L}}{D\eta_k} & = \frac{D\mathcal{J}(\underline{z}_k, \underline{v}_k, Z_k)}{D\eta_k} + \left. \frac{d \left(\int_{T_k}^{T_{k+1}} \underline{p}_k^T (M_h \dot{\underline{z}}_k - A_h \underline{z}_k - B_h \underline{v}_k - c_h) dt \right)}{d\lambda} \right|_{\lambda=0} \\
 & + \left. \frac{d \left((\eta_k + \lambda \phi_k)^T (\underline{z}_k(T_k^+) - Z_k) \right)}{d\lambda} \right|_{\lambda=0}, \\
 & = \phi_k^T (\underline{z}_k(T_k^+) - Z_k), \tag{B.3}
 \end{aligned}$$

que podemos expresar como:

$$\frac{D\mathcal{L}}{D\eta_k} = \left(\phi_k, \frac{\partial \mathcal{L}}{\partial \eta_k} \right),$$

aplicando la condición de optimalidad $\left. \frac{\partial \mathcal{L}}{\partial \eta_k} \right|_* = 0$ tenemos:

$$\left. \frac{\partial \mathcal{L}}{\partial \eta_k} \right|_* = (\underline{z}_k(T_k^+) - Z_k)|_* = 0. \tag{B.4}$$

Las ecuaciones (B.2) y (B.4) son las llamadas Ecuaciones de Estado.

Cálculo de $\frac{\partial \mathcal{L}}{\partial z_k}$:

Primeramente integramos por partes $\int_{T_k}^{T_{k+1}} \underline{p}_k^T M_h \dot{z}_k dt$:

$$\int_{T_k}^{T_{k+1}} \underline{p}_k^T M_h \dot{z}_k dt = - \int_{T_k}^{T_{k+1}} \dot{\underline{p}}_k^T M_h z_k dt + (\underline{p}_k^T M_h z_k) \Big|_{T_k}^{T_{k+1}}. \quad (\text{B.5})$$

Entonces reemplazando (B.5) en (3.7) y aplicando la definición (2.1) tenemos:

$$\begin{aligned} \frac{D\mathcal{L}}{Dz_k} &= \frac{D\mathcal{J}(z_k, v_k, Z_k)}{Dz_k} - \frac{d \left(\int_{T_k}^{T_{k+1}} \dot{\underline{p}}_k^T M_h (z_k + \lambda \phi_k) dt \right)}{d\lambda} \Bigg|_{\lambda=0} \\ &+ \frac{d \left((\underline{p}_k^T M_h (z_k + \lambda \phi_k)) \Big|_{T_k}^{T_{k+1}} - \int_{T_k}^{T_{k+1}} \underline{p}_k^T (A_h (z_k + \lambda \phi_k)) dt \right)}{d\lambda} \Bigg|_{\lambda=0} \\ &- \frac{d \left(\int_{T_k}^{T_{k+1}} \underline{p}_k^T (B_h v_k + c_h) dt - \eta_k^T ((z_k + \lambda \phi_k)(T_k^+) - Z_k) \right)}{d\lambda} \Bigg|_{\lambda=0} \end{aligned} \quad (\text{B.6})$$

Necesitamos calcular por separado $\frac{D\mathcal{J}}{Dz_k}$, donde los módulos respectivos son dados por las ecuaciones (2.2), (2.3) y (2.4) entonces:

$$\begin{aligned} \frac{D\mathcal{J}}{Dz_k} &= \frac{1}{2} \frac{d \int_{T_k}^{T_{k+1}} (r \|v_k\|_{R_h}^2 + q \|(z_k + \lambda \phi_k) - \tilde{y}_k\|_{M_h}^2) dt}{d\lambda} \Bigg|_{\lambda=0} \\ &+ \frac{d \frac{1}{2\varepsilon \Delta T} \|(z_k + \lambda \phi_k)(T_{k+1}^-) - Z_{k+1}\|_{M_h}^2}{d\lambda} \Bigg|_{\lambda=0}, \quad \forall k = 0, 1, \dots, \hat{k} - 2, \\ &= \frac{1}{2} \int_{T_k}^{T_{k+1}} q (\phi_k^T M_h (z_k - \tilde{y}_k) + (z_k - \tilde{y}_k)^T M_h \phi_k) dt \\ &+ \frac{1}{2\varepsilon \Delta T} ((\phi_k(T_{k+1}^-))^T M_h (z_k(T_{k+1}^-) - Z_{k+1})) \\ &+ \frac{1}{2\varepsilon \Delta T} ((z_k(T_{k+1}^-) - Z_{k+1})^T M_h \phi_k(T_{k+1}^-)), \end{aligned} \quad (\text{B.7})$$

$$\begin{aligned}
 \frac{D\mathcal{J}}{D\underline{z}_{\hat{k}-1}} &= \frac{1}{2} \frac{d \int_{T_{\hat{k}-1}}^{T_{\hat{k}}} \left(r \|\underline{v}_{\hat{k}-1}\|_{R_h}^2 + q \|(z_{\hat{k}-1} + \lambda \phi_{\hat{k}-1}) - \underline{\tilde{y}}_{\hat{k}-1}\|_{M_h}^2 \right) dt}{d\lambda} \Bigg|_{\lambda=0} \\
 &\quad + \frac{d^s \|(z_{\hat{k}-1} + \lambda \phi_{\hat{k}-1})(T_{\hat{k}}) - \underline{\tilde{y}}_{\hat{k}-1}(T_{\hat{k}})\|_{M_h}^2}{d\lambda} \Bigg|_{\lambda=0} \\
 &= \frac{1}{2} \int_{T_{\hat{k}-1}}^{T_{\hat{k}}} q \left(\phi_{\hat{k}-1}^T M_h (z_{\hat{k}-1} - \underline{\tilde{y}}_{\hat{k}-1}) + (z_{\hat{k}-1} - \underline{\tilde{y}}_{\hat{k}-1})^T M_h \phi_{\hat{k}-1} \right) dt \\
 &\quad + \frac{s}{2} (\phi_{\hat{k}-1}(T_{\hat{k}}))^T M_h (z_{\hat{k}-1}(T_{\hat{k}}) - \underline{\tilde{y}}_{\hat{k}-1}(T_{\hat{k}})) \\
 &\quad + \frac{s}{2} (z_{\hat{k}-1}(T_{\hat{k}}) - \underline{\tilde{y}}_{\hat{k}-1}(T_{\hat{k}}))^T M_h \phi_{\hat{k}-1}(T_{\hat{k}}) \tag{B.8}
 \end{aligned}$$

notemos que

$$\phi_k^T M_h (\underline{z}_k - \underline{\tilde{y}}_k) = (\underline{z}_k - \underline{\tilde{y}}_k)^T M_h \phi_k, \quad \mathbf{y}$$

$$(\phi_k(T_{k+1}^-))^T M_h (\underline{z}_k(T_{k+1}^-) - Z_{k+1}) = (\underline{z}_k(T_{k+1}^-) - Z_{k+1})^T M_h \phi_k(T_{k+1}^-),$$

entonces las ecuaciones (B.7) y (B.8) quedan:

$$\begin{aligned}
 \frac{D\mathcal{J}}{D\underline{z}_k} &= q \int_{T_k}^{T_{k+1}} \phi_k^T M_h (\underline{z}_k - \underline{\tilde{y}}_k) dt + \frac{1}{\varepsilon \Delta T} (\phi_k(T_{k+1}^-))^T M_h (\underline{z}_k(T_{k+1}^-) - Z_{k+1}), \\
 \forall k &= 0, 1, \dots, \hat{k} - 2. \tag{B.9}
 \end{aligned}$$

$$\begin{aligned}
 \frac{D\mathcal{J}}{D\underline{z}_{\hat{k}-1}} &= q \int_{T_{\hat{k}-1}}^{T_{\hat{k}}} \phi_{\hat{k}-1}^T M_h (\underline{z}_{\hat{k}-1} - \underline{\tilde{y}}_{\hat{k}-1}) dt + s (\phi_{\hat{k}-1}(T_{\hat{k}}))^T M_h (\underline{z}_{\hat{k}-1}(T_{\hat{k}}) - \underline{\tilde{y}}_{\hat{k}-1}(T_{\hat{k}})). \\
 &\tag{B.10}
 \end{aligned}$$

Reemplazando (B.9) y (B.10) en (B.6) tenemos:

$$\begin{aligned}
 \frac{D\mathcal{L}}{D\underline{z}_k} &= q \int_{T_k}^{T_{k+1}} \phi_k^T M_h (\underline{z}_k - \underline{\tilde{y}}_k) dt + \frac{1}{\varepsilon \Delta T} (\phi_k(T_{k+1}^-))^T M_h (\underline{z}_k(T_{k+1}^-) - Z_{k+1}) \\
 &\quad - \int_{T_k}^{T_{k+1}} (\phi_k^T M_h \underline{\dot{p}}_k + \phi_k^T A_h \underline{p}_k) dt + (\phi_k^T M_h \underline{p}_k) \Big|_{T_k}^{T_{k+1}} + (\phi_k(T_k^+))^T \eta_k \\
 \forall k &= 0, 1, \dots, \hat{k} - 2. \tag{B.11}
 \end{aligned}$$

DESARROLLO DE UNA APLICACIÓN PARALELA EN EL TIEMPO PARA LA SIMULACIÓN Y CONTROL DE ECUACIONES DIFERENCIALES PARCIALES PARABÓLICAS

$$\begin{aligned} \frac{D\mathcal{L}}{Dz_{\hat{k}-1}} &= q \int_{T_{\hat{k}-1}}^{T_{\hat{k}}} \phi_{\hat{k}-1}^T M_h(\underline{z}_{\hat{k}-1} - \underline{\tilde{y}}_{\hat{k}-1}) dt + s(\phi_{\hat{k}-1}(T_{\hat{k}}))^T M_h(\underline{z}_{\hat{k}-1}(T_{\hat{k}}) - \underline{\tilde{y}}_{\hat{k}-1}(T_{\hat{k}})) \\ &\quad - \int_{T_{\hat{k}-1}}^{T_{\hat{k}}} \left(\phi_{\hat{k}-1}^T M_h \underline{\dot{p}}_{\hat{k}-1} + \phi_{\hat{k}-1}^T A_h \underline{p}_{\hat{k}-1} \right) dt + \left(\phi_{\hat{k}-1}^T M_h \underline{p}_{\hat{k}-1} \right) \Big|_{T_{\hat{k}-1}}^{T_{\hat{k}}} \\ &\quad + (\phi_{\hat{k}-1}(T_{\hat{k}-1}^+))^T \eta_{\hat{k}-1}. \end{aligned} \quad (\text{B.12})$$

reordenando los términos de (B.11) y (B.12), tenemos:

$$\begin{aligned} \frac{D\mathcal{L}}{Dz_k} &= \int_{T_k}^{T_{k+1}} \phi_k^T (q M_h(\underline{z}_k - \underline{\tilde{y}}_k) - M_h \underline{\dot{p}}_k - A_h \underline{p}_k) dt \\ &\quad + (\phi_k(T_{k+1}^-))^T \left(\frac{1}{\varepsilon \Delta T} M_h(\underline{z}_k(T_{k+1}^-) - Z_{k+1}) + M_h \underline{p}_k(T_{k+1}^-) \right) \\ &\quad + (\phi_k(T_k^+))^T (\eta_k - M_h \underline{p}_k(T_k^+)) \quad \forall k = 0, 1, \dots, \hat{k} - 2. \end{aligned} \quad (\text{B.13})$$

