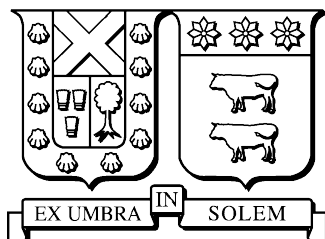


UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA  
DEPARTAMENTO DE MATEMÁTICA



---

## Turnpike Under Uncertainty

---

Tesis presentada por:

**Martín Sebastián Hernández Salinas**

*como requisito parcial para optar al grado académico de  
Magíster en Ciencias, Mención Matemática  
y al título profesional de  
Ingeniero Civil Matemático*

*Profesores Guías:*

Dr. Rodrigo Lecaros Lira  
Dr. Sebastián Zamorano Aliaga

Marzo, 2021

*«No cualquiera puede convertirse en un gran artista, pero un gran artista puede provenir de cualquier lugar.»*

Anton Ego.

## *Agradecimientos*

Quiero comenzar dando gracias a mis profesores directores a Rodrigo y a Sebastián, los cuales han sido excelentes profesores ayudándome y enseñándome, pero más importante aún, han sido excelentes personas guiándome más allá de lo académico y transmitiéndome sus experiencias de vidas. Sin duda la reunión de los jueves, junto a las risas y a la matemática, hizo que este día se convirtiera en uno de los más esperados de la semana.

A mis hermanos José y Luis, aun estando lejos, siempre se sintieron presentes. A mi padre que me dio la posibilidad de continuar mis estudios, además de siempre motivarme para seguir caminando. A mi madre quien sin su presencia emocional quizás no lo hubiera logrado llegar a este punto, llamándome sin falta durante cada noche los últimos seis años. A mi hermana quien fue como una segunda madre, pero más permisiva, la cual durante estos seis años se preocupó tanto de mi bienestar como de mi salud, dándome alegrías y buenos momentos que nos solo los guardaré en mi mente, si no también en mi corazón.

Por último, pero no menos importante a mis amigos. Tanto a ese pequeño grupo que nos conocemos de toda la vida y sabemos que, aunque no hablemos o nos vemos siempre vamos a estar ahí, Ana, Ricky, Pablo, Elías y Laura, siempre en el cora. Como también aquellas lindas personas que he conocido durante estos últimos seis años en especial a mi grupito de matemáticos, Fernando, Cristóbal, Lalo, Fabimath, Reco, Tomás, Gonzalo, Gustavo, Rogelio, Richy, Ike entre muchos y muchas más que han sido parte de esta etapa, a todo ellos y ellas muchas gracias.



UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA

## *Abstract*

Departamento de Matemática

Ingeniería Civil Matemática  
Magíster en Ciencia, Mención Matemática

### **Turnpike Under Uncertainty**

by Martín Sebastián Hernández Salinas

In this document we will analyze how Random Differential Equations, in particular Differential Equations with Random Coefficients, are related to Control Theory. We will focus on analyzing the Exponential Turnpike Property in the context of ordinary differential equations, partial differential equations, and nonlinear ordinary differential equations, all of these with random coefficients. In particular, we will analyze how nonlinear differential equations with random coefficients are related to deep learning, and this related with the Turnpike Property. We will conclude this document by studying the dynamics underlying the average solution of differential equations with random coefficients.



UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA

## *Resumen*

Departamento de Matemática

Ingeniería Civil Matemática  
Magíster en Ciencia, Mención Matemática

### **Turnpike Under Uncertainty**

por Martín Sebastián Hernández Salinas

En el presente documento analizaremos como se relacionan las ecuaciones diferenciales aleatorias, en particular ecuaciones diferenciales con coeficientes aleatorios, con la teoría de control. Nos enfocaremos en analizar la Propiedad de Turnpike Exponencial tanto en el contexto de ecuaciones diferenciales ordinarias, ecuaciones diferenciales parciales y ecuaciones diferenciales ordinarias no lineales, todas estas con coeficientes aleatorios. En particular, analizaremos como las ecuaciones diferenciales no lineales con coeficientes aleatorios, se relacionan con el deep learning, y este con la propiedad de turnpike. Concluiremos este documento estudiando la dinámica que subyace la solución promedio de ecuaciones diferenciales con coeficientes aleatorios.





# Índice general

<b>Agradecimientos</b>	<b>III</b>
<b>Introducción</b>	<b>1</b>
<b>1. Teoría de Control en Promedio Finito Dimensional</b>	<b>9</b>
1.1. Introducción . . . . .	9
1.2. Controlabilidad en Promedio . . . . .	9
1.3. Observabilidad en Promedio . . . . .	19
1.4. Equivalencia entre Controlabilidad y Observabilidad . . . . .	24
1.5. Comentarios y observaciones . . . . .	25
1.5.1. Más nociones de controlabilidad en promedio . . . . .	25
1.5.2. Estabilización de Sistemas con Incertidumbre . . . . .	27
1.5.3. Control en Promedio para Ecuaciones Diferenciales Parciales . . . . .	31
1.5.4. Algunos Problemas Abiertos . . . . .	35
<b>2. Propiedad de Turnpike para Sistemas Finito Dimensionales</b>	<b>37</b>
2.1. Introducción . . . . .	37
2.1.1. Motivación: . . . . .	37
2.1.2. Problema a estudiar . . . . .	38
2.2. Estimaciones previas . . . . .	40
2.3. Teoremas Principales de la Sección . . . . .	47
2.4. Ejemplos Numéricos . . . . .	66
2.5. Problemas Abiertos . . . . .	70
<b>3. Propiedad de Turnpike para la Ecuación del Calor con Incertidumbre</b>	<b>73</b>
3.1. Introducción . . . . .	73
3.1.1. Problema a estudiar . . . . .	73
3.2. Desarrollo y Estimaciones Previas . . . . .	75
3.3. Teoremas Principales . . . . .	87
<b>4. Deep Learning y Propiedad de Turnpike</b>	<b>99</b>
4.1. Introducción . . . . .	99
4.1.1. Redes Neuronales Residuales y Control Óptimo . . . . .	99

4.2.	Turnpike en Deep learning con Incerteza . . . . .	103
4.2.1.	Motivación . . . . .	103
4.2.2.	Planteo del Problema . . . . .	104
4.2.3.	Continuación del Problema . . . . .	111
<b>5.</b>	<b>Soluciones en Promedio</b>	<b>123</b>
5.1.	Introducción . . . . .	123
5.2.	Operadores conmutativos . . . . .	126
5.2.1.	Estimaciones Principales . . . . .	126
5.2.2.	Ejemplos operadores conmutativos . . . . .	128
5.3.	Operadores no conmutativos . . . . .	129
5.3.1.	Estimaciones Principales . . . . .	129
5.3.2.	Ejemplos operadores no conmutativos . . . . .	134
5.4.	Ecuación de Burger's . . . . .	135
5.4.1.	Promedio en presencia de Choque . . . . .	136
5.4.2.	Promedio en presencia de refracción . . . . .	136
5.4.3.	Promedio de una solución suave . . . . .	138
	Ejemplo (Promedio Solución Suave) . . . . .	139
5.4.4.	Otra representación del promedio . . . . .	141
	<b>Bibliografía</b>	<b>149</b>

---

# Introducción

Cuando se desea modelar un fenómeno en física o ingeniería, usualmente este es traducido a ecuaciones, o a sistemas de ecuaciones. Durante la modelación de un fenómeno usualmente se realiza una simplificación de la realidad, a veces perdiendo mucha información de lo que realmente se espera modelar. Dicha simplificación nace desde el hecho que muchas veces se desconoce el origen de ciertas perturbaciones, producidas por el alto número de variables que se involucran en el sistema. Si se intentase incorporar estas variables al modelo, podrían generar una enorme complejidad, dificultando así poder obtener conclusiones de este. Por lo anterior, surge naturalmente la idea de incorporar factores con incertidumbre o aleatoriedad al sistema, esto con la intención de recuperar los efectos producidos por las variables desconocidas. En consecuencia, se esperaría que el modelo representase de una forma más certera el sistema físico. Los sistemas evolutivos con incertidumbre nacen desde los estudios sobre el movimiento browniano realizados por los físicos Sutherland, Smoluchowski, Langevin y Einstein entre 1905-1908, otorgando a este último el premio nobel de física en 1921. El desarrollo del movimiento browniano ha dado grandes avances tanto en física cuántica, economía, medicina, dentro de otras áreas (ver [48]). En nuestro estudio de sistemas con incertidumbre, nos centraremos en ecuaciones con coeficientes aleatorios autónomos (como en [48, Sección 8.3]). En este contexto a lo largo de este trabajo, nos interesarán las propiedades que satisface la solución promedio del sistema.

## Control en Promedio

Motivados inicialmente por la noción de controlabilidad en promedio para sistemas con coeficientes aleatorios, noción introducida en [58], es que en el primer Capítulo de este trabajo analizaremos la controlabilidad de sistemas lineales con coeficientes aleatorios, es decir, sistemas del tipo

$$\begin{cases} x_t + A(\omega)x = B(\omega)u & t \in (0, T), \\ x(0) = x_0, \end{cases} \quad (1)$$

donde  $x_0 \in \mathbb{R}^n$  corresponde a la condición inicial,  $A(\omega), B(\omega)$  matrices que depende de un parámetro aleatorio  $\omega \in \Omega$  y  $u \in L^2(0, T; \mathbb{R}^m)$  corresponde al control del sistema, el cual será considerado independiente del parámetro aleatorio  $\omega \in \Omega$ . Como es de imaginar la noción clásica de controlabilidad exacta, como es mostrada por ejemplo en el libro [10], no tiene mucho

sentido en este caso. Por ello es introducida la noción de controlabilidad en promedio, concepto que busca la existencia de un control  $u$  de tal manera que la trayectoria promedio final sea un valor dado. Formalmente diremos que el sistema (1) es controlable en promedio en tiempo  $T$ , si para cada par de condiciones iniciales y finales  $x_0, x_1 \in \mathbb{R}^n$  existe un control  $u \in L^2(0, T; \mathbb{R}^m)$  de tal forma que  $x$  la solución del sistema (1) verifica que

$$\mathbb{E}(x(T)) = x_1,$$

donde  $\mathbb{E}$  corresponde a la esperanza matemática (promedio). Además de analizar esta noción, también introduciremos el concepto de observabilidad en promedio de un sistema con incertidumbre. Esta noción es motivada por la definición de observabilidad para sistemas lineales sin incertidumbre como es dada en [47, Capítulo 6]. De igual forma, como con la controlabilidad en promedio, caracterizaremos esta noción. También en este capítulo mostraremos como se relacionan los conceptos de controlabilidad y observabilidad en promedio introducidos, probando que estas nociones son equivalentes.

Además, comentaremos sobre las distintas nociones de controlabilidad en promedio (exacta, a cero y aproximada) y como a diferencia del contexto sin incertidumbre, estas no son equivalentes. Por otro lado, también hablaremos sobre la noción de estabilización de un sistema con incertidumbre, comentado sobre distintas definiciones de estabilidad que se han dado en la literatura. Probando además el sistema (1) es estabilizable, en algún sentido, vía un control  $u$  dado. Finalmente concluiremos este capítulo con algunos problemas abiertos con relación al tópico de controlabilidad en promedio.

## Propiedad de Turnpike

Como bien ocurre en física, principalmente termodinámica, ciertos fenómenos que evolucionan en el tiempo tienden a minimizar su gasto energético, y con ello buscar estados de mínima energía. Este efecto en un sistema general ocurre cuando se penaliza lo suficiente la energía del sistema, y en consecuencia se espera que las trayectorias asociadas al sistema sean lo más estacionarias posibles. Un ejemplo cotidiano de esto es cuando se espera viajar desde una ciudad a otra, en tal caso como bien sabemos la distancia más corta es aquella recta que une las dos ciudades, la cual suele ser la autopista más cercana a ambas ciudades. Por tanto, si esperamos reducir nuestro gasto energético para viajar entre dos ciudades, la mejor opción es tomar la autopista. Lo anteriormente descrito motiva a la llamada Propiedad de Turnpike (propiedad de autopista de peaje).

Las ideas e intuiciones tras la Propiedad de Turnpike, comenzaron a desarrollarse a mediados de los años 40's en [42] por John Von Neumann. No obstante, el uso de la palabra turnpike no fue utilizado hasta [12]. Esta propiedad nos señala que la evolución de ciertos fenómenos de crecimiento tiende a ser constante cuando su punto de partida y su punto final están suficientemente lejanos. Lo anteriormente descrito puede ser pensado como un problema de control óptimo, con

un horizonte temporal  $T$  grande, en donde existe una dinámica evolutiva y otra estacionaria (autopista), las cuales se acercan durante un intervalo de tiempo para luego alejarse, siempre y cuando el horizonte de tiempo sea lo suficientemente grande.

En el segundo capítulo de este trabajo nos enfocaremos en la Propiedad de Turnpike. En la literatura es posible encontrar distintas definiciones de la Propiedad de Turnpike. Nosotros nos enfocaremos en la propiedad introducida en [44] llamada Propiedad de Turnpike exponencial. Esta señala que los pares  $(x^T, u^T)$  y  $(\bar{x}, \bar{u})$  satisfacen la propiedad, si se verifica la siguiente estimación

$$\|x^T - \bar{x}\| + \|u^T - \bar{u}\| \leq C(e^{-\delta t} + e^{-\delta(T-t)}), \quad (2)$$

donde el par  $(x^T, u^T)$  es la solución de un problema de control óptimo evolutivo, mientras el par  $(\bar{x}, \bar{u})$  es solución de un problema de control óptimo estacionario. Además,  $C$  y  $\delta$  son constantes positivas independientes del horizonte de tiempo  $T$ . No obstante, a diferencia de lo desarrollado en [44], en el Capítulo 1 nos enfocaremos en analizar la propiedad de Turnpike para el sistema con coeficientes aleatorio. Para ello consideraremos el siguiente problema de minimización

$$\min_{u \in L^2(0,T;\mathbb{R}^m)} \left\{ J^T(u) = \frac{1}{2} \left( \int_0^T \|u(t)\|_{\mathbb{R}^m}^2 dt + \int_0^T \int_{\Omega} \|C(\omega)x(t,\omega) - z\|_{\mathbb{R}^n}^2 d\mu dt \right) \right\}, \quad (3)$$

donde  $x(t, \omega)$  es la solución de (1),  $C(\omega)$  una matriz que depende del parámetro aleatorio  $\omega \in \Omega$  y  $z \in \mathbb{R}^n$  una trayectoria fija. Además, consideraremos el problema estacionario dado por

$$\min_{u \in \mathbb{R}^m} \left\{ J_s(u) = \frac{1}{2} \left( \|u\|_{\mathbb{R}^m}^2 + \int_{\Omega} \|C(\omega)x(\omega) - z\|_{\mathbb{R}^n}^2 d\mu \right) \right\}, \quad (4)$$

donde  $x(\omega)$  soluciona el problema  $A(\omega)x = B(\omega)u$ . Así motivados por la definición de Turnpike exponencial (2), es que probaremos la Propiedad de Turnpike, pero para las trayectorias promedio. Es decir, bajo ciertas hipótesis que serán especificadas en el Capítulo 2, si los problemas (3) y (4) tienen solución  $(u^T, x^T)$  y  $(\bar{u}, \bar{x})$  respectivamente, se probará que existen constantes  $\mathcal{C}, \delta > 0$  independientes de  $T$ , tales que

$$\|\mathbb{E}(x^T) - \mathbb{E}(\bar{x})\|_{\mathbb{R}^n} + \|u^T - \bar{u}\|_{\mathbb{R}^m} \leq \mathcal{C}(e^{-(T-t)\delta} + e^{-t\delta}),$$

para todo  $t \in [0, T]$ .

Adicionalmente, probaremos que el promedio temporal del par óptimo  $(u^T, x^T)$ , converge al par óptimo  $(\bar{u}, \bar{x})$ , es decir

$$\begin{aligned}\frac{1}{T} \int_0^T x^T(t, \omega) dt &\longrightarrow \bar{x}(\omega), \\ \frac{1}{T} \int_0^T u^T(t) dt &\longrightarrow \bar{u},\end{aligned}$$

cuando  $T \rightarrow \infty$ , en  $L^2(\Omega; \mathbb{R}^n)$  y  $\mathbb{R}^m$  respectivamente, donde  $L^2(\Omega; \mathbb{R}^n)$  corresponde al espacio de las funciones cuadrado integrable, definidas en  $\Omega$  a valores en  $\mathbb{R}^n$ .

En el Capítulo 3 también probaremos la Propiedad de Turnpike, pero esta vez para una ecuación del calor con incertidumbre, y condiciones de frontera Dirichlet homogénea. Esta ecuación incorporará un parámetro de difusión  $\alpha$  aleatorio. En este caso consideraremos el siguiente problema de minimización

$$\min_{u \in L^2(0, T; L^2(G_0))} \left\{ J^T(u) = \frac{1}{2} \left( \int_0^T \|u(t)\|_{L^2(G_0)}^2 dt + \int_0^T \int_{\Omega} \|C(\omega)y(t, \omega) - z\|_{L^2(G)}^2 d\mu dt \right) \right\}, \quad (5)$$

con  $z \in L^2(G)$  una trayectoria fija,  $G_0 \subset G \subset \mathbb{R}^n$  e  $y(t, \omega)$  solución de la ecuación de calor

$$\begin{cases} y_t - \alpha(\omega)\Delta y = u\chi_{G_0} & \text{en } G \times (0, T), \\ y = 0 & \text{sobre } \partial G \times (0, T), \\ y(0) = y_0 & \text{en } G, \end{cases}$$

con condición inicial  $y_0 \in L^2(G)$ , donde supondremos que  $\alpha$  es una variable aleatoria estrictamente positiva. Además,  $u \in L^2(0, T; L^2(G_0))$  un control independiente del parámetro aleatorio.

Por otro lado, consideraremos el siguiente problema estacionario

$$\min_{u \in L^2(G_0)} \left\{ J^s(u) = \frac{1}{2} \left( \|u\|_{L^2(G_0)}^2 + \int_{\Omega} \|C(\omega)y(\omega) - z\|_{L^2(G)}^2 d\mu \right) \right\}, \quad (6)$$

con  $y(\omega)$  solución de la ecuación elíptica

$$\begin{cases} -\alpha(\omega)\Delta y = \chi_{G_0} u & \text{en } x \in G, \\ y = 0 & \text{sobre } \partial G. \end{cases}$$

Así sí los problemas (5) y (6), tienen solución  $(u^T, y^T)$  y  $(\bar{u}, \bar{y})$  respectivamente, entonces en el Capítulo 3 será probado que existen constantes positivas  $\mathcal{C}, \delta$  independientes del horizonte de tiempo  $T$ , tales que

$$\left\| \mathbb{E}(y^T) - \mathbb{E}(\bar{y}) \right\|_{L^2(G)} + \left\| u^T - \bar{u} \right\|_{L^2(G_0)} \leq \mathcal{C}(e^{-(T-t)\delta} + e^{-t\delta}),$$

para todo  $t \in [0, T]$ .

Adicionalmente, probaremos que el promedio temporal del par óptimo  $(u^T, x^T)$  converge al par óptimo  $(\bar{u}, \bar{x})$ , es decir

$$\begin{aligned} \frac{1}{T} \int_0^T y^T(t, \omega) dt &\longrightarrow \bar{y}(\omega), \\ \frac{1}{T} \int_0^T u^T(t) dt &\longrightarrow \bar{u}, \end{aligned}$$

cuando  $T \rightarrow \infty$  en  $L^2(\Omega; L^2(G))$  y  $L^2(G_0)$  respectivamente, donde  $L^2(\Omega; L^2(G))$  corresponde al espacio de las funciones cuadrado integrable, definidas en  $\Omega$  a valores en  $L^2(G)$ .

## Deep Learning y Turnpike con Incerteza

En inteligencia artificial (AI), existen distintas áreas de las cuales nos centraremos en el Machine Learning. Esta área es la encargada del aprendizaje automático, basado en ajustes de parámetros según la efectividad del aprendizaje. Para llevar a cabo el aprendizaje es necesario introducir datos de entrenamiento, los cuales están compuestos por un conjunto de datos y etiquetas predefinidas. Una pregunta primordial es como ajustar esto parámetros para que el entrenamiento sea correcto, es decir, como escogemos los parámetros para que, dado un dato de entrada, el modelo pueda predecir la etiqueta esperada, en este caso la herramienta a utilizar suele ser el Deep Learning. Aprendizaje profundo o Deep Learning, es un conjunto de algoritmos los cuales lleva a cabo el proceso de aprendizaje, usando una red neuronal artificial como arquitecturas computacionales que admiten transformaciones no lineales múltiples e iterativas de datos, expresados en forma matricial o tensorial.

Cuando decimos red neuronal, podemos pensar en un modelo discreto el cual está conformado por neuronas, donde una neurona puede entenderse simplemente como una variable real. Formalmente una red neuronal consiste en un conjunto de unidades, llamadas neuronas artificiales, conectadas entre sí para transmitirse señales. La información de entrada atraviesa la red neuronal (donde se somete a diversas operaciones muchas veces no lineales) produciendo valores de salida. La generalidad de la definición anterior da el paso a distintas posibles arquitecturas o estructuras que puede llegar a tener una red neuronal, por supuesto, saber que arquitecturas es mejor para cada aplicación, es un tema de investigación actual. En lo que sigue nos centraremos en una arquitectura de red neuronal residual (ResNet). Una ResNet es una red neuronal en la cual la información de la capa actual puede llegar a depender de las capas anteriores, lo cual puede ir cambiando para cada arquitectura. Consideremos los datos de entrenamiento  $\{x_i, y_i\}_{i=1}^{i=n} \subset \mathbb{R}^d \times \mathbb{R}^m$  y la siguiente ResNet cuya arquitectura es descrita por

$$\begin{cases} x_i^{k+1} = x_i^k + \sigma(w^k x_i^k + b^k), \\ x_i^0 = x_i, \end{cases} \quad (7)$$

para cada  $k \in \{0, \dots, n_{\text{capas}} - 1\}$  y cada  $i \in \{1, \dots, n\}$ . Denotamos como  $x_i^k \in \mathbb{R}^d$  al  $i$ -ésimo estado o neurona de la capa  $k$ . El número  $n_{\text{capas}}$  corresponde a la profundidad de la ResNet (depth). Además, la función  $\sigma$  llamada función de activación, es usualmente considerada como la función sigmoide, la función tangente hiperbólica, o la función Re-Lu. Las cantidades  $w^k \in \mathbb{R}^{d \times d}$  y  $b^k \in \mathbb{R}^d$  son los parámetros de ajuste de la red neuronal. Como fue mencionado para ajustar los parámetros, se buscan aquellos que entrenen de mejor manera la red neuronal. Por lo anterior para estima  $w^k$  y  $b^k$ , en general se busca resolver el siguiente problema

$$\min_{\{w^k, b^k\}_{k=0}^{n_{\text{capas}}-1}} \frac{1}{n} \sum_{i=1}^n \text{loss} \left( \varphi(x_i^{n_{\text{capas}}}), y_i \right), \quad (8)$$

donde la función  $\text{loss} : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$  es una función continua que cuantifica la distancia entre  $y_i$  y  $\varphi(x_i^{n_{\text{capas}}})$ , mientras que  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}^m$  es una función continua, que para efectos prácticos puede ser entendida como una proyección. Observemos que la arquitectura anteriormente descrita, puede ser entendida como un problema de control óptimo, con control  $[w^k, b^k]$ . Además, la ecuación (7) puede ser vista como la desratización del modelo continuo

$$\begin{cases} \dot{x}_i(t) = \sigma(w(t)x_i(t) + b(t)) & t \in (0, T), \\ x_i(0) = x_i, \end{cases} \quad (9)$$

utilizando el método de Euler. Esta ecuación es llamada en la literatura como Ecuación diferencial Neuronal. En [17] es probada la propiedad de turnpike para los siguientes problemas de minimización (problema de aprendizaje supervisado)

$$\inf_{\substack{[w, b]^\top \in H^k(0, T; \mathbb{R}^{du}) \\ \text{Sujeto a (9)}}} \frac{1}{n} \sum_{i=1}^n \text{loss}(\varphi(x_i(T)), \bar{y}_i) + \frac{\alpha}{2} \|[w, b]^\top\|_{H^1(0, T; \mathbb{R}^{du})}^2 + \frac{\beta}{2} \int_0^T \|\mathbf{x}(t) - \bar{x}\|^2 dt. \quad (10)$$

donde  $\mathbf{x} = [x_1, \dots, x_n]$ . Definiendo  $f(w, x, b) = \sigma(w(t)x_i(t) + b(t))$  consideramos problema estacionario

$$\inf_{\substack{[w^s, b^s]^\top \in \mathbb{R}^{du} \\ \text{Sujeto a} \\ f(w, x, b) = 0}} \frac{\alpha}{2} \|[w^s, b^s]^\top\|^2 + \frac{\beta}{2} \|x^s - x_d\|^2. \quad (11)$$

Motivados por todo lo anterior, es que en el Capítulo 4 se analizará este problema pero incorporando incertidumbre. Para ello suponemos que los datos de entrenamiento dependen de un parámetro aleatorio, es decir, consideramos los datos de entrenamiento  $\{x_i(\omega), y_i\} \subset L^2(\Omega; \mathbb{R}^d) \times \mathbb{R}^m$  y consideramos la ecuación diferencial neuronal (9). Como es de esperar, si la ecuación (9) incorpora datos iniciales aleatorios, entonces sus trayectorias también serán aleatorias, por tanto, nos encontramos en presencia de una Ecuación Neuronal con Incertidumbre. Además, inspirados en el problema de minimización (10) es que consideramos el siguiente problema de



aprendizaje supervisado

$$J^T(w, b) = \frac{1}{2} \left( \left\| [w, b]^\top \right\|_{H^k(0, T; \mathbb{R}^{d_u})}^2 + \int_0^T \left\| \mathbb{E}(x(t)) - \bar{x} \right\|_{\mathbb{R}^{d_x}}^2 dt + \left\| \varphi \left( \mathbb{E}(x(T, \omega)) \right) - y \right\|_{\mathbb{R}^{m \times n}}^2 \right).$$

La motivación, interpretación y posibles aplicaciones de este tipo de estructura, es comentado con más detalle en el Capítulo 4. Además, en el Capítulo 4 probaremos una serie de lemas motivados por los trabajos [17, 13], los cuales serán una guía para probar la propiedad de Turnpike en este contexto, no obstante, este resultado será dejado para trabajos futuros.

## Soluciones promedio

Finalmente, el capítulo 5, motivados por entender la naturaleza que subyace la dinámica de promedio de ciertos sistemas, es que analizaremos el sistema finito dimensional con incertidumbre

$$\begin{cases} u_t + Au + \alpha Bu = 0 & t \in (0, T), \\ u(0) = u_0, \end{cases} \quad (12)$$

donde  $\alpha$  es una variable aleatoria acotada. Estaremos interesados en analizar como es posible caracterizar la solución promedio del sistema (12), es decir, que ocurre con

$$\bar{u}(t) = \mathbb{E}(u(t)).$$

Dividiremos este problema en dos partes. La primera cuando  $A, B$  son matrices que conmutan, en tal caso la pregunta sobre como caracterizar la solución promedio es respondida exitosamente, dando una forma explícita de esta, la cual depende solo de soluciones de ecuaciones asociadas a los valores extremos de la variable aleatoria. Es decir, probaremos que existe una función  $\Phi(x)$  tal que si  $\alpha \in [a, b]$ , entonces

$$\bar{u} = \Phi(b) - \Phi(a). \quad (13)$$

Por otro lado, también será analizado el caso cuando  $A, B$  son matrices no conmutativas. A diferencia del caso anterior, no es posible utilizar directamente las mismas técnicas, por lo cual es necesario incorporar resultado de álgebras no conmutativas. Este caso las dificultades algebraicas son mucho mayores. Aun así es posible obtener una expresión como (13).

Por último, con relación a ecuaciones diferenciales no lineales en particular a ecuaciones relacionadas con las leyes de conservación, en este capítulo analizaremos la ecuación de Burger's sin viscosidad

$$\begin{cases} u_t + \alpha uu_x = 0 & (t, x) \in (0, T) \times \mathbb{R} \\ u(0, x) = u^0(x) & \text{en } x \in \mathbb{R}, \end{cases}$$

para distintos tipos de condiciones iniciales. Analizaremos tanto el caso de choques como de refracción. Además, analizaremos el caso particular cuando se tienen condiciones iniciales suaves. Al igual que con la ecuación (12), tendremos una expresión que depende de los valores extremos de la variable aleatoria, pero en este caso una dependencia no lineal.

Finalizaremos la sección comentando los resultados obtenidos y dejaremos algunos problemas abiertos, en relación con soluciones en promedio de ecuaciones diferenciales no lineales.

## Capítulo 1

# Teoría de Control en Promedio Finito Dimensional

### 1.1. Introducción

En este capítulo introduciremos la noción de Controlabilidad en Promedio para sistemas con coeficientes aleatorios. En las primeras tres secciones, nos centraremos en el control y la observabilidad en promedio en el caso finito dimensional. En lo que sigue comentaremos sobre otras nociones de controlabilidad en promedio en el contexto finito dimensional, y hablaremos sobre la estabilización de sistemas con incertidumbre. Finalmente introduciremos la noción de controlabilidad en promedio para el caso infinito dimensional, haciendo un breve recuento bibliográfico.

La noción de controlabilidad en promedio fue introducida en [58], donde se define y se dan las primeras caracterizaciones para la controlabilidad en promedio. En lo que sigue mostraremos tanto resultados conocidos, como resultados y definiciones nuevas.

### 1.2. Controlabilidad en Promedio

Para fijar ideas comenzaremos introduciendo el problema y las definiciones básicas. Consideremos un espacio de probabilidad  $(\Omega, \mathcal{F}, \mu)$ , donde  $\Omega$  corresponde al espacio muestral,  $\mathcal{F}$  una  $\sigma$ -álgebra, y  $\mu$  la medida asociada a este espacio. Es importante recordar que  $\mu(\Omega) = 1$ . Consideremos el siguiente sistema

$$\begin{cases} x_t + A(\omega)x = B(\omega)u & t \in (0, T), \\ x(0) = x_0, \end{cases} \quad (1.1)$$

donde  $\omega \in \Omega$  corresponde al parámetro aleatorio,  $x(\cdot, \omega) \in C([0, T]; \mathbb{R}^n)$  corresponde al estado, para cada  $\omega \in \Omega$ ,  $A(\omega) \in \mathcal{L}(\mathbb{R}^n)$  y  $B(\omega) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$  corresponden a dos matrices para cada  $\omega$ . Además, el control  $u \in L^2(0, T; \mathbb{R}^m)$  y  $x_0 \in \mathbb{R}^n$  serán independientes de  $\omega$ .

Supondremos que  $A \in C(\Omega; \mathcal{L}(\mathbb{R}^n))$  y  $B \in C(\Omega; \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n))$ , son matrices uniformemente acotadas con respecto a  $\omega \in \Omega$ <sup>1</sup>. Además, consideraremos  $K, \hat{K}$  dos constantes positivas que pueden cambiar en cada línea.

Antes de introducir la noción de controlabilidad, definimos la matriz fundamental asociada a  $A$ , como la matriz  $R(t, \omega) \in \mathcal{L}(\mathbb{R}^n)$  para cada  $\omega \in \Omega$ , definida para todo  $t \in \mathbb{R}$  tal que

$$\begin{cases} R_t(t, \omega) + A(\omega)R(t, \omega) = 0 & t \in (0, T), \\ R(0, \omega) = I, \end{cases}$$

donde  $I$  denota la matriz identidad en  $\mathcal{L}(\mathbb{R}^n)$ .

**Observación 1.2.1.** Notemos que  $R(t, \omega) = e^{-A(\omega)t}$  para cada  $t \in [0, T]$  y  $\omega \in \Omega$ . En consecuencia, la matriz  $R(t, \omega)$  satisface las propiedades de la matriz exponencial, es decir  $R(0, \omega) = I$ , como también

$$R(t-s, \omega) = R(t, \omega)R^{-1}(s, \omega) = R^{-1}(s, \omega)R(t, \omega) = R^{-1}(s, \omega)R^{-1}(-t, \omega) = R^{-1}(s-t, \omega).$$

Además, como  $A \in C(\Omega; \mathcal{L}(\mathbb{R}^n))$  tenemos que  $R(t, \cdot) \in C(\Omega; \mathcal{L}(\mathbb{R}^n))$  para cada  $t \in [0, T]$ , por tanto existe  $K > 0$  independiente de  $t \in [0, T]$  tal que

$$\|R(t)\|_{C(\Omega; \mathcal{L}(\mathbb{R}^n))} \leq K \quad \forall t \in [0, T].$$

En consecuencia de que  $K > 0$  es independiente de  $t \in [0, T]$ , tomando supremo en  $t$ , se sigue que

$$\|R\|_{C([0, T]; C(\Omega; \mathcal{L}(\mathbb{R}^n)))} \leq K. \quad (1.2)$$

Es más, desde la continuidad de la matriz exponencial tenemos que  $R \in C([0, T]; C(\Omega; \mathcal{L}(\mathbb{R}^n)))$ . Ahora consideramos el siguiente espacio de Hilbert

$$L^2(\Omega; \mathbb{R}^n; \mu) = \left\{ y(\omega) \in \mathbb{R}^n : \|y\|_{L^2(\Omega; \mathbb{R}^n; \mu)}^2 = \int_{\Omega} \|y(\omega)\|_{\mathbb{R}^n}^2 d\mu < \infty \right\},$$

con el producto interno

$$\langle x, y \rangle_{L^2(\Omega; \mathbb{R}^n; \mu)} = \int_{\Omega} \langle x(\omega), y(\omega) \rangle_{\mathbb{R}^n} d\mu.$$

En lo que sigue denotaremos  $L^2(\Omega; \mathbb{R}^n) = L^2(\Omega; \mathbb{R}^n; \mu)$ . Así tenemos los siguientes resultados de regularidad, para la solución de (1.1).

---

<sup>1</sup>Es decir, que existe  $C > 0$  tal que  $\sup_{\omega \in \Omega} \|A\|_{\mathcal{L}(\mathbb{R}^n)} < C$ .

**Teorema 1.2.1.** *La solución  $x(\cdot, \omega) \in C([0, T]; \mathbb{R}^n)$  de (1.1), satisface que  $x(t, \cdot) \in L^2(\Omega; \mathbb{R}^n)$  para cada  $t \in [0, T]$ . Además, existe  $K > 0$  tal que tenemos la estimación*

$$\|x(t)\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq K \left( \|x_0\|_{\mathbb{R}^n}^2 + \|u\|_{L^2(0, T; \mathbb{R}^m)}^2 \right), \quad (1.3)$$

para cada  $t \in [0, T]$ .

*Demostración.* Observemos que desde la fórmula de Duhamel, para cada  $\omega \in \Omega$ , tenemos que

$$x(t, \omega) = R(t, \omega)x_0 + \int_0^t R(t-s, \omega)B(\omega)u(s)ds.$$

Así por la estimación (1.2) y la regularidad de  $B(\cdot)$ , tenemos que

$$\begin{aligned} \|x(t, \omega)\|_{\mathbb{R}^n}^2 &\leq 2 \left( \|R(t, \omega)x_0\|_{\mathbb{R}^n}^2 + \left\| \int_0^t R(t-s, \omega)B(\omega)u(s)ds \right\|_{\mathbb{R}^n}^2 \right) \\ &\leq K \left( \|x_0\|_{\mathbb{R}^n}^2 + \int_0^T \|u(s)\|_{\mathbb{R}^m}^2 ds \right) = K \left( \|x_0\|_{\mathbb{R}^n}^2 + \|u\|_{L^2(0, T; \mathbb{R}^m)}^2 \right). \end{aligned}$$

Como  $\|x_0\|_{\mathbb{R}^n}$  y  $\|u\|_{L^2(0, T; \mathbb{R}^m)}$  son independientes de  $\omega$ , podemos integrar sobre  $\Omega$  obteniendo

$$\|x\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq K \left( \|x_0\|_{\mathbb{R}^n}^2 + \|u\|_{L^2(0, T; \mathbb{R}^m)}^2 \right).$$

□

**Teorema 1.2.2.** *La solución  $x$  de (1.1) pertenece a  $C([0, T], L^2(\Omega; \mathbb{R}^n))$ . Además, existe  $K > 0$  tal que*

$$\|x\|_{C([0, T], L^2(\Omega; \mathbb{R}^n))}^2 \leq K \left( \|x_0\|_{\mathbb{R}^n}^2 + \|u\|_{L^2(0, T; \mathbb{R}^m)}^2 \right). \quad (1.4)$$

*Demostración.* Procedemos directamente y consideramos  $0 \leq t_2 < t_1 \leq T$ , así usando la fórmula de Duhamel y considerando que  $R(t, \omega)$  y  $B(\omega)$  son uniformemente acotados tenemos que

$$\begin{aligned} \|x(t_1) - x(t_2)\|_{L^2(\Omega; \mathbb{R}^n)} &\leq \|(R(t_1, \omega) - R(t_2, \omega))x_0\|_{\mathbb{R}^n} + K \int_{t_2}^{t_1} \|u(s)\|_{\mathbb{R}^m} ds \\ &\quad + K \int_0^{t_2} \|(R(t_1-s, \omega) - R(t_2-s, \omega))\|_{\mathcal{L}(\mathbb{R}^n)} \|u(s)\|_{\mathbb{R}^m} ds. \end{aligned}$$

Observar que por la continuidad de  $R(t, \omega)$  podemos acotar el primer y tercer término del lado derecho de la desigualdad anterior. Para estimar el término restante, basta utilizar el Teorema

de Diferenciación de Lebesgue, obteniendo así que

$$(t_1 - t_2) \left( \frac{1}{t_1 - t_2} \int_{t_2}^{t_1} \|u(s)\|_{\mathbb{R}^m} ds \right) \rightarrow 0 \quad \text{cuando } t_2 \rightarrow t_1.$$

Probando la continuidad. Finalmente, para obtener la estimación (1.4), basta tomar supremo en tiempo a (1.3).  $\square$

A continuación, damos la definición de controlabilidad en promedio

**Definición 1.2.1** ([58, Sección 2]). Diremos que el sistema (1.1) satisface la propiedad de control exacto en promedio, o simplemente control en promedio, si dado  $T > 0$ , para cada  $x_0$  trayectoria inicial y  $x_1$  trayectoria final en  $\mathbb{R}^n$ , existe un control  $u \in L^2(0, T; \mathbb{R}^m)$  tal que

$$\int_{\Omega} x(T, \omega) d\mu = x_1.$$

La definición anterior nos otorga una noción de controlabilidad para sistemas aleatorios, en donde se espera que la trayectoria promedio en tiempo  $T$ , sea  $x_1$ . Gráficamente esta noción puede ser ilustrada como en la Figura 1.1

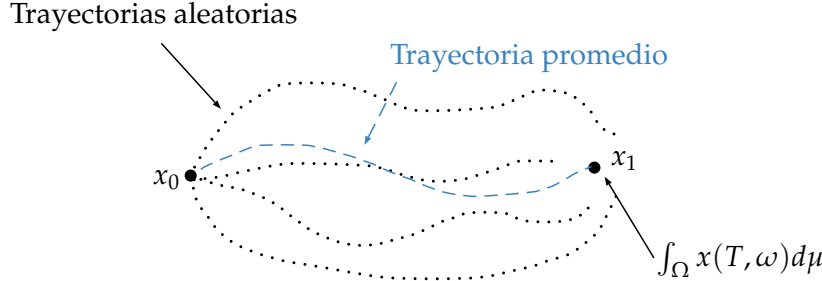


FIGURA 1.1: Representación de la Controlabilidad en Promedio.

Podemos caracterizar la noción de controlabilidad con el siguiente teorema.

**Teorema 1.2.3** ([58, Teorema 1.3]). Las siguientes proposiciones son equivalentes:

I) El sistema

$$\begin{cases} \dot{x}_t + A(\omega)x = B(\omega)u & t \in (0, T) \\ x(0) = x_0, \end{cases}$$

es controlable en promedio.

II) Se satisface la condición de rango en promedio<sup>2</sup>

$$\text{Rank} \left[ \int_{\Omega} [A(\omega)]^j B(\omega) d\mu; \quad j \geq 0 \right] = n.$$

III) Se satisface la desigualdad de observabilidad

$$\|\varphi_T\|_{\mathbb{R}^n}^2 \leq K \int_0^T \left\| \int_{\Omega} B^*(\omega) \varphi(t, \omega) d\mu \right\|_{\mathbb{R}^m}^2 dt, \quad (1.5)$$

para  $K > 0$ , y  $\varphi$  solución de

$$\begin{cases} -\varphi_t + A^*(\omega) \varphi = 0 & t \in (0, T), \\ \varphi(T) = \varphi_T, \end{cases} \quad (1.6)$$

donde  $\varphi_T \in \mathbb{R}^n$  es independiente del parámetro  $\omega \in \Omega$ .

Además, cuando se tiene alguna de las anteriores, el control de norma mínima en  $L^2(0, T; \mathbb{R}^m)$  viene caracterizado por

$$u(t) = - \int_{\Omega} B^*(\omega) \tilde{\varphi}(t, \omega) d\mu, \quad (1.7)$$

donde  $\tilde{\varphi}$  es la solución de la ecuación adjunta (1.6) cuya condición final  $\varphi_T$  minimiza el funcional

$$J(\varphi_T) = \frac{1}{2} \int_0^T \left\| \int_{\Omega} B^*(\omega) \varphi(t, \omega) d\mu \right\|_{\mathbb{R}^m}^2 dt - \langle x_1, \varphi_T \rangle_{\mathbb{R}^n} + \left\langle x_0, \int_{\Omega} \varphi(0, \omega) d\mu \right\rangle_{\mathbb{R}^n}.$$

**Observación 1.2.2.** Es importante destacar que a partir de II) del Teorema 1.2.3, tenemos una caracterización para la controlabilidad en promedio que depende únicamente de las matrices  $(A, B)$ . En particular esta caracterización es independiente del horizonte de tiempo  $T$ .

**Observación 1.2.3** ([58, Remark 2]). La condición de rango II) del Teorema 1.2.3 para las matrices  $A(\omega)$  y  $B(\omega)$ , puede ser simplificada cuando se tiene que  $A(\omega) = A\alpha(\omega)$  y  $B(\omega) = B\beta(\omega)$  con  $A, B$  matrices constantes y  $\alpha, \beta$  variables aleatorias<sup>3</sup>. En tal caso

$$\int_{\Omega} [A(\omega)]^j B(\omega) d\mu = A^j B \int_{\Omega} \alpha^j(\omega) \beta(\omega) d\mu.$$

<sup>2</sup>La matriz asociada a la condición del rango, corresponde a una matriz extendida, cuya  $j$ -ésima columna es  $\int_{\Omega} [A(\omega)]^j B(\omega) d\mu$ .

<sup>3</sup>Dado un espacio de probabilidad  $(\Omega, \mathcal{F}, \mu)$  y un espacio medible  $(S, \Sigma)$ , una aplicación  $\alpha : \Omega \rightarrow S$  es una variable aleatoria si es una aplicación  $\mathcal{F}, \Sigma$ -medible. Usualmente  $(S, \Sigma) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ , donde  $\mathcal{B}(\mathbb{R})$  corresponden a los Boreliandos.

Por tanto la matriz extendida del Teorema 1.2.3, queda de la forma

$$\left[ \int_{\Omega} B(\omega) d\mu, \int_{\Omega} A(\omega)B(\omega) d\mu, \dots \right] = \left[ B \int_{\Omega} \beta(\omega) d\mu, AB \int_{\Omega} \alpha(\omega)\beta(\omega) d\mu, \dots \right].$$

Ahora si se verifica que

$$\int_{\Omega} [\alpha(\omega)]^j \beta(\omega) d\mu \neq 0 \quad j = 0, \dots, n-1,$$

la condición II) del Teorema 1.2.3 es equivalente a la condición del rango de Kalman, es decir,

$$\text{Rank} \left[ B, AB, \dots, A^{n-1}B \right] = n.$$

En lo que sigue introduciremos la noción y definición, de que un par  $(A, B) = (A(\omega), B(\omega))$  sea controlable en promedio. Probaremos algunas implicancias que esto tiene.

**Observación 1.2.4.** Desde la fórmula de Duhamel, sabemos que la solución del sistema (1.1) para cada  $\omega$ , viene dada por

$$x(t, \omega) = R(t, \omega)x_0 + \int_0^t R(t-s, \omega)B(\omega)u(s)ds. \quad (1.8)$$

Tomando  $t = T$  en (1.8), e integrando con respecto a  $\Omega$  obtenemos

$$\int_{\Omega} x(T, \omega) d\mu = \int_{\Omega} R(T, \omega)x_0 d\mu + \int_{\Omega} \int_0^T R(T-s, \omega)B(\omega)u(s)ds d\mu. \quad (1.9)$$

Luego podemos definir el operador

$$\begin{aligned} \Phi : L^2([0, T], \mathbb{R}^m) &\rightarrow \mathbb{R}^n \\ u &\mapsto \int_0^T \int_{\Omega} R(T-s, \omega)B(\omega)u(s) d\mu ds. \end{aligned} \quad (1.10)$$

Si (1.1) satisface la propiedad de controlabilidad en promedio, entonces desde (1.9) tenemos

$$\int_{\Omega} x(T, \omega) d\mu = \int_{\Omega} R(T, \omega) d\mu x_0 + \Phi(u) = x_1,$$

donde el factor  $\int_{\Omega} R(T, \omega) d\mu x_0$  es una traslación de la imagen de  $\Phi$ . Luego se tiene la siguiente equivalencia

$$\text{Im}(\Phi(u)) = \mathbb{R}^n \quad \text{si y solo si,} \quad (1.1) \text{ es controlable en promedio.}$$



Observemos que si (1.1) no es controlable en promedio, entonces  $\text{Im}(\Phi(u)) \neq \mathbb{R}^n$ . En tal caso, para cada  $u \in L^2(0, T; \mathbb{R}^m)$ , existe  $v \in \mathbb{R}^n \setminus \{0\}$  tal que

$$\langle v, \Phi(u) \rangle_{\mathbb{R}^n} = v^* \Phi(u) = 0.$$

Es decir

$$v^* \int_0^T \int_{\Omega} R(T-s, \omega) B(\omega) u(s) d\mu ds = 0.$$

Ahora como  $u$  es independiente de  $\omega \in \Omega$  tenemos

$$\int_0^T \left( \int_{\Omega} v^* R(T-s, \omega) B(\omega) d\mu \right) u(s) ds = 0 \quad \forall u \in L^2((0, T), \mathbb{R}^m).$$

Así como consecuencia de que la estimación anterior, es para todo  $u \in L^2((0, T), \mathbb{R}^m)$ , se sigue que

$$\int_{\Omega} v^* R(T-s, \omega) B(\omega) d\mu = 0 \quad \text{en } L^2((0, T), \mathbb{R}^m),$$

trasponiendo la expresión anterior obtenemos

$$\int_{\Omega} B^*(\omega) R^*(T-s, \omega) v d\mu = 0 \quad \text{en } L^2((0, T), \mathbb{R}^n).$$

Por tanto, tomando norma en  $L^2((0, T), \mathbb{R}^n)$  tenemos

$$\int_0^T \left\| \int_{\Omega} B^*(\omega) R^*(T-s, \omega) v d\mu \right\|_{\mathbb{R}^n}^2 ds = 0.$$

Así

$$\int_0^T \left( \int_{\Omega} B^*(\omega) R^*(T-s, \omega) v d\mu \right)^* \left( \int_{\Omega} B^*(\omega) R^*(T-s, \omega) v d\mu \right) ds = 0.$$

Lo que es equivalente a

$$v^* \cdot \int_0^T \left( \int_{\Omega} R(T-s, \omega) B(\omega) d\mu \right) \left( \int_{\Omega} B^*(\omega) R^*(T-s, \omega) d\mu \right) ds \cdot v = 0. \quad (1.11)$$

Finalmente definimos

$$\begin{aligned} \mathcal{G} &:= \int_0^T \left( \int_{\Omega} R(T-s, \omega) B(\omega) d\mu \right) \left( \int_{\Omega} B^*(\omega) R^*(T-s, \omega) d\mu \right) ds \\ &= \int_0^T \left( \int_{\Omega} R^{-1}(s, \omega) R(T, \omega) B(\omega) d\mu \right) \left( \int_{\Omega} B^*(\omega) R^*(T, \omega) R^{-*}(s, \omega) d\mu \right) ds, \end{aligned}$$

matriz que motiva el siguiente teorema.

**Teorema 1.2.4.** *Las siguientes proposiciones son equivalentes:*

- a) El sistema (1.1) es controlable en promedio.
- b) La matriz  $\mathcal{G}$  es invertible.
- c) Para todo  $v \in \mathbb{R}^n$  existe  $K > 0$  tal que

$$\|v\|^2 \leq K \int_0^T \left\| \int_{\Omega} B^*(\omega) R^*(T-s, \omega) v d\mu \right\|_{\mathbb{R}^n}^2 ds. \quad (1.12)$$

*Demostración.* b)  $\Rightarrow$  a) Supongamos que  $\mathcal{G}$  es invertible, luego el control

$$u^*(s) = \int_{\Omega} B^*(\omega) R^*(T-s, \omega) d\mu \cdot \mathcal{G}^{-1} \cdot \int_{\Omega} \left( x_T - R(T, \omega) x_0 \right) d\mu,$$

para cada  $x_0, x_T \in \mathbb{R}^n$  está bien definido, donde  $x_T$  es arbitrario. Ahora evaluando en  $t = T$  (1.8) e integrado con respecto a  $\Omega$ , obtenemos

$$\int_{\Omega} x(T, \omega) d\mu = \int_{\Omega} R(T, \omega) x_0 d\mu + \int_0^T \int_{\Omega} R(T-s, \omega) B(\omega) u(s) d\mu ds.$$

Así remplazando  $u$  por  $u^*$  tenemos

$$\begin{aligned} \int_{\Omega} x(T, \omega) d\mu &= \int_{\Omega} R(T, \omega) x_0 d\mu + \int_0^T \left[ \int_{\Omega} R(T-s, \omega) B(\omega) d\mu \cdot \right. \\ &\quad \left. \cdot \int_{\Omega} B^*(\omega) R^*(T-s, \omega) d\mu \cdot \mathcal{G}^{-1} \cdot \int_{\Omega} \left( x_T - R(T, \omega) x_0 \right) d\mu \right] ds \\ &= \int_{\Omega} R(T, \omega) x_0 d\mu + \int_0^T \left( \int_{\Omega} R(T-s, \omega) B(\omega) d\mu \right) \cdot \\ &\quad \cdot \left( \int_{\Omega} B^*(\omega) R^*(T-s, \omega) d\mu \right) ds \cdot \mathcal{G}^{-1} \cdot \int_{\Omega} \left( x_T - R(T, \omega) x_0 \right) d\mu \\ &= \int_{\Omega} R(T, \omega) x_0 d\mu + \int_{\Omega} x_T - R(T, \omega) x_0 d\mu = x_T. \end{aligned}$$

Dado que  $x_T$  es arbitrario, se sigue que (1.1) es controlable en promedio.

**a)  $\Rightarrow$  b)** Supongamos que (1.1) es controlable y procedamos por contradicción suponiendo que  $\mathcal{G}$  no es invertible. Si esto es así, entonces existe  $v \in \mathbb{R}^n \setminus \{0\}$  tal que  $v^* G v = 0$ , así

$$\begin{aligned} 0 &= v^* \cdot \int_0^T \left( \int_{\Omega} R(T-s, \omega) B(\omega) d\mu \right) \left( \int_{\Omega} B^*(\omega) R(T-s, \omega) d\mu \right) dt \cdot v \\ &= \int_0^T \left\| \int_{\Omega} v^* R(T-s, \omega) B(\omega) d\mu \right\|_{\mathbb{R}^n}^2 ds, \end{aligned}$$

lo que implica que

$$\int_{\Omega} v^* R(T-s, \omega) B(\omega) d\mu = 0 \quad \text{c.t.p. en } (0, T).$$

Por otro lado, utilizando (1.8), evaluando en  $t = T$  e integrando con respecto a  $\Omega$  obtenemos

$$\int_{\Omega} x(T, \omega) d\mu = \int_{\Omega} R(T, \omega) x_0 d\mu + \int_0^T \int_{\Omega} R(T-s, \omega) B(s) u(s) ds,$$

multiplicando por  $v^*$  por la izquierda tenemos

$$v^* \cdot \int_{\Omega} x(T, \omega) d\mu = v^* \int_{\Omega} R(T, \omega) x_0 d\mu + \int_0^T \underbrace{\left( \int_{\Omega} v^* R(T-s, \omega) B(s) d\mu \right)}_{=0} u(s) ds.$$

Así

$$v^* \cdot \int_{\Omega} x(T, \omega) d\mu = v^* \int_{\Omega} R(T, \omega) x_0 d\mu.$$

Observemos que en particular si tomamos  $x_0 = 0$ , entonces  $\langle v, \int_{\Omega} x(T, \omega) d\mu \rangle_{\mathbb{R}^n} = 0$ , lo que implica que  $v$  es perpendicular a  $\int_{\Omega} x(T, \omega) d\mu$  para cualquier control. Por tanto no existe control tal que  $\int_{\Omega} x(T, \omega) d\mu = v$  cuando  $x_0 = 0$ , contradiciendo la controlabilidad de (1.1).

**b)  $\Rightarrow$  c)** Supongamos que  $\mathcal{G}$  es invertible. Claramente  $\mathcal{G}$  es semi-definida positiva por definición, así junto a la invertibilidad de esta, se sigue que  $\mathcal{G}$  es definida positiva. Luego  $\mathcal{G}$  tiene todos sus valores propios  $\lambda_i$  positivos. Desde la simetría e invertibilidad de  $\mathcal{G}$ , sus vectores propios  $w_i$  definen una base ortogonal de  $\mathbb{R}^n$ . Por tanto, dado  $v \in \mathbb{R}^n$ , se puede representar como

$$v = \sum_{i=1}^n w_i \alpha_i,$$

para  $\alpha_i$  constantes reales. Así

$$v^* \mathcal{G} v = \sum_{i=1}^n \sum_{j=1}^n w_i \alpha_i \mathcal{G} w_j \alpha_j = \sum_{i=1}^n \sum_{j=1}^n w_i \alpha_i \lambda_j w_j \alpha_j = \sum_{i=1}^n \alpha_i^2 \lambda_i \geq \min_i \lambda_i \sum_{i=1}^n \alpha_i^2 = K \|v\|_{\mathbb{R}^n}^2,$$

con  $K = \min_i \lambda_i > 0$ . Finalmente

$$\begin{aligned}
 \|v\|_{\mathbb{R}^n}^2 K &\leq v^* \mathcal{G} v = v^* \cdot \int_0^T \left( \int_{\Omega} R^{-1}(s, \omega) R(T, \omega) B(\omega) d\mu \right) \left( \int_{\Omega} B^*(\omega) R^*(T, \omega) R^{-*}(s, \omega) d\mu \right) ds \cdot v \\
 &= \int_0^T \left( \int_{\Omega} B^*(\omega) R^*(T, \omega) R^{-*}(s, \omega) v d\mu \right)^* \left( \int_{\Omega} B^*(\omega) R^*(T, \omega) R^{-*}(s, \omega) v d\mu \right) ds \\
 &= \int_0^T \left\| \int_{\Omega} B^*(\omega) R^*(T, \omega) R^{-*}(s, \omega) v d\mu \right\|_{\mathbb{R}^m}^2 ds \\
 &= \int_0^T \left\| \int_{\Omega} B^*(\omega) R^*(T-s, \omega) v d\mu \right\|_{\mathbb{R}^m}^2 ds,
 \end{aligned}$$

probando la desigualdad (1.12).

**c)  $\Rightarrow$  b)** Observemos que por definición la matriz  $\mathcal{G}$  es simétrica y semi-definida positiva. Si tenemos la desigualdad (1.12), entonces esta implica que para todo  $v \in \mathbb{R}^n$

$$\|v\|_{\mathbb{R}^n}^2 K \leq v^* \mathcal{G} v,$$

obteniendo directamente que  $\mathcal{G}$  es definida positiva, y por tanto invertible.  $\square$

**Observación 1.2.5.** La desigualdad (1.12) es equivalente a la desigualdad

$$\|v\|_{\mathbb{R}^n}^2 \leq K \int_0^T \left\| \int_{\Omega} B^*(\omega) R^*(t, \omega) v d\mu \right\|_{\mathbb{R}^m}^2 dt,$$

para todo  $v \in \mathbb{R}^n$ . Para probar lo anterior basta utilizar el cambio de variable  $t = T - s$ .

**Observación 1.2.6.** Desde el Teorema 1.2.4, notemos que la propiedad de controlabilidad en promedio del sistema (1.1), depende directamente de la matriz  $\mathcal{G}$ , la cual a su vez está relacionada únicamente con el par  $(A, B)$ , es decir, la invertibilidad de  $\mathcal{G}$  depende solo del par  $(A, B)$ .

**Definición 1.2.2.** Diremos que el par  $(A, B)$  es controlable en promedio si el sistema (1.1) es controlable en promedio.

**Corolario 1.2.4.1.** Si  $(A, B)$  es controlable en promedio, entonces existe  $K > 0$  tal que

$$\|\varphi(0)\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq K \int_0^T \left\| \int_{\Omega} B^*(\omega) \varphi(t, \omega) d\mu \right\|_{\mathbb{R}^m}^2 dt,$$

para  $\varphi$  solución del sistema adjunto

$$\begin{cases} -\varphi_t + A^*(\omega) \varphi = 0, \\ \varphi(T) = \varphi_T, \end{cases} \quad (1.13)$$

con  $\varphi_T \in \mathbb{R}^n$ .

*Demostración.* La solución del sistema adjunto (1.13), viene dada por  $\varphi(t, \omega) = R^*(T - t, \omega)\varphi_T$ . Evaluando en  $t = 0$  obtenemos

$$\varphi(0, \omega) = R^*(T, \omega)\varphi_T.$$

Así tomando norma en  $L^2(\Omega; \mathbb{R}^n)$  tenemos

$$\|\varphi(0)\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq \hat{K} \|\varphi_T\|_{L^2(\Omega; \mathbb{R}^n)}^2 = \hat{K} \|\varphi_T\|_{\mathbb{R}^n}^2.$$

Como el par  $(A, B)$  es controlable, por el teorema anterior existe  $K > 0$  tal que

$$\|v\|_{\mathbb{R}^n}^2 \leq K \int_0^T \left\| \int_{\Omega} B^*(\omega) R^*(T - t, \omega) v d\mu \right\|_{\mathbb{R}^m}^2 dt,$$

para todo  $v \in \mathbb{R}^n$ . En particular tomando  $v = \varphi_T$  en el Teorema 1.2.4, tenemos que

$$\begin{aligned} \|\varphi(0)\|_{L^2(\Omega; \mathbb{R}^n)}^2 &\leq \hat{K} \|\varphi_T\|_{\mathbb{R}^n}^2 \leq K \int_0^T \left\| \int_{\Omega} B^*(\omega) R^*(T - t, \omega) \varphi_T d\mu \right\|_{\mathbb{R}^m}^2 dt \\ &= K \int_0^T \left\| \int_{\Omega} B^*(\omega) \varphi(t, \omega) d\mu \right\|_{\mathbb{R}^m}^2 dt. \end{aligned} \quad (1.14)$$

□

### 1.3. Observabilidad en Promedio

En esta sección introduciremos la noción de observabilidad en promedio. Para ello consideremos el siguiente sistema

$$\begin{cases} y_t(t) + A(\omega)y(t) = B(\omega)u(t) & t \in (0, T), \\ y(0) = y_0, \\ z_u(t, y_0) = \int_{\Omega} C(\omega)y(t, \omega) d\mu, \end{cases} \quad (1.15)$$

donde  $A, C \in C(\Omega; \mathcal{L}(\mathbb{R}^n))$ ,  $B \in C(\Omega; \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n))$  uniformemente acotadas con respecto a  $\omega \in \Omega$ , y  $y_0 \in \mathbb{R}^n$  es independiente del parámetro  $\omega$ . Además, consideremos un control  $u \in L^2(0, T; \mathbb{R}^m)$ .

Llamaremos a  $C(\omega)$  como la matriz de observabilidad asociada al sistema (1.15), y a  $z_u(\cdot, y_0)$  como la observación de (1.15) a partir de  $y_0 \in \mathbb{R}^n$ .

**Definición 1.3.1.** Diremos que (1.15) es observable en promedio si para todo par de condiciones iniciales  $y_1, y_0 \in \mathbb{R}^n$ , y para todo control  $u \in L^2([0, T], \mathbb{R}^m)$  tales que

$$z_u(t, y_0) = z_u(t, y_1) \quad \forall t \in [0, T],$$

se satisface que  $y_0 = y_1$ .

**Observación 1.3.1.** Observemos que esta noción coincide con la noción de observabilidad clásica, introducida en [47, Sección 6] para sistemas lineales, cuando la tripleta  $(A, B, C)$  es independiente del parámetro aleatorio.

**Observación 1.3.2.** Notemos que desde la fórmula de Duhamel, se tiene que

$$y(t, \omega) = R(t, \omega)y_0 + \int_0^t R(t-s, \omega)B(\omega)u(s)ds. \quad (1.16)$$

Así multiplicando (1.16) por  $C(\omega)$ , e integrando sobre  $\Omega$  tenemos

$$z_u(t, y_0) = \int_{\Omega} C(\omega)y(t, \omega)d\mu = \int_{\Omega} C(\omega)R(t, \omega)y_0d\mu + \hat{z}_u(t),$$

donde

$$\hat{z}_u(t) = \int_{\Omega} \left( C(\omega) \int_0^t R(t-s, \omega)B(\omega)u(s)ds \right) d\mu.$$

Como el control  $u$  es conocido, al igual que  $B(\omega)$ ,  $R(t, \omega)$  y  $C(\omega)$  tras integrar sobre  $\Omega$  (pues serán independientes de  $\omega$ ), es que  $\hat{z}_u(t)$  es una cantidad conocida, que a su vez es independiente de la condición inicial  $y_0$ .

Por lo tanto la observabilidad en promedio del sistema (1.15) con  $z_u(t, y_0)$ , es equivalente a la observabilidad en promedio utilizando  $z(t, y_0) := z_u(t, y_0) - \hat{z}_u(t)$ . Además, notemos que  $z(t, y_0)$  corresponde a la observación de

$$\begin{cases} y_t + A(\omega)y = 0, \\ y(0) = y_0, \\ z(t, y_0) = \int_{\Omega} C(\omega)y(t, \omega)d\mu, \end{cases} \quad (1.17)$$

a partir de  $y_0 \in \mathbb{R}^n$ . Concluyendo así que la observabilidad en promedio de (1.15), es equivalente a la observabilidad en promedio de (1.17). Lo anterior, al igual que en el caso de controlabilidad en promedio, se puede entretenerse como una traslación de la imagen de un operador (en este caso trasladado por  $\hat{z}_u(t)$ ).

Observemos que si el sistema (1.15) (o el sistema (1.17)) es observable en promedio, entonces se satisface que

$$\begin{aligned} z(t, y_0) - z(t, y_1) &= \int_{\Omega} C(\omega) R(t, \omega) y_0 d\mu - \int_{\Omega} C(\omega) R(t, \omega) y_1 d\mu \\ &= \int_{\Omega} C(\omega) R(t, \omega) (y_0 - y_1) d\mu = 0. \end{aligned}$$

Luego tenemos que

$$\int_0^T \left\| \int_{\Omega} C(\omega) R(s, \omega) (y_0 - y_1) d\mu \right\|_{\mathbb{R}^n}^2 ds = 0,$$

equivalente a

$$\int_0^T \left( \int_{\Omega} C(\omega) R(s, \omega) (y_0 - y_1) d\mu \right)^* \left( \int_{\Omega} C(\omega) R(s, \omega) (y_0 - y_1) d\mu \right) ds = 0,$$

y por tanto

$$(y_0 - y_1)^* \cdot \int_0^T \left( \int_{\Omega} R^*(s, \omega) C^*(\omega) d\mu \right) \left( \int_{\Omega} C(\omega) R(s, \omega) d\mu \right) ds \cdot (y_0 - y_1) = 0.$$

Así definiendo

$$\mathcal{L} = \int_0^T \left( \int_{\Omega} R^*(s, \omega) C^*(\omega) d\mu \right) \left( \int_{\Omega} C(\omega) R(s, \omega) d\mu \right) ds,$$

tenemos el siguiente teorema.

**Teorema 1.3.1.** *Las siguientes proposiciones son equivalentes:*

- a) El sistema (1.15) es observable en promedio.
- b) La matriz  $\mathcal{L}$  es invertible.
- c) Para todo  $v \in \mathbb{R}^n$  existe  $K > 0$  tal que

$$\|v\|_{\mathbb{R}^n}^2 \leq K \int_0^T \left\| \int_{\Omega} C(\omega) R(s, \omega) v d\mu \right\|_{\mathbb{R}^n}^2 ds. \quad (1.18)$$

*Demostración.* a)  $\Rightarrow$  b) Claramente por construcción  $\mathcal{L}$  es simétrica y semi-definida positiva. Luego basta probar que es definida positiva para concluir la invertibilidad de esta.

Supongamos que  $\mathcal{L}$  no es definida positiva, por tanto debe existir  $y_0 \in \mathbb{R}^n \setminus \{0\}$  tal que  $y_0^* \cdot \mathcal{L} y_0 = 0$ . Como el sistema (1.15) es observable en promedio, se tiene que si  $z_u(t, y_0) = 0$ , entonces  $y_0 = 0$ . Por la Observación 1.3.2 sabemos que la observabilidad en promedio del sistema (1.15)

es equivalente a la observabilidad en promedio del sistema (1.17). Por tanto tenemos que si  $z(t, y_0) = 0$  entonces  $y_0 = 0$ .

Así como la solución del sistema (1.17) es  $y(t, \omega) = R(t, \omega)y_0$ , tenemos que

$$\begin{aligned} \int_0^T \|z(t, y_0)\|^2 dt &= \int_0^T \left\| \int_{\Omega} C(\omega)y(t, \omega)d\mu \right\|_{\mathbb{R}^n}^2 dt = \int_0^T \left\| \int_{\Omega} C(\omega)R(t, \omega)y_0d\mu \right\|_{\mathbb{R}^n}^2 dt \\ &= y_0^* \cdot \mathcal{L}y_0 = 0. \end{aligned}$$

Por tanto  $y_0 = 0$ , lo cual es una contradicción pues tomamos  $y_0 \neq 0$ .

**b)  $\Rightarrow$  a)** Queremos probar que si  $z(t, y_0) = 0$  entonces  $y_0 = 0$ . Supongamos que  $z(t, y_0) = 0$ , entonces

$$y_0^* \cdot \mathcal{L}y_0 = \int_0^T \left\| \int_{\Omega} C(\omega)y(t, \omega)d\mu \right\|_{\mathbb{R}^n}^2 dt = \int_0^T \|z(t, y_0)\|^2 dt = 0.$$

Luego como  $\mathcal{L}$  es invertible y semi-definida positiva, tenemos que es definida positiva (en otro caso existiría un valor propio nulo). Así por la igualdad anterior concluimos que  $y_0 = 0$ , probando la observabilidad.

**b)  $\Rightarrow$  c)** Si  $\mathcal{L}$  es invertible entonces como es semi-definida positiva, se tiene que es definida positiva, en particular esto implica que todos sus valores propios  $\lambda_i$ 's son positivos. Además, como  $\mathcal{L}$  es simétrica e invertible, se tiene que sus vectores propios  $w_i$ 's definen una base ortogonal del  $\mathbb{R}^n$ . Así para  $v \in \mathbb{R}^n$ , existe  $\{\alpha_i\}_{i=1}^{i=n} \subset \mathbb{R}$  tal que

$$v = \sum_{i=1}^n w_i \alpha_i.$$

Por tanto, tenemos que

$$v^* \cdot \mathcal{L}v = \sum_{i=1}^n \sum_{j=1}^n w_i \alpha_i \mathcal{L}w_j \alpha_j = \sum_{i=1}^n \sum_{j=1}^n w_i \alpha_i \lambda_j w_j \alpha_j = \sum_{i=1}^n \alpha_i^2 \lambda_i \geq \min_i \lambda_i \sum_{i=1}^n \alpha_i^2 = K \|v\|_{\mathbb{R}^n}^2, \quad (1.19)$$

donde  $K = \min_i \{\lambda_i\}$  en la línea anterior. Finalmente reescribiendo (1.19) obtenemos

$$\|v\|_{\mathbb{R}^n}^2 K \leq v^* \cdot \mathcal{L}v = \int_0^T \left\| \int_{\Omega} C(\omega)R(t, \omega)v d\mu \right\|_{\mathbb{R}^n}^2 dt.$$

Probando así (1.18).



c)  $\Rightarrow$  b) Como  $\mathcal{L}$  es simétrica y semi-definida positiva, basta probar que es definida positiva, para tener la invertibilidad. Claramente desde (1.18) tenemos que

$$\|v\|_{\mathbb{R}^n}^2 K \leq \int_0^T \left\| \int_{\Omega} C(\omega) R(t, \omega) v d\mu \right\|_{\mathbb{R}^n}^2 dt = v^* \cdot \mathcal{L}v.$$

Se sigue que  $v^* \cdot \mathcal{L}v > 0$  para todo  $v \in \mathbb{R}^n \setminus \{0\}$ . Probando la invertibilidad de  $\mathcal{L}$ .  $\square$

**Observación 1.3.3.** La Observación 1.3.2 junto al teorema anterior, nos muestran que la propiedad de observabilidad en promedio solo depende del par  $(A, C)$ .

Lo anterior motiva la siguiente definición.

**Definición 1.3.2.** Diremos que el par  $(A, C)$  es observable en promedio, si el sistema (1.15) es observable en promedio.

**Corolario 1.3.1.1.** Si  $(A, C)$  es observable en promedio entonces  $y$ , la solución de (1.17) satisface que

$$\|y(T)\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq K \int_0^T \left\| \int_{\Omega} C(\omega) y(t, \omega) d\mu \right\|_{\mathbb{R}^n}^2 dt,$$

para algún  $K > 0$ .

*Demostración.* Primero notemos que la solución de (1.17) viene dada por la expresión  $y(t, \omega) = R(t, \omega)y_0$ , entonces

$$y(T, \omega) = R(T, \omega)y_0,$$

luego tomando norma en  $L^2(\Omega; \mathbb{R}^n)$  a ambos lados

$$\|y(T)\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq \hat{K} \|y_0\|_{L^2(\Omega; \mathbb{R}^n)}^2 = \hat{K} \|y_0\|_{\mathbb{R}^n}^2.$$

Por otro lado, gracias al Teorema 1.3.1 sabemos que si  $(A, C)$  es observable, entonces se tiene la desigualdad (1.18) para todo  $v \in \mathbb{R}^n$ , en particular si tomamos  $v = y_0$ , obtenemos junto a lo anterior

$$\begin{aligned} \|y(T)\|_{L^2(\Omega; \mathbb{R}^n)}^2 &\leq \hat{K} \|y_0\|_{\mathbb{R}^n}^2 \leq K \int_0^T \left\| \int_{\Omega} C(\omega) R(t, \omega) y_0 d\mu \right\|_{\mathbb{R}^n}^2 dt \\ &= K \int_0^T \left\| \int_{\Omega} C(\omega) y(t, \omega) d\mu \right\|_{\mathbb{R}^n}^2 dt. \end{aligned}$$

$\square$

## 1.4. Equivalencia entre Controlabilidad y Observabilidad

Con el fin de unir las nociones introducidas en las secciones anteriores, tenemos el siguiente teorema.

**Teorema 1.4.1.** *El par  $(A, B)$  es controlable en promedio si y solo si,  $(A^*, B^*)$  es observable en promedio.*

*Demostración.* En virtud del Teorema 1.2.4, el par  $(A, B)$  es controlable en promedio si y solamente si la matriz

$$\begin{aligned}\mathcal{G} &= \int_0^T \left( \int_{\Omega} R(T-s, \omega) B(\omega) d\mu \right) \left( \int_{\Omega} B^*(\omega) R^*(T-s, \omega) d\mu \right) ds \\ &= \int_0^T \left( \int_{\Omega} R^{-1}(s, \omega) R(T, \omega) B(\omega) d\mu \right) \left( \int_{\Omega} B^*(\omega) R^*(T, \omega) R^{-*}(s, \omega) d\mu \right) ds,\end{aligned}$$

es invertible. Además, sabemos que gracias al Teorema 1.3.1, el par  $(A^*, B^*)$  es observable en promedio si y solo si, la matriz (cambiando el rol de  $(A, C)$  por  $(A^*, B^*)$ )

$$\mathcal{L} = \int_0^T \left( \int_{\Omega} R(t, \omega) B(\omega) d\mu \right) \left( \int_{\Omega} B^*(\omega) R^*(t, \omega) d\mu \right) ds,$$

es invertible. Luego utilizando el cambio de variable  $t = T - s$  la igualdad anterior tenemos

$$\begin{aligned}\mathcal{L} &= - \int_T^0 \left( \int_{\Omega} R(T-s, \omega) B(\omega) d\mu \right) \left( \int_{\Omega} B^*(\omega) R^*(T-s, \omega) d\mu \right) ds \\ &= \int_0^T \left( \int_{\Omega} R(T, \omega) R^{-1}(s, \omega) B(\omega) d\mu \right) \left( \int_{\Omega} B^*(\omega) R^*(T, \omega) R^{-*}(s, \omega) d\mu \right) ds = \mathcal{G}.\end{aligned}$$

□

De forma análoga al Teorema 1.4.1, cambiando los roles del par  $(A, B)$  para la controlabilidad, podemos probar el siguiente teorema.

**Teorema 1.4.2.** *El par  $(A, C)$  es observable en promedio si solo sí,  $(A^*, C^*)$  es controlable en promedio.*

Finalmente podemos relacionar los resultados que nos entrega [58] (Teorema 1.2.3), con los obtenidos en este capítulo, obteniendo una caracterización de la observabilidad en promedio, vía el análisis del rango de una matriz extendida.

**Teorema 1.4.3.** *El par  $(A, C)$  es observable en promedio si y solamente si, la matriz extendida*

$$\left[ \int_{\Omega} (A^*(\omega))^i C^*(\omega) d\mu, \quad i \geq 0 \right], \quad (1.20)$$

*tiene rango completo.*

*Demostración.* La condición II) del Teorema 1.2.3 nos señala que  $(A^*, C^*)$  es controlable en promedio si y solo si, (1.20) tiene rango completo. Luego utilizando el Teorema 1.4.2 se concluye la demostración.  $\square$

**Observación 1.4.1.** Igualmente como en la Observación 1.2.3, cuando las matrices  $A(\omega), C(\omega)$  tiene la forma  $A\alpha(\omega)$  y  $C\zeta(\omega)$  respectivamente, si se verifica que

$$\int_{\Omega} (\alpha(\omega))^j \zeta(\omega) d\mu \neq 0, \quad j = 0, \dots, n-1,$$

entonces la condición del rango (1.20) para la observabilidad en promedio, es equivalente a la condición

$$\text{Rank} \left[ C^*, A^* C^*, \dots, (A^*)^{n-1} C^* \right] = n.$$

Correspondiente a la condición del rango de Kalman.

## 1.5. Comentarios y observaciones

En esta sección comentaremos distintas nociones de controlabilidad en promedio para sistemas finitos dimensionales. Además, introduciremos nociones de controlabilidad en promedio, para sistemas infinitos dimensionales (EDP's) con incertidumbre, y comentaremos sobre la bibliografía asociada a este tópico. Finalizaremos esta sección comentando algunos problemas abiertos relacionados con el control en promedio.

### 1.5.1. Más nociones de controlabilidad en promedio

Comencemos introduciendo la noción de controlabilidad a cero en promedio para sistemas finito-dimensionales con incertidumbre. Para ello consideramos el sistema (1.1), es decir

$$\begin{cases} x_t + A(\omega)x = B(\omega)u & t \in (0, T), \\ x(0) = x_0, \end{cases}$$

Así la definición de controlabilidad a cero en promedio corresponde a la siguiente.

**Definición 1.5.1.** Diremos que el sistema (1.1) es controlable en promedio si dado  $T > 0$ , para cada  $x_0$  trayectoria inicial en  $\mathbb{R}^n$ , existe un control  $u \in L^2(0, T; \mathbb{R}^n)$  tal que

$$\int_{\Omega} x(T, \omega) d\mu = 0.$$

La definición anterior puede ser caracterizada por un tipo de desigualdad de observabilidad en promedio. No es difícil probar que la controlabilidad a cero del sistema (1.1), es equivalente a la

continuidad del operador lineal

$$\Lambda : \int_{\Omega} B^*(\omega) \varphi(t, \omega) d\mu \in L^2(0, T; \mathbb{R}^m) \mapsto \int_{\Omega} \varphi(0, \omega) d\mu \in \mathbb{R}^n,$$

con  $\varphi$  solución de (1.5). Además, la continuidad del operador lineal  $\Lambda$ , es equivalente a la siguiente desigualdad de observabilidad en promedio

$$\left\| \int_{\Omega} \varphi(0, \omega) d\mu \right\|_{\mathbb{R}^n}^2 \leq K \int_0^T \left\| \int_{\Omega} B^*(\omega) \varphi(t, \omega) d\mu \right\|_{\mathbb{R}^m}^2 dt, \quad (1.21)$$

para  $K > 0$ . Por tanto la controlabilidad a cero en promedio, puede ser caracterizada por la desigualdad de observabilidad en promedio (1.21).

Desde la desigualdad de observabilidad en promedio (1.21), se puede apreciar que no necesariamente esta implica la desigualdad de observabilidad en promedio para el control exacto en promedio (1.5). Es más, como bien señala [58, Observación 4] podemos considerar la ecuación del oscilador armónico para  $\Omega = \{\omega_1, \omega_2\}$ , en tal caso el correspondiente sistema adjunto tiene como solución

$$\begin{aligned} \varphi_1(t) &= \varphi_T \cos(t - T), \\ \varphi_2(t) &= \varphi_T \cos\left(\left(\frac{\pi}{T} + 1\right)(t - T)\right). \end{aligned}$$

Así observemos que

$$\int_{\Omega} \varphi(0, \omega) d\mu = \frac{1}{2} \sum_{i=1}^2 \varphi_i(0) = \varphi_T (\cos(T) + \cos(\pi + T)) = 0.$$

Por tanto  $\int_{\Omega} \varphi(0, \omega) d\mu = 0$  independiente del valor de  $\varphi_T$ . Concluyendo así que la noción de controlabilidad a cero en promedio es una noción más débil que la controlabilidad exacta en promedio.

Por otro lado, podemos definir ahora una noción de control aproximado en promedio de la siguiente forma

**Definición 1.5.2.** Diremos que el sistema (1.1) es aproximadamente controlable en promedio, si para cada  $x_0, x_1 \in \mathbb{R}^n$  y  $\varepsilon > 0$ , existe un control  $u_{\varepsilon} \in L^2(0, T; \mathbb{R}^m)$ , tal que el promedio de la solución de (1.1) satisfice que

$$\left\| \int_{\Omega} x(T, \omega) d\mu - x_1 \right\|_{\mathbb{R}^n} < \varepsilon.$$

Como es comentado en [29, Observación 1.1], si (1.1) es aproximadamente controlable en promedio, entonces el operado lineal continuo  $\Phi$  definido en (1.10), tiene imagen densa en  $\mathbb{R}^n$ , es decir,

$\overline{Im(\Phi)} = \mathbb{R}^n$ . Más aún, como  $\mathbb{R}^n$  es un espacio finito dimensional, se sigue que  $Im(\Phi) = \mathbb{R}^n$ , y en consecuencia se tiene la controlabilidad exacta en promedio.

En conclusión, las nociones definidas anteriormente, control a cero en promedio, y control aproximado en promedio, son equivalentes a la controlabilidad exacta en promedio.

### 1.5.2. Estabilización de Sistemas con Incertidumbre

La estabilización de sistemas con incertidumbre no es algo nuevo. Existen diversos artículos, principalmente publicados en los años 70's en donde se aborda esta noción.

Si consideramos un sistema que depende de algún parámetro aleatorio, es natural pensar que su solución depende de este parámetro. En el libro de R. Khasminskii [24] se recopila una gran cantidad de artículos referidos al tema de estabilización de sistemas aleatorio. En lo que sigue nos centraremos en la noción de exponencialmente  $p$ -estabilizable.

**Definición 1.5.3.** Consideremos el espacio de probabilidad  $(\Omega, \mathcal{F}, \mu)$  y el sistema

$$\begin{cases} x_t(t, \omega) = f(x(t, \omega), t, \omega) & (t, \omega) \in (0, T) \times \Omega. \\ x(0) = x_0, \end{cases} \quad (1.22)$$

para  $x_0 \in \mathbb{R}^n$  y  $f$  conocida. Diremos que el sistema (1.22) es exponencialmente  $p$ -estabilizable, si existen constantes  $K, \gamma > 0$  tales que

$$\mathbb{E}(\|x(t)\|_{\mathbb{R}^n}^p) \leq K \|x_0\|_{\mathbb{R}^n}^p e^{-\gamma t},$$

para todo  $t \geq 0$ .

La definición anterior se puede encontrar en [24, Sección 1.5], junto a otras 6 definiciones de estabilidad para sistemas con incertidumbre.

En nuestro contexto de control en promedio, introducimos las siguientes definiciones. Para ello consideremos el sistema

$$\begin{cases} x_t(t, \omega) + A(\omega)x(t, \omega) = B(\omega)u(t), \\ x(0) = x_0. \end{cases} \quad (1.23)$$

**Definición 1.5.4.** Diremos que el sistema (1.23) es exponencialmente  $p$ -controlable en promedio si existen un control  $u \in L^2(0, T; \mathbb{R}^m)$  tal que (1.23) es exponencialmente  $p$ -estabilizable.

**Definición 1.5.5.** Sea  $A$  una matriz aleatoria sobre el espacio de probabilidad  $(\Omega, \mathcal{F}, \mu)$ . Diremos que  $A$  es anti-adjunta si para casi todo  $\omega \in \Omega$  se tiene que  $A^*(\omega) = -A(\omega)$ .

**Observación 1.5.1.** Si  $A$  es anti-adjunta, entonces se verifica que

$$\langle Av, v \rangle_{L^2(\Omega; \mathbb{R}^n)} = 0, \quad \forall v \in L^2(\Omega; \mathbb{R}^n).$$

En lo que sigue probaremos que el sistema (1.23) es exponencialmente 2-controlable en promedio cuando  $A$  es anti-adjunta. Para probar esto, consideremos el siguiente lema.

**Lema 1.5.1.** Supongamos que  $A$  es anti-adjunta, además, supongamos que el par  $(A, B)$  es controlable en promedio. Luego para  $x$  solución del sistema

$$\begin{cases} x_t(t, \omega) + A(\omega)x(t, \omega) = B(\omega)\mathbb{E}(B^*x(t)), \\ x(0) = x_0, \end{cases} \quad (1.24)$$

se satisface que

$$\|x_0\|_{\mathbb{R}^n}^2 \leq K \int_0^T \|\mathbb{E}(B^*x)\|_{\mathbb{R}^m}^2 dt. \quad (1.25)$$

*Demostración.* Para  $x$  solución de (1.24), consideremos  $x = y - \varphi$ , donde  $\varphi$  y  $y$  soluciones de

$$\begin{cases} \varphi_t(t, \omega) + A(\omega)\varphi(t, \omega) = 0, \\ \varphi(0) = -x_0, \end{cases} \quad (1.26)$$

y

$$\begin{cases} y_t(t, \omega) + A(\omega)y(t, \omega) = B(\omega)\mathbb{E}(B^*x(t)), \\ y(0) = 0. \end{cases} \quad (1.27)$$

Observemos que (1.26) corresponde a la ecuación adjunta. En efecto, utilizando el cambio de variable  $t \mapsto T - t$  y el hecho que  $A$  es anti-adjunta (1.26) es equivalente a

$$\begin{cases} -\varphi_t(t, \omega) + A^*(\omega)\varphi(t, \omega) = 0, \\ \varphi(T) = -x_0. \end{cases}$$

Así desde la hipótesis de la controlabilidad en promedio del par  $(A, B)$  se tiene la desigualdad de observabilidad

$$\|x_0\|_{\mathbb{R}^n}^2 \leq K \int_0^T \left\| \int_{\Omega} B^* \varphi d\mu \right\|_{\mathbb{R}^m}^2 dt = K \int_0^T \|\mathbb{E}(B^* \varphi)\|_{\mathbb{R}^m}^2 dt,$$

para  $K > 0$ . Por tanto

$$\|x_0\|_{\mathbb{R}^n}^2 \leq K \left( \int_0^T \|\mathbb{E}(B^*x)\|_{\mathbb{R}^m}^2 dt + \int_0^T \|\mathbb{E}(B^*y)\|_{\mathbb{R}^m}^2 dt \right). \quad (1.28)$$

Por otro lado, multiplicando (1.27) por  $y$ , tenemos que

$$\frac{1}{2} \frac{d}{dt} \|y(t)\|_{\mathbb{R}^n}^2 + \langle Ay, y \rangle_{\mathbb{R}^n} = \langle \mathbb{E}(B^*x), B^*y \rangle_{\mathbb{R}^n}.$$

Integrando con respecto a  $\Omega$  obtenemos

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|y(t)\|_{L^2(\Omega; \mathbb{R}^n)}^2 &= - \langle \mathbb{E}(B^*x), B^*y \rangle_{L^2(\Omega; \mathbb{R}^m)} \leq \|\mathbb{E}(B^*x)\|_{\mathbb{R}^m} \|B^*\|_{C(\Omega; \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))} \|y\|_{L^2(\Omega; \mathbb{R}^n)} \\ &\leq \frac{1}{2} (\|y\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \|\mathbb{E}(B^*x)\|_{\mathbb{R}^m}^2 \|B^*\|_{C(\Omega; \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))}^2). \end{aligned}$$

Así por consecuencia de la desigualdad de Gronwall tenemos que

$$\|y(t)\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq K \|B^*\|_{C(\Omega; \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))}^2 \int_0^T \|\mathbb{E}(B^*x)\|_{\mathbb{R}^m}^2 dt.$$

Además, como consecuencia tenemos que

$$\begin{aligned} \int_0^T \|\mathbb{E}(B^*y)\|_{\mathbb{R}^m}^2 dt &\leq \|B^*\|_{C(\Omega; \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))}^2 \int_0^T \|\mathbb{E}(y)\|_{\mathbb{R}^n}^2 dt \leq \|B^*\|_{C(\Omega; \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))}^2 \int_0^T \|y\|_{L^2(\Omega; \mathbb{R}^n)}^2 dt \\ &\leq TK \|B^*\|_{C(\Omega; \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))}^2 \int_0^T \|\mathbb{E}(B^*x)\|_{\mathbb{R}^m}^2 dt. \end{aligned}$$

Finalmente remplazando en (1.28)

$$\begin{aligned} \|x_0\|_{\mathbb{R}^n}^2 &\leq K \left( \int_0^T \|\mathbb{E}(B^*x)\|_{\mathbb{R}^m}^2 dt + TK \|B^*\|_{C(\Omega; \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))}^2 \int_0^T \|\mathbb{E}(B^*x)\|_{\mathbb{R}^m}^2 dt \right) \\ &\leq K \int_0^T \|\mathbb{E}(B^*x)\|_{\mathbb{R}^m}^2 dt, \end{aligned}$$

para  $K > 0$ , concluyendo la demostración.  $\square$

Así tenemos el siguiente teorema.