$$\begin{aligned} \frac{D\mathcal{L}}{Dz_{\hat{k}-1}} &= \int_{T_{\hat{k}-1}}^{T_{\hat{k}}} \phi_{\hat{k}-1}^T \left(M_h(\underline{z}_{\hat{k}-1} - \underline{\tilde{y}}_{\hat{k}-1}) - M_h \underline{\dot{p}}_{\hat{k}-1} - A_h \underline{p}_{\hat{k}-1} \right) dt \\ &\quad + (\phi_{\hat{k}-1}(T_{\hat{k}}))^T \left(s M_h(\underline{z}_{\hat{k}-1}(T_{\hat{k}}) - \underline{\tilde{y}}_{\hat{k}-1}(T_{\hat{k}})) + M_h \underline{p}_{\hat{k}-1}(T_{\hat{k}}) \right) \\ &\quad + (\phi_{\hat{k}-1}(T_{\hat{k}-1}^+))^T \left(\eta_{\hat{k}-1} - M_h \underline{p}_{\hat{k}-1}(T_{\hat{k}-1}^+) \right). \end{aligned} \quad (\text{B.14})$$

la definición (2.1) nos permite expresar las ecuaciones (B.13) y (B.14) como:

$$\frac{D\mathcal{L}}{Dz_k} = \left(\phi_k, \frac{\partial \mathcal{L}}{\partial z_k} \right)$$

y aplicando la condición de optimalidad $\left. \frac{\partial \mathcal{L}}{\partial z_k} \right|_* = 0$ tenemos:

$$\left. \frac{\partial \mathcal{L}}{\partial z_k} \right|_* = \begin{cases} (M_h \dot{\underline{p}}_k + A_h \underline{p}_k - q M_h (z_k - \tilde{y}_k)) \Big|_* = 0, & t \in [T_k, T_{k+1}] \\ (\underline{p}_k(T_{k+1}^-) + \frac{1}{\varepsilon \Delta T} (z_k(T_{k+1}^-) - Z_{k+1})) \Big|_* = 0, \\ (\eta_k - M_h \underline{p}_k(T_k^+)) \Big|_* = 0, \\ \forall k = 0, 1, \dots, \hat{k} - 2. \end{cases} \quad (\text{B.15})$$

$$\left. \frac{\partial \mathcal{L}}{\partial z_{\hat{k}-1}} \right|_* = \begin{cases} (M_h \dot{\underline{p}}_{\hat{k}-1} + A_h \underline{p}_{\hat{k}-1} - q M_h (z_{\hat{k}-1} - \tilde{y}_{\hat{k}-1})) \Big|_* = 0, & t \in [T_{\hat{k}-1}, T_{\hat{k}}] \\ (\underline{p}_{\hat{k}-1}(T_{\hat{k}}) + s (z_{\hat{k}-1}(T_{\hat{k}}) - \tilde{y}_{\hat{k}-1}(T_{\hat{k}}))) \Big|_* = 0, \\ (\eta_{\hat{k}-1} - M_h \underline{p}_{\hat{k}-1}(T_{\hat{k}-1}^+)) \Big|_* = 0. \end{cases} \quad (\text{B.16})$$

Las dos primeras relaciones de las ecuaciones (B.15) y (B.16) son las que llamamos de ecuaciones de Estado Adjunto.

Cálculo de $\frac{\partial \mathcal{L}}{\partial v_k}$:

$$\frac{D\mathcal{L}}{Dv_k} = \frac{D\mathcal{J}(z_k, v_k, Z_k)}{Dv_k} + \frac{d \left(\int_{T_k}^{T_{k+1}} \underline{p}_k^T (M_h \dot{z}_k - A_h z_k - c_h) dt \right)}{d\lambda} \Bigg|_{\lambda=0} - \frac{d \left(\int_{T_k}^{T_{k+1}} \underline{p}_k^T B_h (v_k + \lambda \phi_k) dt - \eta_k^T (z_k(T_k^+) - Z_k) \right)}{d\lambda} \Bigg|_{\lambda=0}. \quad (\text{B.17})$$

Necesitamos calcular $\frac{D\mathcal{J}}{Dv_k}$, entonces aplicamos la definición (2.1) a la ecuación

(3.6):

$$\begin{aligned}
 \frac{D\mathcal{J}}{D\underline{v}_k} &= \frac{1}{2} \left. \frac{d \int_{T_k}^{T_{k+1}} (r \|\underline{v}_k + \lambda \phi_k\|_{R_h}^2 + q \|\underline{z}_k - \tilde{\underline{y}}_k\|_{M_h}^2) dt}{d\lambda} \right|_{\lambda=0} \\
 &+ \left. \frac{d \frac{1}{2\varepsilon \Delta T} \|\underline{z}_k(T_{k+1}^-) - Z_{k+1}\|_{M_h}^2}{d\lambda} \right|_{\lambda=0}, \quad \forall k = 0, 1, \dots, \hat{k} - 2, \\
 &= \frac{1}{2} \int_{T_k}^{T_{k+1}} r (\phi_k^T R_h \underline{v}_k + \underline{v}_k^T R_h \phi_k) dt, \tag{B.18}
 \end{aligned}$$

$$\begin{aligned}
 \frac{D\mathcal{J}}{D\underline{v}_{\hat{k}-1}} &= \frac{1}{2} \left. \frac{d \int_{T_{\hat{k}-1}}^{T_{\hat{k}}} (r \|\underline{v}_{\hat{k}-1} + \lambda \phi_{\hat{k}-1}\|_{R_h}^2 + q \|\underline{z}_{\hat{k}-1} - \tilde{\underline{y}}_{\hat{k}-1}\|_{M_h}^2) dt}{d\lambda} \right|_{\lambda=0} \\
 &+ \left. \frac{d \frac{\varepsilon}{2} \|\underline{z}_{\hat{k}-1}(T_{\hat{k}}) - \tilde{\underline{y}}_{\hat{k}-1}(T_{\hat{k}})\|_{M_h}^2}{d\lambda} \right|_{\lambda=0} \\
 &= \frac{1}{2} \int_{T_{\hat{k}-1}}^{T_{\hat{k}}} r (\phi_{\hat{k}-1}^T R_h \underline{v}_{\hat{k}-1} + \underline{v}_{\hat{k}-1}^T R_h \phi_{\hat{k}-1}) dt, \tag{B.19}
 \end{aligned}$$

notemos que:

$$\phi_k^T R_h \underline{v}_k = \underline{v}_k^T R_h \phi_k,$$

entonces (B.18) y (B.19) quedan:

$$\frac{D\mathcal{J}}{D\underline{v}_k} = r \int_{T_k}^{T_{k+1}} \phi_k^T R_h \underline{v}_k dt \quad \forall k = 0, 1, \dots, \hat{k} - 1, \tag{B.20}$$

reemplazando (B.20) en (B.17) tenemos:

$$\begin{aligned}
 \frac{D\mathcal{L}}{D\underline{v}_k} &= r \int_{T_k}^{T_{k+1}} \phi_k^T R_h \underline{v}_k dt - \int_{T_k}^{T_{k+1}} \underline{p}_k^T B_h \phi_k dt, \\
 &= \int_{T_k}^{T_{k+1}} (\phi_k^T r R_h \underline{v}_k - \phi_k^T B_k^T \underline{p}_k) dt, \\
 &= \int_{T_k}^{T_{k+1}} \phi_k^T (r R_h \underline{v}_k - B_k^T \underline{p}_k) dt, \tag{B.21}
 \end{aligned}$$

el cual podemos expresar como:

$$\frac{D\mathcal{L}}{D\underline{v}_k} = \left(\phi_k, \frac{\partial \mathcal{L}}{\partial \underline{v}_k} \right),$$

y aplicando la condición de optimalidad $\left. \frac{\partial \mathcal{L}}{\partial \underline{v}_k} \right|_* = 0$ tenemos:

$$\left. \frac{\partial \mathcal{L}}{\partial \underline{v}_k} \right|_* = (r R_h \underline{v}_k - B_k^T \underline{p}_k) \Big|_* = 0. \quad (\text{B.22})$$

Cálculo de $\frac{\partial \mathcal{L}}{\partial Z_k}$:

$$\begin{aligned} \frac{D\mathcal{L}}{DZ_k} &= \frac{D\mathcal{J}(z_k, v_k, Z_k)}{DZ_k} + \frac{d \left(\int_{T_k}^{T_{k+1}} \underline{p}_k^T (M_h \dot{z}_k - A_h z_k - B_h v_k - c_h) dt \right)}{d\lambda} \Bigg|_{\lambda=0} \\ &+ \frac{d \left(\eta_k^T (z_k(T_k^+) - (Z_k + \lambda \phi_k)) \right)}{d\lambda} \Bigg|_{\lambda=0}. \end{aligned} \quad (\text{B.23})$$

Calculamos $\frac{D\mathcal{J}}{DZ_k}$:

$$\begin{aligned} \frac{D\mathcal{J}}{DZ_k} &= \frac{1}{2} \frac{d \int_{T_k}^{T_{k+1}} (r \|\underline{v}_k\|_{R_h}^2 + q \|z_k - \tilde{y}_k\|_{M_h}^2) dt}{d\lambda} \Bigg|_{\lambda=0} \\ &+ \frac{d \frac{1}{2\varepsilon \Delta T} \|z_{k-1}(T_k^-) - (Z_k + \lambda \phi_k)\|_{M_h}^2}{d\lambda} \Bigg|_{\lambda=0}, \quad (\text{B.24}) \\ &= -\frac{1}{2\varepsilon \Delta T} \left(\phi_k^T M_h (z_{k-1}(T_k^-) - Z_k) + (z_{k-1}(T_k^-) - Z_k)^T M_h \phi_k \right), \end{aligned}$$

notemos que:

$$\phi_k^T M_h (z_{k-1}(T_k^-) - Z_k) = (z_{k-1}(T_k^-) - Z_k)^T M_h \phi_k,$$

entonces la ecuación (B.24) queda:

$$\frac{D\mathcal{J}}{DZ_k} = -\frac{1}{\varepsilon \Delta T} \left(\phi_k^T M_h (z_{k-1}(T_k^-) - Z_k) \right). \quad (\text{B.25})$$

Reemplazando (B.25) en (B.23) tenemos:

$$\begin{aligned}\frac{D\mathcal{L}}{DZ_k} &= -\frac{1}{\varepsilon \Delta T} \left(\phi_k^T M_h(z_{k-1}(T_k^-) - Z_k) \right) - \eta_k^T \phi_k, \\ &= \phi_k^T \left(-\eta_k - \frac{1}{\varepsilon \Delta T} M_h(z_{k-1}(T_k^-) - Z_k) \right),\end{aligned}\quad (\text{B.26})$$

el cual podemos expresar como:

$$\frac{D\mathcal{L}}{DZ_k} = \left(\phi_k, \frac{\partial \mathcal{L}}{\partial Z_k} \right),$$

y aplicando la condición de optimalidad $\left. \frac{\partial \mathcal{L}}{\partial Z_k} \right|_* = 0$ tenemos:

$$\left. \frac{\partial \mathcal{L}}{\partial Z_k} \right|_* = \left(-\eta_k - \frac{1}{\varepsilon \Delta T} M_h(z_{k-1}(T_k^-) - Z_k) \right) \Big|_* = 0. \quad (\text{B.27})$$

En la solución óptima, de la ecuación (B.15) tenemos que:

$$\left(\eta_k - M_h \underline{p}_k(T_k^+) \right) \Big|_* = 0, \quad (\text{B.28})$$

reemplazando en (B.27) tenemos:

$$\begin{aligned}\left. \frac{\partial \mathcal{L}}{\partial Z_k} \right|_* &= \left(-M_h \underline{p}_k(T_k^+) - \frac{1}{\varepsilon \Delta T} M_h(z_{k-1}(T_k^-) - Z_k) \right) \Big|_* \\ &\left(\underline{p}_k(T_k^+) + \frac{1}{\varepsilon \Delta T} (z_{k-1}(T_k^-) - Z_k) \right) \Big|_* = 0.\end{aligned}\quad (\text{B.29})$$

Las ecuaciones (B.22) y (B.29) son conocidas como Condición de Euler.