**Teorema 1.5.1.** *Supongamos que  $A$  es anti-adjunta, y además, supongamos el par  $(A, B)$  es controlable en promedio. Luego el sistema (1.23) es exponencialmente 2-controlable en promedio.*

*Demostración.* Consideremos el control  $u(t) = -\mathbb{E}(B^*x(t))$ . Multiplicando la ecuación (1.23) tenemos que

$$\frac{1}{2} \frac{d}{dt} \|x(t)\|_{\mathbb{R}^n}^2 + \langle Ax, x \rangle_{\mathbb{R}^n} = -\langle B\mathbb{E}(B^*x), x \rangle_{\mathbb{R}^n}$$

Así integrando en sobre  $\Omega$  tenemos que

$$\frac{1}{2} \frac{d}{dt} \|x(t)\|_{L^2(\Omega; \mathbb{R}^n)}^2 = - \int_{\Omega} \langle \mathbb{E}(B^*x), B^*x \rangle_{\mathbb{R}^m} = - \|\mathbb{E}(B^*x)\|_{\mathbb{R}^m}^2. \quad (1.29)$$

Desde aquí podemos observar que  $\|x(t)\|_{L^2(\Omega; \mathbb{R}^n)}^2$  es una función decreciente en tiempo. Ahora gracias al Lema 1.5.1, sabemos que existe  $K > 0$  tal que

$$\|x_0\|_{\mathbb{R}^n}^2 \leq K \int_0^T \|\mathbb{E}(B^*x)\|_{\mathbb{R}^m}^2 dt. \quad (1.30)$$

Entonces

$$- \int_0^T \|\mathbb{E}(B^*x)\|_{\mathbb{R}^m}^2 dt \leq -\frac{1}{K} \|x_0\|_{\mathbb{R}^n}^2.$$

Por tanto, integrando (1.29) en  $(0, T)$  y remplazando lo anterior tenemos que

$$\|x(T)\|_{L^2(\Omega; \mathbb{R}^n)}^2 - \|x_0\|_{\mathbb{R}^n}^2 \leq -\frac{2}{K} \|x_0\|_{\mathbb{R}^n}^2.$$

Así existe  $\lambda \in [0, 1)$  tal que

$$\|x(T)\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq \lambda \|x_0\|_{\mathbb{R}^n}^2$$

Luego tenemos que

$$\|x(kT)\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq \lambda^k \|x_0\|_{\mathbb{R}^n}^2 \leq e^{k \ln(\lambda)} \|x_0\|_{\mathbb{R}^n}^2 \quad \forall k \in \mathbb{N}.$$

Observemos que para cada  $t \geq 0$  se tiene que existir  $\delta \in [0, T)$  tal que  $t = kT + \delta$  para  $k \in \mathbb{N}$ . Entonces utilizando el hecho que  $\|x(t)\|_{L^2(\Omega; \mathbb{R}^n)}^2$  es decreciente, y que  $\lambda \in [0, 1)$  tenemos

$$\begin{aligned} \|x(t)\|_{L^2(\Omega; \mathbb{R}^n)}^2 &\leq \|x(kT)\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq e^{-k|\ln(\lambda)|} \|x_0\|_{\mathbb{R}^n}^2 = e^{-(\frac{t}{T} - \frac{\delta}{T})|\ln(\lambda)|} \|x_0\|_{\mathbb{R}^n}^2 \\ &= e^{-\frac{t}{T}|\ln(\lambda)|} e^{\frac{\delta}{T}|\ln(\lambda)|} \|x_0\|_{\mathbb{R}^n}^2 \leq \frac{1}{\lambda} e^{-t\frac{|\ln(\lambda)|}{T}} \|x_0\|_{\mathbb{R}^n}^2 = Ke^{-\gamma t} \|x_0\|_{\mathbb{R}^n}^2. \end{aligned}$$

Por tanto

$$\mathbb{E}(\|x(t)\|_{\mathbb{R}^n}^2) = \|x(t)\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq Ke^{-\gamma t} \|x_0\|_{\mathbb{R}^n}^2,$$



para  $K, \gamma > 0$ , concluyendo la demostración.  $\square$

**Observación 1.5.2.** Si bien el resultado anterior, nos prevé de una noción de estabilidad para el sistema (1.23), es posible considerar otras nociones como las introducidas en [24, Sección 1.5]. Conjeturamos que posiblemente si introducimos una noción que involucre controlabilidad en el sentido de probabilidad, posiblemente sea necesario considerar alguna noción de estabilidad en el sentido de probabilidad (ver Sección 1.5.4).

### 1.5.3. Control en Promedio para Ecuaciones Diferenciales Parciales

A continuación, introduciremos un breve resumen de controlabilidad en promedio para sistemas infinito-dimensionales. La mayor parte de los siguientes resultados, se encuentran en [30].

Sea  $H, U, V$  tres espacios de Hilbert, donde  $V \subset H$  es denso. Además, sea  $T > 0$  y el conjunto Lebesgue medible  $E \subset [0, T]$ . Consideremos el espacio de probabilidad  $(\Omega, \mathcal{F}, \mu)$ , e introducimos el siguiente problema abstracto de control

$$\begin{cases} y_t(t) + A(\omega)y(t) = \chi_E B(\omega)u(t) & \text{en } (0, T], \\ y(0) = y_0, \end{cases} \quad (1.31)$$

donde  $y_0 \in V$ ,  $u \in L^2(E; U)$  es el control,  $B \in L^2(\Omega; \mathcal{L}(U, V))$  y  $A$  satisface que

- $A \in L^2(\Omega; \mathcal{L}(D(A), H))$ .
- $A(\omega) : D(A) \rightarrow H$  generará un  $C_0$ -semigrupo  $\{\mathcal{T}(t, \omega)\}_{t \geq 0}$  tanto en  $H$  como en  $V$ , para cada  $\omega \in \Omega$ .
- $\mathcal{T}(t, \cdot)u \in L^2(\Omega; V)$  para cada  $u \in V$  y  $t \in [0, T]$ .

**Observación 1.5.3.** Observemos que para cada  $\omega \in \Omega$  c.t.p. la solución  $y(\cdot, \omega) \in C([0, T]; V)$ <sup>4</sup>. Además, el promedio verifica que  $\int_{\Omega} y(\cdot, \omega) d\mu \in C([0, T]; V)$

Al igual como ocurre en el contexto determinista, existen diferentes nociones de controlabilidad en sistemas infinito-dimensionales, las cuales no necesariamente son equivalentes. A continuación, presentamos algunas nociones de controlabilidad.

**Definición 1.5.6.** Diremos que el sistema (1.31) es exactamente controlable en promedio en  $E$  con costo de control  $K > 0$ , si dado cualquier  $y_0, y_1 \in V$  existe un control  $u \in L^2(E; U)$  tal que

$$\|u\|_{L^2(E; V)} \leq K(\|y_0\|_V + \|y_1\|_V),$$

<sup>4</sup>Para cada  $\omega$  c.t.p. el problema (1.31) corresponde a un problema abstracto de Cauchy, recuperando los resultados de semigrupo [9].

y el promedio de la solución de (1.31) con condición inicial  $y_0$  satisface que

$$\int_{\Omega} y(T, \omega) d\mu = y_1.$$

**Definición 1.5.7.** Diremos que el sistema (1.31) es controlable a cero en promedio en  $E$  con costo de control  $K > 0$ , si dado cualquier  $y_0 \in V$  existe un control  $u \in L^2(E; U)$  tal que

$$\|u\|_{L^2(E; V)} \leq K \|y_0\|_V,$$

y el promedio de la solución de (1.31) con condición inicial  $u_0$  satisface que

$$\int_{\Omega} y(T, \omega) d\mu = 0.$$

**Definición 1.5.8.** Diremos que el sistema (1.31) es aproximadamente controlable en promedio en  $E$ , si para cualquier  $y_0, y_1 \in V$  y  $\varepsilon > 0$  existe un control  $u_\varepsilon \in L^2(E; U)$  tal que el promedio de la solución de (1.31) con condición inicial  $y_0$  satisface que

$$\left\| \int_{\Omega} y(T, \omega) d\mu - y_1 \right\|_V < \varepsilon.$$

De igual forma de como en el contexto finito dimensional, podemos caracterizar las nociones de controlabilidad definidas anteriormente, a partir de una desigualdad de observabilidad. Para ello consideremos el sistema adjunto dependiente del parámetro  $\omega$

$$\begin{cases} -z_t(t) + A^*(\omega)z(t) = 0 & \text{en } [0, T) \\ z(T) = z_0 \end{cases} \quad (1.32)$$

donde  $z_0 \in V'$ , con  $V'$  el dual del espacio  $V$  con espacio pivote  $H$ .

Para caracterizar las nociones de controlabilidad introducidas anteriormente, tenemos las siguientes definiciones

**Definición 1.5.9.** El sistema (1.32) es exactamente observable en promedio en  $E$ , si existe una constante  $K > 0$  tal que para cualquier  $z_0 \in V'$  se tiene que

$$\|z_0\|_{V'}^2 \leq K \int_0^T \chi_E \left\| \int_{\Omega} B^*(\omega)z(t, \omega) d\mu \right\|_U^2 dt.$$

**Definición 1.5.10.** El sistema (1.32) es observable a cero en promedio en  $E$ , si existe una constante  $K > 0$  tal que para cualquier  $z_0 \in V'$  se tiene que

$$\left\| \int_{\Omega} z(0, \omega) d\mu \right\|_{V'}^2 \leq K \int_0^T \chi_E \left\| \int_{\Omega} B^*(\omega) z(t, \omega) d\mu \right\|_U^2 dt.$$

**Definición 1.5.11.** El sistema (1.32) satisface la propiedad de continuación única en promedio en  $E$ , si el hecho que  $\chi_E \int_{\Omega} B^*(\omega) z(t, \omega) d\mu = 0$  en  $L^2(0, T; U)$  implica que  $z_0 = 0$  en  $V'$ .

Finalmente, los siguientes teoremas introducidos en [31], relacionan las nociones de controlabilidad y observabilidad en promedio introducidas anteriormente.

**Teorema 1.5.2.** El sistema (1.31) es exactamente controlable en promedio en  $E$ , si y solo si, el sistema adjunto (1.32) es exactamente observable en promedio.

El teorema anterior nos permite caracterizar la controlabilidad exacta a través de una desigualdad de observabilidad. En relación a la controlabilidad a cero y aproximada se tiene los siguientes teoremas

**Teorema 1.5.3.** El sistema (1.31) es controlable a cero en promedio en  $E$ , si y solo si, el sistema adjunto (1.32) es observable a cero en promedio.

**Teorema 1.5.4.** El sistema (1.31) es aproximadamente controlable en  $E$ , si y solo si, el sistema adjunto (1.32) satisface la propiedad de continuación en promedio en  $E$ .

Otra posible referencia relacionada a la controlabilidad en promedio analizada para sistemas abstractos puede ser [28] artículo que además, analiza la Ecuación de Onda y la Ecuación de Schrödinger unidimensional.

En relación con la observabilidad de sistemas no autónomos, en [21] los autores introducen una condición de Hautus promedio, probando así una caracterización de observabilidad exacta en promedio. Finalmente, estos resultados son aplicados a la ecuación de Schrödinger con potencial dependiente del tiempo y la ecuación de onda amortiguada con un coeficiente de amortiguación dependiente del tiempo.

A continuación, mostraremos algunos resultados para la controlabilidad en promedio de la ecuación del calor con incertidumbre.

Sea  $G \subset \mathbb{R}^d$  con  $d > 0$ , un conjunto acotado con frontera  $\partial G$  de clase  $C^2$  y consideremos el siguiente sistema

$$\begin{cases} y_t - \alpha \Delta y = \chi_{G_0 \times E} u & \text{en } G \times (0, T), \\ u = 0 & \text{sobre } \partial G \times (0, T), \\ y(0) = y_0 & \text{en } G. \end{cases} \quad (1.33)$$

Aquí  $y_0 \in L^2(G)$ ,  $G_0 \subset G$  no vacío y  $E \subset [0, T]$  un conjunto Lebesgue medible, con medida positiva. La constante de difusión  $\alpha : \Omega \rightarrow \mathbb{R}^+$  corresponde a una variable aleatoria.

**Observación 1.5.4.** Como nos señala la teoría de semigrupos, para que la ecuación (1.33) esté bien puesta, para cada realización de la variable aleatoria  $\alpha$ , es necesario que el operador  $\alpha\Delta$  tenga el signo correcto. En otras palabras es necesario que  $\alpha\Delta$  sea un operador  $M -$  Disipativo, y en consecuencia para tener un buen planteo del problema, la variable aleatoria  $\alpha$  solo puede tener soporte en  $\mathbb{R}^+$ .

En [31] fue probado que la ecuación (1.33), es aproximadamente controlable en promedio y controlable a cero en promedio, cuando  $\alpha$  se distribuye uniforme o exponencial. No obstante, recientemente en [11] fue probada la controlabilidad aproximada y a cero en promedio, cuando  $\alpha$  es una variable aleatoria arbitraria con soporte en  $\mathbb{R}^+$ .

Además, de lo anterior, en este contexto con incertidumbre, siguen estando presente los efectos regularizantes generados por la ecuación del calor. Por lo tanto, la propiedad de control exacto en promedio no se verifica para la ecuación del calor con incertidumbre.

**Observación 1.5.5.** La principal herramienta usada tanto en [31], como en [11], es la descomposición en series de Fourier de la solución de (1.33), descomposición proporcionada por las buenas propiedades del operador Laplaciano Dirichlet (ver [14, Capítulo 6.5]). En consecuencia, de lo anterior se pudo obtener una fórmula en series, para la solución de la ecuación adjunta.

Por otro lado, también puede ser encontrada en la literatura, trabajos relacionados a distintas ecuaciones. Entre algunos se encuentran [1], artículo que trata la controlabilidad a cero en promedio de ecuaciones hiperbólicas, como también [32], artículo que trata la controlabilidad exacta en promedio de la ecuación Bernoulli-Euler beam, en donde se busca un control robusto, es decir, un control asociado a un funcional que penaliza la dispersión del sistema en tiempo final.

Además, en el artículo [40] se considera el problema de control óptimo para ecuaciones del tipo parabólico, el documento se basa en la noción de “no-regret” y “low-regret control”, introducida por Lions [26] adaptando esta noción al contexto con incertidumbre. También el libro [34] introduce las nociones básicas de la teoría de control óptimo para EDP’s con incerteza, tanto la teoría como los conceptos numéricos, contraponiendo lo desarrollado en este libro con los resultados existentes para el contexto sin incertidumbre. Finalmente en relación con problemas de control en promedio para EDO’s fraccionarias y EDP’s fraccionarias (del tipo calor fraccionario), puede ser visto al artículo [38].

### 1.5.4. Algunos Problemas Abiertos

En esta breve sección comentamos algunos problemas abiertos en relación al tema principal de este capítulo, el cual corresponde a la controlabilidad de sistemas con coeficientes aleatorios.

- **Control en promedio para ecuaciones con coeficientes evolutivos:** Como bien se ha visto, en este capítulo, solo nos enfocamos cuando el parámetro aleatorio es estacionario, es decir, no evoluciona en el tiempo. Por lo anterior es que sería bastante interesante analizar que sucede cuando  $A, B$  son procesos estocásticos, es decir, ecuaciones de la forma

$$\begin{cases} \dot{\xi}_t + A(\omega(t))\xi = B(\omega(t))u(t), \\ \xi(0) = \xi_0. \end{cases}$$

Además, analizar cómo se puede caracterizar la controlabilidad en promedio en este caso. Quizás algunas referencias para intentar introducirse en este problema pueden ser el libro de T.Soong [48] y el libro de R.Khasiminskii [24].

- **Control en Promedio para Ecuaciones no lineales:** De forma análoga al problema de controlabilidad en promedio para ecuaciones lineales, podemos considerar este problema para ecuaciones aleatorias no lineales. Posiblemente para EDP's este problema se debe basar en las herramientas utilizadas en el contexto sin incertidumbre, como el teorema de la función inversa o alguno de punto fijo, quizás introduciendo nociones de controlabilidad en promedio de forma local. Mientras que para el contexto de EDO's posiblemente se deba abordar utilizando álgebra de Lie [10, Parte 2]. En este caso, desde el punto de vista de aplicaciones y modelación de fenómenos físicos, sería de interés saber que ocurre con la controlabilidad en promedio de ecuaciones en derivadas parciales del tipo Navier-Stokes, Burger's o incluso una ecuación difusión convección. En el capítulo 4 se planteará nuevamente esta pregunta, pues estaremos trabajando con una EDO no lineal con incertidumbre del tipo

$$\begin{cases} x_t(t, \omega) = f(u(t), x(t, \omega)) & t \in (0, T), \\ x(0) = x_0(\omega). \end{cases}$$

- **Control en Probabilidad:** Dentro de los problemas propuestos durante el desarrollo de este documento, se planteó introducir la noción de Control en Probabilidad. Esta noción puede ser introducida como sigue

**Definición 1.5.12.** Diremos que el sistema

$$\begin{cases} \dot{x}_t + A(\omega)x(t) = B(\omega)u(t) \\ x(0) = x_0 \end{cases} \quad (1.34)$$

es controlable en probabilidad en tiempo  $T > 0$ , si existe  $u \in L^2(0, T; \mathbb{R}^m)$  tal que para cada  $x_0, x_1 \in \mathbb{R}^n$  y para cada  $\delta, \varepsilon > 0$  se satisface que

$$\mu(\|x(T, \omega) - x_1\|_{\mathbb{R}^n} > \varepsilon) < \delta, \quad (1.35)$$

donde  $x$  corresponde a la solución de (1.34) con condición inicial  $x_0$ .

Observemos que por la desigualdad de Chebyshov nos entrega una relación entre (1.35) y la esperanza matemática, señalando que existe  $K > 0$  tal que

$$\mu(\|x(T, \omega) - x_1\|_{\mathbb{R}^n} > \varepsilon) \leq K\mathbb{E}(\|x(T, \omega) - x_1\|_{\mathbb{R}^n}),$$

en donde es claro que si  $\mathbb{E}(\|x(T, \omega) - x_1\|_{\mathbb{R}^n}) < \delta$  para cada  $\delta > 0$ , se tiene que el sistema (1.34) es controlable en probabilidad.

También es posible introducir otras nociones como querer que un sistema sea controlable en una probabilidad  $\alpha$  dada, es decir, que existe  $u \in L^2(0, T; \mathbb{R}^m)$  tal que para todo  $\varepsilon > 0$  se tiene que

$$\mu(\|x(T, \omega) - x_1\|_{\mathbb{R}^n} < \varepsilon) = \alpha.$$

La definición anterior nos señala que el “ $\alpha$  % de la veces”, la trayectoria final del sistema (1.34) será  $x_1$ . Además, de lo anterior, como bien se comenta en la Observación 1.5.2 puede ser también abordada la pregunta de la estabilización en el sentido de probabilidad.

## Capítulo 2

# Propiedad de Turnpike para Sistemas Finito Dimensionales

### 2.1. Introducción

#### 2.1.1. Motivación:

La propiedad de turnpike, es una propiedad relacionada con problemas de control óptimo. Las ideas e intuiciones tras la propiedad de turnpike, comenzaron a desarrollándose a mediados de los años 40's en [42] por John Von Neumann. No obstante, el uso de la palabra turnpike no fue utilizado hasta [12], donde señalan que la evolución de ciertos fenómenos de crecimiento tiende a ser constante cuando su punto de partida y su punto final, están suficientemente lejos. Los autores asocian esto a que cuando se quiere viajar una larga distancia, uno prefiere usar una autopista de peaje y llegar así más rápido al destino que ir por una vía menor, las cuales suelen ser más costosas en el sentido energético.

Es natural pensar desde la descripción anterior que nos podríamos encontrar frente a un problema de control óptimo, con un horizonte temporal  $T$  grande. En donde existe una dinámica evolutiva y otra estacionaria, las cuales se acercan durante un intervalo de tiempo para luego alejarse, siempre y cuando el horizonte de tiempo sea lo suficientemente grande. Esta propiedad se ha formalizado de diferentes formas, algunas más débiles que otras, pero manteniendo el espíritu del concepto.

Existen variados artículos que prueban la propiedad de turnpike, siendo uno de los primeros [56] en el cual se trata la propiedad de turnpike exponencial, pero desde una perspectiva del cálculo de variaciones. No obstante, nos centraremos en lo realizado en [44]. En este artículo se formaliza la propiedad de turnpike a través de una estimación para las trayectorias y los controles óptimos asociados a problemas de control óptimo, probando la propiedad tanto en el contexto finito dimensional, como en el contexto infinito dimensional. Desde entonces se han publicado diversos documentos en donde se analiza la propiedad de turnpike. Algunos de ellos son [55, 45, 51, 22, 16, 59, 49, 50, 20, 6, 15, 19] entre otros. El camino que toma la mayoría de

estos artículos para obtener la propiedad de turnpike es suponer alguna hipótesis de controlabilidad, o también algunos consideran algún tipo de estabilidad del sistema. Para el caso no lineal se busca linealizar las ecuaciones y por tanto se hacen necesarias condiciones de pequeñez, concluyendo así la propiedad de turnpike de forma local.

Siendo más precisos, en el artículo [44] se señala que si los pares  $(x^T, u^T)$  y  $(\bar{x}, \bar{u})$  satisfacen la estimación

$$\|x^T - \bar{x}\| + \|u^T - \bar{u}\| \leq C(e^{-\delta t} + e^{-\delta(T-t)}), \quad (2.1)$$

entonces verifican la propiedad de turnpike, donde el par  $(x^T, u^T)$  es la solución de un problema de control óptimo evolutivo, mientras el par  $(\bar{x}, \bar{u})$  es solución de un problema de control óptimo estacionario. Además,  $C$  y  $\delta$  son constantes positivas independiente del horizonte de tiempo  $T$ . Esta desigualdad llamada propiedad de turnpike exponencial, nos permite señalar que para un horizonte de tiempo  $T$  suficientemente grande, la trayectoria estacionaria y evolutiva se encuentran cercanas en una vecindad del centro del intervalo  $[0, T]$ .

No obstante, lo señalado, como se ha desarrollado en el capítulo anterior, supondremos que las dinámicas que gobiernan a las trayectorias  $x^T$  y  $\bar{x}$  dependen de un parámetro aleatorio  $\omega$ . Por lo cual surge naturalmente la siguiente pregunta ¿La dinámica promedio satisface la propiedad de Turnpike? En este capítulo nos centraremos en probar la propiedad de Turnpike, pero en un contexto con incertidumbre.

### 2.1.2. Problema a estudiar

Comenzaremos introduciendo el problema y las definiciones básicas para tener un buen entendimiento de este. Para ello al igual que en el Capítulo 1, consideramos un espacio de probabilidad  $(\Omega, \mathcal{F}, \mu)$ , donde  $\Omega$  corresponde al espacio muestral,  $\mathcal{F}$  la  $\sigma$ -álgebra y  $\mu$  la medida definida sobre  $\mathcal{F}$ . Consideremos el siguiente sistema

$$\begin{cases} x_t + A(\omega)x = B(\omega)u & t \in (0, T), \\ x(0) = x_0, \end{cases} \quad (2.2)$$

donde  $\omega \in \Omega$  corresponde al parámetro aleatorio,  $x(\cdot, \omega) \in C([0, T]; \mathbb{R}^n)$  corresponde al estado, para cada  $\omega \in \Omega$ ,  $A(\omega) \in \mathcal{L}(\mathbb{R}^n)$  y  $B(\omega) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$  corresponden a dos matrices. Supondremos que las matrices  $A(\omega), B(\omega)$  son uniformemente acotadas con respecto al parámetro  $\omega$ . Además, el control  $u \in L^2(0, T; \mathbb{R}^m)$  y  $x_0 \in \mathbb{R}^n$  serán independientes de  $\omega$ .

Consideremos el siguiente problema de minimización

$$\min_{u \in L^2(0, T; \mathbb{R}^m)} \left\{ J^T(u) = \frac{1}{2} \left( \int_0^T \|u(t)\|_{\mathbb{R}^m}^2 dt + \int_0^T \int_{\Omega} \|C(\omega)x(t, \omega) - z\|_{\mathbb{R}^n}^2 d\mu dt \right) \right\}, \quad (2.3)$$



con  $x(t, \omega)$  solución de (2.2),  $C(\omega) \in \mathcal{L}(\mathbb{R}^n)$  una matriz que depende del parámetro aleatorio  $\omega \in \Omega$  y  $z \in \mathbb{R}^n$  una trayectoria fija. Además, consideraremos el problema estacionario dado por

$$\min_{u \in \mathbb{R}^m} \left\{ J_s(u) = \frac{1}{2} \left( \|u\|_{\mathbb{R}^m}^2 + \int_{\Omega} \|C(\omega)x(\omega) - z\|_{\mathbb{R}^n}^2 d\mu \right) \right\}, \quad (2.4)$$

con  $x(\omega)$  solución de  $A(\omega)x = B(\omega)u$ . Consideramos el siguiente espacio de Hilbert

$$L^2(\Omega; \mathbb{R}^n; \mu) = \left\{ y(\omega) \in \mathbb{R}^n : \|y\|_{L^2(\Omega; \mathbb{R}^n; \mu)}^2 = \int_{\Omega} \|y(\omega)\|_{\mathbb{R}^n}^2 d\mu < \infty \right\},$$

con el producto interno

$$\langle x, y \rangle_{L^2(\Omega; \mathbb{R}^n; \mu)} = \int_{\Omega} \langle x(\omega), y(\omega) \rangle_{\mathbb{R}^n} d\mu.$$

Por simplicidad denotaremos  $L^2(\Omega; \mathbb{R}^n) = L^2(\Omega; \mathbb{R}^n; \mu)$ . Junto a lo anterior, y con ciertas hipótesis que serán especificadas en la Sección 2.3, si los problemas (2.3) y (2.4) tienen solución  $(u^T, x^T)$  y  $(\bar{u}, \bar{x})$  respectivamente, se probará que existen constantes  $\mathcal{C}, \delta > 0$  independientes de  $T$ , tales que

$$\left\| \mathbb{E}(x^T) - \mathbb{E}(\bar{x}) \right\|_{\mathbb{R}^n} + \left\| \mathbb{E}(\varphi^T) - \mathbb{E}(\bar{\varphi}) \right\|_{\mathbb{R}^n} + \left\| u^T - \bar{u} \right\|_{\mathbb{R}^m} \leq \mathcal{C}(e^{-(T-t)\delta} + e^{-t\delta}),$$

para todo  $t \in [0, T]$ , donde  $(\varphi^T, \bar{\varphi})$  corresponden a las variables duales que representa la caracterización de los minimizadores, vía las condiciones de optimalidad de primer orden y donde  $\mathbb{E}$  corresponde a la esperanza matemática<sup>1</sup>.

Adicionalmente, probaremos que el promedio temporal del par óptimo  $(u^T, x^T)$ , converge al par óptimo  $(\bar{u}, \bar{x})$ , es decir

$$\begin{aligned} \frac{1}{T} \int_0^T x^T(t, \omega) dt &\longrightarrow \bar{x}(\omega), \\ \frac{1}{T} \int_0^T u^T(t) dt &\longrightarrow \bar{u}, \end{aligned}$$

cuando  $T \rightarrow \infty$ , en  $L^2(\Omega; \mathbb{R}^n)$  e  $\mathbb{R}^m$  respectivamente.

---

<sup>1</sup>Es decir,  $\mathbb{E}(u) = \int_{\Omega} u(\omega) d\mu$ .

## 2.2. Estimaciones previas

En esta sección probaremos una serie de lemas los cuales nos permitirán concluir nuestros teoremas principales. Para ello es necesario probar algunas estimaciones previas. Estas estimaciones se obtendrán a partir de hipótesis sobre las matrices aleatorias  $A(\omega)$ ,  $B(\omega)$  y  $C(\omega)$ , hipótesis motivadas por la noción de V-exponentially stabilizable y V-exponentially detectable introducidas en [19]. La parte principal de la estimación corresponden a desigualdades tanto para  $x^T$ , la trayectoria óptima asociada al problema (2.3) en tiempo final  $T$ , como para  $\varphi^T$ , la solución de un sistema adjunto (el cual será mostrado en la siguiente sección) en tiempo 0, es decir, estimaciones para  $x^T(T, \omega)$  y  $\varphi^T(0, \omega)$  respectivamente.

Supondremos que  $A \in C(\Omega, \mathcal{L}(\mathbb{R}^n))$ ,  $B \in C(\Omega, \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n))$  y  $C \in C(\Omega, \mathcal{L}(\mathbb{R}^n))$  son uniformemente acotadas con respecto a  $\omega \in \Omega$ . Además, sean  $K, \hat{K}$  dos constantes positivas que pueden cambiar en cada línea. Así consideraremos la ecuación

$$y_t + A(\omega)y = f(t, \omega) \quad t \in (0, T), \quad (2.5)$$

donde  $f \in L^2(0, T, L^2(\Omega; \mathbb{R}^n))$ . A continuación presentamos nuestra primera hipótesis

**Hipótesis 1.** Para el par  $(A, C)$  suponemos la siguiente condición. Existe un operador feedback  $K_C \in C(\Omega; \mathcal{L}(\mathbb{R}^n))$  uniformemente acotado con respecto a  $\omega$  y una constantes  $\zeta > 0$  tal que

$$\langle (A + K_C C)v, v \rangle_{L^2(\Omega; \mathbb{R}^n)} \geq \zeta \|v\|_{L^2(\Omega; \mathbb{R}^n)}^2 \quad \forall v \in L^2(\Omega; \mathbb{R}^n).$$

**Observación 2.2.1.** La hipótesis anterior nos señala que existe un operador (matriz)  $K_C$  feedback tal que hace al operador conjunto  $(A + K_C C)$  elíptico. La hipótesis anterior es motivada por la noción de v-exponencialmente detectable.

Bajo la hipótesis anterior tenemos el siguiente resultado.

**Lema 2.2.1.** Supongamos que se satisface la hipótesis 1. Entonces existe  $K > 0$  tal que

$$\|y(T)\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \|y\|_{L^2(0, T; L^2(\Omega; \mathbb{R}^n))}^2 \leq K \left( \int_0^T \|f\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \|Cy\|_{L^2(\Omega; \mathbb{R}^n)}^2 dt + \|y(0)\|_{L^2(\Omega; \mathbb{R}^n)}^2 \right), \quad (2.6)$$

para y solución de (2.5).

*Demostración.* Procedemos multiplicando (2.5) por  $y$  e integrando en  $(0, T) \times \Omega$ , obtenemos

$$\begin{aligned} \frac{\|y(T)\|_{L^2(\Omega; \mathbb{R}^n)}^2}{2} - \frac{\|y(0)\|_{L^2(\Omega; \mathbb{R}^n)}^2}{2} + \int_0^T \left\{ \langle Ay, y \rangle_{L^2(\Omega; \mathbb{R}^n)} + \langle K_C Cy, y \rangle_{L^2(\Omega; \mathbb{R}^n)} \right\} dt = \\ \int_0^T \left\{ \langle f, y \rangle_{L^2(\Omega; \mathbb{R}^n)} + \langle \alpha K_C Cy, y \rangle_{L^2(\Omega; \mathbb{R}^n)} \right\} dt. \end{aligned}$$

Usando la hipótesis 1 se tiene

$$\frac{\|y(T)\|_{L^2(\Omega;\mathbb{R}^n)}^2}{2} \leq \int_0^T \left\{ \langle f, y \rangle_{L^2(\Omega;\mathbb{R}^n)} + \langle K_C C y, y \rangle_{L^2(\Omega;\mathbb{R}^n)} - \zeta \|y\|_{L^2(\Omega;\mathbb{R}^n)}^2 \right\} dt + \frac{\|y(0)\|_{L^2(\Omega;\mathbb{R}^n)}^2}{2}.$$

Así gracias a la desigualdad de Young para  $\varepsilon_1, \varepsilon_2 > 0$ , se tiene

$$\begin{aligned} \frac{\|y(T)\|_{L^2(\Omega;\mathbb{R}^n)}^2}{2} &\leq \int_0^T \left\{ \frac{\|f\|_{L^2(\Omega;\mathbb{R}^n)}^2}{2\varepsilon_1} + \frac{\varepsilon_1 \|y\|_{L^2(\Omega;\mathbb{R}^n)}^2}{2} + \langle K_C C y, y \rangle_{L^2(\Omega;\mathbb{R}^n)} - \zeta \|y\|_{L^2(\Omega;\mathbb{R}^n)}^2 \right\} dt \\ &\quad + \frac{\|y(0)\|_{L^2(\Omega;\mathbb{R}^n)}^2}{2} \\ &\leq \int_0^T \left\{ \frac{\|f\|_{L^2(\Omega;\mathbb{R}^n)}^2}{2\varepsilon_1} + \frac{\varepsilon_1 \|y\|_{L^2(\Omega;\mathbb{R}^n)}^2}{2} + \alpha \frac{\|C y\|_{L^2(\Omega;\mathbb{R}^n)}^2}{2\varepsilon_2} \right. \\ &\quad \left. + \frac{\varepsilon_2 \|y\|_{L^2(\Omega;\mathbb{R}^n)}^2}{2} - \zeta \|y\|_{L^2(\Omega;\mathbb{R}^n)}^2 \right\} dt + \frac{\|y(0)\|_{L^2(\Omega;\mathbb{R}^n)}^2}{2}, \end{aligned} \quad (2.7)$$

donde  $\alpha$  corresponde a una constante positiva que acota la norma de  $K_C$ . Reordenando los términos tenemos

$$\begin{aligned} \frac{\|y(T)\|_{L^2(\Omega;\mathbb{R}^n)}^2}{2} + \left( \zeta - \frac{(\varepsilon_1 + \varepsilon_2)}{2} \right) \|y\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^n))}^2 \\ \leq \int_0^T \left\{ \frac{\|f\|_{L^2(\Omega;\mathbb{R}^n)}^2}{2\varepsilon_1} + \alpha \frac{\|C y\|_{L^2(\Omega;\mathbb{R}^n)}^2}{2\varepsilon_2} \right\} dt + \frac{\|y(0)\|_{L^2(\Omega;\mathbb{R}^n)}^2}{2}. \end{aligned}$$

Luego tomando  $\varepsilon_1 + \varepsilon_2 < 2\zeta$ , existe  $K > 0$  tal que

$$\|y(T)\|_{L^2(\Omega;\mathbb{R}^n)}^2 + \|y\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^n))}^2 \leq K \left( \int_0^T \left\{ \|f\|_{L^2(\Omega;\mathbb{R}^n)}^2 + \|C y\|_{L^2(\Omega;\mathbb{R}^n)}^2 \right\} dt + \|y(0)\|_{L^2(\Omega;\mathbb{R}^n)}^2 \right).$$

□

**Observación 2.2.2.** Bajo la hipótesis 1 podemos probar que existe  $K > 0$  tal que se verifica la desigualdad estacionaria

$$\|x\|_{L^2(\Omega;\mathbb{R}^n)}^2 \leq K(\|Ax\|_{L^2(\Omega;\mathbb{R}^n)}^2 + \|Cx\|_{L^2(\Omega;\mathbb{R}^n)}^2). \quad (2.8)$$

Para ello tomando  $y = x \in L^2(\Omega; \mathbb{R}^n)$ , y reemplazando  $f(t, \omega) = A(\omega)x$  en la desigualdad (2.7) obtenemos

$$\begin{aligned} \frac{\|x\|_{L^2(\Omega; \mathbb{R}^n)}^2}{2} &\leq T \frac{\|Ax\|_{L^2(\Omega; \mathbb{R}^n)}^2}{2\varepsilon_1} + T \frac{\varepsilon_1 \|x\|_{L^2(\Omega; \mathbb{R}^n)}^2}{2} + T\alpha \frac{\|Cx\|_{L^2(\Omega; \mathbb{R}^n)}^2}{2\varepsilon_2} \\ &\quad + T \frac{\varepsilon_2 \|x\|_{L^2(\Omega; \mathbb{R}^n)}^2}{2} - T\gamma \|x\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \frac{\|x\|_{L^2(\Omega; \mathbb{R}^n)}^2}{2}. \end{aligned}$$

Así reagrupando

$$\left(\gamma - \frac{\varepsilon_1}{2} - \frac{\varepsilon_2}{2}\right) \|x\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq \frac{\|Ax\|_{L^2(\Omega; \mathbb{R}^n)}^2}{2\varepsilon_1} + \alpha \frac{\|Cx\|_{L^2(\Omega; \mathbb{R}^n)}^2}{2\varepsilon_2}.$$

Finalmente tomando  $\varepsilon_1, \varepsilon_2 > 0$  tal que  $\varepsilon_1 + \varepsilon_2 \leq 2\gamma$ , se concluye que

$$\|x\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq K(\|Ax\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \|Cx\|_{L^2(\Omega; \mathbb{R}^n)}^2).$$

La desigualdad (2.2.2), nos permitirá concluir que existe una única trayectoria óptima  $\bar{x}$ , asociada al control óptimo del problema estacionario (2.4), demostración que será detallada más adelante. Además, esta desigualdad será de suma importancia para probar nuestros teoremas principales.

Ahora consideremos el siguiente sistema

$$-\varphi_t + A^*(\omega)\varphi = g(t, \omega) \quad t \in (0, T), \quad (2.9)$$

donde  $g \in L^2(0, T; L^2(\Omega; \mathbb{R}^n))$ . Así consideramos la siguiente hipótesis que nos permitirá estimar  $\varphi^T(0, \omega)$ .

**Hipótesis 2.** Para el par  $(A^*, B^*)$  suponemos la siguiente condición. Existen dos constantes  $\lambda > 0$  y  $\kappa > 0$  tales que

$$\langle A^*v, v \rangle_{L^2(\Omega; \mathbb{R}^n)} + \kappa \left\| \int_{\Omega} B^*v d\mu \right\|_{\mathbb{R}^m}^2 \geq \lambda \|v\|_{L^2(\Omega; \mathbb{R}^n)}^2 \quad \forall v \in L^2(\Omega; \mathbb{R}^n).$$

**Observación 2.2.3.** La hipótesis anterior es motivada por la noción  $v$ -exponencialmente estabilizable. Además, en el caso que la matriz  $A^*$  sea definida positiva, esta hipótesis siempre será verdad, con  $\lambda$  su menor valor propio.

Bajo la hipótesis anterior podemos concluir el siguiente resultado.

**Lema 2.2.2.** *Supongamos que se satisface la hipótesis 2. Entonces existe una constante  $K > 0$  tal que*

$$\|\varphi(0)\|_{L^2(\Omega;\mathbb{R}^n)}^2 + \|\varphi\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^n))}^2 \leq K \left( \int_0^T \|g\|_{L^2(\Omega;\mathbb{R}^n)}^2 + \left\| \int_{\Omega} B^* \varphi d\mu \right\|_{\mathbb{R}^m}^2 dt + \|\varphi(T)\|_{L^2(\Omega;\mathbb{R}^n)}^2 \right),$$

para  $\varphi$  solución de (2.9).

*Demostración.* Multiplicando (2.9) por  $\varphi$  e integrando en  $(0, T) \times \Omega$ , obtenemos

$$\begin{aligned} \frac{1}{2} \left( \|\varphi(0)\|_{L^2(\Omega;\mathbb{R}^n)}^2 - \|\varphi(T)\|_{L^2(\Omega;\mathbb{R}^n)}^2 \right) + \int_0^T \left\{ \langle A\varphi, \varphi \rangle_{L^2(\Omega;\mathbb{R}^n)} + \left\| \int_{\Omega} B^* \varphi d\mu \right\|_{\mathbb{R}^m}^2 \right\} dt = \\ \int_0^T \left\{ \langle g, \varphi \rangle_{L^2(\Omega;\mathbb{R}^n)} + \left\| \int_{\Omega} B^* \varphi d\mu \right\|_{\mathbb{R}^m}^2 \right\} dt. \end{aligned}$$

Ahora utilizando la hipótesis 2 junto a la desigualdad de Young para  $\varepsilon_1 > 0$

$$\begin{aligned} \frac{\|\varphi(0)\|_{L^2(\Omega;\mathbb{R}^n)}^2}{2} &\leq \int_0^T \left\{ \langle g, \varphi \rangle_{L^2(\Omega;\mathbb{R}^n)} + \kappa \left\| \int_{\Omega} B^* \varphi d\mu \right\|_{\mathbb{R}^m}^2 - \lambda \|\varphi\|_{L^2(\Omega;\mathbb{R}^n)}^2 \right\} dt \\ &\quad + \frac{\|\varphi(T)\|_{L^2(\Omega;\mathbb{R}^n)}^2}{2} \\ &\leq \int_0^T \left\{ \frac{\|g\|_{L^2(\Omega;\mathbb{R}^n)}^2}{2\varepsilon_1} + \varepsilon_1 \frac{\|\varphi\|_{L^2(\Omega;\mathbb{R}^n)}^2}{2} + \kappa \left\| \int_{\Omega} B^* \varphi d\mu \right\|_{\mathbb{R}^m}^2 - \lambda \|\varphi\|_{L^2(\Omega;\mathbb{R}^n)}^2 \right\} dt \\ &\quad + \frac{\|\varphi(T)\|_{L^2(\Omega;\mathbb{R}^n)}^2}{2}. \end{aligned} \quad (2.10)$$

Así tomando  $\varepsilon_1 < 2\lambda$ , existe  $K > 0$  tal que

$$\|\varphi(0)\|_{L^2(\Omega;\mathbb{R}^n)}^2 + \|g\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^n))}^2 \leq K \left( \int_0^T \|g\|_{L^2(\Omega;\mathbb{R}^n)}^2 + \left\| \int_{\Omega} B^* \varphi d\mu \right\|_{\mathbb{R}^m}^2 dt + \|\varphi(T)\|_{L^2(\Omega;\mathbb{R}^n)}^2 \right). \quad (2.11)$$

□

**Observación 2.2.4.** *Bajo la hipótesis 2 podemos probar que existe  $K > 0$  tal que se verifica la desigualdad estacionaria*

$$\|p\|_{L^2(\Omega;\mathbb{R}^n)}^2 \leq K \left( \|A^* p\|_{L^2(\Omega;\mathbb{R}^n)}^2 + \left\| \int_{\Omega} B^* p d\mu \right\|_{\mathbb{R}^m}^2 \right).$$

Para ello basta tomar  $\varphi = p \in L^2(\Omega; \mathbb{R}^n)$  y remplazar  $g(t, \omega) = A^*(\omega)p$  en la desigualdad (2.10), obteniendo

$$\begin{aligned} \frac{\|p\|_{L^2(\Omega; \mathbb{R}^n)}^2}{2} &\leq T \frac{\|A^*p\|_{L^2(\Omega; \mathbb{R}^n)}^2}{2\varepsilon_1} + T\varepsilon_1 \frac{\|p\|_{L^2(\Omega; \mathbb{R}^n)}^2}{2} + T\kappa \left\| \int_{\Omega} B^*p d\mu \right\|_{\mathbb{R}^m}^2 \\ &\quad - T\lambda \|p\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \frac{\|p\|_{L^2(\Omega; \mathbb{R}^n)}^2}{2}. \end{aligned}$$

Reagrupando los términos

$$\left(\lambda - \frac{\varepsilon_1}{2}\right) \|p\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq \frac{\|A^*p\|_{L^2(\Omega; \mathbb{R}^n)}^2}{2\varepsilon_1} + \kappa \frac{\left\| \int_{\Omega} B^*p d\mu \right\|_{\mathbb{R}^m}^2}{2\varepsilon_2}.$$

Finalmente tomando  $\varepsilon_1 > 0$  tal que  $\varepsilon_1 \leq 2\lambda$ , se concluye que existe  $K > 0$  tal que

$$\|p\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq K \left( \|A^*p\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \left\| \int_{\Omega} B^*p d\mu \right\|_{\mathbb{R}^m}^2 \right).$$

Esta desigualdad nos permitirá probar la unicidad de  $\bar{\varphi}$ , solución de un sistema adjunto asociado a la ecuación estacionaria  $A(\omega)x = B(\omega)u$ .