B.2. Cálculo del $\nabla \mathcal{J}$

Asumiendo que tenemos las funciones de control \underline{v}_k^j y las condiciones iniciales Z_k^i para cada sub-intervalo $k = 0, 1, \dots, \hat{k} - 1$, podemos encontrar los estados

\underline{z}_k^j resolviendo las ecuaciones (B.2) y (B.4):

$$\begin{aligned} M_h \underline{\dot{z}}_k^j &= A_h \underline{z}_k^j + B_h \underline{v}_k^j + c_h \quad t \in [T_k, T_{k+1}], \\ \underline{z}_k^j(T_k) &= Z_k^i. \end{aligned} \quad (\text{B.30})$$

Luego podemos hallar los estados adjuntos \underline{p}_k^j resolviendo las ecuaciones (B.15) y (B.16):

$$\left. \begin{aligned} M_h \underline{\dot{p}}_k^j &= -A_h \underline{p}_k^j + q M_h (\underline{z}_k^j - \tilde{y}_k) \quad t \in [T_k, T_{k+1}] \\ \underline{p}_k^j(T_{k+1}^-) &= -\frac{1}{\varepsilon \Delta T} (\underline{z}_k^j(T_{k+1}^-) - Z_{k+1}^i) \end{aligned} \right\} \quad \forall k = 0, 1, \dots, \hat{k} - 2.$$

$$\begin{aligned} M_h \underline{\dot{p}}_{\hat{k}-1}^j &= -A_h \underline{p}_{\hat{k}-1}^j + q M_h (\underline{z}_{\hat{k}-1}^j - \tilde{y}_{\hat{k}-1}) \quad t \in [T_{\hat{k}-1}, T_{\hat{k}}] \\ \underline{p}_{\hat{k}-1}^j(T_{\hat{k}}) &= -s (\underline{z}_{\hat{k}-1}^j(T_{\hat{k}}) - \tilde{y}_{\hat{k}-1}(T_{\hat{k}})) \end{aligned} \quad (\text{B.31})$$

Al resolver las ecuaciones (B.30) y (B.31) estamos condicionando que:

$$\left. \frac{\partial \mathcal{L}}{\partial \underline{p}_k} \right|_{\underline{z}_k^j, \underline{v}_k^j, Z_k^i, \underline{p}_k^j} = \left. \frac{\partial \mathcal{L}}{\partial \eta_k} \right|_{\underline{z}_k^j, \underline{v}_k^j, Z_k^i, \underline{p}_k^j} = \left. \frac{\partial \mathcal{L}}{\partial \underline{z}_k} \right|_{\underline{z}_k^j, \underline{v}_k^j, Z_k^i, \underline{p}_k^j} = 0. \quad (\text{B.32})$$

Entonces la variación total de primer de \mathcal{L} en cada sub-intervalo queda:

$$\delta \mathcal{L} = \left. \frac{\partial \mathcal{L}}{\partial \underline{v}_k} \right|_j \delta \underline{v}_k + \left. \frac{\partial \mathcal{L}}{\partial Z_k} \right|_j \delta Z_k, \quad (\text{B.33})$$

donde la evaluación en $\left|_{\underline{z}_k^j, \underline{v}_k^j, Z_k^i, \underline{p}_k^j} \right|_j$ representamos por $|_j$.

Notemos que en (3.6) tenemos dos variables independientes \underline{v}_k y Z_k , entonces calcularemos las derivadas totales de \mathcal{J} respecto a las variables independientes.

Primeramente calcularemos la derivada total de \mathcal{J} respecto de \underline{v}_k :

$$\begin{aligned}\frac{D\mathcal{J}}{D\underline{v}_k} &= \frac{D\mathcal{J}}{D\underline{v}_k} + \frac{D\mathcal{J}}{D\underline{z}_k} \frac{\partial \underline{z}_k}{\partial \underline{v}_k}, \\ &= \left(\nabla_{\underline{v}_k} \mathcal{J}, \phi_k \right).\end{aligned}\quad (\text{B.34})$$

Al cumplirse la relación (B.32), de la ecuación (B.6) obtenemos:

$$\begin{aligned}\left. \frac{D\mathcal{J}}{D\underline{z}_k} \right|_j &= - \left. \frac{D \int_{T_k}^{T_{k+1}} \underline{p}_k^T (M_h \dot{\underline{z}}_k - A_h \underline{z}_k - B_h \underline{v}_k - c_h) dt}{D\underline{z}_k} \right|_j \\ &\quad - \left. \frac{D\eta_k^T(\underline{z}_k(T_k^+) - Z_k)}{D\underline{z}_k} \right|_j.\end{aligned}\quad (\text{B.35})$$

Reemplazando (B.35) en (B.34) tenemos:

$$\begin{aligned}\left. \frac{D\mathcal{J}}{D\underline{v}_k} \right|_j &= \left. \frac{D\mathcal{J}}{D\underline{v}_k} \right|_j - \left. \frac{D \int_{T_k}^{T_{k+1}} \underline{p}_k^T (M_h \dot{\underline{z}}_k - A_h \underline{z}_k - B_h \underline{v}_k - c_h) dt}{D\underline{z}_k} \right|_j \frac{\partial \underline{z}_k}{\partial \underline{v}_k} \\ &\quad - \left. \frac{D\eta_k^T(\underline{z}_k(T_k^+) - Z_k)}{D\underline{z}_k} \right|_j \frac{\partial \underline{z}_k}{\partial \underline{v}_k}.\end{aligned}\quad (\text{B.36})$$

Además si calculamos el producto interno de \underline{p}_k^i y η_k^i por las ecuaciones de estado y las condiciones iniciales de (B.30) respectivamente, definiendo al mismo como φ tenemos:

$$\varphi = \int_{T_k}^{T_{k+1}} (\underline{p}_k^j)^T (M_h \dot{\underline{z}}_k^j - A_h \underline{z}_k^j - B_h \underline{v}_k^j - c_h) dt + \eta_k^T(\underline{z}_k^j(T_k^+) - Z_k^i) = 0, \quad (\text{B.37})$$

calculando la derivada total de φ respecto de \underline{v}_k tenemos:

$$\frac{D\varphi}{D\underline{v}_k} = \frac{D\varphi}{D\underline{v}_k} + \frac{D\varphi}{D\underline{z}_k} \frac{\partial \underline{z}_k}{\partial \underline{v}_k} = 0, \quad (\text{B.38})$$

entonces:

$$\begin{aligned} \frac{D\varphi}{D\underline{v}_k} = & - \frac{D \int_{T_k}^{T_{k+1}} (\underline{p}_k^j)^T (M_h \underline{\dot{z}}_k^j - A_h \underline{z}_k^j - B_h \underline{v}_k^j - c_h) dt}{D\underline{z}_k} \frac{\partial \underline{z}_k}{\partial \underline{v}_k} \\ & - \frac{D(\eta_k^j)^T (\underline{z}_k^j(T_k^+) - Z_k^i)}{D\underline{z}_k} \frac{\partial \underline{z}_k}{\partial \underline{v}_k}, \\ & - \frac{D \int_{T_k}^{T_{k+1}} \underline{p}_k^T (M_h \underline{\dot{z}}_k - A_h \underline{z}_k - B_h \underline{v}_k - c_h) dt}{D\underline{z}_k} \Bigg|_j \frac{\partial \underline{z}_k}{\partial \underline{v}_k} \\ & - \frac{D\eta_k^T (\underline{z}_k(T_k^+) - Z_k)}{D\underline{z}_k} \Bigg|_j \frac{\partial \underline{z}_k}{\partial \underline{v}_k}. \end{aligned} \quad (\text{B.39})$$

Reemplazando (B.39) en (B.36):

$$\begin{aligned} \frac{D\mathcal{J}}{D\underline{v}_k} \Bigg|_j &= \frac{D\mathcal{J}}{D\underline{v}_k} \Bigg|_j + \frac{D\varphi}{D\underline{v}_k}, \\ &= \frac{D\mathcal{J}}{D\underline{v}_k} \Bigg|_j + \frac{d \int_{T_k}^{T_{k+1}} (\underline{p}_k^j)^T (M_h \underline{\dot{z}}_k^j - A_h \underline{z}_k^j - B_h (\underline{v}_k^j + \lambda \phi_k) - c_h) dt}{d\lambda} \Bigg|_{\lambda=0}, \\ &= \frac{D\mathcal{J}}{D\underline{v}_k} \Bigg|_j - \int_{T_k}^{T_{k+1}} \underline{p}_k^T B_h \phi_k dt \Bigg|_j. \end{aligned} \quad (\text{B.40})$$

Utilizando la relación (B.20) en (B.40) tenemos:

$$\begin{aligned} \frac{D\mathcal{J}}{D\underline{v}_k} \Bigg|_j &= r \int_{T_k}^{T_{k+1}} \phi_k^T R_h \underline{v}_k dt \Bigg|_j - \int_{T_k}^{T_{k+1}} \underline{p}_k^T B_h \phi_k dt \Bigg|_j, \\ &\quad \int_{T_k}^{T_{k+1}} \phi_k^T (r R_h \underline{v}_k - B_h^T \underline{p}_k) dt \Bigg|_j. \end{aligned} \quad (\text{B.41})$$

Comparando (B.21) con (B.41) y utilizando la ecuación (B.34) tenemos:

$$\nabla_{\underline{v}_k} \mathcal{J} \Bigg|_j = \frac{\partial \mathcal{L}}{\partial \underline{v}_k} \Bigg|_j = (r R_h \underline{v}_k - B_h^T \underline{p}_k) \Bigg|_j. \quad (\text{B.42})$$

Ahora calcularemos la derivada total de \mathcal{J} respecto de Z_k :

$$\begin{aligned}\frac{D\mathcal{J}}{DZ_k} &= \frac{D\mathcal{J}}{DZ_k} + \frac{D\mathcal{J}}{Dz_k} \frac{\partial z_k}{\partial Z_k}, \\ &= (\nabla_{Z_k} \mathcal{J}, \phi_k).\end{aligned}\tag{B.43}$$

Reemplazando (B.35) en (B.43) tenemos:

$$\begin{aligned}\left. \frac{D\mathcal{J}}{DZ_k} \right|_j &= \left. \frac{D\mathcal{J}}{DZ_k} \right|_j - \left. \frac{D \int_{T_k}^{T_{k+1}} \underline{p}_k^T (M_h \dot{z}_k - A_h z_k - B_h v_k - c_h) dt}{Dz_k} \right|_j \frac{\partial z_k}{\partial Z_k} \\ &\quad - \left. \frac{D\eta_k^T(z_k(T_k^+) - Z_k)}{Dz_k} \right|_j \frac{\partial z_k}{\partial Z_k}.\end{aligned}\tag{B.44}$$

Calculando la derivada total de φ ecuación (B.37) respecto de Z_k tenemos:

$$\frac{D\varphi}{Dv_k} = \frac{D\varphi}{Dv_k} + \frac{D\varphi}{Dz_k} \frac{\partial z_k}{\partial Z_k} = 0,\tag{B.45}$$

entonces:

$$\begin{aligned}\frac{D\varphi}{DZ_k} &= - \frac{D \int_{T_k}^{T_{k+1}} (\underline{p}_k^j)^T (M_h \dot{z}_k^j - A_h z_k^j - B_h v_k^j - c_h) dt}{Dz_k} \frac{\partial z_k}{\partial Z_k} \\ &\quad - \frac{D(\eta_k^j)^T(z_k^j(T_k^+) - Z_k^j)}{Dz_k} \frac{\partial z_k}{\partial Z_k}, \\ &\quad - \left. \frac{D \int_{T_k}^{T_{k+1}} \underline{p}_k^T (M_h \dot{z}_k - A_h z_k - B_h v_k - c_h) dt}{Dz_k} \right|_j \frac{\partial z_k}{\partial Z_k} \\ &\quad - \left. \frac{D\eta_k^T(z_k(T_k^+) - Z_k)}{Dz_k} \right|_j \frac{\partial z_k}{\partial Z_k}.\end{aligned}\tag{B.46}$$

Reemplazando (B.46) en (B.44)

$$\begin{aligned} \left. \frac{\mathcal{D}\mathcal{J}}{\mathcal{D}Z_k} \right|_j &= \left. \frac{D\mathcal{J}}{DZ_k} \right|_j + \frac{D\varphi}{DZ_k}, \\ & \left. \frac{D\mathcal{J}}{DZ_k} \right|_j + \frac{d(\eta_k^j)^T \left(z_k^j(T_k^+) - (Z_k^i + \lambda\phi_k) \right)}{d\lambda} \Bigg|_{\lambda=0}, \\ & \left. \frac{D\mathcal{J}}{DZ_k} \right|_j - \eta_k^T \phi_k \Big|_j. \end{aligned} \quad (\text{B.47})$$

Entonces reemplazando (B.25) en (B.47) tenemos:

$$\begin{aligned} \left. \frac{\mathcal{D}\mathcal{J}}{\mathcal{D}Z_k} \right|_j &= -\frac{1}{\varepsilon \Delta T} \left(\phi_k^T M_h(z_{k-1}(T_k^-) - Z_k) \right) \Big|_j - \eta_k^T \phi_k \Big|_j, \\ & \phi_k^T \left(-\eta_k - \frac{1}{\varepsilon \Delta T} M_h(z_{k-1}(T_k^-) - Z_k) \right) \Big|_j. \end{aligned} \quad (\text{B.48})$$

Como condicionamos que se cumpla la ecuación (B.32), notemos que también se cumple la relación (B.28). Entonces (B.48) queda:

$$\left. \frac{\mathcal{D}\mathcal{J}}{\mathcal{D}Z_k} \right|_j = \phi_k^T \left(-M_h \underline{p}_k(T_k^+) - \frac{1}{\varepsilon \Delta T} M_h(z_{k-1}(T_k^-) - Z_k) \right) \Big|_j. \quad (\text{B.49})$$

Comparando (B.26) con (B.49) tenemos:

$$\nabla_{Z_k} \mathcal{J} \Big|_j = \frac{\partial \mathcal{L}}{\partial Z_k} \Big|_j = \left(-M_h \underline{p}_k(T_k^+) - \frac{1}{\varepsilon \Delta T} M_h(z_{k-1}(T_k^-) - Z_k) \right) \Big|_j. \quad (\text{B.50})$$

Las expresiones (B.42) y (B.50) son las componentes del gradiente de \mathcal{J} , es decir:

$$\nabla \mathcal{J} = \left[\nabla_{\underline{v}_k} \mathcal{J}^T \quad \nabla_{Z_k} \mathcal{J}^T \right]^T. \quad (\text{B.51})$$

B.3. Cálculo de Hd_k^i

La matriz hessiana del funcional es $H := \nabla^2 \mathcal{J}$, y por la definición (2.8) podemos expresar como sigue:

$$H = \begin{bmatrix} \frac{\partial \nabla_{\underline{v}_k} \mathcal{J}}{\partial \underline{v}_k} & \frac{\partial \nabla_{\underline{v}_k} \mathcal{J}}{\partial Z_k} \\ \frac{\partial \nabla_{Z_k} \mathcal{J}}{\partial \underline{v}_k} & \frac{\partial \nabla_{Z_k} \mathcal{J}}{\partial Z_k} \end{bmatrix} \quad (\text{B.52})$$

entonces Hd_k^i podemos expresar como:

$$H \begin{bmatrix} d\underline{v}_k \\ dz_k \end{bmatrix} = \begin{bmatrix} \frac{\partial \nabla_{\underline{v}_k} \mathcal{J}}{\partial \underline{v}_k} d\underline{v}_k + \frac{\partial \nabla_{\underline{v}_k} \mathcal{J}}{\partial Z_k} dz_k \\ \frac{\partial \nabla_{Z_k} \mathcal{J}}{\partial \underline{v}_k} d\underline{v}_k + \frac{\partial \nabla_{Z_k} \mathcal{J}}{\partial Z_k} dz_k \end{bmatrix} \quad (\text{B.53})$$

Para obtener el gradiente del funcional $\nabla \mathcal{J}$, debemos calcular previamente el estado adjunto \underline{p}_k y para obtener este, debemos calcular el estado \underline{z}_k .

Para obtener los estados \underline{z}_k debemos resolver la ecuación de estados:

$$\begin{aligned} \underline{z}_k(T_k^+) &= Z_k \\ M_h \dot{\underline{z}}_k &= A_h \underline{z}_k + B_h \underline{v}_k + c_h \end{aligned} \quad (\text{B.54})$$

Del mismo modo para obtener los estados adjuntos \underline{p}_k debemos resolver la ecuación de estado adjunto:

$$\begin{aligned} \underline{p}_k(T_{k+1}^-) &= -\frac{1}{\varepsilon \Delta T} (\underline{z}_k(T_{k+1}^-) - Z_{k+1}) \quad \forall k = 0, 1, \dots, \hat{k} - 2, \\ \underline{p}_{\hat{k}-1}(T_{\hat{k}}) &= -s \left(\underline{z}_{\hat{k}-1}(T_{\hat{k}}) - \tilde{y}_{\hat{k}-1}(T_{\hat{k}}) \right), \\ M_h \dot{\underline{p}}_k &= -A_h \underline{p}_k + q M_h (\underline{z}_k - \tilde{y}_k). \end{aligned} \quad (\text{B.55})$$

Notemos que las ecuaciones (B.54) y (B.55), nos permiten expresar \underline{z}_k y \underline{p}_k como sigue:

$$\begin{aligned}\underline{z}_k &= \mathcal{Z}(Z_k, \underline{v}_k, c_h), \\ \underline{p}_k &= \mathcal{P}(\underline{z}_k, Z_k, \tilde{y}_k),\end{aligned}\tag{B.56}$$

donde \mathcal{Z} y \mathcal{P} son operadores lineales dependientes del tiempo.

Expresando $\nabla \mathcal{J}$ en función de la ecuación (B.56) tenemos:

$$\nabla_{\underline{v}_k} \mathcal{J} = r R_h \underline{v}_k - B_h^T \mathcal{P}(\mathcal{Z}(Z_k, \underline{v}_k, c_h), Z_k, \tilde{y}_k)\tag{B.57}$$

$$\begin{aligned}\nabla_{Z_k} \mathcal{J} &= -M_h \mathcal{P}(\mathcal{Z}(Z_k, \underline{v}_k, c_h), Z_k, \tilde{y}_k)(T_k^+) \\ &\quad - \frac{1}{\varepsilon \Delta T} M_h \left(\mathcal{Z}(Z_{k-1}, \underline{v}_{k-1}, c_h)(T_k^-) - Z_k \right).\end{aligned}$$

Ahora calculamos $\frac{D\nabla \mathcal{J}}{Dv_k}$ utilizando la definición (2.1), entonces:

$$\begin{aligned}
 \frac{D\nabla_{\underline{v}_k} \mathcal{J}}{D\underline{v}_k} &= \left. \frac{d r R_h(\underline{v}_k + \lambda \phi_k)}{d\lambda} \right|_{\lambda=0} - \left. \frac{d B_h^T \mathcal{P}(\mathcal{Z}(Z_k, \underline{v}_k + \lambda \phi_k, c_h), Z_k, \tilde{y}_k)}{d\lambda} \right|_{\lambda=0}, \\
 &= r R_h \phi_k - B_h^T \mathcal{P}(\mathcal{Z}(0, \phi_k, 0), 0, 0), \tag{B.58} \\
 &= \frac{\partial \nabla_{\underline{v}_k} \mathcal{J}}{\partial \underline{v}_k} \phi_k
 \end{aligned}$$

$$\begin{aligned}
 \frac{D\nabla_{Z_k} \mathcal{J}}{D\underline{v}_k} &= - \left. \frac{d M_h \mathcal{P}(\mathcal{Z}(Z_k, \underline{v}_k + \lambda \phi_k, c_h), Z_k, \tilde{y}_k)(T_k^+)}{d\lambda} \right|_{\lambda=0} \\
 &\quad - \left. \frac{1}{\varepsilon \Delta T} \frac{d M_h \left(\mathcal{Z}(Z_{k-1}, \underline{v}_{k-1} + \lambda \phi_{k-1}, c_h)(T_k^-) - Z_k \right)}{d\lambda} \right|_{\lambda=0}. \\
 &= -M_h \mathcal{P}(\mathcal{Z}(0, \phi_k, 0), 0, 0)(T_k^+) - \frac{1}{\varepsilon \Delta T} M_h \mathcal{Z}(0, \phi_{k-1}, 0)(T_k^-) \tag{B.59} \\
 &= \frac{\partial \nabla_{Z_k} \mathcal{J}}{\partial \underline{v}_k} \phi_k.
 \end{aligned}$$

Del mismo modo calculamos $\frac{D\nabla\mathcal{J}}{DZ_k}$:

$$\begin{aligned}\frac{D\nabla_{\underline{v}_k}\mathcal{J}}{DZ_k} &= \left. \frac{dr R_h \underline{v}_k}{d\lambda} \right|_{\lambda=0} - \left. \frac{dB_h^T \mathcal{P}(\mathcal{Z}(Z_k + \lambda\phi_{z_k}, \underline{v}_k, c_h), Z_k + \lambda\phi_{z_k}, \tilde{y}_k)}{d\lambda} \right|_{\lambda=0}, \\ &= -B_h^T \mathcal{P}(\mathcal{Z}(\phi_{z_k}, 0, 0), \phi_{z_k}, 0), \\ &= \frac{\partial \nabla_{\underline{v}_k} \mathcal{J}}{\partial Z_k} \phi_{z_k}\end{aligned}\tag{B.60}$$

$$\begin{aligned}\frac{D\nabla_{Z_k}\mathcal{J}}{DZ_k} &= - \left. \frac{dM_h \mathcal{P}(\mathcal{Z}(Z_k + \lambda\phi_{z_k}, \underline{v}_k, c_h), Z_k + \lambda\phi_{z_k}, \tilde{y}_k)(T_k^+)}{d\lambda} \right|_{\lambda=0} \\ &\quad - \left. \frac{1}{\varepsilon \Delta T} \frac{dM_h \left(\mathcal{Z}(Z_{k-1} + \lambda\phi_{z_{k-1}}, \underline{v}_{k-1}, c_h)(T_k^-) - (Z_k + \lambda\phi_{z_k}) \right)}{d\lambda} \right|_{\lambda=0}. \\ &= -M_h \mathcal{P}(\mathcal{Z}(\phi_{z_k}, 0, 0), \phi_{z_k}, 0)(T_k^+) \\ &\quad - \frac{1}{\varepsilon \Delta T} M_h \left(\mathcal{Z}(\phi_{z_{k-1}}, 0, 0)(T_k^-) - \phi_{z_k} \right), \\ &= \frac{\partial \nabla_{Z_k} \mathcal{J}}{\partial Z_k} \phi_{z_k}.\end{aligned}\tag{B.61}$$