**Lema 2.2.3.** Consideremos el sistema de ecuaciones

$$\begin{cases} m_t + Am = f, \\ -n_t + A^*m = g. \end{cases}$$

Entonces para cada  $\omega \in \Omega$  se satisface que

$$\langle n(0), m(0) \rangle_{\mathbb{R}^n} - \langle n(T), m(T) \rangle_{\mathbb{R}^n} = \int_0^T \langle g, m \rangle_{\mathbb{R}^n} dt - \int_0^T \langle f, n \rangle_{\mathbb{R}^n} dt.$$

*Demostración.* Si multiplicando la primera ecuación por  $n$ , la segunda ecuación por  $m$ , e integrando en  $(0, T)$  obtenemos

$$\begin{cases} \int_0^T \langle m_t, n \rangle_{\mathbb{R}^n} + \langle Am, n \rangle_{\mathbb{R}^n} dt = \int_0^T \langle f, n \rangle_{\mathbb{R}^n} dt, \\ - \int_0^T \langle n_t, m \rangle_{\mathbb{R}^n} + \langle A^*n, m \rangle_{\mathbb{R}^n} dt = \int_0^T \langle g, m \rangle_{\mathbb{R}^n} dt. \end{cases}$$

Luego integrando por partes la segunda y restando ambas ecuaciones, se concluye.  $\square$

**Lema 2.2.4.** *Supongamos que se tiene la hipótesis 1. Sea  $y$  solución de (2.5) con dato inicial  $y(0) = y_0$ . Supongamos que existe  $K > 0$  tal que tenemos las siguientes estimaciones*

$$\int_0^T \|f\|_{L^2(\Omega; \mathbb{R}^n)}^2 dt \leq KT, \quad (2.12a)$$

$$\|y(T)\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq KT, \quad (2.12b)$$

$$\int_0^T \|Cy\|_{L^2(\Omega; \mathbb{R}^n)}^2 dt \leq KT. \quad (2.12c)$$

Entonces se verifica que  $\frac{1}{T} \int_0^T y(t, \omega) dt$  es acotado en  $L^2(\Omega; \mathbb{R}^n)$ . En particular si, además

$$\frac{1}{T} \int_0^T Cy dt \rightarrow 0, \quad \frac{1}{T} \int_0^T f dt \rightarrow 0, \quad \text{en } L^2(\Omega; \mathbb{R}^n), \quad (2.13)$$

entonces se tiene que

$$\frac{1}{T} \int_0^T y(t, \omega) dt \rightarrow 0, \quad \text{en } L^2(\Omega; \mathbb{R}^n).$$

*Demostración.* Primero observemos que por (2.12b) se tiene

$$\left\| \frac{y(T) - y_0}{T} \right\|_{L^2(\Omega; \mathbb{R}^n)} \leq \frac{\|y(T)\|_{L^2(\Omega; \mathbb{R}^n)} + \|y_0\|_{L^2(\Omega; \mathbb{R}^n)}}{T} \leq \frac{\sqrt{KT} + \|y_0\|_{L^2(\Omega; \mathbb{R}^n)}}{T} \rightarrow 0,$$

cuando  $T \rightarrow \infty$ . Por tanto, existe  $\hat{K} > 0$  tal que

$$\left\| \frac{y(T) - y_0}{T} \right\|_{L^2(\Omega; \mathbb{R}^n)} \leq \hat{K}.$$

Por otro lado, integrando en  $(0, T)$  la ecuación (2.5), y multiplicando por  $\frac{1}{T}$  tenemos

$$\frac{1}{T} \int_0^T Ay dt = \frac{1}{T} \int_0^T f dt - \left( \frac{y(T) - y_0}{T} \right). \quad (2.14)$$

Tomando norma en  $L^2(\Omega; \mathbb{R}^n)$  a ambos lados y aplicando desigualdad de Hölder junto a (2.12a), se tiene que

$$\begin{aligned}
 \left\| \frac{1}{T} \int_0^T Ay \, dt \right\|_{L^2(\Omega; \mathbb{R}^n)} &\leq \left\| \frac{1}{T} \int_0^T f \, dt \right\|_{L^2(\Omega; \mathbb{R}^n)} + \left\| \frac{y(T) - y_0}{T} \right\|_{L^2(\Omega; \mathbb{R}^n)} \\
 &\leq \frac{1}{T} \int_0^T \|f\|_{L^2(\Omega; \mathbb{R}^n)} \, dt + \hat{K} \\
 &\leq \frac{1}{T} \left( \int_0^T \|f\|_{L^2(\Omega; \mathbb{R}^n)}^2 \, dt \right)^{1/2} \left( \int_0^T 1 \, dt \right)^{1/2} + \hat{K} \\
 &\leq \frac{1}{T} \left( \int_0^T \|f\|_{L^2(\Omega; \mathbb{R}^n)}^2 \, dt \right)^{1/2} \sqrt{T} + \hat{K} \\
 &\leq K + \hat{K} = K.
 \end{aligned} \tag{2.15}$$

Así obtenemos que  $\frac{1}{T} \int_0^T Ay \, dt$  es acotado en  $L^2(\Omega; \mathbb{R}^n)$ . Ahora por la hipótesis 1 podemos utilizando la observación 2.2.2, obteniendo

$$\begin{aligned}
 \left\| \frac{1}{T} \int_0^T y \, dt \right\|_{L^2(\Omega; \mathbb{R}^n)}^2 &\leq K \left( \left\| \frac{1}{T} \int_0^T Ay \, dt \right\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \left\| \frac{1}{T} \int_0^T Cy \, dt \right\|_{L^2(\Omega; \mathbb{R}^n)}^2 \right) \\
 &\leq \beta K + \frac{K}{T^2} \int_0^T \|Cy\|_{L^2(\Omega; \mathbb{R}^n)}^2 \, dt \leq \hat{K} + \frac{\beta}{T} \leq K.
 \end{aligned} \tag{2.16}$$

La última parte del lema sigue directamente observando que (2.15), junto a las hipótesis extras (2.13) implica que  $\frac{1}{T} \int_0^T Ay \, dt \rightarrow 0$ . Finalmente reemplazando en la desigualdad (2.16), se concluye la demostración.  $\square$

**Observación 2.2.5.** La conclusión anterior también es válida para  $\varphi$ , solución del sistema (2.9), y suponiendo que

$$\begin{aligned}
 \int_0^T \|g\|_{L^2(\Omega; \mathbb{R}^n)} \, dt &\leq KT, \\
 \|\varphi(0)\|_{L^2(\Omega; \mathbb{R}^n)}^2 &\leq KT, \\
 \int_0^T \left\| \int_{\Omega} B^* \varphi \, d\mu \right\|_{\mathbb{R}^m}^2 \, dt &\leq KT,
 \end{aligned}$$

junto a la hipótesis 2.



## 2.3. Teoremas Principales de la Sección

Consideraremos el sistema (2.2) enunciado en la introducción, es decir

$$\begin{cases} x_t(t, \omega) + A(\omega)x(t, \omega) = B(\omega)u(t) & t \in (0, T), \\ x(0) = x_0, \end{cases} \quad (2.17)$$

donde  $x(\cdot, \omega) \in C([0, T]; \mathbb{R}^n)$  corresponde al estado. Supondremos que  $A \in C(\Omega, \mathcal{L}(\mathbb{R}^n))$ ,  $B \in C(\Omega, \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n))$ . Además,  $u \in L^2(0, T; \mathbb{R}^m)$  es el control. Con lo anterior tenemos que  $x(t, \cdot) \in L^2(\Omega; \mathbb{R}^n)$ , esto desde el Teorema 1.2.1.

Continuando con el problema en estudio, consideramos el siguiente problema de minimización

$$\begin{aligned} \min_{u \in L^2(0, T; \mathbb{R}^m)} \left\{ J^T(u) = \frac{1}{2} \left( \int_0^T \|u(t)\|_{\mathbb{R}^m}^2 dt + \int_0^T \int_{\Omega} \|C(\omega)x(t, \omega) - z\|_{\mathbb{R}^n}^2 d\mu dt \right) \right. \\ \left. = \frac{1}{2} \left( \int_0^T \left\{ \|u(t)\|_{\mathbb{R}^m}^2 + \|C(\omega)x(t, \omega) - z\|_{L^2(\Omega; \mathbb{R}^n)}^2 \right\} dt \right) \right\}, \end{aligned} \quad (2.18)$$

donde  $z \in \mathbb{R}^n$ , y  $C \in C(\Omega, \mathcal{L}(\mathbb{R}^n))$ . En relación con la existencia de un único mínimo para el problema (2.18) tenemos el siguiente resultado.

**Teorema 2.3.1.** *Existe un único control óptimo  $u^T \in L^2(0, T; \mathbb{R}^m)$  para el problema de minimización (2.18), y además existe una única trayectoria óptima  $x^T$  asociada a  $u^T$ .*

*Demostración.* La demostración de este teorema puede ser desarrollada vía el método directo del cálculo variacional. Algunas pruebas análogas pueden ser vistas en [52, 27]. Para este caso particular basta notar que el funcional  $J^T$  es continuo y estrictamente convexo. Además, la dependencia de  $x$  es lineal y continua con respecto a  $u$ , ya que  $u$  es independiente del parámetro aleatorio  $\omega$ . Por tanto, se puede concluir vía dicho método.  $\square$

El siguiente teorema nos permite caracterizar el mínimo del funcional  $J^T$ .

**Teorema 2.3.2.** *El único mínimo  $u^T \in L^2(0, T; \mathbb{R}^m)$  de (2.18) viene caracterizado por*

$$u^T(t) = - \int_{\Omega} B^* \varphi^T d\mu, \quad \begin{cases} -\varphi_t^T + A^*(\omega)\varphi^T = C^*(Cx^T - z), \\ \varphi^T(T) = 0. \end{cases} \quad (2.19)$$

*Demostración.* Como el funcional  $J^T$  es claramente Gateaux diferenciable, procedemos calculando sus condiciones necesarias de primer orden. Para ello observemos que<sup>2</sup>

$$\delta J^T(u)[\delta u] = \int_0^T \left\{ \langle u(t), \delta u(t) \rangle_{\mathbb{R}^m} + \langle C(\omega)x(t, \omega) - z, C(\omega)\delta x(t, \omega) \rangle_{L^2(\Omega; \mathbb{R}^n)} \right\} dt. \quad (2.20)$$

<sup>2</sup>Denotamos como  $\delta J^T(u)[\delta u]$  a la variación del funcional  $J^T(u)$  en la dirección  $\delta u$ .

Por otro lado  $\delta x(t, \omega)$  satisface el siguiente sistema

$$\begin{cases} \delta x_t(t, \omega) + A(\omega)\delta x(t, \omega) = B(\omega)\delta u(t), \\ \delta x(0) = 0. \end{cases} \quad (2.21)$$

Ahora si multiplicamos el sistema (2.21) por  $\varphi^T(t, \cdot) \in L^2(\Omega; \mathbb{R}^n)$ , e integramos entre  $(0, T)$  y  $\Omega$  obtenemos

$$\begin{aligned} \int_0^T \int_{\Omega} \left\{ \left\langle \delta x_t(t, \omega), \varphi^T(t, \omega) \right\rangle_{\mathbb{R}^n} + \left\langle A(\omega)\delta x(t, \omega), \varphi^T(t, \omega) \right\rangle_{\mathbb{R}^n} \right\} d\mu dt \\ = \int_0^T \int_{\Omega} \left\langle B(\omega)\delta u(t), \varphi^T(t, \omega) \right\rangle_{\mathbb{R}^n} d\mu dt. \end{aligned}$$

Así integrando por partes

$$\begin{aligned} \int_0^T \left\{ \left\langle \delta x(t, \omega), -\varphi_t^T(t, \omega) + A^*(\omega)\varphi^T(t, \omega) \right\rangle_{L^2(\Omega; \mathbb{R}^n)} \right\} dt + \left\langle \delta x(T, \omega), \varphi^T(T, \omega) \right\rangle \\ = \int_0^T \int_{\Omega} \left\langle \delta u(t), B^*(\omega)\varphi^T(t, \omega) \right\rangle_{\mathbb{R}^m} d\mu dt. \end{aligned}$$

Como el control  $u$  es independiente del parámetro aleatorio  $\omega \in \Omega$  tenemos

$$\begin{aligned} \int_0^T \left\{ \left\langle \delta x(t, \omega), -\varphi_t^T + A^*(\omega)\varphi^T(t, \omega) \right\rangle_{L^2(\Omega; \mathbb{R}^n)} \right\} dt + \left\langle \delta x(T, \omega), \varphi^T(T) \right\rangle \\ = \int_0^T \left\langle \delta u(t), \int_{\Omega} B^*(\omega)\varphi^T(t, \omega) d\mu \right\rangle_{\mathbb{R}^m} dt. \quad (2.22) \end{aligned}$$

En particular tomamos  $u^T$  como

$$u^T(t) = - \int_{\Omega} B^*(\omega)\varphi^T(t, \omega) d\mu, \quad (2.23)$$

y  $\varphi^T(T, \omega) = 0$  para todo  $\omega \in \Omega$ . Remplazando (2.22) en (2.20) se satisface que

$$\begin{aligned} \delta J^T(u^T)[\delta u] &= \int_0^T \left\{ \left\langle u^T, \delta u \right\rangle_{\mathbb{R}^m} + \left\langle C(\omega)x(t, \omega) - z, C(\omega)\delta x \right\rangle \right\} dt \\ &= \int_0^T \left\{ \left\langle \delta x(t, \omega), \varphi_t^T - A^*(\omega)\varphi^T(t, \omega) \right\rangle_{L^2(\Omega; \mathbb{R}^n)} \right. \\ &\quad \left. + \left\langle C^*(\omega)(C(\omega)x(t, \omega) - z), \delta x \right\rangle_{L^2(\Omega; \mathbb{R}^n)} \right\} dt. \end{aligned}$$

Así tenemos que

$$\delta J^T(u^T)[\delta u] = \int_0^T \left\{ \left\langle \delta x(t, \omega), \varphi_t^T - A^*(\omega)\varphi^T(t, \omega) + C^*(\omega)(C(\omega)x - z) \right\rangle_{L^2(\Omega; \mathbb{R}^n)} \right\} dt.$$

Por lo tanto, si  $\varphi^T$  es solución del sistema

$$\begin{cases} -\varphi_t^T + A^*(\omega)\varphi^T = C^*(\omega)(Cx^T - z), \\ \varphi^T(T) = 0. \end{cases}$$

Entonces se verifica que

$$\delta J^T(u^T)[\delta u] = 0.$$

Concluyendo que  $u^T$  dado por (2.23) es un extremal del funcional  $J^T$ . Por el Teorema 2.3.1 sabemos que este extremal es el único mínimo del funcional  $J^T$ .  $\square$

**Observación 2.3.1.** Antes de continuar observemos que bajo las hipótesis 1 y 2, a partir de los lemas 2.2.1 y 2.2.2, podemos obtener las siguientes estimaciones para  $\varphi^T(0, \omega)$  y  $x^T(T, \omega)$  <sup>3</sup>

1. Desigualdad evolutiva primal

$$\|x^T(T)\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq \hat{K} \left( \int_0^T \|u\|_{\mathbb{R}^m}^2 + \|C(\omega)x^T\|_{L^2(\Omega; \mathbb{R}^n)}^2 dt + \|x_0\|_{\mathbb{R}^n}^2 \right). \quad (2.24)$$

2. Desigualdad evolutiva dual

$$\begin{aligned} \|\varphi^T(0)\|_{L^2(\Omega; \mathbb{R}^n)}^2 &\leq \hat{K} \left( \int_0^T \|C(\omega)x^T - z\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \left\| \int_{\Omega} B^* \varphi^T d\mu \right\|_{\mathbb{R}^m}^2 dt \right) \\ &= \hat{K} \left( \int_0^T \|C(\omega)x^T - z\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \|u\|_{\mathbb{R}^m}^2 dt \right). \end{aligned} \quad (2.25)$$

Además, podemos concluir las siguientes estimaciones estacionarias desde las observaciones (2.2.4) y (2.2.2), estas son

3. Desigualdad estacionaria primal

$$\|x\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq K \left( \|Ax\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \|Cx\|_{L^2(\Omega; \mathbb{R}^n)}^2 \right). \quad (2.26)$$

<sup>3</sup>Remplazando  $f(t, \omega) = B(\omega)u(t)$  y  $g(t, \omega) = C^*(\omega)(C(\omega)x^T(t, \omega) - z)$ , junto a las condiciones iniciales  $x^T(0) = x_0$  y  $\varphi^T(T) = 0$ .

## 4. Desigualdad estacionaria dual

$$\|p\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq K \left( \|A^*p\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \left\| \int_{\Omega} B^* p d\mu \right\|_{\mathbb{R}^m}^2 \right). \quad (2.27)$$

Ahora analicemos el problema estacionario ya mostrado en la introducción

$$\min_{u \in \mathbb{R}^m} \left\{ J_s(u) = \frac{1}{2} \left( \|u\|_{\mathbb{R}^m}^2 + \|Cx - z\|_{L^2(\Omega; \mathbb{R}^n)}^2 \right) \right\}, \quad (2.28)$$

donde  $x$  soluciona el problema estacionario  $A(\omega)x = B(\omega)u$ . Así tenemos el siguiente resultado.

**Lema 2.3.1.** *Supongamos que tenemos la hipótesis 1. Entonces  $J_s$  admite un único  $\bar{u} \in \mathbb{R}^m$  tal que soluciona (2.28), y además existe una única trayectoria óptima  $\bar{x}$  asociada a  $\bar{u}$ .*

*Demostración.* Claramente el funcional  $J_s$  es convexo y continuo, por tanto, existe  $\bar{u} \in \mathbb{R}^m$  que soluciona el problema (2.28). Es más, desde la convexidad estricta existe un único  $\bar{u} \in \mathbb{R}^m$ , y una única observación  $Cx$  asociada a  $\bar{u}$  en  $L^2(\Omega; \mathbb{R}^n)$ . Ahora si existiesen  $x_1$  y  $x_2$  dos trayectorias óptimas asociadas a  $\bar{u}$ , por la convexidad estricta de  $J_s$  se tiene que  $Cx_1 = Cx_2$  en  $L^2(\Omega; \mathbb{R}^n)$ . Además, como  $\bar{u}$  soluciona (2.28) se satisface que  $Ax_1 = B\bar{u} = Ax_2$ . Como tenemos la hipótesis 1, se verifica la desigualdad (2.26). Así se satisface que

$$\|x_1 - x_2\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq K \left( \|A(x_1 - x_2)\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \|C(x_1 - x_2)\|_{L^2(\Omega; \mathbb{R}^n)}^2 \right) = 0,$$

lo cual es una contradicción. Se sigue que existe una única trayectoria óptima  $\bar{x}$  asociada a  $\bar{u}$ .  $\square$

El siguiente resultado nos permite caracterizar la trayectoria óptima. Para ello denotemos como  $\mathcal{R}(A(\omega))$  el rango de  $A(\omega)$  para cada  $\omega$ , y definamos

$$D = \{u \in \mathbb{R}^m : B(\omega)u \in \text{Im}(A(\omega)), \text{ para cada } \omega \in \Omega\}.$$

**Lema 2.3.2.** *Supongamos que tenemos la hipótesis 1. Entonces existe  $\bar{\varphi} \in L^2(\Omega; \mathbb{R}^n)$  tal que para cada  $\omega \in \Omega$  se satisface*

$$A^*(\omega)\bar{\varphi} = C^*(\omega)(C(\omega)\bar{x} - z). \quad (2.29)$$

Además, para  $\bar{u}$  solución de (2.28), se tiene la igualdad

$$\langle \bar{u}, v \rangle_{\mathbb{R}^m} + \langle \bar{\varphi}, Bv \rangle_{L^2(\Omega; \mathbb{R}^n)} = 0 \quad \forall v \in D. \quad (2.30)$$

*Demostración.* Como  $\bar{u}$  es el único mínimo de (2.28) (por el lema 2.3.1) se satisface que

$$\langle \bar{u}, v \rangle_{\mathbb{R}^m} + \langle C\bar{x} - z, C\varphi \rangle_{L^2(\Omega; \mathbb{R}^n)} = 0, \quad (2.31)$$

para cada  $v \in D$  y cada  $\varphi \in L^2(\Omega; \mathbb{R}^n)$  tal que  $A(\omega)\varphi = B(\omega)v$ . Observemos que si tomamos  $v = 0$ , entonces  $A(\omega)\varphi = 0$ , por tanto  $\varphi \in \text{Ker}(A(\omega))$ . Además, de (2.31) se verifica que

$$\langle C^*(C\bar{x} - z), \varphi \rangle_{L^2(\Omega; \mathbb{R}^n)} = \langle C\bar{x} - z, C\varphi \rangle_{L^2(\Omega; \mathbb{R}^n)} = 0.$$

Por tanto  $C^*(\omega)(C(\omega)\bar{x} - z) \in \text{Ker}(A(\omega))^\perp$  para cada  $\omega \in \Omega$  y así  $C^*(\omega)(C(\omega)\bar{x} - z) \in \text{Im}(A^*)$ . Es decir existe  $\bar{\varphi} \in \text{Dom}(A^*(\omega))$  tal que

$$A^*(\omega)\bar{\varphi} = C^*(\omega)(C(\omega)\bar{x} - z).$$

Remplazando lo anterior en (2.31), obtenemos que

$$\langle \bar{u}, v \rangle_{\mathbb{R}^m} + \langle \bar{\varphi}, Bv \rangle_{L^2(\Omega; \mathbb{R}^n)} = 0,$$

para cada  $v \in D$ . □

**Observación 2.3.2.** Desde el lema 2.2.3 podemos remplazar  $m = x^T$ ,  $n = \varphi^T$ ,  $f = Bu$ ,  $g = C^*(Cx^T - z)$ , junto a las condiciones iniciales  $x^T(0) = x_0$  y  $\varphi^T(T) = 0$ , para obtener que para cada  $\omega \in \Omega$   $\varphi^T$  y  $x^T$  satisfacen

$$\langle \varphi^T(0), x_0 \rangle_{\mathbb{R}^n} = \int_0^T \langle C^*(Cx^T - z), x^T \rangle_{\mathbb{R}^n} dt - \int_0^T \langle Bu^T, \varphi^T \rangle_{\mathbb{R}^n} dt. \quad (2.32)$$

De la misma forma es fácil ver que si consideramos  $m = x^T - \bar{x}$  y  $n = \varphi^T - \bar{\varphi}$ <sup>4</sup>, en el lema 2.2.3 obtenemos

$$\begin{aligned} \langle n(0), m(0) \rangle_{\mathbb{R}^n} - \langle n(T), m(T) \rangle_{\mathbb{R}^n} &= \int_0^T \langle C^*Cm, m \rangle_{\mathbb{R}^n} dt - \int_0^T \langle B(u^T - \bar{u}), \varphi^T - \bar{\varphi} \rangle_{\mathbb{R}^n} dt \\ &= \int_0^T \|Cm\|_{\mathbb{R}^n}^2 dt - \int_0^T \langle u^T - \bar{u}, B^*\varphi^T - B^*\bar{\varphi} \rangle_{\mathbb{R}^m} dt, \end{aligned} \quad (2.33)$$

para cada  $\omega \in \Omega$ .

El siguiente teorema corresponde a nuestro primer resultado principal, el cual establece la convergencia en promedio temporal del par óptimo  $(u^T, x^T)$ .

---

<sup>4</sup>Con  $\bar{\varphi} = \bar{p}$  del lema 2.3.2.

**Teorema 2.3.3.** *Supongamos que se satisfacen las hipótesis 1 y 2. Entonces se verifica que*

$$\begin{aligned}\frac{1}{T} \int_0^T x^T(t, \omega) dt &\longrightarrow \bar{x}(\omega) \quad \text{en } L^2(\Omega; \mathbb{R}^n), \\ \frac{1}{T} \int_0^T u^T(t) dt &\longrightarrow \bar{u} \quad \text{en } \mathbb{R}^m,\end{aligned}$$

donde  $(x^T, u^T)$  corresponde al par óptimo asociado al problema evolutivo (2.18), mientras que el par  $(\bar{x}, \bar{u})$  corresponde al par óptimo asociado al problema estacionario (2.28).

*Demostración.* Realizaremos la demostración en varios pasos.

**Paso 1:** Probaremos que  $x^T$  trayectoria asociada al problema de control óptimo (2.18) y  $\varphi^T$  solución del sistema adjunto (2.19), satisfacen

$$\begin{aligned}\|x^T(T)\|_{L^2(\Omega; \mathbb{R}^n)}^2 &\leq KT, \\ \|\varphi^T(0)\|_{L^2(\Omega; \mathbb{R}^n)}^2 &\leq \hat{K}T,\end{aligned}$$

donde  $\hat{K}, K$  son dos constantes positivas independientes del horizonte de tiempo  $T$ . Para probar esto procedemos utilizando (2.32) de la observación 2.3.2. Así para cada  $\omega \in \Omega$  obtenemos que

$$\begin{aligned}\langle \varphi^T(0), x_0 \rangle_{\mathbb{R}^n} &= \int_0^T \langle C^*(Cx^T - z), x^T \rangle_{\mathbb{R}^n} dt - \int_0^T \langle Bu^T, \varphi^T \rangle_{\mathbb{R}^n} dt \\ &= \int_0^T \langle Cx^T - z, Cx^T \rangle_{\mathbb{R}^n} dt - \int_0^T \langle u^T, B^* \varphi^T \rangle_{\mathbb{R}^m} dt \\ &= \int_0^T \left( \|Cx^T\|_{\mathbb{R}^n}^2 - \langle z, Cx^T \rangle_{\mathbb{R}^n} \right) dt - \int_0^T \langle u^T, B^* \varphi^T \rangle_{\mathbb{R}^m} dt \\ &= \int_0^T \left( \|Cx^T - z\|_{\mathbb{R}^n}^2 + \langle z, Cx^T \rangle_{\mathbb{R}^n} - \|z\|^2 \right) dt - \int_0^T \langle u^T, B^* \varphi^T \rangle_{\mathbb{R}^m} dt \\ &= \int_0^T \left( \|Cx^T - z\|_{\mathbb{R}^n}^2 + \langle Cx^T - z, z \rangle_{\mathbb{R}^n} \right) dt - \int_0^T \langle u^T, B^* \varphi^T \rangle_{\mathbb{R}^m} dt.\end{aligned}$$

Ahora integrando con respecto a  $\Omega$ , y utilizando que  $u^T = - \int_{\Omega} B^* \varphi^T d\mu$ , se tiene

$$\begin{aligned}\langle \varphi^T(0), x_0 \rangle_{L^2(\Omega; \mathbb{R}^n)} &= \int_0^T \left( \|Cx^T - z\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \langle Cx^T - z, z \rangle_{L^2(\Omega; \mathbb{R}^n)} \right) dt \\ &\quad - \int_0^T \left\langle u^T, \int_{\Omega} B^* \varphi^T d\mu \right\rangle_{\mathbb{R}^m} dt \\ &= \int_0^T \left( \|Cx^T - z\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \langle Cx^T - z, z \rangle_{L^2(\Omega; \mathbb{R}^n)} + \|u^T\|_{\mathbb{R}^m}^2 \right) dt.\end{aligned}$$

Luego

$$\int_0^T \left( \|Cx^T - z\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \|u^T\|_{\mathbb{R}^m}^2 \right) dt = \left\langle \varphi^T(0), x_0 \right\rangle_{L^2(\Omega; \mathbb{R}^n)} - \int_0^T \left\langle Cx^T - z, z \right\rangle_{L^2(\Omega; \mathbb{R}^n)} dt. \quad (2.34)$$

Observemos que el lado izquierdo de (2.34) se asemeja a las estimaciones recopiladas en la observación 2.3.1, las cuales estiman superiormente  $x^T(T, \omega)$  y  $\varphi^T(0, \omega)$ . Por lo anterior es que nos enfocaremos en estimar el lado derecho de (2.34) superiormente por el factor  $CT$ , y así poder concluir el paso 1.

Dicho lo anterior, gracias a la desigualdad de Cauchy Schwarz y la desigualdad de Young para  $\varepsilon_3 > 0$  tenemos que

$$\int_0^T \left\langle Cx^T - z, z \right\rangle_{L^2(\Omega; \mathbb{R}^n)} dt \leq \int_0^T \varepsilon_3 \frac{\|Cx^T - z\|_{L^2(\Omega; \mathbb{R}^n)}^2}{2} dt + T \frac{\|z\|_{\mathbb{R}^n}^2}{2\varepsilon_3}. \quad (2.35)$$

Estimando el segundo término del lado derecho de (2.34). Además, utilizando la desigualdad (2.25) de la observación 2.3.1, y nuevamente las desigualdades de Cauchy-Schwarz y Young para  $\varepsilon_4 > 0$  obtenemos

$$\begin{aligned} \left\langle \varphi^T(0), x_0 \right\rangle_{L^2(\Omega; \mathbb{R}^n)} &\leq \frac{\varepsilon_4 \|\varphi^T(0)\|_{L^2(\Omega; \mathbb{R}^n)}^2}{2} + \frac{\|x_0\|_{\mathbb{R}^n}^2}{2\varepsilon_4} \\ &\leq \frac{\gamma\varepsilon_4}{2} \left( \int_0^T \|Cx^T - z\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \|u^T\|_{\mathbb{R}^m}^2 dt \right) + \frac{\|x_0\|_{\mathbb{R}^n}^2}{2\varepsilon_4}. \end{aligned} \quad (2.36)$$

Estimando el primer término del lado derecho de (2.34). Ahora remplazando (2.35) y (2.36) en (2.34), y agrupando términos semejantes obtenemos que

$$\begin{aligned} \int_0^T \left( 1 - \frac{\gamma\varepsilon_4}{2} - \varepsilon_3 \right) \|Cx^T - z\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \left( 1 - \frac{\gamma\varepsilon_4}{2} \right) \|u^T\|_{\mathbb{R}^m}^2 dt &\leq T \frac{\|z\|_{\mathbb{R}^n}^2}{2\varepsilon_3} + \frac{\|x_0\|_{\mathbb{R}^n}^2}{2\varepsilon_4} \\ &= \frac{T \|z\|_{\mathbb{R}^n}^2 \varepsilon_4 + \varepsilon_3 \|x_0\|_{\mathbb{R}^n}^2}{2\varepsilon_3 \varepsilon_4} \leq \max \left\{ \|z\|_{\mathbb{R}^n}^2, \|x_0\|_{\mathbb{R}^n}^2 \right\} \frac{T\varepsilon_4 + \varepsilon_3}{2\varepsilon_3}. \end{aligned} \quad (2.37)$$

Así tomando  $\varepsilon_3, \varepsilon_4 > 0$  tales que  $1 - \frac{\gamma\varepsilon_4}{2} > \varepsilon_3$ , podemos encontrar  $K > 0$  independiente de  $T$ , tal que

$$\int_0^T \left( \|Cx^T - z\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \|u^T\|_{\mathbb{R}^m}^2 \right) dt \leq KT, \quad (2.38)$$

Utilizando nuevamente la desigualdad (2.25) de la observación 2.3.1, se verifica que

$$\left\| \varphi^T(0) \right\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq \hat{K}T. \quad (2.39)$$

Ahora volviendo a (2.37) y sumando a ambos lados  $\|x_0\|_{\mathbb{R}^n}^2$ , obtenemos

$$\begin{aligned} \int_0^T \left(1 - \frac{\gamma\epsilon_4}{2} - \epsilon_3\right) \left\| Cx^T - z \right\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \left(1 - \frac{\gamma\epsilon_4}{2}\right) \left\| u^T \right\|_{\mathbb{R}^m}^2 dt + \|x_0\|_{\mathbb{R}^n}^2 \\ \leq \max \left\{ \|z\|_{\mathbb{R}^n}^2, \|x_0\|_{\mathbb{R}^n}^2 \right\} \frac{T\epsilon_4 + \epsilon_3(1 + 2\epsilon_4)}{2\epsilon_3}. \end{aligned}$$

Así escogiendo  $\epsilon_3, \epsilon_4 > 0$  tales que  $1 - \frac{\gamma\epsilon_4}{2} > \epsilon_3$ , tenemos que existe  $K > 0$  independiente de  $T$  tal que

$$\int_0^T \left\| Cx^T - z \right\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \left\| u^T \right\|_{\mathbb{R}^m}^2 dt + \|x_0\|_{\mathbb{R}^n}^2 \leq KT. \quad (2.40)$$

Remplazando (2.40) en la desigualdad (2.24) de la observación 2.3.1, obtenemos que

$$\left\| x^T(T) \right\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq KT, \quad (2.41)$$

concluyendo el Primer paso.

**Paso 2:** Probaremos que

$$\frac{1}{T} \int_0^T x^T dt \quad y \quad \frac{1}{T} \int_0^T u^T dt,$$

son acotados en  $L^2(\Omega; \mathbb{R}^n)$  e  $\mathbb{R}^m$  respectivamente. Para ello observemos que a partir de (2.40) y (2.41) obtenemos

$$\int_0^T \left\| Bu^T \right\|_{L^2(\Omega; \mathbb{R}^n)}^2 dt \leq K \int_0^T \left\| u^T \right\|_{\mathbb{R}^m}^2 dt \leq KT \quad (2.42a)$$

$$\left\| x^T(T) \right\|_{\mathbb{R}^n}^2 \leq KT \quad (2.42b)$$

$$\int_0^T \left\| Cx^T \right\|_{L^2(\Omega; \mathbb{R}^n)}^2 dt \leq KT \quad (2.42c)$$

Así podemos utilizar el lema 2.2.4, concluyendo que

$$\frac{1}{T} \int_0^T x^T dt, \quad \text{es acotado en } L^2(\Omega; \mathbb{R}^n).$$



Finalmente, por la estimación (2.40) tenemos

$$\begin{aligned} \left\| \int_0^T u^T dt \right\|_{\mathbb{R}^m} &\leq \int_0^T \|u^T\|_{\mathbb{R}^m} dt \leq \left( \int_0^T \|u^T\|_{\mathbb{R}^m}^2 dt \right)^{1/2} \left( \int_0^T dt \right)^{1/2} \\ &\leq \sqrt{KT} \sqrt{T} = \hat{K}T, \end{aligned}$$

obteniendo que

$$\frac{1}{T} \int_0^T u^T dt \quad \text{es acotado en } \mathbb{R}^m. \quad (2.43)$$

Concluyendo el paso 2.

**Paso 3:** En este paso probaremos que

$$\frac{1}{T} \left( \int_0^T \left\{ \|C(x^T - \bar{x})\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \|u^T - \bar{u}\|_{\mathbb{R}^m}^2 \right\} dt \right) \rightarrow 0,$$

cuando  $T \rightarrow \infty$ . Para ello definamos  $m = x^T - \bar{x}$  y  $n = \varphi^T - \bar{\varphi}$ . De (2.33) obtenemos

$$\langle n(0), m(0) \rangle_{\mathbb{R}^n} - \langle n(T), m(T) \rangle_{\mathbb{R}^n} = \int_0^T \|Cm\|_{\mathbb{R}^n}^2 dt - \int_0^T \left\langle u^T - \bar{u}, B^* \varphi^T - B^* \bar{\varphi} \right\rangle_{\mathbb{R}^m} dt.$$

Integrando con respecto a  $\Omega$  y completando el cuadrado para obtener el término  $\|u^T - \bar{u}\|_{\mathbb{R}^m}^2$  al lado derecho, se tiene que

$$\begin{aligned} &\langle n(0), m(0) \rangle_{L^2(\Omega; \mathbb{R}^n)} - \langle n(T), m(T) \rangle_{L^2(\Omega; \mathbb{R}^n)} \\ &= \int_0^T \|Cm\|_{L^2(\Omega; \mathbb{R}^n)}^2 dt - \int_0^T \left\langle u^T - \bar{u}, \int_{\Omega} B^* \varphi^T d\mu - \int_{\Omega} B^* \bar{\varphi} d\mu \right\rangle_{\mathbb{R}^m} dt \\ &= \int_0^T \|Cm\|_{L^2(\Omega; \mathbb{R}^n)}^2 dt + \int_0^T \|u^T - \bar{u}\|_{\mathbb{R}^m}^2 dt + \int_0^T \left\langle u^T, \bar{u} + \int_{\Omega} B^* \bar{\varphi} d\mu \right\rangle_{\mathbb{R}^m} dt \\ &\quad - \int_0^T \left( \langle \bar{u}, \bar{u} \rangle_{\mathbb{R}^m} + \langle B\bar{u}, \bar{\varphi} \rangle_{L^2(\Omega; \mathbb{R}^n)} \right) dt. \quad (2.44) \end{aligned}$$

Observemos que el término  $\langle \bar{u}, \bar{u} \rangle_{\mathbb{R}^m} + \langle B\bar{u}, \bar{\varphi} \rangle_{L^2(\Omega; \mathbb{R}^n)}$  es nulo, esto pues al ser  $(\bar{u}, \bar{\varphi})$  el par óptimo asociado al problema estacionario (2.28), en particular satisface (2.30). Así tomando  $v = \bar{u}$  se verifica que

$$\langle \bar{u}, \bar{u} \rangle_{\mathbb{R}^m} + \langle B\bar{u}, \bar{\varphi} \rangle_{L^2(\Omega; \mathbb{R}^n)} = 0.$$

Por tanto, reordenando (2.44) y utilizando la desigualdad Cauchy-Schwarz

$$\begin{aligned}
& \int_0^T \|Cm\|_{L^2(\Omega; \mathbb{R}^n)}^2 dt + \int_0^T \|u^T - \bar{u}\|_{\mathbb{R}^m}^2 dt \\
&= \langle n(0), m(0) \rangle_{L^2(\Omega; \mathbb{R}^n)} + \langle n(T), m(T) \rangle_{L^2(\Omega; \mathbb{R}^n)} - \int_0^T \left\langle u^T, \bar{u} + \int_{\Omega} B^* \bar{\varphi} d\mu \right\rangle_{\mathbb{R}^m} dt \\
&\leq \|n(0)\|_{L^2(\Omega; \mathbb{R}^n)} \|m(0)\|_{L^2(\Omega; \mathbb{R}^n)} + \|n(T)\|_{L^2(\Omega; \mathbb{R}^n)} \|m(T)\|_{L^2(\Omega; \mathbb{R}^n)} \\
&\quad - \int_0^T \left\langle u^T, \bar{u} + \int_{\Omega} B^* \bar{\varphi} d\mu \right\rangle_{\mathbb{R}^m} dt. \tag{2.45}
\end{aligned}$$

Ahora nos centraremos en probar que

$$\frac{1}{T} \int_0^T \left\langle u^T, \bar{u} + \int_{\Omega} B^* \bar{\varphi} d\mu \right\rangle_{\mathbb{R}^m} dt \rightarrow 0.$$

Primero observemos que de (2.43) en particular, salvo subsucesión, se tiene que existe  $v \in \mathbb{R}^m$  tal que

$$\frac{1}{T} \int_0^T u^T dt \rightarrow v \quad \text{en } \mathbb{R}^m.$$

Por otro lado, integrando entre  $(0, T)$  y multiplicando por  $\frac{1}{T}$  la ecuación (2.17) (como en el lema 2.2.4), obtenemos

$$\frac{1}{T} \int_0^T Ax^T dt = \frac{1}{T} \int_0^T Bu^T dt - \left( \frac{x^T(T) - x_0}{T} \right). \tag{2.46}$$

Así junto a la estimación (2.41), el término  $\left( \frac{x^T(T) - x_0}{T} \right)$  converge a cero en  $L^2(\Omega; \mathbb{R}^n)$ , verificándose que

$$A \left( \frac{1}{T} \int_0^T x^T dt \right) = B \left( \frac{1}{T} \int_0^T u^T dt \right) - \left( \frac{x^T(T) - x_0}{T} \right) \rightarrow Bv, \quad \text{en } L^2(\Omega; \mathbb{R}^n).$$

Lo anterior implica en particular que  $v \in D$ , es decir,  $Bv$  es un elemento de la imagen de  $A$ . Así podemos utilizar el lema 2.3.2, precisamente (2.30) obteniendo que

$$\langle \bar{u}, v \rangle_{\mathbb{R}^m} + \langle \bar{\varphi}, Bv \rangle_{L^2(\Omega; \mathbb{R}^n)} = 0.$$

En consecuencia, tenemos que

$$\begin{aligned} \frac{1}{T} \int_0^T \left\langle u^T, \bar{u} + \int_{\Omega} B^* \bar{\varphi} d\mu \right\rangle_{\mathbb{R}^m} dt &\rightarrow \left\langle v, \bar{u} + \int_{\Omega} B^* \bar{\varphi} d\mu \right\rangle_{\mathbb{R}^m} = \langle \bar{u}, v \rangle_{\mathbb{R}^m} + \langle \bar{\varphi}, Bv \rangle_{L^2(\Omega; \mathbb{R}^n)} \\ &= 0. \end{aligned}$$

Finalmente, gracias a (2.41) y (2.39) existen constantes  $K, \hat{K} > 0$  independientes de  $T$ , tales que

$$\begin{aligned} \|m(T)\|_{L^2(\Omega; \mathbb{R}^n)} &\leq K\sqrt{T} + \|\bar{x}\|_{L^2(\Omega; \mathbb{R}^n)} = K\sqrt{T} + \hat{K}, \\ \|n(0)\|_{L^2(\Omega; \mathbb{R}^n)} &\leq K\sqrt{T} + \|\bar{\varphi}\|_{L^2(\Omega; \mathbb{R}^n)} = K\sqrt{T} + \hat{K}. \end{aligned}$$

Ahora volviendo a la estimación (2.45) y utilizando las estimaciones anteriores, tenemos que

$$\begin{aligned} \frac{1}{T} \left( \int_0^T \|Cm\|_{L^2(\Omega; \mathbb{R}^n)}^2 dt + \int_0^T \|u^T - \bar{u}\|_{\mathbb{R}^m}^2 dt \right) \\ \leq \frac{K}{T} \left( \sqrt{T} + \sqrt{T} + \hat{K} \right) + \frac{1}{T} \int_0^T \left\langle u^T, \bar{u} - \int_{\Omega} B^* \bar{\varphi} d\mu \right\rangle_{\mathbb{R}^m} dt \rightarrow 0. \end{aligned} \quad (2.47)$$

Concluyendo así el paso 3.

**Paso 4:** En este paso concluimos la demostración del teorema. Observemos que desde (2.47) tenemos que

$$\left\| \frac{1}{T} \int_0^T (u^T - \bar{u}) dt \right\|_{\mathbb{R}^m}^2 \leq \frac{1}{T^2} \int_0^T \|u^T - \bar{u}\|_{\mathbb{R}^m}^2 dt \rightarrow 0.$$

De igual forma tenemos que

$$\left\| \frac{1}{T} \int_0^T Cm dt \right\|_{L^2(\Omega; \mathbb{R}^n)}^2 \leq \frac{1}{T^2} \int_0^T \|Cm\|_{L^2(\Omega; \mathbb{R}^n)}^2 dt \rightarrow 0.$$

Por tanto

$$\frac{1}{T} \int_0^T u^T dt \rightarrow \bar{u} = v, \quad \frac{1}{T} \int_0^T Cm dt \rightarrow 0, \quad (2.48)$$

en  $\mathbb{R}^m$  y en  $L^2(\Omega; \mathbb{R}^n)$  respectivamente. Además, utilizando nuevamente el lema 2.2.4 con  $f = B(u^T - \bar{u})$  e  $y = m$  se verifica que

$$\left\| \frac{1}{T} \int_0^T B(u^T - \bar{u}) dt \right\|_{L^2(\Omega; \mathbb{R}^n)} \leq K \left\| \frac{1}{T} \int_0^T (u^T - \bar{u}) dt \right\|_{\mathbb{R}^m} \rightarrow 0. \quad (2.49)$$

Por tanto, desde el mismo lema 2.2.4, considerando (2.49) y que  $\frac{1}{T} \int_0^T C m dt \rightarrow 0$  fuertemente en  $L^2(\Omega; \mathbb{R}^n)$ , se tiene

$$\frac{1}{T} \int_0^T (x^T - \bar{x}) dt = \frac{1}{T} \int_0^T m dt \rightarrow 0.$$

Concluyendo la demostración. □

**Corolario 2.3.3.1.** *Supongamos que se satisface las hipótesis 1 y 2. Entonces  $\bar{\varphi}$  la solución de la ecuación (2.29) es única. Además, se tiene la caracterización*

$$\bar{u} = \int_{\Omega} B^* \bar{\varphi} d\mu,$$

para  $\bar{u}$  la única solución del problema (2.28).