Como ϕ_k y ϕ_{z_k} de las ecuaciones (B.58), (B.59), (B.60) y (B.61) son direcciones arbitrarias, tomamos a los mismos como las direcciones de búsqueda de las funciones de control y de las condiciones iniciales respectivamente, es decir:

$$\begin{aligned}\phi_k &= \underline{dv}_k, \\ \phi_{z_k} &= dz_k.\end{aligned}\tag{B.62}$$

Luego si sumamos las ecuaciones (B.58) y (B.60) tenemos:

$$\begin{aligned}\frac{\partial \nabla_{\underline{v}_k} \mathcal{J}}{\partial \underline{v}_k} \underline{dv}_k + \frac{\partial \nabla_{Z_k} \mathcal{J}}{\partial Z_k} dz_k &= r R_h \underline{dv}_k - B_h^T \mathcal{P}(\mathcal{Z}(0, \underline{dv}_k, 0), 0, 0) \\ &\quad - B_h^T \mathcal{P}(\mathcal{Z}(dz_k, 0, 0), dz_k, 0),\end{aligned}\tag{B.63}$$

como ya mencionamos \mathcal{Z} y \mathcal{P} son operadores lineales, entonces sumando los términos de la ecuación (B.63) tenemos:

$$\frac{\partial \nabla_{\underline{v}_k} \mathcal{J}}{\partial \underline{v}_k} \underline{d}v_k + \frac{\partial \nabla_{\underline{z}_k} \mathcal{J}}{\partial \underline{z}_k} dz_k = r R_h \underline{d}v_k - B_h^T \mathcal{P}(\mathcal{Z}(dz_k, \underline{d}v_k, 0), dz_k, 0). \quad (\text{B.64})$$

Además si sumamos las ecuaciones (B.59) y (B.61) tenemos:

$$\begin{aligned} \frac{\partial \nabla_{\underline{z}_k} \mathcal{J}}{\partial \underline{v}_k} \underline{d}v_k + \frac{\partial \nabla_{\underline{z}_k} \mathcal{J}}{\partial \underline{z}_k} dz_k &= -M_h \mathcal{P}(\mathcal{Z}(0, \underline{d}v_k, 0), 0, 0)(T_k^+) \\ &\quad - \frac{1}{\varepsilon \Delta T} M_h \mathcal{Z}(0, \underline{d}v_{k-1}, 0)(T_k^-) \\ &\quad - M_h \mathcal{P}(\mathcal{Z}(dz_k, 0, 0), dz_k, 0)(T_k^+) \quad (\text{B.65}) \\ &\quad - \frac{1}{\varepsilon \Delta T} M_h (\mathcal{Z}(dz_{k-1}, 0, 0)(T_k^-) - dz_k), \end{aligned}$$

al sumar los operadores lineales resulta:

$$\begin{aligned} \frac{\partial \nabla_{\underline{z}_k} \mathcal{J}}{\partial \underline{v}_k} \underline{d}v_k + \frac{\partial \nabla_{\underline{z}_k} \mathcal{J}}{\partial \underline{z}_k} dz_k &= -M_h \mathcal{P}(\mathcal{Z}(dz_k, \underline{d}v_k, 0), dz_k, 0)(T_k^+) \quad (\text{B.66}) \\ &\quad - \frac{1}{\varepsilon \Delta T} M_h (\mathcal{Z}(dz_{k-1}, \underline{d}v_{k-1}, 0)(T_k^-) - dz_k). \end{aligned}$$

Las ecuaciones (B.64) y (B.66) nos indican que para obtener Hd_k^i necesitamos calcular previamente $\mathcal{Z}(dz_k, \underline{d}v_k, 0)$ el cual calculamos a partir de las ecuaciones de estado, haciendo las condiciones iniciales y las funciones de control iguales a sus direcciones de búsqueda respectivos, es decir:

$$\begin{aligned} \underline{\psi}_k(T_k^+) &= dz_k, \\ M_h \underline{\dot{\psi}}_k &= A_h \underline{\psi}_k + B_h \underline{d}v_k, \quad (\text{B.67}) \end{aligned}$$

donde $\underline{\psi}_k = \mathcal{Z}(dz_k, \underline{d}v_k, 0)$.

Luego necesitamos calcular y $\mathcal{P}(\mathcal{Z}(dz_k, \underline{dv}_k, 0), dz_k, 0)$, el cual calculamos a partir de los estados adjuntos, con los siguientes cambios:

$$\begin{aligned}\underline{\psi}_{p_k}(T_{k+1}^-) &= -\frac{1}{\varepsilon \Delta T}(\underline{\psi}_k(T_{k+1}^-) - dz_{k+1}) \quad \forall k = 0, 1, \dots, \hat{k} - 2, \\ \underline{\psi}_{p_{\hat{k}-1}}(T_{\hat{k}}) &= -s\underline{\psi}_{\hat{k}-1}(T_{\hat{k}}), \\ M_h \dot{\underline{\psi}}_{p_k} &= -A_h \underline{\psi}_{p_k} + q M_h \underline{\psi}_k,\end{aligned}\tag{B.68}$$

donde $\underline{\psi}_{p_k} = \mathcal{P}(\mathcal{Z}(dz_k, \underline{dv}_k, 0), dz_k, 0)$.

Luego las ecuaciones (B.64) y (B.66) quedan:

$$\begin{aligned}\frac{\partial \nabla_{v_k} \mathcal{J}}{\partial v_k} \underline{dv}_k + \frac{\partial \nabla_{v_k} \mathcal{J}}{\partial Z_k} dz_k &= r R_h \underline{dv}_k - B_h^T \underline{\psi}_{p_k}. \\ \frac{\partial \nabla_{z_k} \mathcal{J}}{\partial v_k} \underline{dv}_k + \frac{\partial \nabla_{z_k} \mathcal{J}}{\partial Z_k} dz_k &= -M_h \underline{\psi}_{p_k}(T_k^+) - \frac{1}{\varepsilon \Delta T} M_h (\underline{\psi}_{k-1}(T_k^-) - dz_k).\end{aligned}\tag{B.69}$$

Reemplazando (B.69) en (B.53) tenemos:

$$H \begin{bmatrix} \underline{dv}_k \\ dz_k \end{bmatrix} = \begin{bmatrix} r R_h \underline{dv}_k - B_h^T \underline{\psi}_{p_k} \\ -M_h \underline{\psi}_{p_k}(T_k^+) - \frac{1}{\varepsilon \Delta T} M_h (\underline{\psi}_{k-1}(T_k^-) - dz_k) \end{bmatrix}\tag{B.70}$$

Separando por componentes tenemos:

$$\begin{aligned}H_v \underline{dv}_k &= r R_h \underline{dv}_k - B_h^T \underline{\psi}_{p_k} \\ H_z dz_k &= -M_h \underline{\psi}_{p_k}(T_k^+) - \frac{1}{\varepsilon \Delta T} M_h (\underline{\psi}_{k-1}(T_k^-) - dz_k).\end{aligned}\tag{B.71}$$

B.4. Matriz de Precondicionamiento

Para deducir la matriz P propuesta en [1] primeramente obtenemos de la ecuación de estados en forma discreta, un método para hallar el valor de $z_{k,\hat{i}}$ a partir de Z_k .

El esquema iterativo para hallar el estado z_k es:

$$z_{k,0} = Z_k$$

$$F_1 z_{k,l} = F_0 z_{k,l-1} + \tau B_h v_{k,l} + \tau c_h; \quad \text{para } l = 1 \text{ hasta } \hat{l} \quad (\text{B.72})$$

con el cual se obtiene cada valor de $z_{k,l}$ de la siguiente forma:

$$z_{k,1} = F_1^{-1} F_0 z_{k,0} + \tau F_1^{-1} (B_h v_{k,1} + c_h) = F_1^{-1} F_0 Z_k + \tau F_1^{-1} (B_h v_{k,1} + c_h).$$

$$z_{k,2} = F_1^{-1} F_0 z_{k,1} + \tau F_1^{-1} (B_h v_{k,2} + c_h),$$

$$= F_1^{-1} F_0 (F_1^{-1} F_0 Z_k + \tau F_1^{-1} (B_h v_{k,1} + c_h)) + \tau F_1^{-1} (B_h v_{k,2} + c_h),$$

$$= (F_1^{-1} F_0)^2 Z_k + \tau (F_1^{-1} F_0) F_1^{-1} (B_h v_{k,1} + c_h) + \tau F_1^{-1} (B_h v_{k,2} + c_h).$$

$$z_{k,3} = F_1^{-1} F_0 z_{k,2} + \tau F_1^{-1} (B_h v_{k,3} + c_h),$$

$$= (F_1^{-1} F_0)^3 Z_k + \tau (F_1^{-1} F_0)^2 F_1^{-1} (B_h v_{k,1} + c_h) + \tau (F_1^{-1} F_0) F_1^{-1} (B_h v_{k,2} + c_h)$$

$$+ \tau F_1^{-1} (B_h v_{k,3} + c_h),$$

$$= (F_1^{-1} F_0)^3 Z_k + \tau \sum_{i=1}^3 (F_1^{-1} F_0)^{(3-i)} F_1^{-1} (B_h v_{k,i} + c_h).$$

$$\vdots$$

$$\vdots$$

$$z_{k,\hat{l}} = (F_1^{-1} F_0)^{\hat{l}} Z_k + \tau \sum_{i=1}^{\hat{l}} (F_1^{-1} F_0)^{(\hat{l}-i)} F_1^{-1} (B_h v_{k,i} + c_h), \quad (\text{B.73})$$

entonces podemos expresar $z_{k,\hat{l}}$ como:

$$z_{k,\hat{l}} = \mathcal{F}_{\Delta\tau} Z_k + S_k \quad (\text{B.74})$$

donde $\mathcal{F}_{\Delta\tau} = (F_1^{-1} F_0)^{\hat{l}}$ y $S_k = \tau \sum_{i=1}^{\hat{l}} (F_1^{-1} F_0)^{(\hat{l}-i)} F_1^{-1} (B_h v_{k,i} + c_h)$.

Para resolver el problema de control paralelo se introduce la condición (3.4),

el cual expresando en su forma discreta establece:

$$\begin{aligned} Z_0 &= z_{0,0} = z(t_0), & \text{si } k &= 0; \\ Z_k &= z_{k-1,\hat{l}}, & \text{si } k &= 1, 2, \dots, \hat{k} - 1, \end{aligned} \quad (\text{B.75})$$

entonces podemos hallar las condiciones iniciales $Z_k, \forall k = 1, 2, \dots, \hat{k} - 1$ utilizando la ecuación (B.74) mediante la siguiente ecuación matricial:

$$\begin{bmatrix} I & & & & \\ -\mathcal{F}_{\Delta\tau} & I & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & -\mathcal{F}_{\Delta\tau} & I \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_{\hat{k}-1} \end{bmatrix} = \begin{bmatrix} S_0 + \mathcal{F}_{\Delta\tau} Z_0 \\ S_1 \\ \vdots \\ S_{\hat{k}-2} \end{bmatrix} \quad (\text{B.76})$$

$\mathbf{CZ} = \mathbf{S}.$

Seguidamente hallaremos a partir de la ecuación adjunta discreta, un método para hallar $p_{k,0}$ a partir de $p_{k,\hat{l}}$.