*Demostración.* Primero observemos que utilizando la desigualdad estacionaria (2.27),  $\varphi^T$  solución del sistema (2.19) satisface que

$$\begin{aligned} \left\| \frac{1}{T} \int_0^T \varphi^T dt \right\|_{L^2(\Omega; \mathbb{R}^n)} &\leq \left\| A^* \left( \frac{1}{T} \int_0^T \varphi^T dt \right) \right\|_{L^2(\Omega; \mathbb{R}^n)} + \left\| \int_{\Omega} B^* \left( \frac{1}{T} \int_0^T \varphi^T dt \right) \right\|_{\mathbb{R}^m} \\ &\leq \left\| \frac{1}{T} \int_0^T A^* \varphi^T dt \right\|_{L^2(\Omega; \mathbb{R}^n)} + \left\| \frac{1}{T} \int_0^T \left( \int_{\Omega} B^* \varphi^T d\mu \right) dt \right\|_{\mathbb{R}^m} \\ &\leq \left\| \frac{1}{T} \int_0^T C^* (Cx^T - z) dt - \frac{\varphi^T(0)}{T} \right\|_{L^2(\Omega; \mathbb{R}^n)} + \left\| \frac{1}{T} \int_0^T u^T dt \right\|_{\mathbb{R}^m} \\ &\leq \left\| \frac{1}{T} \int_0^T C^* (Cx^T - z) dt \right\|_{L^2(\Omega; \mathbb{R}^n)} + \left\| \frac{\varphi^T(0)}{T} \right\|_{L^2(\Omega; \mathbb{R}^n)} + \left\| \frac{1}{T} \int_0^T u^T dt \right\|_{\mathbb{R}^m}. \end{aligned}$$

Por el teorema 2.3.3, sabemos que todos los términos del lado derecho de la estimación anterior son convergente, entonces existe  $K > 0$  tal que

$$\left\| \frac{1}{T} \int_0^T \varphi^T dt \right\|_{L^2(\Omega; \mathbb{R}^n)} \leq K.$$

Por tanto  $\left\| \frac{1}{T} \int_0^T \varphi^T dt \right\|_{L^2(\Omega; \mathbb{R}^n)}$  es acotado en  $L^2(\Omega; \mathbb{R}^n)$ , y en consecuencia de [7, Teorema 3.18] existe  $\bar{\varphi} \in L^2(\Omega; \mathbb{R}^n)$  tal que salvo subsucesión se tiene que

$$\frac{1}{T} \int_0^T \varphi^T dt \rightharpoonup \bar{\varphi} \quad \text{en } L^2(\Omega; \mathbb{R}^n). \quad (2.50)$$

Observemos que  $\bar{\varphi}$  soluciona (2.29). En efecto, integrando entre  $(0, T)$  y multiplicando por  $\frac{1}{T}$  la ecuación (2.19), como en el lema 2.2.4 obtenemos

$$\frac{1}{T} \int_0^T A^* \varphi^T dt = \frac{1}{T} \int_0^T C^*(Cx^T - z) dt - \frac{\varphi^T(0)}{T}. \quad (2.51)$$

Por el Teorema 2.3.3 se verifica que  $\frac{1}{T} \int_0^T x^T dt \rightarrow \bar{x}$ , y gracias a la estimación (2.39) tenemos que  $\frac{\|\varphi^T(0)\|_{L^2(\Omega; \mathbb{R}^n)}}{T} \rightarrow 0$  cuando  $T \rightarrow \infty$ , así obtenemos

$$C^* \left( C \left( \frac{1}{T} \int_0^T x^T dt \right) - z \right) - \frac{\varphi^T(0)}{T} \rightarrow C^*(C\bar{x} - z). \quad (2.52)$$

Además, si definimos el operador

$$\begin{aligned} G_{A^*} : L^2(\Omega; \mathbb{R}^n) &\rightarrow L^2(\Omega; \mathbb{R}^n) \\ v &\mapsto F_{B^*}(v) = A^*(\omega)v(\omega), \end{aligned}$$

es lineal y continuo, en particular es continuo en la topología débil [7, Teorema 3.10]. Así

$$G_{A^*} \left( \frac{1}{T} \int_0^T \varphi^T dt \right) \rightharpoonup G_{A^*}(\bar{\varphi}),$$

lo que es equivalente a

$$\frac{1}{T} \int_0^T A^* \varphi^T dt = A^* \left( \frac{1}{T} \int_0^T \varphi^T dt \right) \rightharpoonup A^* \bar{\varphi}.$$

Por tanto, utilizando (2.51), junta a (2.52) y la unicidad del límite, tenemos  $\bar{\varphi}$  satisface la ecuación  $A^* \bar{\varphi} = C^*(C\bar{x} - z)$ .

Ahora para probar que

$$\bar{u} = - \int_{\Omega} B^* \bar{\varphi} d\mu. \quad (2.53)$$

Observemos que desde la igualdad  $u^T = - \int_{\Omega} B^* \varphi^T dt$ , si multiplicando esta por  $\frac{1}{T}$  e integrándola entre  $(0, T)$  tenemos que

$$\frac{1}{T} \int_0^T u^T dt = - \int_{\Omega} B^* \frac{1}{T} \int_0^T \varphi^T dt d\mu.$$

Además, como el funcional

$$F_{B^*} : L^2(\Omega; \mathbb{R}^n) \rightarrow \mathbb{R}$$

$$v \mapsto F_{B^*}(v) = \int_{\Omega} B^*(\omega) v(\omega) d\mu,$$

verifica que  $F_{B^*} \in (L^2(\Omega; \mathbb{R}^n))^*$ , por (2.50) tenemos

$$\int_{\Omega} B^* \left( \frac{1}{T} \int_0^T \varphi^T dt \right) d\mu \rightarrow \int_{\Omega} B^*(\bar{\varphi}) d\mu.$$

Como  $\frac{1}{T} \int_0^T u^T dt \rightarrow \bar{u}$ , concluimos (2.53) por la unicidad del límite.

Finalmente, si suponemos por contradicción que existen dos soluciones  $\bar{\varphi}_1$  y  $\bar{\varphi}_2$  de (2.29), por lo anterior tenemos que satisfacen también que

$$\bar{u} = - \int_{\Omega} B^* \bar{\varphi}_i d\mu,$$

para  $i = 1, 2$ . Entonces definiendo  $p = \bar{\varphi}_2 - \bar{\varphi}_1$ , tenemos por la desigualdad estacionaria (2.27) de la observación 2.3.1 que

$$\begin{aligned} \|p\|_{L^2(\Omega; \mathbb{R}^n)}^2 &\leq \gamma \left( \|A^* p\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \left\| \int_{\Omega} B^* p d\mu \right\|_{\mathbb{R}^m}^2 \right) \\ &= \gamma \left( \|A^*(\bar{\varphi}_2 - \bar{\varphi}_1)\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \left\| \int_{\Omega} B^*(\bar{\varphi}_2 - \bar{\varphi}_1) d\mu \right\|_{\mathbb{R}^m}^2 \right) \\ &= \gamma \left( \|C^*(C\bar{x} - z) - C^*(C\bar{x} - z)\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \|\bar{u} - \bar{u}\|_{\mathbb{R}^m}^2 \right) = 0. \end{aligned}$$

Concluyendo la unicidad de  $\bar{\varphi}$ .

□

A continuación, enunciaremos y probaremos nuestro segundo resultado principal.

**Teorema 2.3.4.** *Supongamos que se satisface las hipótesis 1 y 2. Entonces existen  $\mathcal{C}, \delta > 0$ , independientes de  $T$ , tales que*

$$\left\| x^T(t) - \bar{x} \right\|_{L^2(\Omega; \mathbb{R}^n)} + \left\| \varphi^T - \bar{\varphi} \right\|_{L^2(\Omega; \mathbb{R}^n)} + \left\| u^T(t) - \bar{u} \right\|_{\mathbb{R}^m} \leq \mathcal{C}(e^{-\delta(T-t)} + e^{-\delta t}),$$

para todo  $t \in [0, T]$ .

*Demostración.* Dividiremos la demostración en cuatro pasos.

**Paso 1:** Sea  $m = x^T - \bar{x}$  y  $n = \varphi^T - \bar{\varphi}$ , soluciones del sistema

$$\begin{cases} m_t + Am = -B \left( \int_{\Omega} B^* n \, d\mu \right), \\ m(0) = m_0 = x_0 - \bar{x}, \\ -n_t + A^* n = C^* C m, \\ n(T) = n_T = -\bar{\varphi}. \end{cases} \quad (2.54)$$

Así aplicando los Lemas 2.2.1 y 2.2.2 a  $m$  y a  $n$  solución de (2.54), y la caracterización de  $\bar{u}$  entregada por el Corolario 2.3.3.1 tenemos que

$$\begin{aligned} \|m(T)\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \|m\|_{L^2(0,T;L^2(\Omega; \mathbb{R}^n))}^2 &\leq K \left( \int_0^T \left( \|u^T - \bar{u}\|_{\mathbb{R}^m}^2 + \|Cm\|_{L^2(\Omega; \mathbb{R}^n)}^2 \right) dt + \|m_0\|_{L^2(\Omega; \mathbb{R}^n)}^2 \right), \\ \|n(0)\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \|n\|_{L^2(0,T;L^2(\Omega; \mathbb{R}^n))}^2 &\leq K \left( \int_0^T \left( \|Cm\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \|u^T - \bar{u}\|_{\mathbb{R}^m}^2 \right) dt + \|n_T\|_{L^2(\Omega; \mathbb{R}^n)}^2 \right). \end{aligned} \quad (2.55)$$

Por otro lado, desde la observación 2.3.2 y la desigualdad de Young obtenemos que

$$\begin{aligned} \int_0^T \left( \|Cm\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \|u^T - \bar{u}\|_{\mathbb{R}^m}^2 \right) dt &\leq \frac{\varepsilon_1 \|n(0)\|_{L^2(\Omega; \mathbb{R}^n)}^2}{2} + \frac{\|m_0\|_{L^2(\Omega; \mathbb{R}^n)}^2}{2\varepsilon_1} \\ &\quad + \frac{\varepsilon_2 \|m(T)\|_{L^2(\Omega; \mathbb{R}^n)}^2}{2} + \frac{\|n_T\|_{L^2(\Omega; \mathbb{R}^n)}^2}{2\varepsilon_2}. \end{aligned}$$

Usando la estimación (2.55) tenemos

$$\begin{aligned} \int_0^T \left( \|Cm\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \|u^T - \bar{u}\|_{\mathbb{R}^m}^2 \right) dt &\leq K \left[ \left( \frac{1}{2\varepsilon_1} + \frac{\varepsilon_2}{2} \right) \|m_0\|_{L^2(\Omega; \mathbb{R}^n)}^2 \right. \\ &\quad \left. + \left( \frac{1}{2\varepsilon_2} + \frac{\varepsilon_1}{2} \right) \|n_T\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \left( \frac{\varepsilon_1 + \varepsilon_2}{2} \right) \int_0^T \left( \|Cm\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \|u^T - \bar{u}\|_{\mathbb{R}^m}^2 \right) dt \right] \end{aligned} \quad (2.56)$$

Así tomado  $\varepsilon_1, \varepsilon_2 > 0$  tal que  $\varepsilon_1 + \varepsilon_2 < 2$ , entonces el último término de (2.56) puede ser absorbido por el lado izquierdo, obteniendo

$$\int_0^T \left( \|Cm\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \|u^T - \bar{u}\|_{\mathbb{R}^m}^2 \right) dt \leq K \left( \|m_0\|_{L^2(\Omega; \mathbb{R}^n)}^2 + \|n_T\|_{L^2(\Omega; \mathbb{R}^n)}^2 \right).$$

En consecuencia, de (2.55) obtenemos la estimación

$$\|m\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^n))}^2 + \|n\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^n))}^2 \leq K \left( \|m_0\|_{L^2(\Omega;\mathbb{R}^n)}^2 + \|n_T\|_{L^2(\Omega;\mathbb{R}^n)}^2 \right).$$

Además, tomando norma a la primera y a la tercera ecuación de (2.54), considerando que  $A, B$  y  $C$  son uniformemente acotados tenemos que

$$\begin{aligned} \|m_t\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^n))}^2 &\leq K \left( \|m\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^n))}^2 + \|n\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^n))}^2 \right), \\ \|n_t\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^n))}^2 &\leq K \left( \|n\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^n))}^2 + \|m\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^n))}^2 \right), \end{aligned}$$

Por tanto utilizando la inyección continua  $H^1(0, T; L^2(\Omega; \mathbb{R}^n)) \hookrightarrow C([0, T]; L^2(\Omega; \mathbb{R}^n))$  concluimos que

$$\|m\|_{\mathcal{X}}^2 + \|n\|_{\mathcal{X}}^2 \leq K \left( \|m_0\|_{L^2(\Omega;\mathbb{R}^n)}^2 + \|n_T\|_{L^2(\Omega;\mathbb{R}^n)}^2 \right), \quad (2.57)$$

donde  $\mathcal{X} = C([0, T]; L^2(\Omega; \mathbb{R}^n))$ .

**Paso 2:** Observemos que el sistema (2.54) puede ser escrito como

$$\begin{pmatrix} -C^*C & -\frac{d}{dt} + A^* \\ 0 & E_T \\ \frac{d}{dt} + A & B \left( \int_{\Omega} B^* \cdot d\mu \right) \\ E_0 & 0 \end{pmatrix} \begin{pmatrix} m \\ n \end{pmatrix} = \begin{pmatrix} 0 \\ n_T \\ 0 \\ m_0 \end{pmatrix}, \quad (2.58)$$

donde  $E_0 m := m(0)$  y  $E_T n := n(T)$ .

Denotamos por  $\mathcal{H}$  operador matricial

$$\mathcal{H} = \begin{pmatrix} -C^*C & -\frac{d}{dt} + A^* \\ 0 & E_T \\ \frac{d}{dt} + A & B \left( \int_{\Omega} B^* \cdot d\mu \right) \\ E_0 & 0 \end{pmatrix},$$

observemos que  $\mathcal{H} : (\mathcal{X})^2 \rightarrow (L^2(\Omega; \mathbb{R}^n))^2$  es invertible. En efecto, primero observemos que el operador  $\mathcal{H}^{-1}$  corresponde al operador solución, que lleva las condiciones iniciales del sistema primal y el adjunto, a sus respectivas soluciones. Como el sistema (2.58) posee una única solución (pues las variables que las componen están únicamente definidas por el Teorema 2.3.1, y el Corolario 2.3.3.1), el operador  $\mathcal{H}$  es invertible, es decir  $\mathcal{H}^{-1} : (L^2(\Omega; \mathbb{R}^n))^2 \rightarrow (\mathcal{X})^2$  está bien definido. Finalmente, desde (2.57) del paso anterior, obtenemos que existe  $K > 0$ , independiente



de  $T$  tal que

$$\left\| \mathcal{H}^{-1} \right\|_{\mathcal{L}((L^2(\Omega; \mathbb{R}^n))^2, (\mathcal{X})^2)} < K. \quad (2.59)$$

**Paso 3:** Ahora consideremos el cambio de variable

$$\hat{m} = \frac{m}{e^{-\delta(T-t)} + e^{-\delta t}}, \quad \hat{n} = \frac{n}{e^{-\delta(T-t)} + e^{-\delta t}},$$

para  $\delta > 0$  por fijar. Notemos que  $\hat{m}$  y  $\hat{n}$  satisfacen el sistema

$$\begin{cases} \hat{m}_t + A\hat{m} = -B \left( \int_{\Omega} B^* \hat{n} d\mu \right) + \hat{m} \delta h(t), \\ -\hat{n}_t + A^* \hat{n} = C^* C \hat{m} - \hat{n} \delta h(t), \\ \hat{m}(0) = \frac{m_0}{1+e^{-\delta T}}, \quad \hat{n}(T) = \frac{n_T}{1+e^{-\delta T}}, \end{cases}$$

donde  $h(t)$  corresponde a

$$h(t) = \frac{e^{-\delta t} - e^{-\delta(T-t)}}{e^{-\delta t} + e^{-\delta(T-t)}}.$$

Usando la definición de  $\mathcal{H}$ , podemos escribir el sistema anterior como

$$\left[ \mathcal{H} - \delta \begin{pmatrix} h(t) & 0 \\ 0 & 0 \\ 0 & -h(t) \\ 0 & 0 \end{pmatrix} \right] \begin{pmatrix} \hat{m} \\ \hat{n} \end{pmatrix} = \frac{1}{1+e^{-\delta T}} \begin{pmatrix} 0 \\ n_T \\ 0 \\ m_0 \end{pmatrix}. \quad (2.60)$$

Además, denotando como  $\mathcal{P}$  el operador matricial

$$\mathcal{P} = \begin{pmatrix} h(t) & 0 \\ 0 & 0 \\ 0 & -h(t) \\ 0 & 0 \end{pmatrix},$$

se satisface que  $\|\mathcal{P}\|_{\mathcal{L}((\mathcal{X})^2)} \leq 1$ , pues  $|h(t)| \leq 1$  para todo  $t \in [0, T]$ .

Por ser  $\mathcal{H}$  invertible, el sistema (2.60) puede ser escrito de la forma

$$(I - \delta \mathcal{H}^{-1} \mathcal{P}) \mathcal{Z} = \mathcal{H}^{-1} \mathcal{Y},$$

donde

$$\mathcal{Z} := \begin{pmatrix} \hat{m} \\ \hat{n} \end{pmatrix}, \quad \mathcal{Y} := \frac{1}{1 + e^{-\delta T}} \begin{pmatrix} 0 \\ n_T \\ 0 \\ m_0 \end{pmatrix}.$$

Ahora, con intenciones de despejar  $\mathcal{Z}$ , deseamos que el operador  $(I - \delta \mathcal{H}^{-1} \mathcal{P})$  sea invertible. Para ello como  $\|\mathcal{P}\|_{\mathcal{L}((\mathcal{X})^2)} \leq 1$ , basta fijar  $\delta > 0$ <sup>5</sup> tal que  $\delta \|\mathcal{H}^{-1}\| < 1$ . En tal caso,  $(I - \delta \mathcal{H}^{-1} \mathcal{P})^{-1}$  existe y su norma esta acotada por la representación de Neumann (ver [35, Teorema 2.14])

$$\left\| (I - \delta \mathcal{H}^{-1} \mathcal{P})^{-1} \right\|_{\mathcal{L}((L^2(\Omega; \mathbb{R}^n))^2, (\mathcal{X})^2)} \leq \frac{1}{1 - \|\delta \mathcal{H}^{-1} \mathcal{P}\|_{\mathcal{L}((L^2(\Omega; \mathbb{R}^n))^2, (\mathcal{X})^2)}}.$$

Luego tomando norma a  $\mathcal{Z} = (I - \delta \mathcal{H}^{-1} \mathcal{P})^{-1} \mathcal{H}^{-1} \mathcal{Y}$  obtenemos

$$\|\hat{m}\|_{\mathcal{X}} + \|\hat{n}\|_{\mathcal{X}} \leq \frac{\|\mathcal{H}^{-1}\|_{\mathcal{L}((L^2(\Omega; \mathbb{R}^n))^2, (\mathcal{X})^2)}}{1 - \|\delta \mathcal{H}^{-1} \mathcal{P}\|_{\mathcal{L}((L^2(\Omega; \mathbb{R}^n))^2, (\mathcal{X})^2)}} \frac{1}{1 + e^{-\delta T}} \left( \|m(0)\|_{L^2(\Omega; \mathbb{R}^n)} + \|n(T)\|_{L^2(\Omega; \mathbb{R}^n)} \right).$$

Observemos que las cantidades del lado derecho pueden ser acotadas superiormente por un constante de  $\mathcal{C}_1 > 0$  independiente de  $T$  tal que

$$\|\hat{m}\|_{\mathcal{X}} + \|\hat{n}\|_{\mathcal{X}} \leq \mathcal{C}_1.$$

Así volviendo a las variables originales

$$\left\| \frac{m}{e^{-\delta(T-t)} + e^{-\delta t}} \right\|_{\mathcal{X}} + \left\| \frac{n}{e^{-\delta(T-t)} + e^{-\delta t}} \right\|_{\mathcal{X}} \leq \mathcal{C}_1.$$

Además, como  $\mathcal{X} = C([0, T]; L^2(\Omega; \mathbb{R}^n))$ , en particular tenemos que

$$\|m(t)\|_{L^2(\Omega; \mathbb{R}^n)} + \|n(t)\|_{L^2(\Omega; \mathbb{R}^n)} \leq \mathcal{C}_1 (e^{-\delta(T-t)} + e^{-\delta t}),$$

para cada  $t \in [0, T]$ .

**Paso 4:** Finalmente observando que

$$\left\| u^T - \bar{u} \right\|_{\mathbb{R}^m}^2 = \left\| \int_{\Omega} B^* n d\mu \right\|_{\mathbb{R}^m}^2 \leq C \left\| \int_{\Omega} n d\mu \right\|^2 \leq C \|n\|_{L^2(\Omega; \mathbb{R}^n)}^2,$$

---

<sup>5</sup>El cual es independiente del horizonte de tiempo  $T$  por lo probado en (2.59).

concluimos que existe una constante positiva  $\mathcal{C}$  independiente de  $T$  tal que

$$\left\|x^T(t) - \bar{x}\right\|_{L^2(\Omega; \mathbb{R}^n)} + \left\|\varphi^T - \bar{\varphi}\right\|_{L^2(\Omega; \mathbb{R}^n)} + \left\|u^T(t) - \bar{u}\right\|_{\mathbb{R}^m} \leq \mathcal{C}(e^{-\delta(T-t)} + e^{-\delta t}),$$

para todo  $t \in [0, T]$ .  $\square$

A continuación, probamos nuestro último resultado principal. Obtendremos este resultado como consecuencia directa del Teorema 2.3.4.

**Teorema 2.3.5.** *Supongamos que se satisface las hipótesis 1 y 2. Entonces existen  $\bar{\mathcal{C}}, \delta > 0$  independientes de  $T$  tales que*

$$\left\|\mathbb{E}(x^T) - \mathbb{E}(\bar{x})\right\|_{\mathbb{R}^n} + \left\|\mathbb{E}(\varphi^T) - \mathbb{E}(\bar{\varphi})\right\|_{\mathbb{R}^n} + \left\|u^T - \bar{u}\right\|_{\mathbb{R}^m} \leq \bar{\mathcal{C}}(e^{-\delta(T-t)} + e^{-\delta t}), \quad (2.61)$$

para todo  $t \in [0, T]$ , donde  $\mathbb{E} : L^2(\Omega; \mathbb{R}^n) \rightarrow \mathbb{R}^n$  denota la esperanza matemática.

*Demostración.* Por las hipótesis del teorema, sabemos que se satisface el teorema 2.3.4, así tenemos

$$\begin{aligned} \left\|\mathbb{E}(x^T) - \mathbb{E}(\bar{x})\right\|_{\mathbb{R}^n} &= \left\|\int_{\Omega} x^T d\mu - \int_{\Omega} \bar{x} d\mu\right\|_{\mathbb{R}^n} \leq \int_{\Omega} \|x^T - \bar{x}\|_{\mathbb{R}^n} d\mu \\ &\leq \left(\int_{\Omega} d\mu\right)^{1/2} \left(\int_{\Omega} \|x^T - \bar{x}\|_{\mathbb{R}^n}^2 d\mu\right)^{1/2} \\ &= \|x^T - \bar{x}\|_{L^2(\Omega; \mathbb{R}^n)} \leq \mathcal{C}(e^{-\delta(T-t)} + e^{-\delta t}). \end{aligned}$$

De forma análoga obtenemos

$$\left\|\mathbb{E}(\varphi^T) - \mathbb{E}(\bar{\varphi})\right\|_{\mathbb{R}^n} \leq \left\|\varphi^T - \bar{\varphi}\right\|_{L^2(\Omega; \mathbb{R}^n)} \leq \mathcal{C}(e^{-\delta(T-t)} + e^{-\delta t}).$$

Por tanto, existe  $\delta > 0$  y  $\bar{\mathcal{C}} > 0$  tales que

$$\left\|\mathbb{E}(x^T) - \mathbb{E}(\bar{x})\right\|_{\mathbb{R}^n} + \left\|\mathbb{E}(\varphi^T) - \mathbb{E}(\bar{\varphi})\right\|_{\mathbb{R}^n} + \left\|u^T - \bar{u}\right\|_{\mathbb{R}^m} \leq \bar{\mathcal{C}}(e^{-\delta(T-t)} + e^{-\delta t}),$$

para todo  $t \in [0, T]$ .  $\square$

## 2.4. Ejemplos Numéricos

Con el fin de ilustrar la propiedad de turnpike exponencial en este contexto aleatorio, consideremos los siguientes ejemplos numéricos. Nos enfocaremos en el caso en que  $A(\omega) = A\alpha(\omega)$  y  $B(\omega) = B\beta(\omega)$ , con  $A, B$  matrices fijas y  $\alpha$  y  $\beta$  funciones escalares dependientes del parámetro aleatorio  $\omega$ . Además, supondremos que  $\Omega$  es un conjunto discreto.

Desde aquí en adelante consideraremos el espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$ .

### Ejemplo 2.4.1.

Sea  $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7\}$ , consideremos la variable aleatoria discreta  $\mathcal{V}$ , la cual se distribuye uniforme, es decir  $\mathcal{V} \sim U(\Omega)$ . Por simpleza supondremos que  $\omega_i = i$ . Sea  $\beta(\omega) = \mathbb{P}(\mathcal{V} \leq \omega)$  la función de cuantía de la variable aleatoria  $\mathcal{V}$ , es decir

$$\beta(\omega) = \frac{1}{7} \sum_{i=1}^7 \chi_{(-\infty, \omega]}(\omega_i).$$

Lo anterior es equivalente a una función escalonada. Consideremos el siguiente problema de control óptimo

$$\min_{u \in L^2(0, T; \mathbb{R}^m)} \left\{ J^T(u) = \frac{1}{2} \int_0^T \left\{ \|u\|_{\mathbb{R}^m}^2 + \|Cx - z\|_{L^2(\Omega; \mathbb{R}^n)}^2 \right\} dt \right\},$$

con  $x$  solución de

$$\begin{cases} x_t + Ax = \beta(\omega)Bu & t \in (0, T), \\ x(0) = x_0, \end{cases}$$

donde

$$A = \begin{pmatrix} 0,2 & 0,5 \\ -0,5 & 0,5 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad z = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad x_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{y} \quad T = 50.$$

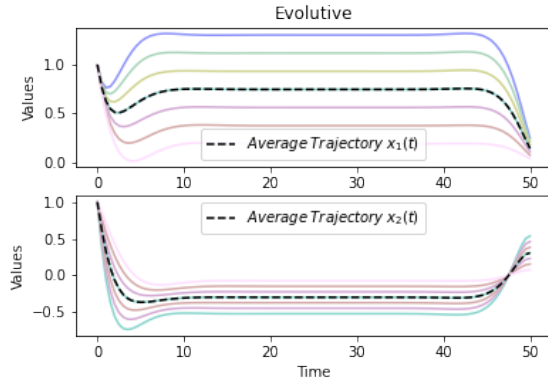
Utilizando las mismas matrices consideramos el problema análogo estacionario

$$\min_{u \in \mathbb{R}^m} \left\{ J_s(u) = \frac{1}{2} \left( \|u\|_{\mathbb{R}^m}^2 + \|Cx - z\|_{L^2(\Omega; \mathbb{R}^n)}^2 \right) \right\},$$

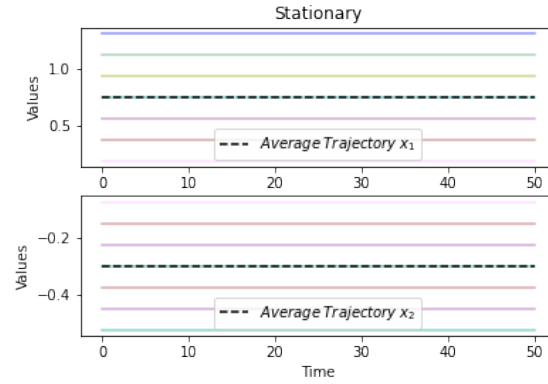
donde  $x$  es solución de  $Ax = \beta(\omega)Bu$ .

**Observación 2.4.1.** Es fácil verificar que  $(A^*, B^*(\omega))$  satisface la hipótesis 2, y que  $(A, C)$  la hipótesis 1, en este último podemos considerar  $K_C = C$ .

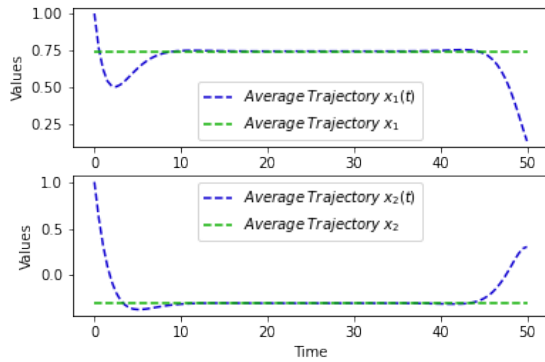
Resolviendo el problema anterior numéricamente, obtenemos las siguientes gráficas



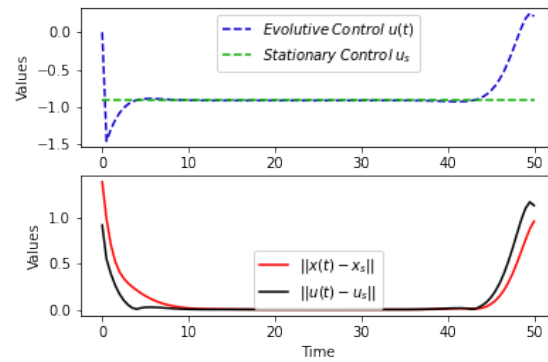
a) Primera componente de las soluciones.



b) Segunda componente de las soluciones.



c) Comparación trayectorias promedio.



d) Comparación control y diferencias de normas.

FIGURA 2.1: Comportamiento sistema estacionario v/s evolutivo.

Observemos que se puede apreciar con claridad la propiedad de turnpike, tanto para el par  $(x^T, \bar{x})$  como para  $(u^T, \bar{u})$ , como nos asegura el Teorema 2.3.5. En el gráfico c) podemos observar que efectivamente las trayectorias promedio, se encuentran cerca una gran cantidad de tiempo. Además, en d) se verifica que la diferencia entre las trayectorias, y los controles es mínima cerca de  $T/2 = 25$ .

### Ejemplo 2.4.2.

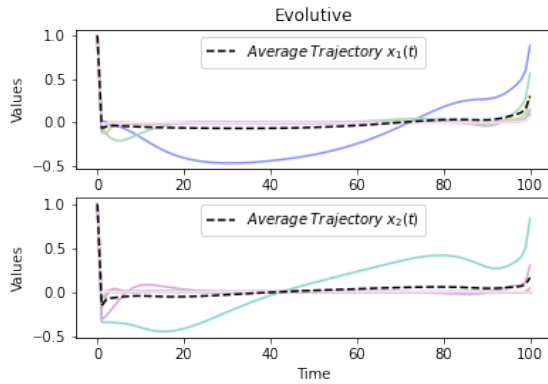
Consideremos el funcional  $J^T$  del ejemplo anterior, pero con  $x$  solución de

$$\begin{cases} x_t + \alpha(\omega)A = Bu & t \in (0, T), \\ x(0) = x_0, \end{cases}$$

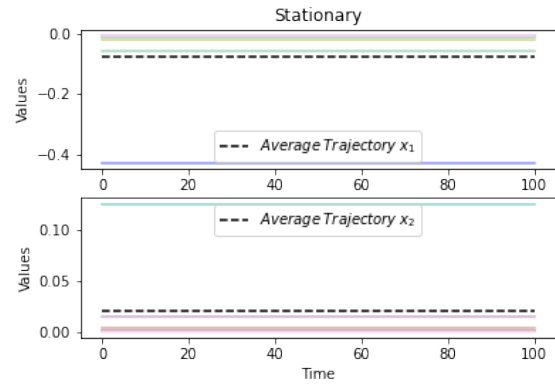
y el sistema estacionario  $\alpha(\omega)Ax = Bu$ , donde

$$A = \begin{pmatrix} 2 & -5 \\ 5 & 0.5 \end{pmatrix}, \quad B = \begin{pmatrix} 5 \\ 7 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad z = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad x_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{y} \quad T = 100.$$

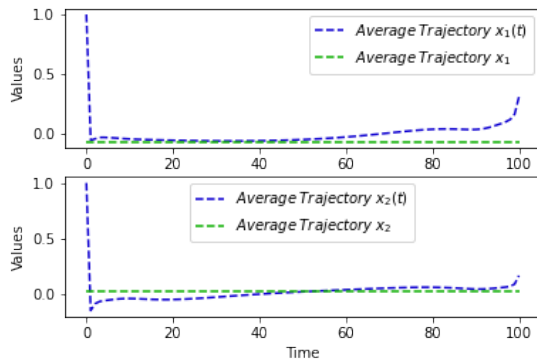
En este caso consideraremos  $\alpha$  como su función de distribución acumulada de la variable aleatoria  $\mathcal{V}$ , que se distribuye binomial con parámetros  $n = 7$  y  $p = 1/2$ , es decir  $\mathcal{V} \sim B(7, 1/2)$ . Al igual como se comenta en la observación 2.4.1, se puede mostrar que  $(A(\omega), C)$  y  $(A^*(\omega), B^*)$  satisfacen las hipótesis 1 y 2 respectivamente. La solución de este problema, numéricamente, corresponde a



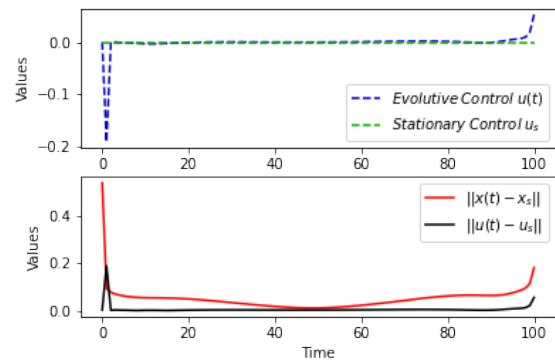
a) Primera componente de las soluciones.



b) Segunda componente de las soluciones.



c) Comparación trayectorias promedio.



d) Comparación control y diferencias de normas.

FIGURA 2.2: Comportamiento sistema estacionario v/s evolutivo.

Notemos que, en este caso, la dinámica del sistema se observa mucho más perturbada, no obstante, se sigue percibiendo la propiedad de turnpike. Lo anterior puede deberse a que la propiedad probada en el teorema 2.3.5 nos asegura la existencia de constantes  $C, \delta > 0$  independientes del horizonte de tiempo  $T$ , tales que satisfacen la propiedad de turnpike (2.61), aun así, las constantes  $C$  y  $\delta$  no necesariamente son independientes de la variable aleatoria, lo cual podría ser la razón de las perturbación en la dinámica promedio.

### Ejemplo 2.4.3.

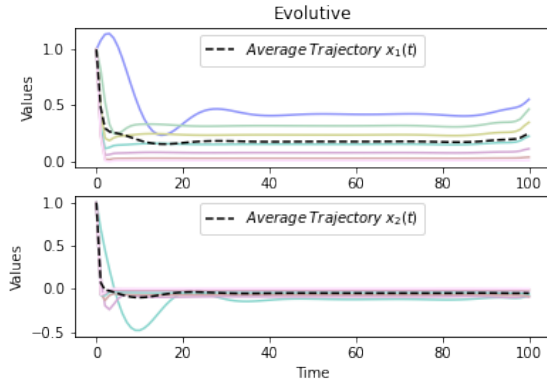
Considerando el funcional  $J^T$  del ejemplo 2.4.2, pero  $x$  es solución de

$$\begin{cases} x_t + \alpha(\omega)A = \beta(\omega)Bu & t \in (0, T), \\ x(0) = x_0, \end{cases}$$

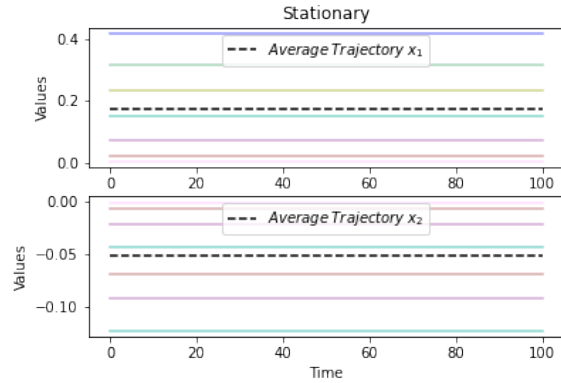
y el problema estacionario, solución de  $\alpha(\omega)Ax = \beta(\omega)Bu$ , donde

$$A = \begin{pmatrix} 2 & -5 \\ 5 & 0.5 \end{pmatrix}, \quad B = \begin{pmatrix} 5 \\ 7 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad z = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad x_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{y} \quad T = 100.$$

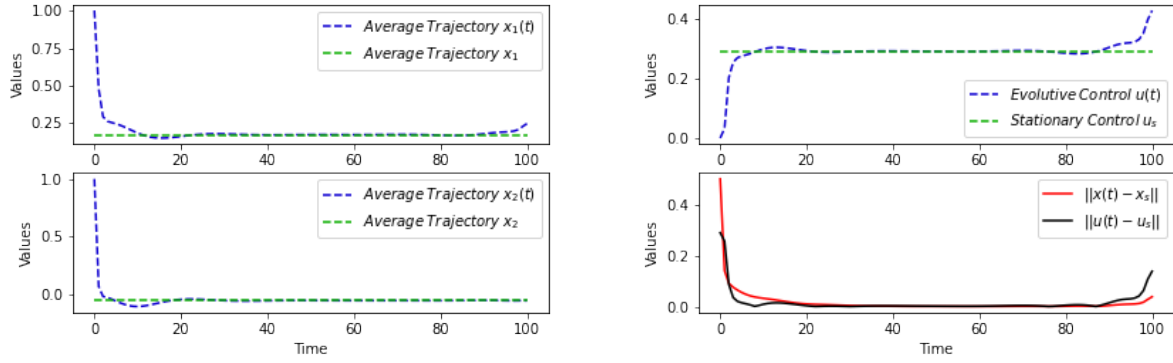
En este caso las matrices  $A, B, C, z$ , la condición inicial  $x_0$ , y la función  $\alpha$  son las mismas al ejemplo 2.4.2, y consideramos la función  $\beta(\omega)$  correspondiente a la función de densidad de la variable aleatoria  $\mathcal{V} \sim B(7, 1/2)$ . Implementado numéricamente este problema, y comparando sus trayectorias óptimas, tenemos



a) Primera componente de las soluciones.



b) Segunda componente de las soluciones.



c) Comparación trayectorias promedio.

Comparación control y diferencias de normas.

FIGURA 2.3: Comportamiento sistema estacionario v/s evolutivo.

En este caso tanto la matriz  $A$  como la matriz  $B$  son perturbadas por factores aleatorios, obteniendo trayectoria más perturbadas. No obstante, al igual que el ejemplo anterior sigue siendo posible observar la propiedad de turnpike, tanto para los promedios de las trayectorias evolutivas y estacionarias, como para el control óptimo del problema evolutivo y el estacionario.

**Observación 2.4.2.** Para obtener las soluciones óptimas evolutivas  $(x^T, u^T)$ , y las soluciones óptimas estacionarias  $(\bar{x}, \bar{u})$ , utilizamos la librería Gekko de Python. Gekko es una librería especializada en problemas de optimización dinámica, como programación en enteros, ecuaciones no lineales y problemas de ecuaciones diferenciales algebraicas (DAE) [3]. En el caso del problema en evolución, utilizamos Gekko en el modo para resolver problemas de control óptimo (imode=6), con 100 pasos de tiempo. Este modo está basado en estrategias simultáneas (simultaneous strategies), para optimización dinámica [5]. Para ser más precisos, el imode 6 de Gekko usa colocación ortogonal de elementos finitos [8], esta es una forma explícita del método de Runge-Kutta.

## 2.5. Problemas Abiertos

En esta breve sección comentamos algunos problemas abiertos en relación con la propiedad de Turnpike en el contexto con incertidumbre

- Posiblemente el problema abierto que más se relaciona con este documento, es probar la Propiedad de Turnpike bajo hipótesis de controlabilidad en promedio y observabilidad en promedio, esto respondería al modo natural de extender los resultados obtenidos en [44].

Por supuesto una posible alternativa para abordar este problema es desarrollar una construcción análoga, incluso es posible replicar los pasos utilizados para estimar  $\|x^T(T)\|_{L^2(\Omega; \mathbb{R}^n)}$  y  $\|\varphi^T(0)\|_{L^2(\Omega; \mathbb{R}^n)}$ , no obstante, como es desarrollado en este artículo es necesario utilizar



la ecuación de Ricatti, ecuación que no sabemos si tiene las mismas propiedades, tanto de estabilidad como incluso de buen planteo cuando las matrices asociadas dependen del parámetro  $\omega$ .

No es difícil observar que la principal propiedad que utilizamos en la demostración de nuestros teoremas principales corresponde a las estimaciones entregadas por los Lemas 2.2.1 y 2.2.2, que a su vez estos son obtenidos gracias a las hipótesis 1 y 2. Por lo anterior, podría ser natural preguntarse si es posible obtener tales lemas, o incluso probar que las hipótesis 1 y 2 son una consecuencia de las hipótesis de controlabilidad y observabilidad en promedio.

- Al igual que en el contexto determinista, nos preguntamos qué ocurre cuando la dinámica que subyace el sistema evolutivo tiene un comportamiento no lineal que incorpora el parámetro  $\omega$ , y si existen condiciones en donde nos sea posible probar la propiedad de turnpike. Quizás para este problema solo es posible dar una respuesta local, como es en el artículo [51] en donde se utilizan entre otras herramientas la ecuación de Ricatti y el Principio del máximo de Pontryagin, para probar la propiedad de turnpike en un sistema no lineal sin incertidumbre. En el Capítulo 4 analizamos este problema en el contexto de Deep Learning.



## Capítulo 3

# Propiedad de Turnpike para la Ecuación del Calor con Incertidumbre

### 3.1. Introducción

En este capítulo probaremos la propiedad de turnpike exponencial para una ecuación del calor con incertidumbre, y condiciones de frontera Dirichlet homogénea. Esta ecuación incorporará un parámetro de difusión  $\alpha$  aleatorio.

Al igual que en el contexto determinista, como bien se comenta en el Capítulo 1, el efecto regularizante de la ecuación del calor se mantiene presente en el contexto con incertidumbre. Por lo anterior, a diferencia del caso finito dimensional con incertidumbre, no es posible probar la propiedad de turnpike utilizando hipótesis de controlabilidad y observabilidad exacta en promedio. Para probar la propiedad de turnpike nos centraremos en las propiedades inherentes de la ecuación del calor.

La ecuación del calor con incertidumbre que analizaremos fue estudiada en la subsección 1.5.3, del Capítulo 1, en donde se comenta que bajo algunos supuestos sobre  $\alpha$ , esta ecuación es aproximadamente controlable en promedio y controlable a cero en promedio.

#### 3.1.1. Problema a estudiar

Consideramos un espacio de probabilidad  $(\Omega, \mathcal{F}, \mu)$ , donde  $\Omega$  corresponde al espacio muestral,  $\mathcal{F}$  la  $\sigma$ -álgebra, y  $\mu$  la medida definida sobre  $\mathcal{F}$ . Recordemos que si  $(\Omega, \mathcal{F}, \mu)$  es un espacio de probabilidad, se verifica que  $\mu(\Omega) = 1$ . Sea  $G \subset \mathbb{R}^n$  acotado y consideremos el siguiente sistema

$$\begin{cases} y_t - \alpha(\omega)\Delta y = u\chi_{G_0} & \text{en } G \times (0, T), \\ y = 0 & \text{sobre } \partial G \times (0, T), \\ y(0) = y_0 & \text{en } G, \end{cases} \quad (3.1)$$

donde  $\omega \in \Omega$  corresponde al parámetro aleatorio,  $y(\cdot, \omega) \in C([0, T]; X)$  corresponde al estado,  $\alpha : \Omega \rightarrow \mathbb{R}^+$  es una variable aleatoria estrictamente positiva<sup>1</sup> y  $\chi_{G_0}$  es la función característica sobre  $G_0 \subset G$ . Además, supondremos que tanto el control  $u \in L^2(0, T; L^2(G_0))$  como la condición inicial  $y_0 \in L^2(G)$  son independientes de  $\omega$ .

Consideremos el siguiente problema de minimización

$$\min_{u \in L^2(0, T; L^2(G_0))} \left\{ J^T(u) = \frac{1}{2} \left( \int_0^T \|u(t)\|_{L^2(G_0)}^2 dt + \int_0^T \int_{\Omega} \|C(\omega)y(t, \omega) - z\|_{L^2(G)}^2 d\mu dt \right) \right\}, \quad (3.2)$$

con  $y(t, \omega)$  solución de (3.1) y  $z \in L^2(G)$  una trayectoria fija. Por otro lado, consideremos el siguiente problema estacionario

$$\min_{u \in L^2(G_0)} \left\{ J^s(u) = \frac{1}{2} \left( \|u\|_{L^2(G_0)}^2 + \int_{\Omega} \|C(\omega)y(\omega) - z\|_{L^2(G)}^2 d\mu \right) \right\}, \quad (3.3)$$

con  $y(\omega)$  solución de la ecuación elíptica

$$\begin{cases} -\alpha(\omega)\Delta y = \chi_{G_0}u & \text{en } x \in G, \\ y = 0 & \text{sobre } \partial G. \end{cases}$$

Además, consideremos el siguiente espacio de Hilbert

$$L^2(\Omega; L^2(G); \mu) = \left\{ y(\omega) \in L^2(G) : \|y\|_{L^2(\Omega; L^2(G); \mu)}^2 = \int_{\Omega} \|y\|_{L^2(G)}^2 d\mu < \infty \right\},$$

con el producto interno

$$\langle x(\omega), y(\omega) \rangle_{L^2(\Omega; L^2(G); \mu)} = \int_{\Omega} \langle x(\omega), y(\omega) \rangle_{L^2(G)} d\mu.$$

Por simplicidad denotamos  $L^2(\Omega; L^2(G)) = L^2(\Omega; L^2(G); \mu)$ . En el presente capítulo probaremos que si los problemas (3.2) y (3.3), tienen solución  $(u^T, y^T)$  y  $(\bar{u}, \bar{y})$  respectivamente, entonces existen constantes positivas  $\mathcal{C}, \delta$  independientes del horizonte de tiempo  $T$ , tales que

$$\left\| \mathbb{E}(y^T) - \mathbb{E}(\bar{y}) \right\|_{L^2(G)} + \left\| \mathbb{E}(\varphi^T) - \mathbb{E}(\bar{\varphi}) \right\|_{L^2(G)} + \left\| u^T - \bar{u} \right\|_{L^2(G_0)} \leq \mathcal{C}(e^{-(T-t)\delta} + e^{-t\delta}),$$

para todo  $t \in [0, T]$ , donde  $\mathbb{E}$  corresponde a la esperanza matemática<sup>2</sup> y donde  $(\varphi^T, \bar{\varphi})$  corresponden a las variables duales que representa la caracterización de los minimizadores, vía las condiciones de optimalidad de primer orden. Esta propiedad es llamada en la literatura como propiedad de turnpike exponencial.

<sup>1</sup>Es decir, que  $\alpha$  es una variable aleatoria tal que satisface que existe  $\delta > 0$  tal que  $\alpha(\omega) \geq \delta > 0$  para todo  $\omega \in \Omega$ .

<sup>2</sup>Es decir,  $\mathbb{E}(u) = \int_{\Omega} u(\omega) d\mu$ .

Adicionalmente, probaremos que el promedio temporal del par óptimo  $(u^T, x^T)$  converge al par óptimo  $(\bar{u}, \bar{x})$ , es decir

$$\begin{aligned}\frac{1}{T} \int_0^T y^T(t, \omega) dt &\longrightarrow \bar{y}(\omega), \\ \frac{1}{T} \int_0^T u^T(t) dt &\longrightarrow \bar{u},\end{aligned}$$

en  $L^2(\Omega; L^2(G))$  y  $L^2(G_0)$  respectivamente, cuando  $T \rightarrow \infty$ .

### 3.2. Desarrollo y Estimaciones Previas

Como es señalado anteriormente, introducimos  $(\Omega, \mathcal{F}, \mu)$  un espacio de probabilidad. Además, supondremos que  $G_0 \subset G \subset \mathbb{R}^n$ , con  $G$  un conjunto acotado con frontera  $C^2$ . Así consideremos el sistema lineal (*Ecuación del calor aleatoria*)

$$\begin{cases} y_t - \alpha(\omega)\Delta y = u\chi_{G_0} & \text{en } G \times (0, T), \\ y = 0 & \text{sobre } \partial G \times (0, T), \\ y(0) = y_0 & \text{en } G, \end{cases} \quad (3.4)$$

donde  $y$  corresponde al estado, la función  $\chi_{G_0}$  corresponde a indicatriz sobre  $G_0$  y  $\alpha : \Omega \rightarrow \mathbb{R}^+$  es una variable aleatoria estrictamente positiva. Supondremos que  $u \in L^2(0, T; L^2(G_0))$ ,  $y_0 \in L^2(G)$  a menos que se diga lo contrario. Además, supondremos que tanto el control  $u$  como la condición inicial  $y_0$ , son independientes de  $\omega \in \Omega$ . Desde aquí en adelante denotaremos  $U = L^2(G_0)$  y  $X = L^2(G)$ . En lo que sigue  $K$  y  $\hat{K}$  serán constantes positivas independientes de  $T$  que pueden cambiar en cada línea.

**Observación 3.2.1.** Sabemos desde la Teoría de Semigrupos que, para una ecuación del calor sin coeficientes aleatorios, es decir,

$$\begin{cases} w_t - \Delta w = f & \text{en } G \times (0, T), \\ w = 0 & \text{sobre } \partial G \times (0, T), \\ w(0) = x & \text{en } G. \end{cases} \quad (3.5)$$

El operador  $\Delta : H^2(G) \cap H_0^1(G) \subset X \rightarrow X$ <sup>3</sup> es un generador infinitesimal de un semigrupo fuertemente continuo de contracciones  $\{S(t)\}_{t \geq 0}$ , cuando  $f \in L^2(0, T; X)$  y  $x \in X$  (ver [9]). Además, se asegura la existencia de una única solución débil  $w \in C([0, T]; X)$ , caracterizada por la formula ([9],[4,

<sup>3</sup>En este caso el dominio de  $\Delta$  es  $D(\Delta) = H_0^1(G) \cap H^2(G)$ , pues  $\partial G$  es de clase  $C^2$ .

Proposición 3.1])

$$w(t) = S(t)x + \int_0^t S(t-s)f(s)ds. \quad (3.6)$$

En particular, al ser  $S$  un semigrupo de contracciones se verifica que  $\|S(t)\|_{\mathcal{L}(X)} \leq 1$  para todo  $t \geq 0$ , y en consecuencia tenemos que (3.6) implica que existe  $\hat{K} > 0$  tal que

$$\|w\|_{C([0,T];X)}^2 \leq \hat{K}(\|x\|_X^2 + \|f\|_{L^2(0,T;X)}^2). \quad (3.7)$$

Así para nuestro problema, podemos definir el operador  $\alpha(\omega)\Delta : H^2(G) \cap H_0^1(G) \subset X \rightarrow X$ , el cual para cada  $\omega \in \Omega$  fijo, es un generador infinitesimal de un semigrupo fuertemente continuo de contracciones. Por tanto existe una única solución débil  $y(\cdot, \cdot, \omega) \in C([0, T]; X)$ , para cada  $\omega$ . Además, si denotamos el semigrupo generado por  $\alpha(\omega)\Delta$  como  $\{\mathcal{T}(t, \omega)\}_{t \geq 0}$  para cada  $\omega \in \Omega$ , no es difícil verificar que la solución  $y$  de (3.4) satisface

$$\begin{aligned} y(t, \omega) &= \mathcal{T}(t, \omega)y_0 + \int_0^t \mathcal{T}(t-s, \omega)u(s)\chi_{G_0}ds \\ &= S(\alpha(\omega)t)y_0 + \int_0^t S(\alpha(\omega)(t-s))u(s)\chi_{G_0}ds, \end{aligned} \quad (3.8)$$

donde  $\{S(t)\}_{t \geq 0}$  corresponde al semigrupo asociado a la ecuación (3.5) de la observación 3.2.1. Bajo nuestros supuestos iniciales tenemos el siguiente teorema.

**Teorema 3.2.1.** Sea  $u \in L^2(0, T; U)$  e  $y_0 \in X$ . Entonces  $y$  la solución de (3.4) pertenece a  $C([0, T]; L^2(\Omega; X))$ . Además, existe  $K > 0$  tal que se satisface la estimación

$$\|y\|_{C([0,T];L^2(\Omega;X))}^2 \leq K(\|y_0\|_X^2 + \|u\|_{L^2(0,T;U)}^2).$$

*Demostración.* Consideremos  $0 \leq t_2 \leq t_1 \leq T$ . Así procedemos directamente considerando la expresión (3.8), obteniendo que

$$\begin{aligned} \|y(t_1) - y(t_2)\|_{L^2(\Omega;X)} &\leq \left\| \left( S(\alpha(\omega)t_1)y_0 + \int_0^{t_1} S(\alpha(\omega)(t_1-s))u(s)\chi_{G_0}ds \right) \right. \\ &\quad \left. - \left( S(\alpha(\omega)t_2)y_0 + \int_0^{t_2} S(\alpha(\omega)(t_2-s))u(s)\chi_{G_0}ds \right) \right\|_{L^2(\Omega;X)} \\ &\leq \| [S(\alpha(\omega)t_1) - S(\alpha(\omega)t_2)]y_0 \|_{L^2(\Omega;X)} + \left\| \int_{t_2}^{t_1} S(\alpha(\omega)(t_1-s))u(s)\chi_{G_0}ds \right\|_{L^2(\Omega;X)} \\ &\quad + \left\| \int_0^{t_2} [S(\alpha(\omega)(t_1-s)) - S(\alpha(\omega)(t_2-s))]u(s)\chi_{G_0}ds \right\|_{L^2(\Omega;X)}. \end{aligned}$$

Observemos que tanto el primer como tercer término del lado derecho de la desigualdad anterior, van a cero cuando  $t_2 \rightarrow t_1$ , esto desde la continuidad del semigrupo. Para estimar el término restante definamos  $g(s) := S(\alpha(\omega)(t_1 - s))u(s)\chi_{G_0} \in L^2(0, T; L^2(\Omega; X))$ . Así utilizando ([9, Proposición 1.4.29.]) tenemos que

$$|t_1 - t_2| \left( \frac{1}{t_1 - t_2} \int_{t_2}^{t_1} g(s) ds \right) \rightarrow 0 \quad \text{en } L^2(0, T; L^2(\Omega; X)),$$

cuando  $t_2 \rightarrow t_1$ . Concluyendo la continuidad de  $y$  con respecto a  $t$ . Finalmente, desde (3.7) tenemos en particular que para cada  $t \in [0, T]$

$$\|y(t, \omega)\|_X^2 \leq \hat{K}(\|y_0\|_X^2 + \|u\|_{L^2(0, T; U)}^2). \quad (3.9)$$

Así integrando sobre  $\Omega$  y tomado supremo en tiempo, obtenemos

$$\|y\|_{C([0, T]; L^2(\Omega; X))}^2 \leq K(\|y_0\|_X^2 + \|u\|_{L^2(0, T; U)}^2). \quad (3.10)$$

□

**Observación 3.2.2.** La regularidad anterior puede ser razonado como una consecuencia del hecho que el semigrupo  $\{\mathcal{T}(t, \omega)\}_{t \geq 0}$  es una composición entre el semigrupo  $\{S(t)\}_{t \geq 0}$  y la función medible  $\alpha$ . En el contexto general con incertidumbre, la estructura de un semigrupo arbitrario  $\{\mathcal{W}(t, \omega)\}_{t \geq 0}$  no necesariamente es la composición de un semigrupo  $\{\mathcal{B}(t)\}_{t \geq 0}$  generado por la ecuación sin incertidumbre, y la función con incertidumbre. En relación al contexto general, en [54] se analiza desde una perspectiva de semigrupos un operador abstracto  $C(q) : D(C) \subset \mathcal{V} \rightarrow \mathcal{V}$ , el cual depende de un parámetro  $q \in Q_{ad} \subset Q$ , donde  $Q$  y  $\mathcal{V}$  son espacios de Banach.

Consideremos ahora el problema de minimización,

$$\min_{u \in L^2(0, T; U)} \left\{ J^T(u) = \frac{1}{2} \int_0^T \left( \|u(t)\|_U^2 + \int_{\Omega} \|C(\omega)y(t, \omega) - z\|_X^2 d\mu \right) dt, \text{ y solución de (3.4)} \right\}, \quad (3.11)$$

donde supondremos que  $C \in C(\Omega, \mathcal{L}(X))$  y  $z \in X$  es una trayectoria fija. Con respecto a la existencia de un mínimo del problema (3.11) tenemos el siguiente resultado.

**Proposición 3.2.1.** Existe un único control óptimo  $u^T \in L^2(0, T; U)$  para (3.11), y una única trayectoria óptima  $y^T$  asociada al control óptimo, Además, este control óptimo viene caracterizado por

$$u^T = - \int_{\Omega} \chi_{G_0} \varphi^T d\mu, \quad \begin{cases} -\varphi_t^T - \alpha(\omega)\Delta\varphi^T = C^*(Cy^T - z) & \text{en } G \times (0, T), \\ \varphi^T = 0 & \text{sobre } \partial G \times (0, T), \\ \varphi^T(T) = 0 & \text{en } G. \end{cases} \quad (3.12)$$

*Demostración.* Comenzamos observando que  $J^T \geq 0$ , entonces existe un ínfimo

$$P = \inf_{u \in L^2(0,T;U)} \{J^T(u)\}.$$

Así consideremos  $\{u_n\}_{n \in \mathbb{N}}$  una sucesión minimizante en  $L^2(0, T, U)$ , es decir, que satisface que  $J^T(u_n) \rightarrow 0$ . Observemos que  $\{u_n\}_{n \in \mathbb{N}}$  es acotada. En efecto, notemos que para algún  $u_0 \in L^2(0, T, U)$ , por ser  $\{u_n\}_{n \in \mathbb{N}}$  minimizante, existe  $N > 0$  tal que

$$\|u_n\|_{L^2(0,T;U)}^2 \leq 2J^T(u_0) < \infty \quad \forall n \geq N,$$

y en consecuencia  $K > 0$  tal que

$$\|u_n\|_{L^2(0,T;U)}^2 \leq \max \{ \|u_1\|^2, \dots, \|u_N\|^2, 2J^T(u_0) \} = K \quad \forall n \in \mathbb{N}.$$

Luego desde la reflexividad de  $L^2(0, T; U)$  sigue que existe una subsucesión  $\{u_k\}_{k \in \mathbb{N}} \subset L^2(0, T; U)$  tal que  $u_k \rightharpoonup u^T$  para algún  $u^T \in L^2(0, T; U)$ . Ahora utilizando (3.8), obtenemos una sucesión  $\{y_n\}_{n \in \mathbb{N}}$  definida por

$$y_n(t, \omega) = S(\alpha(\omega)t)y_0 + \int_0^t S(\alpha(\omega)(t-s))u_n(s)\chi_{G_0}ds.$$

Claramente  $\{y_n\}_{n \in \mathbb{N}}$  converge puntualmente, es decir,  $y_n(t, \omega) \rightarrow y^T(t, \omega) \in X$  para cada  $t \in [0, T]$  y  $\omega \in \Omega$ , con  $y^T$  solución de (3.4) asociada a  $u^T$ . Probemos que  $\{y_n(t)\}_{n \in \mathbb{N}}$  converge en  $L^2(\Omega; X)$  para cada  $t \in [0, T]$ . Para ello observemos que la sucesión  $\{y_n(t)\}_{n \in \mathbb{N}}$  está dominada. En efecto, si definimos  $\tau_t = \alpha(\omega)t \in [0, \infty)$ , y utilizando el hecho que  $S(\cdot)$  es un semigrupo de contracciones [9, Definición 3.41], tenemos que

$$\begin{aligned} \|y_n(\omega)\|_X^2 &\leq 2(\|S(\tau_t)y_0\|_X^2 + \int_0^t \|S(\tau_s)\chi_{G_0}u_n(t-s)\|_X^2 ds) \leq 2(\|y_0\|_X^2 + \int_0^t \|u_n(t-s)\|_X^2 ds) \\ &\leq 2(\|y_0\|_X^2 + \sqrt{TK}) = C_{T,y_0}. \end{aligned}$$

Como el lado derecho no depende de  $\omega \in \Omega$  y  $\mu(\Omega) = 1$ , se sigue desde el teorema de convergencia dominada que  $y_n(t) \rightarrow y^T(t)$  en  $L^2(\Omega; X)$  para cada  $t \in [0, T]$ . Por otro lado, podemos probar de forma estándar que  $y_n \rightarrow y^T$  en  $C([0, T]; L^2(\Omega; X))$ , gracias al teorema de Arzela-Azcoli. En particular como  $C \in C(\Omega; \mathcal{L}(X))$ , se tiene que  $Cy_n \rightarrow Cy^T$  en  $C([0, T]; L^2(\Omega; X))$ . Finalmente, por la semi-continuidad inferior del funcional  $J^T$  (en particular por la semi-continuidad inferior débil), tenemos que

$$J^T(u^T) \leq \liminf_{k \rightarrow \infty} J^T(u_k) = P.$$

Por tanto, necesariamente  $J^T(u^T) = P$ . La unicidad de  $u^T$  sigue directamente (por contradicción) desde la convexidad estricta del funcional  $J^T$ , mientras que la unicidad de  $y^T$  se obtiene desde



el buen planteo del problema. La caracterización (3.12) del control  $u^T \in L^2(0, T; U)$  se obtienen de las condiciones necesarias de primer orden para el problema (3.11)<sup>4</sup>.  $\square$

**Observación 3.2.3.** Es clásico de la literatura verificar que el sistema (3.12) está bien puesto (en el sentido de Hadamard) para cada  $\omega \in \Omega$  (ver [10]). Para ello basta usar el cambio de variable  $t = T - s$ , el cual lleva la ecuación retrograda (3.12), a una ecuación en el buen sentido del tiempo, como el sistema (3.4). Además, observando que  $C^*(Cy^T - z) \in L^2(0, T; L^2(\Omega; X))$ , y bajo el razonamiento realizado en la observación 3.2.1, podemos verificar que para cada  $\omega \in \Omega$ , existe una única solución  $\varphi^T(\cdot, \cdot, \omega) \in C([0, T]; X)$  de (3.12) caracterizada por

$$\varphi^T(t, \omega) = \int_t^T S(\alpha(\omega)(s - t)) C^*(Cy^T - z) ds.$$

A partir de la caracterización anterior y realizando los mismos pasos que en Teorema 3.2.1, podemos probar que  $\varphi^T \in C([0, T]; L^2(\Omega; X))$  y verifica la siguiente estimación

$$\|\varphi^T\|_{C([0, T]; L^2(\Omega; X))} \leq K \|C^*(Cy^T - z)\|_{L^2(0, T; L^2(\Omega; X))}.$$

Ahora consideremos el siguiente problema estacionario

$$\min_{u \in U} \left\{ J^s(u) = \frac{1}{2} \left( \|u\|_U^2 + \int_{\Omega} \|Cy - z\|_X^2 d\mu \right), \text{ y solución de (3.14)} \right\}, \quad (3.13)$$

donde  $u \in U$  es independiente de  $\omega \in \Omega$ ,  $z \in X$  es la trayectoria fija definida en el problema (3.11), e  $y$  es solución del sistema

$$\begin{cases} -\alpha(\omega)\Delta y = \chi_{G_0} u & \text{en } x \in G, \\ y = 0 & \text{sobre } \partial G. \end{cases} \quad (3.14)$$

**Observación 3.2.4.** Como  $\alpha \geq \delta > 0$  podemos probar que existe una única solución débil  $y(\omega) \in H_0^1(G)$  del sistema (3.14), para cada  $\omega \in \Omega$ . Esta viene caracterizada por la formulación variacional

$$\alpha(\omega) \langle \nabla y, \nabla v \rangle_X = \langle u, v \rangle_X \quad \forall v \in H_0^1(G), \quad (3.15)$$

para cada  $\omega$ . Lo anterior vía al teorema de Lax-Milgram [7, Corollary 5.8]. Además, para  $\hat{K} > 0$  se satisface la siguiente estimación

$$\|y(\omega)\|_{H_0^1(G)}^2 \leq \hat{K} \|u\|_U^2. \quad (3.16)$$

---

<sup>4</sup>Los pasos son completamente análogos a los realizados en el Teorema 2.3.2, del Capítulo 2

Así integrando con respecto a  $\Omega$  la desigualdad (3.16) obtenemos que

$$\|y\|_{L^2(\Omega; H_0^1(G))}^2 \leq \hat{K} \|u\|_U^2. \quad (3.17)$$

Por tanto  $y \in L^2(\Omega; H_0^1(G))$ .

Con respecto a la existencia de un mínimo del problema (3.13) tenemos el siguiente resultado

**Proposición 3.2.2.** *Existe un único control óptimo para (3.13), y una única trayectoria óptima  $\bar{y}$  asociada al control óptimo. Además, este control óptimo viene caracterizado por*

$$\bar{u} = - \int_{\Omega} \chi_{G_0} \bar{\varphi} d\mu, \quad \begin{cases} -\alpha(\omega) \Delta \bar{\varphi} = C^*(C\bar{y} - z) & \text{en } G, \\ \bar{\varphi} = 0 & \text{sobre } \partial G. \end{cases} \quad (3.18)$$

*Demostración.* Como  $J^s(u) \geq 0$  entonces existe  $P \geq 0$  tal que

$$P = \inf_{u \in U} \{J^s(u)\}.$$

Así consideremos una sucesión minimizante  $\{u_n\}_{n \in \mathbb{N}}$ , y notemos que para algún  $u_0 \in U$  existe  $N > 0$  tal que

$$\|u_n\|_U^2 \leq 2J^s(u_0) < \infty \quad \forall n \geq N,$$

y en consecuencia  $K > 0$  tal que

$$\|u_n\|_U^2 \leq \max \left\{ \|u_1\|^2, \dots, \|u_N\|^2, 2J^T(u_0) \right\} = K \quad \forall n \in \mathbb{N}.$$

Por lo tanto  $\{u_n\}_{n \in \mathbb{N}}$  es una sucesión acotada en  $U$ . Como  $U$  es reflexivo existe una subsucesión de  $\{u_n\}_{n \in \mathbb{N}}$  que tal que  $u_n \rightharpoonup \bar{u}$  para algún  $\bar{u} \in U$ .

Sabemos que cada  $u_n$  define una solución de  $y_n$  de (3.14), para cada  $\omega$ . Esta solución viene caracterizada por la formulación variacional (3.15) para cada  $\omega$ , en consecuencia, tenemos que

$$\alpha(\omega) \langle \nabla y_n, \nabla v \rangle_X = \langle u_n, v \rangle_X \quad \forall v \in H_0^1(\Omega).$$

En particular por la convergencia débil

$$\alpha(\omega) \langle \nabla y_n, \nabla v \rangle_X = \langle u_n, v \rangle_X \rightarrow \langle \bar{u}, v \rangle_X \quad \forall v \in H_0^1(\Omega). \quad (3.19)$$

Como cada  $u \in U$  define una solución de (3.14), entonces se sigue de (3.19) que existe  $\bar{y}(\omega)$  tal que  $y_n(\omega) \rightharpoonup \bar{y}(\omega)$  en  $H^2(G)$ . Utilizando el teorema de Rellich-Kondrachov se tiene que  $y_n(\omega) \rightarrow \bar{y}(\omega)$  en  $H^1(G)$  para cada  $\omega \in \Omega$ . Además, utilizando la continuidad del operador traza, obtenemos que  $y_n(\omega) \rightarrow \bar{y}(\omega)$  en  $H_0^1(G)$  para cada  $\omega \in \Omega$ .

Por la observación 3.2.4 podemos considerar a la sucesión  $\{y_n(\omega)\} \subset L^2(\Omega; H_0^1(G))$ , así por la estimación (3.17) tenemos que la sucesión está dominada. Luego por el teorema de convergencia dominada, se sigue que  $y_n \rightarrow y$  en  $L^2(\Omega; H_0^1(G))$ .

Finalmente, desde la semi-continuidad inferior del funcional  $J^s$ , se verifica que

$$J^s(\bar{u}) \leq \liminf_{k \rightarrow \infty} J^s(u_n) = P.$$

Se sigue que  $J^s(u) = P$ , concluyendo la existencia. La unicidad de  $\bar{u}$  se obtiene directamente (por contradicción) desde la convexidad estricta del funcional  $J^s$ , mientras que la unicidad de  $\bar{y}$ , se obtiene desde el buen planteo del problema. La caracterización (3.18) de  $\bar{u} \in U$ , obtiene directamente de las condiciones necesarias de primer orden del problema (3.13).  $\square$

**Observación 3.2.5.** *Observemos que podemos obtener las mismas estimaciones que en la observación 3.2.4, pero para  $\bar{\varphi}$  solución del sistema (3.18). Utilizando el teorema de Lax-Milgram podemos probar que existe una única solución débil para cada  $\omega$  de (3.18) caracterizada por la formulación variacional*

$$\alpha(\omega) \langle \nabla \bar{\varphi}, \nabla v \rangle_X = \langle C^*(C\bar{y} - z), v \rangle_X \quad \forall v \in H_0^1(G). \quad (3.20)$$

para cada  $\omega$ . Además, satisface que

$$\|\bar{\varphi}\|_{L^2(\Omega; H_0^1(G))}^2 \leq K \|C^*(C\bar{y} - z)\|_{L^2(\Omega; X)}^2.$$

En lo que sigue denotemos como “ $A \xrightarrow{D} B$ ” cuando  $A$  se inyecte de forma densa en  $B$ . Así consideremos la siguiente definición

**Definición 3.2.1.** *Sea  $V$  un espacio de Banach reflexivo y separable, y sea  $H$  un espacio de Hilbert tal que  $V \xrightarrow{D} H$  de forma continua. Además, supongamos que forman un tripleta de Gelfand (o tripleta de evolución), es decir,  $V \xrightarrow{D} H \xrightarrow{D} V^*$  continuamente. Denotemos como  $Z = L^p(0, T; V)$ , así definimos el espacio*

$$W_p^1(0, T; V; H) := \{z \in Z : z_t \in Z^*\},$$

con norma

$$\|z\|_{W_p^1} = \|z\|_{L^p(0, T; V)} + \|z_t\|_{L^p(0, T; V^*)},$$

donde  $z_t$  denota la derivada en la variable  $t$  de  $z$  en un sentido distribucional.

**Observación 3.2.6.** *Denotando  $H = L^2(\Omega; L^2(G))$  y  $V = L^2(\Omega; H_0^1(G))$ , tenemos que  $V \hookrightarrow H \hookrightarrow V^*$ . Esto es directo desde el hecho que  $H_0^1(G) \hookrightarrow L^2(G) \hookrightarrow H^{-1}(G)$ , junto al hecho de que si  $Q$  es un espacio de Banach reflexivo, entonces  $(L^2(\Omega; Q))^* \cong L^2(\Omega; Q^*)$ . Por lo tanto, podemos tomar  $H = L^2(\Omega; L^2(G))$  y  $V = L^2(\Omega; H_0^1(G))$  en la Definición 3.2.1.*

Así tenemos el siguiente teorema de regularidad.

**Teorema 3.2.2.** Denotemos como  $H = L^2(\Omega; X)$  y  $V = L^2(\Omega; H_0^1(G))$ . Supongamos que  $y_0 \in X$  y  $u \in L^2(0, T; U)$ , luego la solución de (3.4) pertenece a  $W_2^1(0, T; V; H)$ .

*Demostración.* Observemos que si definimos  $Z = L^2(0, T; V)$  con  $V$  como en el enunciado del teorema, entonces  $Z = L^2(0, T; L^2(\Omega; H_0^1(G)))$ . Además, por lo comentado en la observación 3.2.6, tenemos que  $(Z)^* = (L^2(0, T; L^2(\Omega; H_0^1(G))))^* = L^2(0, T; L^2(\Omega; H^{-1}(G)))$ . Así dividiremos la demostración en dos pasos

**Paso 1:** Probaremos que  $y$  solución de (3.4) satisface que  $y \in Z$ .

Para ello sabemos desde la Teoría de semigrupos que si  $y_0 \in D := H_0^1(G) \cap H^2(G) \subset X$  y  $u\chi_{G_0} \in C^1([0, T]; X)$ , entonces  $y(\omega) \in C^1([0, T]; D) \cap C([0, T]; X)$  para cada  $\omega \in \Omega$ . Por lo anterior, supondremos que  $y_0 \in D \subset X$  y  $u\chi_{G_0} \in C^1([0, T]; X) \subset L^2(0, T; U)$ , y concluiremos por densidad este paso.

Procedemos multiplicando (3.4) por  $y$  e integrando en  $G$ , denotando como  $x$  la variable espacial obtenemos

$$\frac{1}{2} \frac{d \|y\|_X^2}{dt} - \int_G \langle \alpha \Delta y, y \rangle dx = \langle \chi_{G_0} u, y \rangle_X.$$

Así integrando en  $(0, T)$  e integrando por partes en espacio tenemos

$$\frac{\|y(T)\|_X^2}{2} - \frac{\|y_0\|_X^2}{2} + \alpha(\omega) \int_0^T \int_G \|\nabla y\|^2 dx dt \leq \int_0^T \langle \chi_{G_0} u, y \rangle_X dt.$$

Utilizando la desigualdad de Young para  $\varepsilon > 0$  obtenemos

$$\frac{\|y(T)\|_X^2}{2} + \alpha(\omega) \int_0^T \int_G \|\nabla y\|^2 dx dt \leq \int_0^T \left\{ \frac{\|u\|_U^2}{2\varepsilon} + \frac{\varepsilon \|y\|_X^2}{2} \right\} dt + \frac{\|y_0\|_X^2}{2}. \quad (3.21)$$

Ahora por la regularidad de  $y$  es válida la desigualdad de Poincaré existe  $C_{pp} > 0$  tal que

$$\int_0^T \int_G \frac{\varepsilon \|y\|^2}{2} dx dt \leq \int_0^T \int_G \frac{\varepsilon C_{pp} \|\nabla y\|^2}{2} dx dt.$$

Luego remplazando la estimación anterior en (3.21), y utilizando la existencia de  $\delta > 0$  tal que  $\alpha \geq \delta > 0$  obtenemos

$$\left( \delta - \frac{C_{pp}\varepsilon}{2} \right) \int_0^T \int_G \|\nabla y\|^2 dx dt \leq \int_0^T \frac{\|u\|_U^2}{2\varepsilon} dt + \frac{\|y_0\|_X^2}{2}.$$

Tomando  $\varepsilon < 2\delta/C_{pp}$  se tiene que existe  $K > 0$  tal que

$$\int_0^T \int_G \|\nabla y\|^2 dx dt \leq K \left( \int_0^T \|u\|_U^2 dt + \|y_0\|_X^2 \right).$$

Ahora integrando con respecto a  $\Omega$ , y utilizando la independencia tanto del control  $u$  como de la condición inicial  $y_0$  con respecto a  $\omega \in \Omega$ , tenemos<sup>5</sup>

$$\|y\|_Z^2 = \int_\Omega \int_0^T \int_G \|\nabla y\| dx dt d\mu \leq K \left( \int_0^T \|u\|_U^2 dt + \|y_0\|_X^2 \right).$$

En consecuencia  $y \in Z$ . El resultado general cuando  $y_0 \in X$  y  $u \in L^2(0, T; U)$  se concluye por densidad utilizando la continuidad de la norma.

**Paso 2:** Probaremos que  $y$  solución de (3.4) satisface que  $y \in Z^* = L^2(0, T; L^2(\Omega; H^{-1}(G)))$ . Para ello observemos que  $\xi \in L^2(0, T; L^2(\Omega; H^{-1}(G)))$  si existe  $K > 0$  tal que

$$\left\| \langle \xi, v \rangle_{L^2(G)} \right\|_{L^2(0, T; L^2(\Omega))} \leq K \|v\|_{H_0^1(G)} \quad \forall v \in H_0^1(G).$$

Con lo anterior en mente observemos que si multiplicamos por  $v \in H_0^1(G)$  la ecuación (3.4), e integramos sobre  $G$  tenemos que

$$\langle y_t, v \rangle_X - \langle \Delta y, v \rangle_X = \langle \chi_{G_0} u, v \rangle_X$$

Como del paso anterior sabemos que  $y \in L^2(0, T; L^2(\Omega; H_0^1(G)))$ , integrando por partes tenemos que

$$\langle y_t, v \rangle_X + \langle \nabla y, \nabla v \rangle_X = \langle \chi_{G_0} u, v \rangle_X$$

Ahora tomando norma en  $L^2(0, T; L^2(\Omega))$  y aplicando la desigualdad de Cauchy-Schwarz tenemos que

$$\begin{aligned} \left\| \langle y_t, v \rangle_{L^2(G)} \right\|_{L^2(0, T; L^2(\Omega))} &\leq \|y\|_{L^2(0, T; L^2(\Omega; H_0^1(G)))} \|v\|_{H_0^1(G)} + \|u\|_{L^2(0, T; L^2(\Omega; X))} \|v\|_X \\ &\leq \max \{1, C_{pp}\} \left( \|u\|_{L^2(0, T; L^2(\Omega; X))} + \|y\|_{L^2(0, T; L^2(\Omega; H_0^1(G)))} \right) \|v\|_{H_0^1(G)}, \end{aligned}$$

donde  $C_{pp}$  corresponde a la constante de Poincaré. □

**Observación 3.2.7.** Hasta el momento hemos probado que  $y$  solución de (3.4) satisface que  $y \in C(0, T; H) \cap W_p^1(0, T; V; H)$  con  $H$  y  $V$  definidos como en el Teorema 3.2.2, no obstante, desde el libro [57, Capítulo 23] sabemos que  $W_p^1(0, T; V; H) \hookrightarrow C(0, T; H)$  continuamente. Por tanto basta considerar  $y \in$

<sup>5</sup>Podemos utilizar el Teorema de Fubini pues  $y(\cdot, \omega, x)$  es medible con respecto a  $\omega$  y  $x$ .

$W_p^1(0, T; V; H) \hookrightarrow C(0, T; H)$  y tomaremos el respectivo representante continuo definido casi en todo punto <sup>6</sup>, cuando lo necesitemos.

Además de lo anterior, observemos que  $\varphi$  solución de (3.12), vía un cambio de variable, satisface la misma ecuación que  $y$ , por tanto de igual forma tenemos que  $\varphi \in W_p^1(0, T; V; H)$ .

Los siguientes lemas serán de gran importancia para concluir nuestros resultados principales.

**Lema 3.2.1.** *Sea  $y$  solución de (3.4). Entonces existe  $K > 0$  tal que*

$$\|y(T)\|_{L^2(\Omega; X)}^2 \leq K \left( \int_0^T \|u\|_U^2 dt + \|y_0\|_X^2 \right). \quad (3.22)$$

*Demostración.* Multiplicando (3.4) por  $y$  e integrando en espacio, obtenemos que

$$\frac{1}{2} \frac{d \|y\|_X^2}{dt} - \langle \alpha \Delta y, y \rangle_X = \langle \chi_{G_0} u, y \rangle_X.$$

Luego integrando en  $(0, T)$ , integrando por partes en espacio, y aplicando la desigualdad de Young para  $\varepsilon > 0$  tenemos

$$\frac{\|y(T)\|_X^2}{2} + \alpha \int_0^T \|\nabla y\|_X^2 dt \leq \int_0^T \left\{ \frac{1}{2\varepsilon} \|u\|_U^2 + \frac{\varepsilon}{2} \|y\|_X^2 \right\} dt + \frac{\|y_0\|_X^2}{2}. \quad (3.23)$$

Ahora como en particular  $y \in L^2(0, T; H_0^1(G))$  para cada  $\omega \in \Omega$  y  $G$  es acotado, entonces es válida la desigualdad de Poincaré. Así utilizando la desigualdad de Poincaré, y el hecho que existe  $\delta > 0$  tal que  $\alpha \geq \delta$ , obtenemos

$$\frac{\|y(T)\|_X^2}{2} + C_{pp}\delta \int_0^T \|y\|_X^2 dt \leq \int_0^T \left\{ \frac{1}{2\varepsilon} \|u\|_U^2 + \frac{\varepsilon}{2} \|y\|_X^2 \right\} dt + \frac{\|y_0\|_X^2}{2}.$$

Tomando  $\varepsilon < 2C_{pp}\delta$ , e integrando con respecto a  $\Omega$  existe  $K > 0$  tal que

$$\|y(T)\|_{L^2(\Omega; X)}^2 \leq K \left( \int_0^T \|u\|_U^2 dt + \|y_0\|_X^2 \right).$$

□

**Lema 3.2.2.** *Sea  $\varphi$  la solución de (3.12). Entonces existe  $K > 0$  tal que*

$$\|\varphi(0)\|_{L(\Omega; X)}^2 \leq K \left( \int_0^T \|Cy^T - z\|_{L^2(\Omega; X)}^2 dt \right). \quad (3.24)$$

<sup>6</sup>Que además es único c.t.p. [57, Teorema 23.23].

*Demostración.* Multiplicando (3.12) por  $\varphi$  obtenemos que

$$-\frac{1}{2} \frac{d \|\varphi\|_X}{dt} - \langle \alpha \Delta \varphi, \varphi \rangle_X = \left\langle C^*(Cy^T - z), \varphi \right\rangle_X.$$

Ahora integrando por partes, integrando con respecto al tiempo, aplicando la desigualdad de Young, utilizando la existencia de  $\delta > 0$  tal que  $\alpha \geq \delta$  y la desigualdad de Poincaré obtenemos

$$\frac{\|\varphi(0)\|_X^2}{2} + C_{pp}\delta \int_0^T \|\varphi\|_X^2 dt \leq \int_0^T \left\{ \frac{\|Cy^T - z\|_X^2}{2\varepsilon} + \frac{\varepsilon \|\varphi\|_X^2}{2} \right\} dt.$$

Finalmente tomando  $\varepsilon < 2C_{pp}\delta$  e integrando sobre  $\Omega$  se concluye.  $\square$

**Lema 3.2.3.** Sean  $y^T, \bar{y}, \varphi^T$  y  $\bar{\varphi}$  soluciones de los sistemas (3.4), (3.14), (3.12) y (3.18) respectivamente. Definamos  $m = y^T - \bar{y}$  y  $n = \varphi^T - \bar{\varphi}$ . Luego se satisface que

$$\begin{aligned} \langle n(0), m(0) \rangle_{L^2(\Omega; X)} - \langle n(T), m(T) \rangle_{L^2(\Omega; X)} &= \\ &= \int_0^T \|Cm\|_{L^2(\Omega; X)}^2 dt + \int_0^T \|u^T - \bar{u}\|_{L^2(\Omega; X)}^2 dt. \end{aligned}$$

*Demostración.* Observemos que  $m$  y  $n$  satisfacen el siguiente sistema

$$\begin{cases} m_t - \alpha(\omega) \Delta m = -\chi_{G_0} \int_{\Omega} n d\mu & \text{en } G \times (0, T), \\ m = 0 & \text{sobre } \partial G \times (0, T), \\ m(0) = m_0 & \text{en } G, \\ -n_t - \alpha(\omega) \Delta n = C^* C m & \text{en } G \times (0, T), \\ n = 0 & \text{sobre } \partial G \times (0, T), \\ n(T) = n_T & \text{en } G. \end{cases} \quad (3.25)$$

Multiplicando la primera ecuación del sistema (3.25) por  $n$ , y la cuarta ecuación del sistema (3.25) por  $m$ , e integrando ambas sobre  $G \times \Omega$

$$\begin{cases} \langle m_t, n \rangle_{L^2(\Omega; X)} - \langle \alpha(\omega) \Delta m, n \rangle_{L^2(\Omega; X)} = \left\langle -\int_{\Omega} \chi_{G_0} n d\mu, n \right\rangle_{L^2(\Omega; X)}, \\ -\langle n_t, m \rangle_{L^2(\Omega; X)} - \langle \alpha(\omega) \Delta n, m \rangle_{L^2(\Omega; X)} = \langle C^* C m, m \rangle_{L^2(\Omega; X)}. \end{cases} \quad (3.26)$$

Así integrando por parte en  $(0, T)$  la segunda ecuación del sistema (3.26), y restando las ecuaciones del sistema (3.26) resultantes, tenemos

$$\begin{aligned}
 \langle n(0), m(0) \rangle_{L^2(\Omega; X)} - \langle n(T), m(T) \rangle_{L^2(\Omega; X)} &= \\
 &= \int_0^T \langle C^* C m, m \rangle_{L^2(\Omega; X)} dt + \int_0^T \left\langle \int_{\Omega} \chi_{G_0} n d\mu, n \right\rangle_{L^2(\Omega; X)} dt \\
 &= \int_0^T \|Cm\|_{L^2(\Omega; X)}^2 dt + \int_0^T \left\| \int_{\Omega} \chi_{G_0} n d\mu \right\|_X^2 dt \\
 &= \int_0^T \|Cm\|_{L^2(\Omega; X)}^2 dt + \int_0^T \|u^T - \bar{u}\|_U^2 dt.
 \end{aligned}$$

□

**Lema 3.2.4.** Consideremos  $y^T, \bar{y}$  trayectorias óptimas asociadas a (3.11) y (3.13) respectivamente. Entonces existe  $K > 0$  tal que

$$\left\| \frac{1}{T} \int_0^T (y^T - \bar{y}) dt \right\|_{L^2(\Omega, X)}^2 \leq K \left( \left\| \int_0^T \frac{(u^T - \bar{u})}{T} dt \right\|_U^2 + \left\| \frac{y^T(T) - y^T(0)}{T} \right\|_{L^2(\Omega, X)}^2 \right).$$

*Demostración.* Primero notemos que

$$\left\| \int_0^T y^T dt \right\|_{L^2(\Omega; H_0^1(G))}^2 \leq \int_0^T \|y^T\|_{L^2(\Omega; H_0^1(G))}^2 dt = \|y^T\|_{L^2(0, T; L^2(\Omega; H_0^1(G)))}^2.$$

Por el Teorema 3.2.2, sabemos que en particular  $y^T \in L^2(0, T; L^2(\Omega; H_0^1(G)))$ , se sigue que  $\int_0^T y^T dt \in L^2(\Omega; H_0^1(G))$ . Sea  $l \in L^2(\Omega; H_0^1(G))$ , así para cada  $\omega$  tenemos

$$\begin{aligned}
 \left\langle \alpha \nabla \left( \frac{1}{T} \int_0^T (y^T - \bar{y}) dt \right), \nabla l \right\rangle_X &= \left\langle -\alpha \Delta \left( \frac{1}{T} \int_0^T (y^T - \bar{y}) dt \right), l \right\rangle_X \\
 &= \left\langle \frac{1}{T} \int_0^T \left\{ (\chi_{G_0} u^T - y_t^T) - (\chi_{G_0} \bar{u}) \right\} dt, l \right\rangle_X \\
 &= \left\langle \frac{1}{T} \int_0^T \chi_{G_0} (u^T - \bar{u}) dt - \frac{y^T(T) - y^T(0)}{T}, l \right\rangle_X. \quad (3.27)
 \end{aligned}$$

Luego tomando  $l = \frac{1}{T} \int_0^T y^T - \bar{y} dt \in L^2(\Omega; H_0^1(G))$  en (3.27), e integrando en  $\Omega$  obtenemos

$$\langle -\alpha \nabla l, \nabla l \rangle_{L^2(\Omega, X)} = \left\langle l, \frac{1}{T} \int_0^T \chi_{G_0} (u^T - \bar{u}) dt \right\rangle_{L^2(\Omega, X)} - \left\langle l, \frac{y^T(T) - y^T(0)}{T} \right\rangle_{L^2(\Omega, X)}.$$



Usando la desigualdad de Poincaré al lado izquierdo, junto a la existencia de  $\delta > 0$  tal que  $\alpha \geq \delta > 0$ , tenemos que

$$\delta C_{pp} \|l\|_{L^2(\Omega, X)}^2 \leq \left\langle l, \frac{1}{T} \int_0^T \chi_{G_0}(u^T - \bar{u}) dt \right\rangle_{L^2(\Omega, X)} - \left\langle l, \frac{y^T(T) - y^T(0)}{T} \right\rangle_{L^2(\Omega, X)}.$$

Ahora ocupando la desigualdad de Young al lado derecho, existen  $\varepsilon_1, \varepsilon_2 > 0$  tales que

$$\begin{aligned} C_{pp} \delta \|l\|_{L^2(\Omega, X)}^2 &\leq \frac{\varepsilon_1 \|l\|_{L^2(\Omega, X)}^2}{2} + \frac{1}{2\varepsilon_1} \left\| \int_0^T \frac{(u^T - \bar{u})}{T} dt \right\|_U^2 \\ &\quad + \frac{\varepsilon_2 \|l\|_{L^2(\Omega, X)}^2}{2} + \frac{1}{2\varepsilon_2} \left\| \frac{y^T(T) - y^T(0)}{T} \right\|_{L^2(\Omega, X)}^2. \end{aligned}$$

Así escogiendo  $\varepsilon_1, \varepsilon_2 > 0$  de tal forma que  $2C_{pp}\delta > \varepsilon_1 + \varepsilon_2$ , obtenemos que existe  $K > 0$  tal que

$$\|l\|_{L^2(\Omega, X)}^2 \leq K \left( \left\| \int_0^T \frac{(u^T - \bar{u})}{T} dt \right\|_U^2 + \left\| \frac{y^T(T) - y^T(0)}{T} \right\|_{L^2(\Omega, X)}^2 \right).$$

□

### 3.3. Teoremas Principales

En esta sección mostraremos y probaremos nuestros resultados principales del capítulo. Estos corresponden a los resultados comentados en la introducción 3.1.