El esquema iterativo para hallar el estado p_k es:

$$\begin{aligned} p_{k,\hat{l}} &= -\frac{1}{\varepsilon \Delta T} (z_{k,\hat{l}} - Z_{k+1}) \quad \forall k = 0, 1, \dots, \hat{k} - 2, \\ p_{\hat{k}-1,\hat{l}} &= -s(z_{\hat{k}-1,\hat{l}} - \tilde{y}_{\hat{k}-1,\hat{l}}), \\ F_1 p_{k,l} &= F_0 p_{k,l+1} + \tau q M_h(z_{k,l} - \tilde{y}_{k,l}); \quad \text{para } l = \hat{l} - 1 \text{ hasta } 0. \end{aligned} \quad (\text{B.77})$$

entonces podemos obtener $p_{k,0}$ de la siguiente forma:

$$\begin{aligned} p_{k,\hat{l}-1} &= F_1^{-1} F_0 p_{k,\hat{l}} + \tau q F_1^{-1} M_h(z_{k,\hat{l}-1} + \tilde{y}_{k,\hat{l}-1}), \\ p_{k,\hat{l}-2} &= F_1^{-1} F_0 p_{k,\hat{l}-1} + \tau q F_1^{-1} M_h(z_{k,\hat{l}-2} + \tilde{y}_{k,\hat{l}-2}), \end{aligned}$$

$$\begin{aligned}
 p_{k,\hat{l}-2} &= F_1^{-1}F_0 \left(F_1^{-1}F_0 p_{k,\hat{l}} + \tau q F_1^{-1} M_h(z_{k,\hat{l}-1} + \tilde{y}_{k,\hat{l}-1}) \right) \\
 &\quad + \tau q F_1^{-1} M_h(z_{k,\hat{l}-2} + \tilde{y}_{k,\hat{l}-2}), \\
 &= (F_1^{-1}F_0)^2 p_{k,\hat{l}} + \tau q (F_1^{-1}F_0) F_1^{-1} M_h(z_{k,\hat{l}-1} + \tilde{y}_{k,\hat{l}-1}) \\
 &\quad + \tau q F_1^{-1} M_h(z_{k,\hat{l}-2} + \tilde{y}_{k,\hat{l}-2}). \\
 p_{k,\hat{l}-3} &= F_1^{-1}F_0 p_{k,\hat{l}-2} + \tau q F_1^{-1} M_h(z_{k,\hat{l}-3} + \tilde{y}_{k,\hat{l}-3}), \\
 &= (F_1^{-1}F_0)^3 p_{k,\hat{l}} + \tau q (F_1^{-1}F_0)^2 F_1^{-1} M_h(z_{k,\hat{l}-1} + \tilde{y}_{k,\hat{l}-1}) \\
 &\quad + \tau q (F_1^{-1}F_0) F_1^{-1} M_h(z_{k,\hat{l}-2} + \tilde{y}_{k,\hat{l}-2}) + \tau q F_1^{-1} M_h(z_{k,\hat{l}-3} + \tilde{y}_{k,\hat{l}-3}), \\
 &= (F_1^{-1}F_0)^3 p_{k,\hat{l}} + \tau q \sum_{i=1}^3 (F_1^{-1}F_0)^{(3-i)} F_1^{-1} M_h(z_{k,\hat{l}-i} + \tilde{y}_{k,\hat{l}-i}). \\
 &\quad \vdots \\
 p_{k,0} &= (F_1^{-1}F_0)^{\hat{l}} p_{k,\hat{l}} + \tau q \sum_{i=1}^{\hat{l}} (F_1^{-1}F_0)^{(\hat{l}-i)} F_1^{-1} M_h(z_{k,\hat{l}-i} + \tilde{y}_{k,\hat{l}-i}). \tag{B.78}
 \end{aligned}$$

Luego podemos expresar $p_{k,0}$ de la siguiente forma:

$$p_{k,0} = \mathcal{F}_{\Delta\tau} p_{k,\hat{l}} + S_{p_k}. \tag{B.79}$$

En el apéndice (B.2) se obtuvo el valor de $\nabla_{Z_k} \mathcal{J}$, dado por la ecuación (B.50), que en forma discreta podemos expresar como sigue:

$$\nabla_{Z_k} \mathcal{J} = M_h \left(-p_{k,0} + p_{k-1,\hat{l}} \right), \quad \forall k = 1, 2, \dots, \hat{k} - 2, \tag{B.80}$$

entonces definimos $\nabla_Z \mathcal{J} := \left[\nabla_{Z_1} \mathcal{J}^T \quad \nabla_{Z_2} \mathcal{J}^T \quad \dots \quad \nabla_{Z_{\hat{k}-1}} \mathcal{J}^T \right]^T$, es decir:

$$\nabla_Z \mathcal{J} = \begin{bmatrix} M_h(p_{0,\hat{i}} - p_{1,0}) \\ M_h(p_{1,\hat{i}} - p_{2,0}) \\ \vdots \\ M_h(p_{\hat{k}-2,\hat{i}} - p_{\hat{k}-1,0}) \end{bmatrix} = \begin{bmatrix} M_h(p_{0,\hat{i}} - \mathcal{F}_{\Delta\tau} p_{1,\hat{i}} - S_{p_1}) \\ M_h(p_{1,\hat{i}} - \mathcal{F}_{\Delta\tau} p_{2,\hat{i}} - S_{p_2}) \\ \vdots \\ M_h(p_{\hat{k}-2,\hat{i}} - \mathcal{F}_{\Delta\tau} p_{\hat{k}-1,\hat{i}} - S_{p_{\hat{k}-1}}) \end{bmatrix} \quad (\text{B.81})$$

teniendo en cuenta que $F_0 = M_h$ y $\mathcal{F}_{\Delta\tau} = (F_1^{-1} F_0)^{\hat{i}}$ en la ecuación (B.81) tenemos:

$$\nabla_Z \mathcal{J} = \begin{bmatrix} F_0 p_{0,\hat{i}} - F_0 (F_1^{-1} F_0)^{\hat{i}} p_{1,\hat{i}} - F_0 S_{p_1} \\ F_0 p_{1,\hat{i}} - F_0 (F_1^{-1} F_0)^{\hat{i}} p_{2,\hat{i}} - F_0 S_{p_2} \\ \vdots \\ F_0 p_{\hat{k}-2,\hat{i}} - F_0 (F_1^{-1} F_0)^{\hat{i}} p_{\hat{k}-1,\hat{i}} - F_0 S_{p_{\hat{k}-1}} \end{bmatrix}$$

$$\nabla_Z \mathcal{J} = \begin{bmatrix} F_0 p_{0,\hat{i}} - (F_0 F_1^{-1})^{\hat{i}} F_0 p_{1,\hat{i}} - F_0 S_{p_1} \\ F_0 p_{1,\hat{i}} - (F_0 (F_1^{-1})^{\hat{i}} F_0) p_{2,\hat{i}} - F_0 S_{p_2} \\ \vdots \\ F_0 p_{\hat{k}-2,\hat{i}} - (F_0 F_1^{-1})^{\hat{i}} F_0 p_{\hat{k}-1,\hat{i}} - F_0 S_{p_{\hat{k}-1}} \end{bmatrix} \quad (\text{B.82})$$

notemos que $(F_0 F_1^{-1})^{\hat{i}} = ((F_1^{-1} F_0)^{\hat{i}})^T = \mathcal{F}_{\Delta\tau}^T$, entonces la ecuación (B.82) queda:

$$\nabla_Z \mathcal{J} = \begin{bmatrix} F_0 p_{0,\hat{i}} - \mathcal{F}_{\Delta\tau}^T F_0 p_{1,\hat{i}} - F_0 S_{p_1} \\ F_0 p_{1,\hat{i}} - \mathcal{F}_{\Delta\tau}^T F_0 p_{2,\hat{i}} - F_0 S_{p_2} \\ \vdots \\ F_0 p_{\hat{k}-2,\hat{i}} - \mathcal{F}_{\Delta\tau}^T F_0 p_{\hat{k}-1,\hat{i}} - F_0 S_{p_{\hat{k}-1}} \end{bmatrix} \quad (\text{B.83})$$

que podemos expresar como:

$$\begin{aligned} \nabla_Z \mathcal{J} &= \begin{bmatrix} I & -\mathcal{F}_{\Delta\tau}^T & & \\ & I & \ddots & \\ & & \ddots & -\mathcal{F}_{\Delta\tau}^T \\ & & & I \end{bmatrix} \begin{bmatrix} F_0 p_{0,\hat{i}} \\ F_0 p_{1,\hat{i}} \\ \vdots \\ F_0 p_{\hat{k}-2,\hat{i}} \end{bmatrix} - \begin{bmatrix} F_0 S_{p_1} \\ F_0 S_{p_2} \\ \vdots \\ F_0 (S_{p_{\hat{k}-1}} + \mathcal{F}_{\Delta\tau}^T p_{\hat{k}-1,\hat{i}}) \end{bmatrix} \\ &= \mathbf{C}^T \mathbf{P}_{F_0} - \mathbf{S}_p \end{aligned} \quad (\text{B.84})$$

utilizando condición final $p_{k,\hat{i}}$ tenemos:

$$\begin{aligned} \begin{bmatrix} F_0 p_{0,\hat{i}} \\ F_0 p_{1,\hat{i}} \\ \vdots \\ F_0 p_{\hat{k}-2,\hat{i}} \end{bmatrix} &= \begin{bmatrix} \frac{1}{\varepsilon \Delta T} F_0 (z_{0,\hat{i}} - Z_1) \\ \frac{1}{\varepsilon \Delta T} F_0 (z_{1,\hat{i}} - Z_2) \\ \vdots \\ \frac{1}{\varepsilon \Delta T} F_0 (z_{\hat{k}-2,\hat{i}} - Z_{\hat{k}-1}) \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{\varepsilon \Delta T} F_0 & & & \\ & \frac{1}{\varepsilon \Delta T} F_0 & & \\ & & \ddots & \\ & & & \frac{1}{\varepsilon \Delta T} F_0 \end{bmatrix} \begin{bmatrix} z_{0,\hat{i}} - Z_1 \\ z_{1,\hat{i}} - Z_2 \\ \vdots \\ z_{\hat{k}-2,\hat{i}} - Z_{\hat{k}-1} \end{bmatrix} \\ \mathbf{P}_{F_0} &= \mathbf{K} (\mathbf{CZ} - \mathbf{S}). \end{aligned} \quad (\text{B.85})$$

Reemplazando (B.85) en (B.84) tenemos:

$$\nabla_Z \mathcal{J} = \mathbf{C}^T \mathbf{K} \mathbf{CZ} - \mathbf{F}, \quad (\text{B.86})$$

donde $\mathbf{F} = \mathbf{C}^T \mathbf{K} \mathbf{S} + \mathbf{S}_p$.

Para poder hallar las condiciones iniciales de cada sub-dominio del problema

P2 (3.6) se debe cumplir que $\nabla_Z \mathcal{J} = 0$, lo cual es equivalente a resolver la

siguiente ecuación matricial:

$$\mathbf{C}^T \mathbf{K} \mathbf{C} \mathbf{Z} = \mathbf{F}. \quad (\text{B.87})$$

Para resolver (B.87) podemos invertir el producto de matrices $\mathbf{C}^T \mathbf{K} \mathbf{C}$, pero ésto presenta un elevado costo computacional, debido a que $\mathcal{F}_{\Delta\tau}$ se obtiene marchando la malla fina en pasos de tamaño τ .