A continuación, presentamos un resultado de convergencia que relaciona los pares óptimos  $(y^T, u^T)$  y  $(\bar{y}, \bar{u})$ .

**Teorema 3.3.1.** *Los pares óptimos  $(y^T, u^T)$  y  $(\bar{y}, \bar{u})$  asociados a (3.11) y a (3.13) respectivamente, satisfacen*

$$\frac{1}{T} \int_0^T y^T dt \rightarrow \bar{y} \quad \text{en } L^2(\Omega; X), \quad \frac{1}{T} \int_0^T u^T dt \rightarrow \bar{u} \quad \text{en } U,$$

cuando  $T \rightarrow \infty$ .

*Demostración.* Realizaremos la demostración en 2 pasos.

**Paso 1:** Probaremos que se tiene la siguiente estimación

$$\|y^T(T)\|_{L^2(\Omega; X)}^2 \leq \hat{K}T, \quad \|\varphi^T(0)\|_{L(\Omega; X)}^2 \leq \hat{K}T,$$

donde  $y^T$  y  $\varphi^T$  corresponden a las soluciones de (3.12) y (3.4) respectivamente. Para ello consideremos el sistema

$$\begin{cases} y_t^T - \alpha(\omega)\Delta y^T = \chi_{G_0} u^T & \text{en } G \times (0, T), \\ y^T = 0 & \text{sobre } \partial G \times (0, T), \\ y^T(0) = y_0 & \text{en } G, \\ u^T = -\int_{\Omega} \chi_{G_0} \varphi^T d\mu, & \\ -\varphi_t^T - \alpha(\omega)\Delta \varphi^T = C^*(Cy^T - z) & \text{en } G \times (0, T), \\ \varphi^T = 0 & \text{sobre } \partial G \times (0, T), \\ \varphi^T(T) = 0 & \text{en } G. \end{cases} \quad (3.28)$$

Si multiplicamos la primera ecuación por  $\varphi^T$ , y la quinta ecuación por  $y^T$ , luego integrando en espacio y en tiempo tenemos

$$\begin{cases} \int_0^T \{ \langle y_t^T, \varphi^T \rangle_X - \langle y^T, \alpha(\omega)\Delta \varphi^T \rangle_X \} dt = \int_0^T \langle \chi_{G_0} u^T, y^T \rangle_X dt, \\ \int_0^T \{ \langle -\varphi_t^T, y^T \rangle_X - \langle \alpha(\omega)\Delta \varphi^T, y^T \rangle_X \} dt = \int_0^T \langle C^*(Cy^T - z), y^T \rangle_X dt. \end{cases}$$

Así integramos por partes en tiempo la primera ecuación obtenemos

$$\begin{cases} \int_0^T \{ -\langle y^T(0), \varphi^T(0) \rangle_X - \langle y^T, \varphi_t^T \rangle_X - \langle y^T, \alpha(\omega)\Delta \varphi^T \rangle_X \} dt = \int_0^T \langle \chi_{G_0} u^T, y^T \rangle_X dt, \\ \int_0^T \{ \langle -\varphi_t^T, y^T \rangle_X - \langle \alpha(\omega)\Delta \varphi^T, y^T \rangle_X \} dt = \int_0^T \langle C^*(Cy^T - z), y^T \rangle_X dt. \end{cases}$$

Luego restando las dos ecuaciones anteriores tenemos

$$\langle \varphi^T(0), y_0 \rangle_X = \int_0^T \left\{ \langle C^*(Cy^T - z), y^T \rangle_X - \langle u^T, \chi_{G_0} \varphi^T \rangle_U \right\} dt.$$

Ahora integrando con respecto a  $\Omega$

$$\begin{aligned} \langle \varphi^T(0), y_0 \rangle_{L^2(\Omega; X)} &= \int_0^T \left\langle Cy^T - z, Cy^T - z + z \right\rangle_{L^2(\Omega; X)} dt \\ &\quad - \int_0^T \left\langle u^T, \int_{\Omega} \chi_{G_0} \varphi^T d\mu \right\rangle_U dt \\ &= \int_0^T \left\{ \|Cy^T - z\|_{L^2(\Omega; X)}^2 + \langle Cy^T - z, z \rangle_{L^2(\Omega; X)} \right\} dt \\ &\quad - \int_0^T \left\langle u^T, \int_{\Omega} \chi_{G_0} \varphi^T d\mu \right\rangle_U dt. \end{aligned}$$

Utilizando la caracterización de  $u^T$  dada por la cuarta ecuación de (3.28), obtenemos

$$\int_0^T \left\{ \|u^T\|_U^2 + \|Cy^T - z\|_{L^2(\Omega;X)}^2 \right\} dt = \langle \varphi^T(0), y_0 \rangle_{L^2(\Omega;X)} - \int_0^T \langle Cy^T - z, z \rangle_{L^2(\Omega;X)} dt. \quad (3.29)$$

Con intenciones de acotar cada término del lado derecho, observemos que por la desigualdad de Young tenemos

$$\int_0^T \langle Cy^T - z, z \rangle_{L^2(\Omega;X)} dt \leq \int_0^T \varepsilon_1 \frac{\|Cy^T - z\|_{L^2(\Omega;X)}^2}{2} dt + T \frac{\|z\|_X^2}{2\varepsilon_1}.$$

Mientras que para el primer término del lado derecho, utilizando la desigualdad de Cauchy-Schwarz, y nuevamente la desigualdad de Young junto al Lema 3.2.2 obtenemos

$$\langle \varphi^T(0), y_0 \rangle_{L^2(\Omega;X)} \leq \frac{\varepsilon_2 \zeta}{2} \left( \int_0^T \|Cy^T - z\|_{L^2(\Omega;X)}^2 dt \right) + \frac{\|y_0\|_X^2}{2\varepsilon_2}.$$

Así remplazando todo lo anterior en (3.29), juntando los términos y sumando a ambos lados  $\|y_0\|_X^2$

$$\int_0^T \left\{ \|u^T\|_U^2 + \left(1 - \frac{\varepsilon_2 \zeta}{2} - \frac{\varepsilon_1}{2}\right) \|Cy^T - z\|_{L^2(\Omega;X)}^2 \right\} dt + \|y_0\|_X^2 \leq \frac{\|y_0\|_X^2 + 2\varepsilon_2}{2\varepsilon_2} + T \frac{\|z\|_X^2}{2\varepsilon_1}.$$

Por tanto, tomando  $\varepsilon_1, \varepsilon_2 > 0$  tales que  $2 > \varepsilon_2 \zeta + \varepsilon_1$ , tenemos que existe  $\hat{K} > 0$  independiente de  $T$ , tal que

$$\int_0^T \left\{ \|u^T\|_U^2 + \|Cy^T - z\|_{L^2(\Omega;X)}^2 \right\} dt + \|y_0\|_X^2 \leq \hat{K}T. \quad (3.30)$$

Luego por los lemas 3.2.1 y 3.2.2 concluimos que

$$\|y^T(T)\|_{L^2(\Omega;X)}^2 \leq \hat{K}T, \quad \|\varphi^T(0)\|_{L(\Omega;X)}^2 \leq \hat{K}T. \quad (3.31)$$

**Paso 2:** Definamos  $m = y^T - \bar{y}$  y  $n = \varphi^T - \bar{\varphi}$ , por el Lema 3.2.3 tenemos que

$$\int_0^T \left\{ \|Cm\|_{L^2(\Omega;X)}^2 + \|u^T - \bar{u}\|_U^2 \right\} dt = \langle n(0), m(0) \rangle_{L^2(\Omega;X)} - \langle n(T), m(T) \rangle_{L^2(\Omega;X)}. \quad (3.32)$$

Observemos que  $m(0) = x_0 - \bar{x}$ , y  $n(T) = -\bar{\varphi}$ . Utilizando la desigualdad de Cauchy-Schwarz al lado derecho de (3.32)

$$\begin{aligned} \frac{1}{T} \int_0^T \left\{ \|Cm\|_{L^2(\Omega; X)}^2 + \|u^T - \bar{u}\|_U^2 \right\} dt &\leq \frac{1}{T} \left( \|n(0)\|_{L^2(\Omega; X)} \|m(0)\|_{L^2(\Omega; X)} \right. \\ &\quad \left. + \|n(T)\|_{L^2(\Omega; X)} \|m(T)\|_{L^2(\Omega; X)} \right) \\ &\leq \frac{K}{T} \left( \|n(0)\|_{L^2(\Omega; X)} + \|m(T)\|_{L^2(\Omega; X)} \right). \end{aligned}$$

Por otro lado, las estimaciones (3.31) implican que

$$\begin{aligned} \frac{K}{T} \left( \|n(0)\|_{L^2(\Omega; X)} + \|m(T)\|_{L^2(\Omega; X)} \right) &= \frac{K}{T} \left( \|\varphi^T(0) - \bar{\varphi}\|_{L^2(\Omega; X)} + \|y^T(T) - \bar{y}\|_{L^2(\Omega; X)} \right) \\ &\leq \frac{K}{T} \left( K_1 \sqrt{T} + K_2 \right) \\ &\leq \frac{\hat{K}}{\sqrt{T}}. \end{aligned}$$

Luego se tiene que

$$\frac{1}{T} \int_0^T \left\{ \|Cm\|_{L^2(\Omega; X)}^2 + \|u^T - \bar{u}\|_U^2 \right\} dt \leq \frac{\hat{K}}{\sqrt{T}}.$$

En particular, por la desigualdad de Hölder obtenemos que

$$\frac{1}{T} \int_0^T u^T dt \rightarrow \bar{u}, \quad (3.33)$$

en  $U$ . Además, desde el Lema 3.2.4 sabemos que existe  $K > 0$  tal que

$$\left\| \frac{1}{T} \int_0^T y^T - \bar{y} dt \right\|_{L^2(\Omega; X)}^2 \leq K \left( \left\| \int_0^T \frac{(u^T - \bar{u})}{T} dt \right\|_U^2 + \left\| \frac{y^T(T) - y_0}{T} \right\|_{L^2(\Omega; X)}^2 \right).$$

Se sigue de (3.33), y de (3.31) que

$$\frac{1}{T} \int_0^T y^T dt \rightarrow \bar{y} \quad \text{en } L^2(\Omega; X),$$

concluyendo la demostración.  $\square$

**Lema 3.3.1.** Sea  $m = y^T - \bar{y}$  y  $n = \varphi^T - \bar{\varphi}$ . Entonces se satisfacen las siguientes estimaciones

$$\begin{aligned} \|m(T)\|_{L^2(\Omega;X)}^2 + \|m\|_{\mathcal{X}}^2 &\leq K \left( \int_0^T \left\{ \left\| \int_{\Omega} \chi_{G_0} n d\mu \right\|_X^2 + \|Cm\|_{L^2(\Omega;X)}^2 \right\} dt + \|m(0)\|_{L^2(\Omega;X)}^2 \right), \\ \|n(0)\|_{L^2(\Omega;X)}^2 + \|n\|_{\mathcal{X}}^2 &\leq K \left( \int_0^T \left\{ \|Cm\|_{L^2(\Omega;X)}^2 + \left\| \int_{\Omega} \chi_{G_0} n d\mu \right\|_X^2 \right\} dt + \|n(T)\|_{L^2(\Omega;X)}^2 \right), \end{aligned}$$

donde  $\mathcal{X} = C([0, T]; L^2(\Omega; X))$ .

*Demostración.* Primero observemos que  $m$  y  $n$  satisfacen el siguiente sistema

$$\begin{cases} m_t - \alpha(\omega)\Delta m = -\chi_{G_0} \int_{\Omega} n d\mu & \text{en } G \times (0, T), \\ m = 0 & \text{sobre } \partial G \times (0, T), \\ m(0) = m_0 & \text{en } G, \\ -n_t - \alpha(\omega)\Delta n = C^* C m & \text{en } G \times (0, T), \\ n = 0 & \text{sobre } \partial G \times (0, T), \\ n(T) = n_T & \text{en } G. \end{cases} \quad (3.34)$$

Sabemos del Lema 3.2.1 (tomando el lado derecho igual a  $-\chi_{G_0} \int_{\Omega} n d\mu$ ) que existe  $K > 0$  tal que

$$\|m(T)\|_{L^2(\Omega;X)}^2 \leq K \left( \int_0^T \left\| \int_{\Omega} \chi_{G_0} n d\mu \right\|_X^2 dt + \|m_0\|_X^2 \right). \quad (3.35)$$

Además, desde la estimación de energía (3.10) podemos obtener para el sistema (3.34) que

$$\|m\|_{\mathcal{X}}^2 \leq K \left( \int_0^T \left\| \int_{\Omega} \chi_{G_0} n d\mu \right\|_X^2 dt + \|m_0\|_X^2 \right). \quad (3.36)$$

Así juntando (3.35) y (3.36) tenemos

$$\|m(T)\|_{L^2(\Omega;X)}^2 + \|m\|_{\mathcal{X}}^2 \leq K \left( \int_0^T \left\| \int_{\Omega} \chi_{G_0} n d\mu \right\|_X^2 dt + \|m_0\|_X^2 \right).$$

Para la segunda desigualdad basta proceder de forma análoga, considerando esta vez el Lema 3.2.2 (con lado derecho igual a  $C^* C m$ ), y la observación 3.2.3 obteniendo así

$$\|n(0)\|_{L^2(\Omega;X)}^2 + \|n\|_{\mathcal{X}}^2 \leq \hat{K} \left( \int_0^T \|Cm\|_X^2 dt + \|n_T\|_X^2 \right).$$

Concluye la demostración.  $\square$

A continuación demostraremos nuestro primer teorema principal.

**Teorema 3.3.2.** Sean los pares óptimos  $(y^T, u^T)$  y  $(\bar{y}, \bar{u})$ , asociados a (3.11) y a (3.13) respectivamente. Entonces existen dos constantes  $\mathcal{K}, \delta > 0$  independientes del horizonte de tiempo  $T$ , tales que

$$\|y^T - \bar{y}\|_{L^2(\Omega; X)} + \|\varphi^T - \bar{\varphi}\|_{L^2(\Omega; X)} + \|u^T - \bar{u}\|_U \leq \mathcal{K}(e^{-t\delta} + e^{-\delta(T-t)}),$$

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

*Demostración.* Separemos la demostración en pasos.

**Paso 1:** Sea  $m = y^T - \bar{y}$  y  $n = \varphi^T - \bar{\varphi}$ . Así se satisface el siguiente sistema

$$\begin{cases} m_t - \alpha(\omega)\Delta m = -\chi_{G_0} \int_{\Omega} n d\mu & \text{en } G \times (0, T), \\ m = 0 & \text{sobre } \partial G \times (0, T), \\ m(0) = m_0 & \text{en } G, \\ -n_t - \alpha(\omega)\Delta n = C^* C m & \text{en } G \times (0, T), \\ n = 0 & \text{sobre } \partial G \times (0, T), \\ n(T) = n_T & \text{en } G. \end{cases} \quad (3.37)$$

Sabemos desde el Lema 3.3.1 que  $m, n$  satisfacen las estimaciones

$$\begin{aligned} \|m(T)\|_{L^2(\Omega; X)}^2 + \|m\|_{\mathcal{X}}^2 &\leq K \left( \int_0^T \left\{ \left\| \int_{\Omega} \chi_{G_0} n d\mu \right\|_X^2 + \|Cm\|_{L^2(\Omega; X)}^2 \right\} dt + \|m_0\|_{L^2(\Omega; X)}^2 \right), \\ \|n(0)\|_{L^2(\Omega; X)}^2 + \|n\|_{\mathcal{X}}^2 &\leq \hat{K} \left( \int_0^T \left\{ \|Cm\|_{L^2(\Omega; X)}^2 + \left\| \int_{\Omega} \chi_{G_0} n d\mu \right\|_X^2 \right\} dt + \|n_T\|_{L^2(\Omega; X)}^2 \right), \end{aligned} \quad (3.38)$$

Por otro lado, utilizando el Lema 3.2.3 y aplicando la desigualdad de Young, tenemos para  $\varepsilon_1, \varepsilon_2 > 0$  que

$$\begin{aligned} \int_0^T \left\{ \|Cm\|_{L^2(\Omega; X)}^2 + \|u^T - \bar{u}\|_U^2 \right\} dt &\leq \frac{\varepsilon_1 \|n(0)\|_{L^2(\Omega; X)}^2}{2} + \frac{\|m_0\|_{L^2(\Omega; X)}^2}{2\varepsilon_1} \\ &\quad + \frac{\varepsilon_2 \|m(T)\|_{L^2(\Omega; X)}^2}{2} + \frac{\|n_T\|_{L^2(\Omega; X)}^2}{2\varepsilon_2} \end{aligned}$$

Luego en consecuencia de la estimación (3.38), tenemos que

$$\begin{aligned} \int_0^T \left\{ \|Cm\|_{L^2(\Omega;X)}^2 + \|u^T - \bar{u}\|_u^2 \right\} dt &\leq K \left[ \|m_0\|_{L^2(\Omega;X)}^2 \left( \frac{1}{2\varepsilon_1} + \frac{\varepsilon_2}{2} \right) \right. \\ &\quad \left. + \|n_t\|_{L^2(\Omega;X)}^2 \left( \frac{1}{2\varepsilon_2} + \frac{\varepsilon_1}{2} \right) + \left( \frac{\alpha\varepsilon_1 + \beta\varepsilon_1}{2} \right) \int_0^T \left\{ \|Cm\|_{L^2(\Omega;X)}^2 + \left\| \int_\Omega \chi_{G_0} n d\mu \right\|_X^2 \right\} dt \right]. \end{aligned} \quad (3.39)$$

Notando que

$$\int_0^T \left\| \int_\Omega \chi_{G_0} n d\mu \right\|_X^2 dt = \int_0^T \left\| \int_\Omega \chi_{G_0} \varphi^T - \chi_{G_0} \bar{\varphi} d\mu \right\|_X^2 dt = \int_0^T \|u^T - \bar{u}\|_u^2 dt, \quad (3.40)$$

podemos tomar  $\varepsilon_1, \varepsilon_2 > 0$ , de tal forma que el ultimo término del lado derecho de (3.39) sea absorbido por el lado izquierdo, obteniendo

$$\int_0^T \left\{ \|Cm\|_{L^2(\Omega;X)}^2 + \left\| \int_\Omega \chi_{G_0} n d\mu \right\|_X^2 \right\} dt \leq K \left( \|m_0\|_{L^2(\Omega;X)}^2 + \|n_T\|_{L^2(\Omega;X)}^2 \right).$$

Así utilizando nuevamente (3.38), obtenemos que existe  $\hat{K} > 0$  independiente de  $T$ , tal que

$$\|m\|_{\mathcal{X}}^2 + \|n\|_{\mathcal{X}}^2 \leq \hat{K} \left( \|n(T)\|_{L^2(\Omega;X)}^2 + \|m(0)\|_{L^2(\Omega;X)}^2 \right), \quad (3.41)$$

**Paso 2:** Observemos que si multiplicamos por  $v \in L^2(\Omega; H_0^1(G))$  la primera y cuarta ecuación de (3.37), e integrando sobre  $G \times \Omega$  tenemos

$$\begin{cases} \langle m_t, v \rangle_{L^2(\Omega;X)} + \langle \alpha \nabla m, \nabla v \rangle_{L^2(\Omega;X)} = \langle -\chi_{G_0} \int_\Omega n d\mu, v \rangle_{L^2(\Omega;X)}, \\ \langle -n_t, v \rangle_{L^2(\Omega;X)} + \langle \alpha \nabla n, \nabla v \rangle_{L^2(\Omega;X)} = \langle C^* C m, v \rangle_{L^2(\Omega;X)}. \end{cases} \quad (3.42)$$

Así se tiene de forma equivalente que (3.42) puede ser escrito

$$\begin{cases} \langle m_t, v \rangle_{L^2(\Omega;X)} + \langle \alpha \nabla m, \nabla v \rangle_{L^2(\Omega;X)} = - \langle n, \chi_{G_0} \int_\Omega v d\mu \rangle_{L^2(\Omega;X)}, \\ \langle -n_t, v \rangle_{L^2(\Omega;X)} + \langle \alpha \nabla n, \nabla v \rangle_{L^2(\Omega;X)} = \langle C^* C m, v \rangle_{L^2(\Omega;X)}, \end{cases} \quad (3.43)$$

Ahora para simplificar la notación e incorporar las condiciones iniciales a la ecuación consideremos

$$\begin{aligned} \langle D_t m, v \rangle &:= \langle m_t, v \rangle_{L^2(\Omega;X)} + \langle m(0), v \rangle_{L^2(\Omega;X)}, \\ \langle D_t^* n, v \rangle &:= - \langle n_t, v \rangle_{L^2(\Omega;X)} + \langle n(T), v \rangle_{L^2(\Omega;X)}. \end{aligned} \quad (3.44)$$

Antes de continuar observemos que por la regularidad de la solución en  $W_2^1(0, T; V; H)$  con  $V = L^2(\Omega; H_0^1(G))$  y  $H = L^2(\Omega; L^2(G))$ , tanto los productos en (3.43), como los operadores  $D_t$  y  $D_t^*$  están bien definidos. Así utilizando la notación anterior (3.44) junto a las condiciones iniciales, el sistema (3.43) puede ser escrito matricialmente como

$$\begin{pmatrix} \langle D_t \cdot, v \rangle + \langle \alpha \nabla \cdot, v \rangle_{L^2(\Omega; X)} & \langle n, \chi_{G_0} \int_{\Omega} v d\mu \rangle_{L^2(\Omega; X)} \\ - \langle C^* C \cdot, v \rangle_{L^2(\Omega; X)} & \langle D_t^* \cdot, v \rangle + \langle \alpha \nabla \cdot, \nabla v \rangle_{L^2(\Omega; X)} \end{pmatrix} \begin{pmatrix} m \\ n \end{pmatrix} = \begin{pmatrix} \langle m_0, v \rangle_{L^2(\Omega; X)} \\ \langle n_T, v \rangle_{L^2(\Omega; X)} \end{pmatrix}. \quad (3.45)$$

Definamos el operador matricial  $\Xi : (\mathcal{X})^2 \rightarrow (L^2(\Omega; X))^2$  dado por

$$\Xi = \begin{pmatrix} \langle D_t \cdot, v \rangle + \langle \alpha \nabla \cdot, \nabla v \rangle_{L^2(\Omega; X)} & \langle n, \chi_{G_0} \int_{\Omega} v d\mu \rangle_{L^2(\Omega; X)} \\ - \langle C^* C \cdot, v \rangle_{L^2(\Omega; X)} & \langle D_t^* \cdot, v \rangle + \langle \alpha \nabla \cdot, \nabla v \rangle_{L^2(\Omega; X)} \end{pmatrix},$$

Observemos que  $\Xi : (\mathcal{X})^2 \rightarrow (L^2(\Omega; X))^2$  es invertible. En efecto, primero observemos que el operador  $\Xi^{-1}$  corresponde al operador solución, que lleva las condiciones iniciales del sistema primal y el adjunto a sus respectivas soluciones. Como el sistema (3.45) posee una única solución (pues las variables que las componen, están únicamente definidas por lo desarrollado en la Sección 3.2), el operador  $\Xi$  es invertible, es decir,  $\Xi^{-1} : (L^2(\Omega; X))^2 \rightarrow (\mathcal{X})^2$  está bien definido. Además, desde (3.41) del paso anterior, obtenemos que existe  $\hat{K} > 0$  independiente de  $T$  tal que

$$\|\Xi^{-1}\|_{\mathcal{L}((L^2(\Omega; X))^2, (\mathcal{X})^2)} < \hat{K}. \quad (3.46)$$

**Paso 3:** Ahora para  $m, n$  definido como el el paso anterior, consideremos el cambio de variable

$$\hat{m} = \frac{m}{e^{-\delta(T-t)} + e^{-\delta t}}, \quad \hat{n} = \frac{n}{e^{-\delta(T-t)} + e^{-\delta t}},$$

para  $\delta > 0$  por fijar. Notemos que  $\hat{m}$  y  $\hat{n}$  satisfacen que

$$\begin{cases} \hat{m}_t - \alpha \Delta \hat{m} = -\chi_{G_0} \int_{\Omega} \hat{n} d\mu + \hat{m} \delta h(t) & \text{en } G \times (0, T), \\ -\hat{n}_t - \alpha \Delta \hat{n} = C^* C \hat{m} - \hat{n} \delta h(t) & \text{en } G \times (0, T), \\ \hat{m}(0) = \frac{m_0}{1+e^{-\delta T}}, \quad \hat{n}(T) = \frac{n_T}{1+e^{-\delta T}} & \text{en } G, \\ \hat{m} = \hat{n} = 0 & \text{sobre } \partial G \times (0, T) \end{cases} \quad (3.47)$$

donde  $h(t)$  corresponde a

$$h(t) = \frac{e^{-\delta t} - e^{-\delta(T-t)}}{e^{-\delta t} + e^{-\delta(T-t)}}.$$



Luego utilizando el operador  $\Xi$ , podemos escribir el sistema (3.47) como

$$\left[ \Xi - \delta \begin{pmatrix} h(t) & 0 \\ 0 & -h(t) \end{pmatrix} \right] \begin{pmatrix} \hat{m} \\ \hat{n} \end{pmatrix} = \frac{1}{1 + e^{-\delta T}} \begin{pmatrix} \langle m_0, v \rangle_{L^2(\Omega; X)} \\ \langle n_T, v \rangle_{L^2(\Omega; X)} \end{pmatrix}. \quad (3.48)$$

Además, denotando como  $\mathcal{J}$  el operador matricial

$$\mathcal{J} = \begin{pmatrix} h(t) & 0 \\ 0 & -h(t) \end{pmatrix},$$

se satisface que  $\|\mathcal{J}\|_{\mathcal{L}((\mathcal{X})^2)} \leq 1$ , pues  $|h(t)| \leq 1$  para todo  $t \in [0, T]$ .

Como  $\Xi$  es invertible, el sistema (3.48), se puede escribir de la forma

$$(I - \delta \Xi^{-1} \mathcal{J}) \mathcal{Z} = \Xi^{-1} \mathcal{Y},$$

donde

$$\mathcal{Z} = \begin{pmatrix} \hat{m} \\ \hat{n} \end{pmatrix}, \quad \mathcal{Y} = \frac{1}{1 + e^{-\delta T}} \begin{pmatrix} \langle m_0, v \rangle_{L^2(\Omega; X)} \\ \langle n_T, v \rangle_{L^2(\Omega; X)} \end{pmatrix}.$$

Con intenciones de despejar  $\mathcal{Z}$ , deseamos que el operador  $(I - \delta \Xi^{-1} \mathcal{P})$  sea invertible. Para ello como  $\|\mathcal{J}\|_{\mathcal{L}(0, T, L^2(\Omega; X))} \leq 1$ , basta fijar  $\delta > 0$  (independiente de  $T$ ) tal que  $\Theta := \delta \|\Xi^{-1}\|_{\mathcal{L}((L^2(\Omega; X))^2, (\mathcal{X})^2)} < 1$ . En tal caso  $(I - \delta \Xi^{-1} \mathcal{P})^{-1}$  existe, y su norma esta acotada mediante la representación de Neumann (ver [35, Teorema 2.14])

$$\|(I - \delta \Xi^{-1} \mathcal{P})^{-1}\|_{\mathcal{L}((L^2(\Omega; X))^2, (\mathcal{X})^2)} \leq \frac{1}{1 - \|\delta \Xi^{-1} \mathcal{P}\|_{\mathcal{L}((L^2(\Omega; X))^2, (\mathcal{X})^2)}} \leq \frac{1}{1 - \Theta}.$$

Luego tomando norma a  $\mathcal{Z} = (I - \delta \Xi^{-1} \mathcal{P})^{-1} \Xi^{-1} \mathcal{Y}$  obtenemos

$$\|\hat{m}\|_{\mathcal{X}} + \|\hat{n}\|_{\mathcal{X}} \leq \frac{\hat{K}}{1 - \Theta} \left( \|m_0\|_{L^2(\Omega; X)} + \|n_T\|_{L^2(\Omega; X)} \right). \quad (3.49)$$

Como el lado derecho de (3.49) es independiente de  $T$ , entonces existe  $K > 0$  independiente de  $T$ , tal que

$$\|\hat{m}\|_{\mathcal{X}} + \|\hat{n}\|_{\mathcal{X}} \leq K.$$

Volviendo a las variables originales, tenemos en particular que

$$\left\| \frac{m}{e^{-\delta(T-t)} + e^{-\delta t}} \right\|_{\mathcal{X}} + \left\| \frac{n}{e^{-\delta(T-t)} + e^{-\delta t}} \right\|_{\mathcal{X}} \leq K,$$

como  $\mathcal{X} = C([0, T]; L^2(\Omega; X))$ , obtenemos

$$\|m(t)\|_{L^2(\Omega; X)} + \|n(t)\|_{L^2(\Omega; X)} \leq K(e^{-\delta(T-t)} + e^{-\delta t}),$$

para cada  $t \in [0, T]$ . Finalmente observando que

$$\|u^T - \bar{u}\|_U^2 = \left\| \int_{\Omega} \chi_{G_0} n d\mu \right\|_X^2 \leq \left\| \int_{\Omega} n d\mu \right\|_X^2 \leq \|n\|_{L^2(\Omega; X)}^2,$$

y que

$$\|\varphi^T - \bar{\varphi}\|_{L^2(\Omega; X)}^2 = \|n\|_{L^2(\Omega; X)}^2.$$

Concluimos que existen dos constantes  $\mathcal{K}, \delta > 0$ , independientes del horizonte de tiempo tales que

$$\|y^T - \bar{y}\|_{L^2(\Omega; X)} + \|\varphi^T - \bar{\varphi}\|_{L^2(\Omega; X)} + \|u^T - \bar{u}\|_U \leq \mathcal{K}(e^{-\delta(T-t)} + e^{-\delta t}).$$

□

Así concluimos este capítulo con nuestro tercer y último teorema principal. Este teorema corresponde a una consecuencia directa del Teorema 3.3.2.

**Teorema 3.3.3.** *Sean los pares óptimos  $(y^T, u^T)$  y  $(\bar{y}, \bar{u})$ , asociados a (3.11) y a (3.13) respectivamente. Entonces existen dos constantes  $\bar{\mathcal{K}}, \delta > 0$  independientes del horizonte de tiempo  $T$ , tales que*

$$\|\mathbb{E}(y^T) - \mathbb{E}(\bar{y})\|_X + \|\mathbb{E}(\varphi^T) - \mathbb{E}(\bar{\varphi})\|_X + \|u^T - \bar{u}\|_U \leq \bar{\mathcal{K}}(e^{-t\delta} + e^{-\delta(T-t)}),$$

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

*Demostración.* Sabemos desde le teorema 3.3.2 que

$$\|y^T - \bar{y}\|_{L^2(\Omega; X)} + \|u^T - \bar{u}\|_U \leq \mathcal{K}(e^{-t\delta} + e^{-\delta(T-t)}).$$

Por otro lado

$$\begin{aligned} \|\mathbb{E}(y^T) - \mathbb{E}(\bar{y})\|_X &\leq \left\| \int_{\Omega} (y^T - \bar{y}) d\mu \right\|_X \leq \int_{\Omega} \|y^T - \bar{y}\|_X d\mu \leq \|y^T - \bar{y}\|_{L^2(\Omega; X)}^2 \\ &\leq \mathcal{K}(e^{-t\delta} + e^{-\delta(T-t)}). \end{aligned}$$

De igual forma para  $\varphi^T - \bar{\varphi}$  obtenemos

$$\left\| \mathbb{E}(\varphi^T) - \mathbb{E}(\bar{\varphi}) \right\|_X \leq \left\| \varphi^T - \bar{\varphi} \right\|_{L^2(\Omega; X)} \leq \mathcal{K}(e^{-t\delta} + e^{-\delta(T-t)}).$$

Concluyendo el resultado. □



## Capítulo 4

# Deep Learning y Propiedad de Turnpike

En este breve capítulo vamos a introducir lo básico de redes neuronales. Comentaremos las relaciones existentes entre redes neuronales residuales, con la teoría de control óptimo. Nos basaremos en los artículos [17, 13] en los cuales se prueba, dentro de otros resultados, la propiedad de turnpike en deep learning, y en un contexto general para sistemas no lineales. Además, comentaremos sobre posibles trabajos a futuro y analizaremos la propiedad de turnpike para deep learning cuando este último involucra incerteza.

### 4.1. Introducción

#### 4.1.1. Redes Neuronales Residuales y Control Óptimo

El desarrollo computacional nos ha permitido implementar diversos algoritmos, los cuales nos han ayudado a resolver interesantes problemas de diversas índoles, problemas que hace algunas décadas hubiera sido difícil abordar con las herramientas existentes. Dentro de estas herramientas se encuentran las redes neuronales, cuyas ideas fundamentales surgen en 1943 en [36]. En las siguientes décadas continuaron los avances, no obstante, la principal dificultad era que los ordenadores no tenían suficiente poder de procesamiento para manejar eficazmente el gran tiempo de ejecución requerido por las grandes redes neuronales. No fue hasta 1975, cuando se introduce el algoritmo de “back-propagation”, algoritmo que permitió resolver de forma rápida problemas de redes neuronales. Las redes neuronales hoy en día son herramientas bastantes utilizadas, principalmente para tareas de clasificación y regresión. Aunque desde una mirada más abstracta, estas pueden ser entendidas como un problema de aproximación de una función desconocida  $f(\cdot)$  a partir de los datos de muestra  $\{x_i, y_i = f(x_i)\}_{i=1}^n$ . En relación con la construcción, las aplicaciones e investigación actuales en redes neuronales, y concretamente en Deep Learning, recomendamos visitar el completo y extenso libro [18].

Por otro lado, existen diversos trabajos que relaciona la teoría del control y las redes neuronales. Antes de comentar estos, es necesario introducir el modelo de red neuronal residual (ResNet). Para ello consideremos una muestra de datos de entrenamiento  $\{x_i\}_{i=1}^{i=n} \subset \mathbb{R}^d$ , entonces una

arquitectura ResNet podría tener la forma

$$\begin{cases} x_i^{j+1} = x_i^j + \sigma(w^j x_i^j + b^j), \\ x_i^0 = x_i, \end{cases} \quad (4.1)$$

para cada  $j \in \{0, \dots, n_{\text{capas}} - 1\}$  y cada  $i \in \{1, \dots, n\}$ . Denotamos como  $x_i^j \in \mathbb{R}^d$  al  $i$ -ésimo estado o neurona de la capa  $j$ . El número  $n_{\text{capas}}$  corresponde a la profundidad de la ResNet (depth). Además, la función  $\sigma$  corresponde a una función globalmente Lipschitz continua, usualmente es considerada la función sigmoide, la función tangente hiperbólica, o la función Re-Lu. Las cantidades  $w^j \in \mathbb{R}^{d \times d}$  y  $b^j \in \mathbb{R}^d$  son parámetros de ajuste de la red neuronal, usualmente llamados peso y sesgo (weights and biases). A modo de ejemplo, arquitectura dada por (4.1) con tres capas de profundidad y dato inicial  $x_i^0 \in \mathbb{R}^d$ , es ilustrada en la siguiente figura

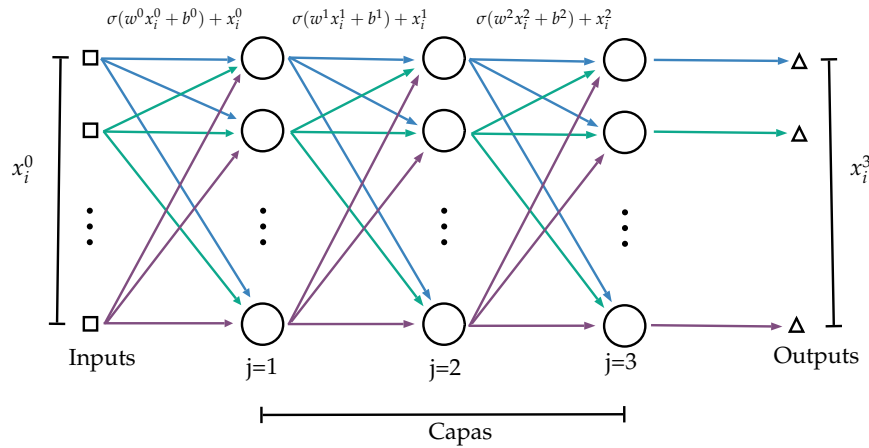


FIGURA 4.1: Arquitectura (4.1) con tres capas de profundidad.

Los parámetros  $w, b$  son “entrenados” a partir de etiquetas ya conocidas  $\{y_i\}_{i=1}^{i=n} \subset \mathbb{R}^m$  asociadas a los datos  $\{x_i\}_{i=1}^{i=n} \subset \mathbb{R}^d$ . Para ajustar los parámetros de peso y sesgo, se buscan aquellos que minimicen la distancia entre  $y_i$  y  $x_i^{n_{\text{capas}}}$ , es decir, aquellos que hacen que el output final de la red neuronal y la etiqueta esperada estén lo más cerca posible. Por tanto, para estimar los parámetros que necesita la red neuronal, usualmente se resuelve el problema

$$\min_{\{w^j, b^j\}_{j=0}^{j=n_{\text{capas}}-1}} \frac{1}{n} \sum_{i=1}^n \text{loss} \left( \varphi(x_i^{n_{\text{capas}}}), y_i \right), \quad (4.2)$$

donde  $x_i^{n_{\text{capas}}}$  corresponde a la última capa del sistema (4.1). La función  $\text{loss} : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$  llamada función de pérdida, es una función continua creciente que cuantifica la distancia entre  $y_i$  y  $\varphi(x_i^{n_{\text{capas}}})$ , mientras que  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}^m$  es una función continua, que para efectos prácticos puede ser entendida como una proyección.

Con la estructura anteriormente definida, es natural pensar en un problema de control óptimo, donde se tiene el problema de minimización (4.2) para los parámetros de control  $\{w^j, b^j\}$ , sujeto a la ecuación (4.1). Es más, la ecuación (4.1) puede ser entendida como una discretización del sistema

$$\begin{cases} \dot{x}_i(t) = \sigma(w(t)x_i(t) + b(t)) & t \in (0, T), \\ x_i(0) = x_i, \end{cases} \quad (4.3)$$

utilizando el método de Euler para  $i \in \{1, \dots, n\}$ , y  $T > 0$  dado. Se han realizado distintos avances en relación al sistema continuo (4.3), llamado a veces Ecuación Diferencial Neuronal (EDN). Algunos artículos importantes son mencionados en [17], dentro de los cuales destacamos los trabajos realizados en [43, 25], donde se relaciona el “back-propagation” con un sistema adjunto al igual que en el contexto de control óptimo. Además, el artículo [46] nos permite dar respuesta sobre la controlabilidad del sistema continuo. Por otro lado [37] realiza un análisis de estabilidad del sistema. En último lugar mencionamos el trabajo realizado en [2], donde se analiza la identificabilidad de los pesos a partir de los datos iniciales.

Como se comenta en [17], el análisis del tiempo  $T > 0$  tiene una gran importancia en sistemas continuos, pero hasta donde se sabe no existe mayor estudio en redes neuronales. En el caso de la EDN (4.3), cada paso de tiempo representa una capa de profundidad de la red, por lo cual, es natural preguntarse sobre qué ocurre con la dinámica del sistema continuo cuando el tiempo es grande, o en otras palabras, preguntarse por la red neuronal cuando es muy profunda.

En tal caso y dependiendo del tipo de penalización, podría emerger alguna propiedad de convergencia o de estabilización de la dinámica sobre algún estado estacionario. Para analizar esta hipótesis consideramos la siguiente EDN o red neuronal residual

$$\begin{cases} \dot{x}_i(t) = w(t)\sigma(x_i(t)) + b(t) & t \in (0, T), \\ x_i(0) = x_i. \end{cases} \quad (4.4)$$

En el artículo [17] se plantea el problema de aprendizaje supervisado

$$\inf_{\substack{[w, b]^T \in H^k(0, T; \mathbb{R}^{d_u}) \\ \text{Sujeto a (4.3) (o (4.4))}}} \frac{1}{n} \sum_{i=1}^n \text{loss}(\varphi(x_i(T)), \vec{y}_i) + \frac{\alpha}{2} \|[w, b]^T\|_{H^k(0, T; \mathbb{R}^{d_u})}^2 + \frac{\beta}{2} \int_0^T \|\mathbf{x}(t) - \bar{x}\|^2 dt. \quad (4.5)$$

donde  $\mathbf{x} = [x_1, \dots, x_n]$ , además,  $k = 1$  para el problema (4.3) y  $k = 0$  para el problema (4.4). Definiendo  $f(w, x, b) = \sigma(w(t)x_i(t) + b(t))$  cuando la red es (4.3), o  $f(w, b, s) = w(t)\sigma(x_i(t)) + b(t)$

cuando es (4.4), en [17] es probado la propiedad de turnpike, asociada al problema estacionario

$$\inf_{\substack{[w^s, b^s]^\top \in \mathbb{R}^{d_u} \\ \text{Sujeto a} \\ f(w, x, b) = 0}} \frac{\alpha}{2} \left\| [w^s, b^s]^\top \right\|^2 + \frac{\beta}{2} \|x^s - x_d\|^2. \quad (4.6)$$

Observemos que si  $f(0) = 0$ , para cualquier  $x \in \mathbb{R}^d$  se tiene que,  $[w^s, b^s]^\top = [0, 0]^\top$  y  $x^s = \bar{x}$  corresponden al único mínimo y trayectoria que resuelven el problema (4.6).

Para ser más precisos, al igual como en el Capítulo 2 y 3, en [17] se prueba que la trayectoria  $x^T$  asociada al mínimo de (4.5), está  $\mathcal{O}(e^{-\delta t} + e^{-\delta(T-t)})$  cerca de  $\bar{x}$  para cada  $t \in [0, T]$ , donde  $\delta$  es una constante positiva independiente de  $T$ . Además, en [17] también es probado que cuando la función  $loss \equiv 0$  y  $\bar{x} = y := [y_1, \dots, y_n]$ , la trayectoria  $x$  está  $\mathcal{O}(e^{-\delta T})$  cerca de  $\bar{x}$ , para todo  $t \in [0, T]$ . También se prueba un resultado de convergencia como el del Capítulo 2 y 3, para  $x^T(T)$  cuando  $T \rightarrow \infty$ , donde en este caso se supone que  $\beta = 0$  en (4.5) y (4.6).

Por otro lado, en este mismo artículo motivados por las arquitecturas convolucionales (ver [18]), arquitectura donde varía no solo la profundidad, sino también la cantidad de neuronas por capa es que se deduce un modelo continuo cuya discretización cae en esta arquitectura. Este modelo continuo, como es de esperar, corresponde a la siguiente Ecuación diferencial Parcial que tiene un término no local

$$\begin{cases} \partial_t u_i(t, x) = \sigma \left( \int_G w(t, x, \xi) u_i(t, \xi) d\xi + b(t, x) \right) & (t, x) \in (0, T) \times G \\ u_i(0, x) = u_i^0(x) & x \in G, \end{cases}$$

donde  $G \subset \mathbb{R}^{d_G}$  es un dominio acotado, con  $d_G \geq 1$  definido según la naturaleza de los datos (en procesamiento de imagen  $d_G = 3$ , es decir matrices de intensidades, por 3 canales de colores). Además, se conjetura que en este caso también es posible probar la propiedad de turnpike, realizando los mismos pasos desarrollados para las ResNet's (4.3) y (4.4).



## 4.2. Turnpike en Deep learning con Incerteza

### 4.2.1. Motivación

Ya con las nociones anteriormente definidas, estamos en condiciones de mostrar nuestro problema. Nuestra principal pregunta es: ¿Es posible probar la propiedad de turnpike cuando una red neuronal ResNet tiene datos de entradas aleatorio? Es decir, consideraremos un espacio de probabilidad  $(\Omega, \mathcal{F}, \mu)$  y una ResNet cuyos datos de entrada  $\{x_i(\omega), y_i\}_{i=1}^{i=n}$  dependan de un parámetro aleatorio  $\omega \in \Omega$ . Supondremos que tanto el peso  $w$  como el sesgo  $b$  son independientes del parámetro aleatorio. Además, en consecuencia de que los datos de entradas dependen del parámetro aleatorio, las siguiente capas dependerán del parámetro aleatorio. Lo anterior se puede ver representar en la siguiente arquitectura

$$\begin{cases} x_i^{j+1}(\omega) = x_i^j(\omega) + \sigma(w^j x_i^j(\omega) + b^j), \\ x_i^0 = x_i(\omega). \end{cases} \quad (4.7)$$

Llamaremos a la arquitectura anterior como red neuronal residual con incertidumbre ResNetI. Al hacer que nuestros parámetros de control  $[w^j, b^j]$  sean independientes de  $\omega$ , se podría seguir dos posibles caminos al momento de entrenar la ResNetI. El primero podría ser esperar que  $[w^j, b^j]$  pueda llevar cada  $x_i(\omega)$  a cada etiqueta asociada, noción cercana a la del control simultáneo (ver [29]). El otro camino es pensar que en promedio los parámetros  $[w^j, b^j]$ , puedan llevar los datos  $x_i(\omega)$  a una etiqueta en particular  $y_i$ , lo cual es más cercano al control en promedio (Capítulo 1).