Podemos obtener \mathbf{Z} de forma aproximada por medio de la discretización de la malla gruesa por el método de Euler Implícito (Backward Euler), que expresamos como sigue:

$$\begin{aligned} M_h \frac{Z_{k+1} - Z_k}{\Delta T} &= A_h Z_{k+1} + B_h \underline{v}_{k,0} + c_h, \\ Z_{k+1} &= (M_h - \Delta T A_h)^{-1} M_h Z_k + (M_h - \Delta T A_h)^{-1} \Delta T (B_h \underline{v}_{k,0} + c_h). \end{aligned} \quad (\text{B.88})$$

Entonces la aproximación de \mathbf{Z} podemos expresar mediante la siguiente ecuación matricial:

$$\tilde{\mathbf{C}} \mathbf{Z} = \tilde{\mathbf{S}},$$

donde:

$$\tilde{\mathbf{C}} = \begin{bmatrix} I & & & & \\ -\mathcal{G}_{\Delta T} & I & & & \\ & & \ddots & \ddots & \\ & & & & -\mathcal{G}_{\Delta T} & I \end{bmatrix}; \quad \mathcal{G}_{\Delta T} = (M_h - \Delta T A_h)^{-1} M_h. \quad (\text{B.89})$$

Notemos que \tilde{C} es una aproximación de C , además invertir \tilde{C} puede calcularse rápidamente por presentar un bajo costo computacional, debido a que $\mathcal{G}_{\Delta T}$ se obtiene marchando la malla gruesa con pasos de tamaño ΔT . Entonces para acelerar la convergencia del algoritmo gradiente conjugado utilizamos la siguiente matriz como preconditionador:

$$\mathbf{P} = \tilde{C}^T \mathbf{K} \tilde{C}. \quad (\text{B.90})$$

APÉNDICE C

APÉNDICE: DISCRETIZACIÓN ESPACIAL

C.1. Ecuación de conducción de calor con condiciones de frontera de Neumann

C.1.1. Formulación Variacional

Primeramente necesitamos obtener la formulación variacional, para ello introducimos las siguientes definiciones:

Definición C.1 (Producto Interno). *Dado un espacio vectorial H sobre K , una función $(\cdot, \cdot) : H \times H \rightarrow K$ se llama producto interno si cumple con:*

$$(a) \quad (x, y) = \overline{(y, x)} \quad \forall x, y \in H,$$

$$(b) \quad (\alpha x + \beta x', y) = \alpha(x, y) + \beta(x', y) \quad \forall x, x', y \in E, \alpha, \beta \in K,$$

$$(c) \quad (x, \alpha y + \beta y') = \bar{\alpha}(x, y) + \bar{\beta}(x, y') \quad \forall x, y, y' \in H, \alpha, \beta \in K,$$

$$(d) \quad (x, x) \geq 0 \quad \forall x \in H \quad \text{y} \quad (x, x) = 0 \Leftrightarrow x = 0,$$

donde $\bar{\alpha}$ es el complejo conjugado de α , para todo $\alpha \in K$.

Con un producto interno, siempre podemos definir una norma en H , poniendo

$$|x| = \sqrt{(x, x)}.$$

Luego, con la norma definida, podemos definir una métrica, según como hemos visto en el ítem anterior, por tanto, un espacio con producto interno siempre puede verse como un espacio métrico, esto es, estamos habilitados a caracterizar la “cercanía” o “parecido” entre los elementos.

Definición C.2 (Ortogonalidad de funciones). *Dadas dos funciones u y v definidas en Ω , decimos que son ortogonales si:*

$$(u, v) = 0 \quad \text{sobre } \Omega. \quad (\text{C.1})$$

Definición C.3 (Espacio de Sobolev $H^1(\Omega)$ de orden 1). *Sea $\Omega \subset R^2$, definimos el espacio de Sobolev $H^1(\Omega)$ de orden 1, que es el espacio de funciones cuyos elementos y sus derivadas son cuadrado integrables.*

$$H^1(\Omega) = \left\{ f \in L^2(\Omega) \text{ y } \frac{\partial f}{\partial x_i} \in L^2(\Omega), i = 1, 2 \right\}. \quad (\text{C.2})$$

El objetivo es obtener una solución numérica aproximada $z \in H^1(\Omega)$, expresada como combinación lineal de una base del espacio de funciones de prueba $w \in H^1$. Nótese que el espacio H^1 es de dimensión infinita.

Luego, definimos r el residuo de la ecuación (4.21) como:

$$r = \rho c_p \partial_t z - \lambda \Delta z - c, \quad \text{en } \Omega \times [t_o, t_f], \quad (\text{C.3})$$

donde si z es la solución de exacta de (4.21) $r = 0$.

Entonces el producto interno entre el residuo (C.3) y las funciones de prueba w es igual a:

$$(r, w) = (\rho c_p \partial_t z - \lambda \Delta z - c, w), \quad \text{en } \Omega \times [t_o, t_f], \quad (\text{C.4})$$

y como buscamos una solución aproximada z tal que $(r, w) = 0 \quad \forall w \in H^1(\Omega)$, es decir que r sea ortogonal al espacio de las funciones de prueba, tenemos:

$$(\rho c_p \partial_t z - \lambda \Delta z - c, w) = 0, \quad \text{en } \Omega \times [t_o, t_f] \quad \forall w \in H^1(\Omega), \quad (\text{C.5})$$

lo cual es equivalente a:

$$(\rho c_p \partial_t z - \lambda \Delta z, w) = (c, w) \quad \text{en } \Omega \times [t_o, t_f] \quad \forall w \in H^1(\Omega). \quad (\text{C.6})$$

El producto interno que utilizamos en la ecuación (C.6) es:

$$(f, g) = \int_{\Omega} f(\Omega)g(\Omega) d\Omega, \quad (\text{C.7})$$

entonces la ecuación (C.6) puede ser re-escrita como:

$$\int_{\Omega} (\rho c_p \partial_t z - \lambda \Delta z) w d\Omega = \int_{\Omega} c w d\Omega \quad \forall w \in H^1(\Omega), \quad (\text{C.8})$$

o bien:

$$\int_{\Omega} \rho c_p \partial_t z w d\Omega = \int_{\Omega} \lambda \Delta z w d\Omega + \int_{\Omega} c w d\Omega \quad \forall w \in H^1(\Omega), \quad (\text{C.9})$$

Además integrando por partes:

$$\int_{\Omega} (\Delta z) w d\Omega = \int_{\Gamma} \frac{\partial z}{\partial \eta} w d\Gamma - \int_{\Omega} \nabla z \nabla w d\Omega, \quad (\text{C.10})$$

reemplazando (C.10) en (C.9) obtenemos:

$$\int_{\Omega} \rho c_p \partial_t z w d\Omega = \int_{\Gamma} v w d\Gamma - \int_{\Omega} \lambda \nabla z \nabla w d\Omega + \int_{\Omega} c w d\Omega \quad w \in H^1(\Omega) \quad (\text{C.11})$$

C.1.2. Discretización por el método de Galerkin

La discretización en el dominio espacial de (C.11) por el método estándar de Galerkin consiste en reemplazar el espacio de dimensión infinita $H^1(\Omega)$ por un sub-espacio de dimensión finita Y_h que aproxima a $H^1(\Omega)$ mediante un conjunto de funciones de prueba linealmente independientes y lineales por partes como en [20].

Sean $\{\eta_1, \eta_2, \dots, \eta_{\hat{q}}\}$ la base del espacio Y_h de dimensión \hat{q} . Entonces la discretización de (C.11) es:

$$\int_{\Omega} \rho c_p \partial_t z \eta_j d\Omega = \int_{\Gamma} v \eta_j d\Gamma - \int_{\Omega} \lambda \nabla z \nabla \eta_j d\Omega + \int_{\Omega} c \eta_j d\Omega \quad \forall j = 1, \dots, \hat{q}, \quad (\text{C.12})$$

luego para obtener una aproximación de $z \in H^1(\Omega) \times [t_o, t_f]$, proyectamos también el espacio $H^1(\Omega)$ sobre Y_h , haciendo la siguiente separación de variables:

$$z(t, \Omega) \approx z_h(t, \Omega) = \sum_{k=1}^{\hat{q}} z_k(t) \eta_k(\Omega) = \sum_{k=1}^{\hat{q}} z_k \eta_k, \quad (\text{C.13})$$

y la derivada temporal es:

$$\partial_t z \approx \partial_t \sum_{k=1}^{\hat{q}} z_k(t) \eta_k(\Omega) = \sum_{k=1}^{\hat{q}} \dot{z}_k \eta_k, \quad (\text{C.14})$$

donde las incógnitas a determinar podemos agrupar en un vector que definimos a continuación:

$$\underline{z} = [z_1(t) \ z_2(t) \ \dots \ z_{\hat{q}}(t)]^T, \quad (\text{C.15})$$

reemplazando (C.13) y (C.14) en (C.12):

$$\int_{\Omega} \sum_{k=1}^{\hat{q}} \rho c_p \dot{z}_k \eta_k \eta_j d\Omega = \int_{\Gamma} v \eta_j d\Gamma - \int_{\Omega} \lambda \nabla \sum_{k=1}^{\hat{q}} z_k \eta_k \nabla \eta_j d\Omega + \int_{\Omega} c \eta_j d\Omega \quad \forall j = 1, \dots, \hat{q}. \quad (\text{C.16})$$

Además como la función de control $v \in L^2(\Gamma) \times [t_o, t_f]$, proyectaremos también éste espacio de dimensión infinita sobre un espacio de dimensión finita U_h de funciones constantes por partes. Sea $\{\varphi_1, \varphi_2, \dots, \varphi_{\hat{p}}\}$ la base del espacio de dimensión \hat{p} . Entonces la proyección de v sobre U_h se puede expresar mediante la siguiente separación de variables:

$$v(t, \Gamma) \approx v_h(t, \Gamma) = \sum_{k=1}^{\hat{p}} v_k(t) \varphi_k(\Gamma) = \sum_{k=1}^{\hat{p}} v_k \varphi_k, \quad (\text{C.17})$$

donde las proyecciones pueden ser agrupadas en un vector que definimos a continuación:

$$\underline{v} = [v_1(t) \ v_2(t) \ \dots \ v_{\hat{p}}(t)]^T, \quad (\text{C.18})$$

y reemplazando (C.17) en (C.12) tenemos:

$$\int_{\Omega} \sum_{k=1}^{\hat{q}} \rho c_p \dot{z}_k \eta_k \eta_j \, d\Omega = \int_{\Gamma} \sum_{k=1}^{\hat{p}} v_k \varphi_k \eta_j \, d\Gamma - \int_{\Omega} \lambda \nabla \sum_{k=1}^{\hat{q}} z_k \eta_k \nabla \eta_j \, d\Omega + \int_{\Omega} c \eta_j \, d\Omega \quad \forall j = 1, \dots, \hat{q}. \quad (\text{C.19})$$

Como los términos z_k y v_k ya solo dependen del tiempo, pueden quedar fuera de la integración:

$$\sum_{k=1}^{\hat{q}} \dot{z}_k \int_{\Omega} \rho c_p \eta_k \eta_j \, d\Omega = \sum_{k=1}^{\hat{p}} v_k \int_{\Gamma} \varphi_k \eta_j \, d\Gamma - \sum_{k=1}^{\hat{q}} z_k \int_{\Omega} \lambda \nabla \eta_k \nabla \eta_j \, d\Omega + \int_{\Omega} c \eta_j \, d\Omega \quad \forall j = 1, \dots, \hat{q}. \quad (\text{C.20})$$

Finalmente nótese que la ecuación (C.20) es un sistema lineal de ecuaciones diferenciales de primer orden y se puede representarlo en forma matricial, como sigue:

$$M_h \dot{\underline{z}} = A_h \underline{z} + B_h \underline{v} + c_h \quad (\text{C.21})$$

donde:

$$M_{h_{jk}} = \int_{\Omega} \rho c_p \eta_k \eta_j d\Omega \quad \forall j, k = 1, \dots, \hat{q}, \quad (\text{C.22})$$

$$A_{h_{jk}} = - \int_{\Omega} \lambda \nabla \eta_k \nabla \eta_j d\Omega \quad \forall j, k = 1, \dots, \hat{q}, \quad (\text{C.23})$$

$$B_{h_{jk}} = \int_{\Gamma} \varphi_k \eta_j d\Gamma \quad \forall j = 1, \dots, \hat{q}, k = 1, \dots, \hat{p}, \quad (\text{C.24})$$

$$c_{h_j} = \int_{\Omega} c \eta_j d\Omega \quad \forall j = 1, \dots, \hat{q}, \quad (\text{C.25})$$

C.2. Funcional de costo

Primeramente re-escribimos el funcional de costo (4.18):

$$\begin{aligned} \mathcal{J}(z, v) = & \frac{1}{2} \int_{t_o}^{t_f} \left(r \|v\|_{L^2(\Omega)}^2 + q \|(z(t) - \tilde{y}(t))\|_{L^2(\Gamma)}^2 \right) dt \\ & + \frac{1}{2} s \|z(t_f) - \tilde{y}(t_f)\|_{L^2(\Gamma)}^2, \end{aligned} \quad (\text{C.26})$$

luego si definimos una separación variables de los valores de referencia como sigue $\tilde{y}_h(t) = \sum_{k=1}^{\hat{q}} \tilde{y}_k \eta_k$, donde $\tilde{y}_k = \tilde{y}$ y utilizamos las aproximaciones de z , v realizadas en la sección anterior ((C.13) y (C.17)) el funcional (C.26) toma la forma:

$$\begin{aligned} \mathcal{J}_h(z_h, v_h) = & \frac{1}{2} \int_{t_o}^{t_f} \left(r \|v_h\|_{L^2(\Omega)}^2 + q \|(z_h(t) - \tilde{y}_h(t))\|_{L^2(\Gamma)}^2 \right) dt \\ & + \frac{1}{2} s \|z_h(t_f) - \tilde{y}_h(t_f)\|_{L^2(\Gamma)}^2, \end{aligned} \quad (\text{C.27})$$

y aplicando la definición de la norma $L^2(\Omega)$ tenemos:

$$\begin{aligned}
 \|z_h(t) - \tilde{y}_h(t)\|_{L^2(\Omega)}^2 &= \int_{\Omega} \left(\sum_{k=1}^{\hat{q}} z_k \eta_k - \sum_{k=1}^{\hat{q}} \tilde{y}_k \eta_k \right)^2 d\Omega \\
 &= \int_{\Omega} \left(\sum_{k=1}^{\hat{q}} z_k \eta_k - \sum_{k=1}^{\hat{q}} \tilde{y}_k \eta_k \right) \left(\sum_{k=1}^{\hat{q}} z_k \eta_k - \sum_{k=1}^{\hat{q}} \tilde{y}_k \eta_k \right) d\Omega \\
 &= \left(\sum_{k=1}^{\hat{q}} (z_k - \tilde{y}_k) \right) \left(\sum_{k=1}^{\hat{q}} \sum_{j=1}^{\hat{q}} \int_{\Omega} \eta_k \eta_j d\Omega \right) \left(\sum_{j=1}^{\hat{q}} (z_j - \tilde{y}_j) \right),
 \end{aligned} \tag{C.28}$$

además si definimos el vector $\underline{\tilde{y}} = [\tilde{y}_1(t) \tilde{y}_2(t) \cdots \tilde{y}_i(t)]^T$ y teniendo en cuenta la definición (2.4), la norma es equivalente a:

$$\|z_h(t) - \tilde{y}_h(t)\|_{L^2(\Omega)}^2 = (\underline{z} - \underline{\tilde{y}})^T M_h (\underline{z} - \underline{\tilde{y}}) = \|\underline{z}(t) - \underline{\tilde{y}}_h(t)\|_{M_h}^2. \tag{C.29}$$

Análogamente:

$$\|z_h(t_f) - \tilde{y}_h(t_f)\|_{L^2(\Omega)}^2 = (\underline{z}(t_f) - \underline{\tilde{y}}(t_f))^T M_h (\underline{z}(t_f) - \underline{\tilde{y}}(t_f)) = \|\underline{z}(t_f) - \underline{\tilde{y}}(t_f)\|_{M_h}^2. \tag{C.30}$$

Luego aplicando la definición de norma $L^2(\Gamma)$

$$\begin{aligned}
 \|v_h\|_{L^2(\Gamma)}^2 &= \int_{\Gamma} \left(\sum_{k=1}^{\hat{p}} v_k \varphi_k \right)^2 d\Gamma, \\
 &= \int_{\Omega} \left(\sum_{k=1}^{\hat{p}} v_k \varphi_k \right) \left(\sum_{k=1}^{\hat{p}} v_k \varphi_k \right) d\Omega \\
 &= \left(\sum_{k=1}^{\hat{p}} (v_k) \right) \left(\sum_{k=1}^{\hat{p}} \sum_{j=1}^{\hat{p}} \int_{\Omega} \varphi_k \varphi_j d\Omega \right) \left(\sum_{j=1}^{\hat{p}} (v_j) \right),
 \end{aligned} \tag{C.31}$$

y definiendo la matriz $R_h \in \mathbb{R}^{\hat{p} \times \hat{p}}$ como sigue:

$$R_{h_{ij}} = \int_{\Omega} \varphi_k \varphi_j d\Omega, \quad \forall j, k = 1, \dots, \hat{q}, \tag{C.32}$$

por lo tanto:

$$\|v_h(t)\|_{L^2(\Gamma)}^2 = \underline{v}^T R_h \underline{v} = \|\underline{v}\|_{R_h}^2. \quad (\text{C.33})$$

Finalmente reemplazando (C.29), (C.30) y (C.33) en (C.27):

$$\mathcal{J}_h(\underline{z}, \underline{v}) = \frac{1}{2} \int_{t_o}^{t_f} (r \|\underline{v}(t)\|_{R_h}^2 + q \|\underline{z} - \tilde{y}(t)\|_{M_h}^2) dt + \frac{1}{2} s \|\underline{z}(t_f) - \tilde{y}(t_f)\|_{M_h}^2. \quad (\text{C.34})$$

Nótese que esta ecuación es equivalente al funcional (2.5).

BIBLIOGRAFÍA

- [1] Maday I., Turinici G., (2002) *A parareal in time procedure for the control of partial differential equations*, Numerical Analysis, p. 387-392.
- [2] Staff G. A., (2003) *The Parareal Algorithm*, Norwegian University of Science and Technology Department of Mathematical Sciences.
- [3] Grama A., Gupta A., Karypis G., Kumar V., (2003) *Introduction to Parallel Computing*, Pearson.
- [4] Blaise B., (2010) *Introduction to Parallel Computing*, Lawrence Livermore Laboratory, 2010.
- [5] Neittaanmäki, P., Tiba D., (1994), *Optimal Control of Nonlinear Parabolic Systems*, Marcel Dekker.
- [6] Intel Corporation, (2010), *ARK — Your source for information on Intel® product*, <http://ark.intel.com/>. Visitado el 9 de Setiembre 2010.
- [7] Kumar R., Zyuban R., Tullsen D.M., (2005), *Interconnections in Multi-core Architectures: Understanding Mechanisms, Overheads and Scaling*, Proceed-

- ings of the 32nd International Symposium on Computer Architecture (ISCA 05) IEEE. 1063-6897/05.
- [8] Fischer P. F., Patera A. T., (1994), *Parallel Simulation of viscous incompressible flow*, Annu. Rev. Fluid Mech. v.26, p. 483-527.
- [9] Flynn, M., (1972), *Some Computer Organizations and Their Effectiveness*, IEEE Trans. Comput. v.21, p. 948.
- [10] Dean J., Ghemawat S., (2008), *MapReduce: Simplified Data Processing on Large Clusters*, Communications of the ACM. v.51, p. 107-113.
- [11] Top 500 Supercomputers, (2010), *Performance Development*, http://www.top500.org/lists/2010/06/performance_development. Visitado el 9 de Setiembre 2010.
- [12] Glowinski R, Neittaanmaki P, (2008) *Partial differential equations, Modeling and numerical simulation*, Modeling and numerical simulation Springer.
- [13] Strang G., (2007) *Computational science and engineering*, Wellesley-Cambridge Press.
- [14] Karniadakis, G., Kirby, R., (2003) *Parallel Scientific Computing in C++ and MPI: A Seamless Approach to Parallel Algorithms and Their Implementation*, Cambridge University Press, USA.

- [15] Schaerer C. E., Mathew T., Sarkis M., (2007) *Block iterative algorithms for the solution of parabolic optimal control problems*, Lecture Notes in Computer Science, 4395, p. 452-465.
- [16] Sobolev, S. L., (1964) *Partial Differential Equations of Mathematical Physics*, Pergamon Press, England.
- [17] Stephenson, G., (1975) *Introducción a las Ecuaciones en Derivadas Parciales*, Editorial Reverté S. A. España.
- [18] Sommerfeld, Arnold, (1967) *Partial Differential Equations in Physics. Lectures on Theoretical Physics Volume VI*, Academic Press, Fourth Printing, USA.
- [19] Liu I., Rincon M., (2001) *Introdução ao Métodos de Elementos Finitos. Análise e Aplicação*, Universidade Federal do Rio de Janeiro.
- [20] Sauer C., (2009) *Circuitos Electrónicos 2D: Simulación, Control y Refrigeración*, Facultad de Ingeniería - Universidad Nacional de Asunción. San Lorenzo, Paraguay. 116 p.
- [21] Carstensen C., Alberty J., Funken S.,(1999) Remarks around 50 lines of Matlab: short finite element Implementation, Numerical Algorithms. v.20, p. 117-137.
- [22] Naidu D. S., (2003) *Optimal Control Systems*, CRC Press LLC.

- [23] Kirk D. E., (2004) *Optimal Control Theory*, Dover Publications.
- [24] Gunzburger M. D., (2002) *Perspectives in Flow Control and Optimization*, Siam.
- [25] Kelley C.T., (1999) *Iterative Methods for Optimization*, Siam.
- [26] Luenberger D. G., Ye Y., (2008) *Linear and Nonlinear Programming*, Springer.
- [27] Bertsekas D., (1999) *Nonlinear Programming*, Athena Scientific.
- [28] Bhaya A., Kaszkurewicz E., (2006) *Control Perspectives on Numerical Algorithms and Matrix Problems*, Siam.
- [29] Bertsekas D., (1974) *Partial Conjugate Gradient for a Class of Optimal Control Problems*, IEEE Transactions on Automatic Control. v. 19, p. 209 - 217.
- [30] Shewchuk J. R., (1994) *An Introduction to the Conjugate Gradient Method Without the Agonizing Pain*, Carnegie Mellon University.