En lo que sigue nos enfocaremos en este último camino, el cual no solo es el más adecuado en relación con los trabajos realizados anteriormente en este trabajo, sino que también creemos que este tipo de arquitectura, pensando los datos  $\{x_i(\omega)\}_{i=1}^{i=n}$  aleatorios, posiblemente cumpliendo alguna relación para cada  $i$ , reduce la posibilidad del overfitting o sobre ajuste. Esto pues, no se espera que el aprendizaje supervisado final pueda predecir en su totalidad los datos de entrenamiento, solo el promedio de ellos, reduciendo así el sobre ajuste.

**Observación 4.2.1.** Si bien en esta construcción informal estamos considerando un espacio de probabilidad, donde los datos de entrada  $x_i(\omega)$  posiblemente son entendidos como variables aleatorias, también pueden ser vistos para cada  $i \in \{1, \dots, n\}$ , como una familia de condiciones iniciales. Esto ocurre cuando a priori ya se tiene una segmentación de los datos. Lo anterior puede ser, por ejemplo, cuando se tienen datos de un país, en tal caso la geografía prevé de una segmentación natural.

Otro ejemplo podría ser cuando se tiene una baja cantidad de datos de entrada (inputs), en este caso se podría estar tentado en generar datos ficticios a partir de nuestros datos reales, posiblemente perturbando de forma aleatoria cada uno de los datos de entrada  $x_i$ , así obteniendo una familia de datos  $x_i(\omega)$ . En tal caso, podría ser natural entrenar con respecto al promedio de los datos  $x_i(\omega)$ .

Por tanto, tiene sentido querer entrenar nuestra red en promedio, no obstante, la eficiencia y la eficacia de lo anteriormente descrito tiene que ser estudiada.

### 4.2.2. Planteo del Problema

Consideremos el espacio de probabilidad  $(\Omega, \mathcal{F}, \mu)$  y consideraremos los datos de entrenamiento  $\{x_i^0(\omega), y_i\}_{i=1}^{i=n} \subset L^2(\Omega; \mathbb{R}^d) \times \mathbb{R}^m$ . Consideramos modelo continuo de (4.7), es decir

$$\begin{cases} \dot{x}_i(t, \omega) = f(w(t), b(t), x_i(t, \omega)) & t \in (0, T), \\ x_i(0) = x_i^0(\omega), \end{cases} \quad (4.8)$$

donde  $w(t) \in \mathbb{R}^{d \times d}$  y  $b(t) \in \mathbb{R}^d$  para cada  $t \in [0, T]$ . Supondremos  $f \in Lip_0(\mathbb{R}^d)$ <sup>1</sup> es una función homogénea de grado 1 con respecto a  $w, b$ <sup>2</sup>. Además, como  $x_i^0 \in L^2(\Omega; \mathbb{R}^d)$  tenemos que  $x_i^0$  corresponde a un vector aleatorio que es cuadrado integrable para cada  $i$ .

Sea  $d_x := d \times n$  y definamos

$$\mathbf{w} := \begin{bmatrix} w & & \\ & \ddots & \\ & & w \end{bmatrix} \in \mathbb{R}^{d_x \times d_x}, \quad \mathbf{b} := \begin{bmatrix} b \\ \vdots \\ b \end{bmatrix} \in \mathbb{R}^{d_x}. \quad (4.9)$$

Así el sistema (4.8) puede ser escrito como

$$\begin{cases} \dot{\mathbf{x}}_t(t, \omega) = f(\mathbf{w}(t), \mathbf{b}(t), \mathbf{x}(t, \omega)) & t \in (0, T), \\ \mathbf{x}(0) = \mathbf{x}_0(\omega), \end{cases} \quad (4.10)$$

donde  $\mathbf{x}_0 = [x_1^0(\omega), \dots, x_n^0(\omega)]^\top \in \mathbb{R}^{d_x}$  y  $\mathbf{x}(t, \omega) = [x_1(t, \omega), \dots, x_n(t, \omega)]^\top \in \mathbb{R}^{d_x}$  para cada  $t \in [0, T]$  y  $\omega \in \Omega$ . Además, por la equivalencia de las normas matriciales, es claro que  $\|\mathbf{w}\|_{\mathbb{R}^{d_x \times d_x}} = \|w\|_{\mathbb{R}^{d \times d}}$  y  $\|\mathbf{b}\|_{\mathbb{R}^{d_x}} = \|b\|_{\mathbb{R}^d}$ .

En este contexto proponemos analizar los siguientes dos casos ya analizados en [17], cuando no se tiene incerteza. Estos son

1. Cuando

$$f(\mathbf{w}(t), \mathbf{b}(t), \mathbf{x}(t, \omega)) = \sigma(\mathbf{w}(t)\mathbf{x}(t, \omega) + \mathbf{b}(t)), \quad (4.11)$$

donde  $\sigma \in Lip_0(\mathbb{R}^{d_x})$  es una función homogénea de grado 1.

2. Cuando

$$f(\mathbf{w}(t), \mathbf{b}(t), \mathbf{x}(t, \omega)) = \mathbf{w}(t)\sigma(\mathbf{x}(t, \omega)) + \mathbf{b}(t), \quad (4.12)$$

donde  $\sigma \in Lip_0(\mathbb{R}^{d_x})$ .

<sup>1</sup>Denotaremos como  $Lip_0(\mathbb{R}^d)$  a las funciones Lipschitz continuas  $f$  tales que  $f(0) = 0$ .

<sup>2</sup>Es decir, que  $f$  satisface que  $f(x, \lambda w, \lambda b) = \lambda f(x, w, b)$

En lo que sigue siempre que no haya confusiones, denotaremos solo como  $w$  a  $\mathbf{w}$ , como  $b$  a  $\mathbf{b}$  y como  $x$  a  $\mathbf{x}$ . Antes de continuar tenemos el siguiente resultado.

**Proposición 4.2.1.** *La solución  $\mathbf{x}(\cdot, \omega) \in C([0, T]; \mathbb{R}^{d_x})$  de (4.10) para  $f$  dado por (4.11) o por (4.12), satisface que  $\mathbf{x}(t, \cdot) \in L^2(\Omega; \mathbb{R}^{d_x})$  para cada  $t \in [0, T]$  y cada  $\mathbf{x}_0 \in L^2(\Omega; \mathbb{R}^{d_x})$ .*

*Demostración.* Procedemos integrando (4.10), así obtenemos

$$x(t, \omega) = x_0(\omega) + \int_0^t f(w(s), b(s), x(s, \omega)) ds. \quad (4.13)$$

Observemos que tanto para (4.11) como para (4.12), como  $f \in Lip_0(\mathbb{R}^{d_x})$  si tomamos norma y elevando al cuadrado la igualdad anterior obtenemos que

$$\|x(t, \omega)\|_{\mathbb{R}^{d_x}}^2 \leq \|x_0(\omega)\|_{\mathbb{R}^{d_x}}^2 + C \left( \int_0^t \|w(s)\|_{\mathbb{R}^{d \times d}} \|x(s, \omega)\|_{\mathbb{R}^{d_x}} ds \right)^2 + \int_0^t \|b(s)\|_{\mathbb{R}^d}^2 ds.$$

Lo anterior gracias a que  $f \in Lip_0(\mathbb{R}^{d_x})$ . Ahora aplicando la desigualdad de Hölder tenemos

$$\begin{aligned} \|x(t, \omega)\|_{\mathbb{R}^{d_x}}^2 &\leq \|x_0(\omega)\|_{\mathbb{R}^{d_x}}^2 + C \|w\|_{L^2(0, t; \mathbb{R}^{d \times d})}^2 \int_0^t \|x(s, \omega)\|_{\mathbb{R}^{d_x}}^2 ds + \|b\|_{L^2(0, t; \mathbb{R}^d)}^2 \\ &\leq \|x_0(\omega)\|_{\mathbb{R}^{d_x}}^2 + C \|w\|_{L^2(0, T; \mathbb{R}^{d \times d})}^2 \int_0^t \|x(s, \omega)\|_{\mathbb{R}^{d_x}}^2 ds + \|b\|_{L^2(0, T; \mathbb{R}^d)}^2. \end{aligned}$$

Aplicando el Lema de Gronwall tenemos que

$$\|x(t, \omega)\|_{\mathbb{R}^{d_x}}^2 \leq (\|x_0(\omega)\|_{\mathbb{R}^{d_x}}^2 + \|b\|_{L^2(0, T; \mathbb{R}^d)}^2) \exp(C \|w\|_{L^2(0, T; \mathbb{R}^{d \times d})}^2).$$

Finalmente integrando en  $\Omega$  obtenemos

$$\|x(t)\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 \leq (\|x_0\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 + \|b\|_{L^2(0, T; \mathbb{R}^d)}^2) \exp(C \|w\|_{L^2(0, T; \mathbb{R}^{d \times d})}^2), \quad (4.14)$$

concluyendo la demostración.  $\square$

Además, tenemos el siguiente teorema en relación a la solución de (4.10).

**Teorema 4.2.1.** *Dado  $x_0 \in L^2(\Omega; \mathbb{R}^{d_x})$ , existe una única solución de (4.10) en  $C([0, T]; L^2(\Omega; \mathbb{R}^{d_x}))$ . Además, satisface que*

$$\|x\|_{C([0, T]; L^2(\Omega; \mathbb{R}^{d_x}))}^2 \leq (\|x_0\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 + \|b\|_{L^2(0, T; \mathbb{R}^d)}^2) \exp(C \|w\|_{L^2(0, T; \mathbb{R}^{d \times d})}^2), \quad (4.15)$$

para  $C > 0$  independiente de  $T$ .

*Demostración.* Sabemos que desde [47, Apendice C.3] que para cada  $\omega$  existe una única solución  $x(\omega)$  de (4.10) en  $C([0, T]; \mathbb{R}^{d_x})$ . Además, desde la Proposición 4.2.1 tenemos que esta solución

$x(t, \cdot)$  pertenece a  $L^2(\Omega; \mathbb{R}^{d_x})$  para cada  $t \in [0, T]$ . Así procedemos a probar que esta solución  $x$  pertenece a  $C([0, T]; L^2(\Omega; \mathbb{R}^{d_x}))$ .

Para ello consideremos  $0 \leq t_2 < t_1 \leq T$ , así utilizando (4.13) tenemos

$$\begin{aligned} \|x(t_1) - x(t_2)\|_{L^2(\Omega; \mathbb{R}^{d_x})} &= \left\| x_0(\omega) + \int_0^{t_1} f(w(s), b(s), x(s, \omega)) ds \right. \\ &\quad \left. - x_0(\omega) - \int_0^{t_2} f(w(s), b(s), x(s, \omega)) ds \right\|_{L^2(\Omega; \mathbb{R}^{d_x})} \\ &= \left\| \int_{t_2}^{t_1} f(w(s), b(s), x(s, \omega)) ds \right\|_{L^2(\Omega; \mathbb{R}^{d_x})}. \end{aligned}$$

Ahora como  $f \in Lip(\mathbb{R}^{d_x})$  y  $f(0) = 0$ , tenemos que  $\|f(x)\| \leq C \|x\|$ . Entonces cuando  $f$  es dado por (4.11) o (4.12), tenemos que

$$\begin{aligned} \|x(t_1) - x(t_2)\|_{L^2(\Omega; \mathbb{R}^{d_x})} &\leq \int_{t_2}^{t_1} (\|x(s)\|_{L^2(\Omega; \mathbb{R}^{d_x})} \|w(s)\|_{\mathbb{R}^{d \times d}} + \|b(s)\|_{\mathbb{R}^d}) ds \\ &= |t_1 - t_2| \left( \frac{1}{|t_1 - t_2|} \int_{t_2}^{t_1} g(s) ds \right), \end{aligned}$$

con  $g(s) = \|x(s)\|_{L^2(\Omega; \mathbb{R}^{d_x})} \|w(s)\|_{\mathbb{R}^{d \times d}} + \|b(s)\|_{\mathbb{R}^d}$ . Como  $g$  es integrable en  $[0, T]$ , se sigue del Teorema de Diferenciación de Lebesgue ([14, Apéndice E]) que

$$|t_1 - t_2| \left( \frac{1}{|t_1 - t_2|} \int_{t_2}^{t_1} g(s) ds \right) \rightarrow 0 \quad \text{cuando } t_1 \rightarrow t_2.$$

Por tanto se sigue la continuidad. Ahora para la unicidad consideramos  $x_1, x_2 \in C([0, T]; L^2(\Omega; \mathbb{R}^{d_x}))$  dos soluciones de (4.10), y observemos que desde (4.13) tanto para  $f$  igual a (4.11) como (4.12) tenemos

$$\begin{aligned} \|x_1 - x_2\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 &\leq \left\| \int_0^T f(w(s), b(s), x_2(t, \omega)) - f(w(s), b(s), x_1(t, \omega)) ds \right\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 \\ &\leq \left( \int_0^T \|f(w(s), b(s), x_2(t, \omega)) - f(w(s), b(s), x_1(t, \omega))\|_{L^2(\Omega; \mathbb{R}^{d_x})} ds \right)^2 \\ &\leq \left( \int_0^T \|w(s)\|_{\mathbb{R}^{d \times d}} \|x_1(s) - x_2(s)\|_{L^2(\Omega; \mathbb{R}^{d_x})} ds \right)^2 \\ &\leq \|w\|_{L^2(0, T; \mathbb{R}^{d \times d})}^2 \int_0^T \|x_1(s) - x_2(s)\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 ds. \end{aligned}$$

Así utilizando el Lema de Gronwall y tomando supremo en tiempo concluimos que  $x_1 = x_2$  en  $C([0, T]; L^2(\Omega; \mathbb{R}^{d_x}))$ . Finalmente para probar (4.15), basta replicar los pasos de la Proposición 4.2.1, y tomar supremos en tiempo en (4.14).  $\square$

Ahora como bien fue señalado en la Sección 4.1.1, para obtener los parámetros de control consideremos un problema de minimización. Así consideramos el siguiente funcional a minimizar

$$\begin{aligned} J^T(w, b) &= \frac{1}{2} \left( \left\| [w, b]^\top \right\|_{H^k(0, T; \mathbb{R}^{d_u})}^2 dt + \int_0^T \left\| \mathbb{E}(x(t)) - \bar{x} \right\|_{\mathbb{R}^{d_x}}^2 dt + \left\| \varphi \left( \mathbb{E}(x(T, \omega)) \right) - y \right\|_{\mathbb{R}^{m \times n}}^2 \right) \\ &= \frac{1}{2} \left( \left\| [w, b]^\top \right\|_{H^k(0, T; \mathbb{R}^{d_u})}^2 dt + \int_0^T \left\| \int_{\Omega} x(t, \omega) d\mu - \bar{x} \right\|_{\mathbb{R}^{d_x}}^2 dt + \left\| \varphi \left( \int_{\Omega} x(T, \omega) d\mu \right) - y \right\|_{\mathbb{R}^{m \times n}}^2 \right), \end{aligned} \quad (4.16)$$

donde  $\mathbb{E}(x) = \int_{\Omega} x(\omega) d\mu$  denota la esperanza matemática,  $\bar{x} \in \mathbb{R}^{d_x}$  es una trayectoria fija,  $y := [y_1, \dots, y_n]^\top \in \mathbb{R}^{m \times n}$  y  $d_u = d \times d + d$ . Además  $\varphi \in C^0(\mathbb{R}^{d_x}, \mathbb{R}^{m \times n})$  es una función continua. Esta función usualmente es considerada como

$$\begin{aligned} \varphi(x) &:= \text{softmax}(\theta_1 x + \theta_2) && \text{Para clasificación,} \\ \varphi(x) &:= \theta_1 x + \theta_2 && \text{Para regresión.} \end{aligned}$$

Además consideramos el problema estacionario

$$\begin{aligned} \inf_{\substack{[w^s, b^s]^\top \in \mathbb{R}^{d_u} \\ \text{Sujeto a} \\ f(w, x, b) = 0}} \frac{\alpha}{2} \left\| [w^s, b^s]^\top \right\|^2 + \frac{\beta}{2} \|x^s - \bar{x}\|^2. \end{aligned} \quad (4.17)$$

En donde como es mencionado anteriormente, observando que si  $f(0) = 0$ , para cualquier  $x \in \mathbb{R}^d$  se tiene que,  $[w^s, b^s]^\top = [0, 0]^\top$  y  $x^s = \bar{x}$  corresponden al único mínimo y trayectoria que resuelven el problema (4.17).

**Observación 4.2.2.** Desde aquí en adelante supondremos que si  $f$  es dado por (4.11) entonces  $k = 1$ , y si  $f$  es dado por (4.12) entonces  $k = 0$ .

En relación la existencia de mínimos del problema  $\min \{J^T(w, b)\}$ , tenemos el siguiente teorema cuya demostración se basa en el método directo del calculo variacional.

**Proposición 4.2.2.** Existe un único mínimo para el problema

$$\min_{\substack{[w, b]^\top \in H^k(0, T; \mathbb{R}^{d_u}) \\ \text{Sujeto a (4.10)}}} \left\{ J^T(w, b) \right\}. \quad (4.18)$$

*Demostración.* Denotemos como

$$\min_{\substack{[w, b]^\top \in H^k(0, T; \mathbb{R}^{d_u}) \\ \text{Sujeto a (4.10)}}} \left\{ J^T(w, b) \right\} = P.$$

Así consideremos una sucesión minimizante, es decir, una sucesión  $\{[w_n, b_n]^\top\}_{n \in \mathbb{N}} \subset H^k(0, T; \mathbb{R}^{d_u})$  tal que

$$J^T((w_n, b_n)) \rightarrow P.$$

En particular observemos que

$$\left\| [w_n, b_n]^\top \right\|_{H^k(0, T; \mathbb{R}^{d_u})}^2 \leq J^T([0, 0]) \leq C(x_0, y). \quad (4.19)$$

Ahora separemos en dos casos

**Caso 1:**  $f$  es dado como en (4.11).

En tal caso  $k = 1$ , así tenemos que existen  $[w^*, b^*]^\top \in H^1(0, T; \mathbb{R}^{d_u})$  tales que satisfacen (4.19). En consecuencia existe una subsucesión (denotada de la misma forma)  $\{[w_n, b_n]^\top\} \subset H^1(0, T; \mathbb{R}^{d_u})$  tal que

$$[w_n, b_n]^\top \rightharpoonup [w^*, b^*]^\top \quad \text{en } H^1(0, T; \mathbb{R}^{d_u}).$$

Luego en utilizando del Teorema de Rellich-Kondrachov se satisface

$$[w_n, b_n]^\top \rightarrow [w^*, b^*]^\top \quad \text{en } L^2(0, T; \mathbb{R}^{d_u}). \quad (4.20)$$

Ahora probemos que  $x^* \in C([0, T]; L^2(\Omega; \mathbb{R}^{d_x}))$  la solución de (4.10) asociada  $[w^*, b^*]^\top$ , verifica que

$$\begin{aligned} x_n &\rightarrow x^* \quad \text{en } C([0, T]; L^2(\Omega; \mathbb{R}^{d_x})), \\ \mathbb{E}(x_n) &\rightarrow \mathbb{E}(x^*) \quad \text{en } C([0, T]; \mathbb{R}^{d_x}), \end{aligned} \quad (4.21)$$

donde  $x_n \in C([0, T]; L^2(\Omega; \mathbb{R}^{d_x}))$  es la solución de (4.10) asociada  $[w_n, b_n]^\top$  para cada  $n \in \mathbb{N}$ . Primero observemos que

$$\begin{aligned} \|x_n(t, \omega) - x^*(t, \omega)\|_{\mathbb{R}^{d_x}} &\leq \int_0^t \|\sigma(w_n(s)x_n(s, \omega) + b_n(s)) - \sigma(w^*(s)x^*(s, \omega) + b(s))\|_{\mathbb{R}^{d_x}} ds \\ &\leq \int_0^t \|(w_n(s)x_n(s, \omega) + b_n(s)) + w_n(s)x^*(s, \omega) \\ &\quad - w_n(s)x^*(s, \omega) - (w^*(s)x^*(s, \omega) + b(s))\|_{\mathbb{R}^{d_x}} ds. \end{aligned}$$

Agrupando términos, elevando al cuadrado y utilizando la desigualdad triangular

$$\begin{aligned} \|x_n(t, \omega) - x^*(t, \omega)\|_{\mathbb{R}^{d_x}} &\leq \int_0^t \|(w_n(s)x_n(s, \omega) - w_n(s)x^*(s, \omega))\|_{\mathbb{R}^{d_x}} \\ &\quad + \|w_n(s)x^*(s, \omega) - w^*(s)x^*(s, \omega)\|_{\mathbb{R}^{d_x}} + \|b_n(s) - b^*(s)\|_{\mathbb{R}^{d_x}} ds \\ &= \int_0^t \|(w_n(s)[x_n(s, \omega) - x^*(s, \omega)]\|_{\mathbb{R}^{d_x}} \\ &\quad + \|[w_n(s) - w^*(s)]x^*(s, \omega)\|_{\mathbb{R}^d} + \|b_n(s) - b^*(s)\|_{\mathbb{R}^d} ds. \end{aligned}$$

Utilizando la desigualdad de Hölder, el hecho de que  $x^*(\omega) \in C([0, T]; \mathbb{R}^n)$  y que  $\{w_n\}_{n \in \mathbb{N}}$  es una sucesión acotada, tenemos

$$\begin{aligned} \|x_n(t, \omega) - x^*(t, \omega)\|_{\mathbb{R}^{d_x}}^2 &\leq C \int_0^t \|x_n(s, \omega) - x^*(s, \omega)\|_{\mathbb{R}^{d_x}}^2 ds + C \int_0^t \|w_n(s) - w^*(s)\|_{\mathbb{R}^{d \times d}}^2 ds \\ &\quad + C \int_0^t \|b_n(s) - b^*(s)\|_{\mathbb{R}^d}^2 ds \\ &\leq C \int_0^t \|x_n(s, \omega) - x^*(s, \omega)\|_{\mathbb{R}^{d_x}}^2 ds + C \|w_n(s) - w^*(s)\|_{L^2(0, T; \mathbb{R}^{d \times d})}^2 \\ &\quad + C \|b_n(s) - b^*(s)\|_{L^2(0, T; \mathbb{R}^d)}^2. \end{aligned}$$

Ahora aplicando el lema de Gronwall obtenemos

$$\|x_n(t, \omega) - x^*(t, \omega)\|^2 \leq C(\|w_n(s) - w^*(s)\|_{L^2(0, T; \mathbb{R}^{d \times d})}^2 + \|b_n(s) - b^*(s)\|_{L^2(0, T; \mathbb{R}^d)}^2)e^C. \quad (4.22)$$

Así integrando en  $\Omega$ , tomando supremo en tiempo y utilizando (4.20) se verifica que

$$\|x_n - x^*\|_{C(0, T; L^2(\Omega; \mathbb{R}^{d_x}))} \xrightarrow{n \rightarrow \infty} 0.$$

Finalmente observemos que si integramos sobre  $\Omega$  la estimación (4.22) tenemos

$$\|x_n(t) - x^*(t)\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 \leq C(\|w_n(s) - w^*(s)\|_{L^2(0, T; \mathbb{R}^{d \times d})}^2 + \|b_n(s) - b^*(s)\|_{L^2(0, T; \mathbb{R}^d)}^2). \quad (4.23)$$

Además, como

$$\|\mathbb{E}(x_n(t)) - \mathbb{E}(x^*(t))\|_{\mathbb{R}^{d_x}}^2 \leq \|x_n(t) - x^*(t)\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2. \quad (4.24)$$

Juntando (4.23) y (4.23) obtenemos que

$$\|\mathbb{E}(x_n(t)) - \mathbb{E}(x^*(t))\|_{\mathbb{R}^{d_x}}^2 \leq C(\|w_n(s) - w^*(s)\|_{L^2(0, T; \mathbb{R}^{d \times d})}^2 + \|b_n(s) - b^*(s)\|_{L^2(0, T; \mathbb{R}^d)}^2).$$

Por tanto, tomando supremo concluimos que  $\mathbb{E}(x_n) \rightarrow \mathbb{E}(x^*)$  en  $C(0, T; \mathbb{R}^n)$ .

**Caso 2:** Si  $f$  es dado como (4.12). En tal caso  $k = 0$ , así la sucesión  $\{[w_n, b_n]^\top\}_{n \in \mathbb{N}} \subset L^2(0, T; \mathbb{R}^{d_u})$  es acotada, por tanto tenemos

$$[w_n, b_n]^\top \rightharpoonup [w^*, b^*]^\top \quad \text{en } L^2(0, T; \mathbb{R}^{d_u}). \quad (4.25)$$

Probemos como el caso anterior que se verifica (4.21). Procediendo de forma análoga, por el hecho que  $\sigma \in Lip_0(\mathbb{R}^{d_x})$  obtenemos

$$\begin{aligned} \|x_n(t, \omega) - x^*(t, \omega)\|_{\mathbb{R}^{d_x}} &= \left\| \int_0^t (w_n(s)\sigma(x_n(s, \omega)) + b_n(s)) - (w^*(s)\sigma(x^*(s, \omega)) + b^*(s))ds \right\|_{\mathbb{R}^{d_x}} \\ &= \left\| \int_0^t (w_n(s)\sigma(x_n(s, \omega)) + b_n(s)) + w_n(s)\sigma(x^*(s, \omega)) \right. \\ &\quad \left. - w_n(s)\sigma(x^*(s, \omega)) - (w^*(s)\sigma(x^*(s, \omega)) + b^*(s))ds \right\|_{\mathbb{R}^{d_x}} \\ &= \left\| \int_0^t w_n(s)[\sigma(x_n(s, \omega)) - \sigma(x^*(s, \omega))] + [w_n(s) - w^*(s)]\sigma(x^*(s, \omega)) \right. \\ &\quad \left. + b_n(s) - b^*(s)ds \right\|_{\mathbb{R}^{d_x}} \\ &\leq \int_0^t \|w_n(s)\|_{\mathbb{R}^{d \times d}} \|\sigma(x_n(s, \omega)) - \sigma(x^*(s, \omega))\|_{\mathbb{R}^{d_x}} ds \\ &\quad + \left\| \int_0^t [w_n(s) - w^*(s)]\sigma(x^*(s, \omega))ds \right\|_{\mathbb{R}^{d_x}} + \left\| \int_0^t b_n(s) - b^*(s)ds \right\|_{\mathbb{R}^d} \\ &\leq C \int_0^t \|w_n(s)\|_{\mathbb{R}^{d \times d}} \|x_n(s, \omega) - x^*(s, \omega)\|_{\mathbb{R}^{d_x}} ds \\ &\quad + \left\| \int_0^t \sigma(x^*(s, \omega))[w_n(s) - w^*(s)]ds \right\|_{\mathbb{R}^{d_x}} + \left\| \int_0^t b_n(s) - b^*(s)ds \right\|_{\mathbb{R}^d}. \end{aligned}$$

Ahora por la desigualdad de Hölder y el hecho que la sucesión  $\{w_n\}_{n \in \mathbb{N}} \subset H^1(0, T; \mathbb{R}^{d \times d})$  es acotada, tomado cuadrado tenemos que

$$\begin{aligned} \|x_n(t, \omega) - x^*(t, \omega)\|^2 &\leq C \int_0^t \|x_n(s, \omega) - x^*(s, \omega)\|^2 ds \\ &\quad + \left\| \int_0^t \sigma(x^*(s, \omega))[w_n(s) - w^*(s)]ds \right\|^2 + \left\| \int_0^t b_n(s) - b^*(s)ds \right\|^2. \end{aligned}$$



Luego utilizando el lema de Gronwall tenemos que

$$\begin{aligned} \|x_n(t, \omega) - x^*(t, \omega)\|^2 \leq C \left( \left\| \int_0^T \chi_{[0,t]} \sigma(x^*(s, \omega)) [w_n(s) - w^*(s)] ds \right\|^2 \right. \\ \left. + \left\| \int_0^T \chi_{[0,t]} (b_n(s) - b^*(s)) ds \right\|^2 \right) e^C, \end{aligned}$$

donde  $\chi_{[0,t]}$  denota la función característica en el intervalo  $[0, t]$ . Así gracias a (4.25) podemos concluimos de forma análoga al caso anterior.

Finalmente utilizando la semicontinuidad inferior del funcional  $J^T$  tenemos que

$$J^T(w^*, b^*) \leq \liminf_{n \rightarrow \infty} J^T(w_n, b_n) = P.$$

Se sigue que  $[w^*, b^*]^\top$  es mínimo de (4.18). Además, se concluye de forma directa la unicidad desde la convexidad estricta del funcional  $J^T$ .  $\square$

### 4.2.3. Continuación del Problema

Ya con lo anterior sabemos que nuestro problema está bien puesto, es decir, (4.18) tiene un único mínimo, y una única trayectoria asociada a este mínimo. En lo que sigue probaremos una serie de lemas que posiblemente nos permitirán concluir la Propiedad de Turnpike, no obstante, la prueba final de este resultado será dejada para trabajos a futuro.

Antes que todo, consideremos la siguiente definición.

**Definición 4.2.1.** Diremos que el sistema (4.10) es controlable en promedio en tiempo  $T > 0$ , si para cada  $x_0 \in L^2(\Omega; \mathbb{R}^{d_x})$  y  $x_1 \in \mathbb{R}^{d_x}$  existe  $u := [w, b]^\top \in H^k(0, T; \mathbb{R}^{d_u})$  tal que  $x$  la solución de (4.10) satisface que

$$\int_{\Omega} x(T, \omega) d\mu = x_1.$$

Además, consideremos la siguiente Hipótesis.

**Hipótesis 1.** Supondremos que (4.10) es controlable para algún tiempo  $T_0$ , es decir, para  $T_0 > 0$  y para cualquier  $x_0 \in L^2(\Omega; \mathbb{R}^n)$  y  $x_1 \in \mathbb{R}^{d_x}$ , existe un control  $u \in H^k(0, T_0; \mathbb{R}^{d_u})$  tal que la correspondiente solución  $x \in C([0, T_0]; L^2(\Omega; \mathbb{R}^{d_x}))$  de (4.10) con  $x(0) = x_0$  satisface que  $\mathbb{E}(x(T_0)) = x_1$ . Además, supondremos que existe  $r > 0$  y  $C > 0$  tal que

$$\inf_{\substack{u \\ \text{tal que} \\ x(0)=x_0, \mathbb{E}(x(T_0))=\bar{x}}} \|u\|_{H^k(0, T_0; \mathbb{R}^{d_u})} \leq C(T_0) \|\mathbb{E}(x_0) - \bar{x}\|_{\mathbb{R}^{d_x}},$$

y

$$\inf_{\substack{u \\ \text{tal que} \\ x(0)=x_3, \quad \mathbb{E}(x(T_0))=\mathbb{E}(x_1)}} \|u\|_{H^k(0,T_0;\mathbb{R}^{d_u})} \leq C(T_0) \|\mathbb{E}(x_1) - \bar{x}\|_{\mathbb{R}^{d_x}}.$$

Lo anterior para cualquier  $x_0, x_1 \in \{x \in L^2(\Omega; \mathbb{R}^{d_x}) : \|\mathbb{E}(x) - \bar{x}\|_{\mathbb{R}^{d_x}} < r\}$  y  $x_3 \in L^2(\Omega; \mathbb{R}^{d_x})$  tal que  $\mathbb{E}(x_3) = \bar{x}$ .

**Observación 4.2.3.** 1. La hipótesis anterior requiere más desarrollo y análisis, en relación a que tan restrictivo es pedir que el sistema (4.10) sea controlable en promedio. En otras palabras, al igual como en la sección 1.5.4, señalamos que es necesario analizar la controlabilidad en promedio de sistemas no lineales con incertidumbre, donde posiblemente [46], como lo realizado en el Capítulo 1 sean referencias y/o puntos de inicio a considerar.

2. La Hipótesis 1 será utilizada parcialmente en los siguientes lemas. Solo en el Lema 4.32 se utilizará en su totalidad.

Ya con lo anterior proponemos el teorema principal de esta sección como sigue.

**Teorema 4.2.2.** Supongamos que se satisface la **Hipótesis 1**. Entonces existe  $T^* > 0$  y constantes  $C_1, C_2, \delta > 0$  tales que para todo  $T \geq T^*$ , el mínimo  $u^T = [w, b]^T$  de (4.16), y su trayectoria asociada  $x^T$  con  $x^T(0) = x_0 \in L^2(\Omega; \mathbb{R}^{d_x})$  satisface que

$$\|\mathbb{E}(x^T(t)) - \bar{x}\|_{\mathbb{R}^{d_x}} \leq C_1 \left( e^{-\delta t} + e^{-(T-t)\delta} \right),$$

para cada  $t \in [0, T]$ . Además, se tiene que  $\|u^T\|_{L^2(0,T;\mathbb{R}^{d_u})} \leq C_2$ .

Con los siguientes lemas, demarcaremos el camino para probar el Teorema 4.2.2. Para ello utilizaremos las herramientas desarrolladas tanto en [13] como en [17]. A continuación, nuestro primero lema.

**Lema 4.2.1.** Sea  $x(\omega) \in C([0, T]; \mathbb{R}^n)$  la única solución de (4.10). Entonces existe  $C > 0$  tal que

$$\left\| \mathbb{E} \left( x^T(t) \right) - \bar{x} \right\|_{\mathbb{R}^{d_x}} \leq C \left( \|\mathbb{E}(x_0) - \bar{x}\|_{\mathbb{R}^{d_x}} + \|u\|_{H^k(0,T;\mathbb{R}^{d_u})} + \|\mathbb{E}(x) - \bar{x}\|_{L^2(0,T;\mathbb{R}^{d_x})} \right).$$

para cada  $t \in [0, T]$ .

*Demostración.* Primero observemos que si integramos entre  $(0, t)$  la ecuación (4.10) tenemos

$$x(t, \omega) = x_0(\omega) + \int_0^t f(w(s), b(s), x(s, \omega)) ds. \quad (4.26)$$

Supondremos por simpleza que  $f$  es dada por (4.11). En caso de que  $f$  este dada por (4.12) los argumentos son completamente análogos.

Así restando el término  $\bar{x}$ , aplicando esperanza y tomando norma a (4.26) tenemos que

$$\begin{aligned} \|\mathbb{E}(x(t)) - \bar{x}\|_{\mathbb{R}^{d_x}} &\leq \|\mathbb{E}(x_0) - \bar{x}\|_{\mathbb{R}^{d_x}} + \left\| \mathbb{E} \left( \int_0^t \sigma(w(s)x(s, \omega) + b(s)) ds \right) \right\|_{\mathbb{R}^{d_x}} \\ &\leq \|\mathbb{E}(x_0) - \bar{x}\|_{\mathbb{R}^{d_x}} + \mathbb{E} \left( \left\| \int_0^t \sigma(w(s)x(s, \omega) + b(s)) ds \right\|_{\mathbb{R}^{d_x}} \right). \end{aligned} \quad (4.27)$$

Ahora separando 2 casos.

**Caso**  $t \leq T \leq 1$ .

Antes de todo observemos que si tomamos norma  $L^2(\Omega; \mathbb{R}^{d_x})$  a (4.26) y aplicando la desigualdad de Hölder tenemos

$$\|x(t)\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 \leq C \left( \|x_0\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 + \|w\|_{L^2(0, T; \mathbb{R}^{d \times d})}^2 \int_0^t \|x(s)\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 ds + C \|b\|_{L^2(0, T; \mathbb{R}^d)}^2 \right).$$

Así aplicando el Lema de Gronwall obtenemos

$$\|x(t)\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 \leq K_1 \left( \|x_0\|^2 + \|b\|_{L^2(0, T; \mathbb{R}^{d_x})}^2 \right) e^{K_2(\|w\|_{L^2(0, T; \mathbb{R}^{d \times d})}^2)t} \quad (4.28)$$

$$\leq K_1 \left( \|x_0\|^2 + \|b\|_{L^2(0, 1; \mathbb{R}^{d_x})}^2 \right) e^{K_2(\|w\|_{L^2(0, 1; \mathbb{R}^{d \times d})}^2)} = K, \quad (4.29)$$

donde  $K$  es una constante positiva independiente de  $T$ . Ahora utilizando el hecho de que  $\sigma \in Lip_0(\mathbb{R}^n)$ , y la desigualdad de Hölder en (4.27) tenemos que

$$\begin{aligned} \|\mathbb{E}(x(t)) - \bar{x}\|_{\mathbb{R}^{d_x}} &\leq \|\mathbb{E}(x_0) - \bar{x}\|_{\mathbb{R}^{d_x}} + C_\sigma \int_0^T \|w(s)\| \|x(s)\|_{L^2(\Omega; \mathbb{R}^n)} + \|b(s)\| ds \\ &\leq C(\|\mathbb{E}(x_0) - \bar{x}\|_{\mathbb{R}^{d_x}} + \|u\|_{L^2(0, T; \mathbb{R}^{d_u})}) \\ &\leq C(\|\mathbb{E}(x_0) - \bar{x}\|_{\mathbb{R}^{d_x}} + \|u\|_{H^k(0, T; \mathbb{R}^{d_u})}). \end{aligned}$$

**Caso**  $T \geq t > 1$ .

Primero observemos que para cada  $t \in (1, T]$  existe  $t^* \in (t-1, t]$  tal que tenemos la siguiente estimación

$$\|\mathbb{E}(x(t^*)) - \bar{x}\|^2 \leq \|\mathbb{E}(x) - \bar{x}\|_{L^2(0, T; \mathbb{R}^n)}^2. \quad (4.30)$$

Para probar lo anterior supongamos que

$$\|\mathbb{E}(x(t^*)) - \bar{x}\|^2 > \|\mathbb{E}(x) - \bar{x}\|_{L^2(0, T; \mathbb{R}^n)}^2,$$

para todo  $t^* \in (t-1, t]$ . Entonces

$$\begin{aligned} \|\mathbb{E}(x) - \bar{x}\|_{L^2(0,T;\mathbb{R}^n)}^2 &= \int_0^T \|\mathbb{E}(x(t)) - \bar{x}\|^2 dt \geq \int_{t-1}^t \|\mathbb{E}(x(s)) - \bar{x}\|^2 ds \\ &> \|\mathbb{E}(x) - \bar{x}\|_{L^2(0,T;\mathbb{R}^n)}^2. \end{aligned}$$

Una contradicción, por tanto, tenemos (4.30). Por otro lado observemos que de forma análoga a (4.28) tenemos para todo  $s \in [t^*, t]$  que

$$\|x(s)\|_{L^2(\Omega;\mathbb{R}^{d_x})}^2 \leq K_1 \left( \|x(t^*)\|^2 + \|b\|_{L^2(t^*,t;\mathbb{R}^{d_x})}^2 \right) e^{K_2(\|w\|_{L^2(t^*,t;\mathbb{R}^{d \times d})}^2)(t^*-s)}.$$

Así observando que  $0 < s - t^* \leq 1$  tenemos que

$$\|x(s)\|_{L^2(\Omega;\mathbb{R}^{d_x})}^2 \leq K,$$

para  $K$  independiente de  $T$ . Ahora si integramos entre  $(t^*, t)$  y aplicamos esperanza a (4.10)

$$\mathbb{E}(x(t)) - \bar{x} = \mathbb{E}(x(t^*)) - \bar{x} + \mathbb{E} \left( \int_{t^*}^t \sigma(w(s), b(s), x^T(s, \omega)) ds \right).$$

Tomando norma, aplicando la estimación (4.30), utilizando la desigualdad de Hölder y observando que  $0 < t - t^* \leq 1$  obtenemos

$$\begin{aligned} \|\mathbb{E}(x(t)) - \bar{x}\| &\leq \|\mathbb{E}(x) - \bar{x}\|_{L^2(0,T;\mathbb{R}^{d_x})} + C \int_{t^*}^t \|w\| \|x\|_{L^2(\Omega;\mathbb{R}^{d_x})} + \|b\| ds \\ &\leq \|\mathbb{E}(x) - \bar{x}\|_{L^2(0,T;\mathbb{R}^{d_x})} + C \|w\|_{L^2(t^*,t;\mathbb{R}^{d \times d})} + (t - t^*) \|b\|_{L^2(t^*,t;\mathbb{R}^d)} \\ &\leq C(\|\mathbb{E}(x) - \bar{x}\|_{L^2(0,T;\mathbb{R}^{d_x})} + \|u\|_{L^2(0,T;\mathbb{R}^{d_u})}) \\ &\leq C(\|\mathbb{E}(x) - \bar{x}\|_{L^2(0,T;\mathbb{R}^{d_x})} + \|u\|_{H^k(0,T;\mathbb{R}^{d_u})}). \end{aligned}$$

Concluyendo la demostración. □

**Lema 4.2.2.** Supongamos que se satisface la **Hipótesis 1** para  $T_0 > 0$ . Sea  $x_0 \in L^2(\Omega;\mathbb{R}^{d_x})$ ,  $T > 0$  fijo y  $u^T \in L^2(0, T;\mathbb{R}^{d_u})$  mínimo de (4.18) y  $x^T$  la trayectoria asociada  $u^T$  con  $x^T(0) = x_0$ . Entonces tenemos que existe una constante  $C > 0$  independiente de  $T$  tal que

$$\|u^T\|_{H^k(0,T;\mathbb{R}^{d_u})} + \|\mathbb{E}(x^T) - \bar{x}\|_{L^2(0,T;\mathbb{R}^{d_x})} + \|\mathbb{E}(x^T(t)) - \bar{x}\|_{\mathbb{R}^{d_x}} \leq C,$$

para todo  $t \in [0, T]$ .

*Demostración.* Separamos la demostración en dos casos

**Caso**  $T \geq T_0$ . Observemos que desde la **Hipótesis 1** tenemos que existe  $u^* = [w, b]^\top \in H^k(0, T; \mathbb{R}^{d_u})$  tal que la  $x^*$  la solución de

$$\begin{cases} x_t^* = f(u, x^*) & \text{en } (0, T_0), \\ x^*(0) = x_0, \end{cases}$$

satisface que  $\mathbb{E}(x^*(T_0)) = \bar{x}$ . Así podemos definir  $u^{aux}$  como sigue

$$u^{aux}(t) = \begin{cases} u^*(t) & \text{en } (0, T_0), \\ 0 & \text{en } (T_0, T). \end{cases}$$

Denotando como  $x^{aux}$  la trayectoria asociada a  $u^{aux}$ , observemos que  $x^{aux}(0) = x_0$  y  $x^{aux}(t) = \bar{x}$  para cada  $t \in [T_0, T]$ . Por otro lado, recordemos que

$$\begin{aligned} J^T(u) &= \frac{1}{2} \left( \|u\|_{H^k(0, T; \mathbb{R}^{d_u})}^2 dt + \int_0^T \|\mathbb{E}(x(t)) - \bar{x}\|_{\mathbb{R}^{d_x}}^2 dt + \left\| \varphi \left( \mathbb{E}(x(T, \omega)) \right) - y \right\|_{\mathbb{R}^{m \times n}}^2 \right) \\ &= \frac{1}{2} \left( \|u\|_{H^k(0, T; \mathbb{R}^{d_u})}^2 dt + \int_0^T \|\mathbb{E}(x(t)) - \bar{x}\|_{\mathbb{R}^{d_x}}^2 dt + \Phi(\mathbb{E}(x(T))) \right). \end{aligned}$$

Como  $\Phi > 0$  y  $J^T(u^T) \leq J^T(u^{aux})$  por definición de mínimo tenemos

$$\left\| \mathbb{E}(x^T) - \bar{x} \right\|_{L^2(0, T; \mathbb{R}^{d_u})}^2 + \|u^T\|_{H^k(0, T; \mathbb{R}^{d_u})}^2 \leq \Phi(\bar{x}) + \|\mathbb{E}(x^*) - \bar{x}\|_{L^2(0, T_0; \mathbb{R}^{d_x})}^2 + \|u^*\|_{H^k(0, T_0; \mathbb{R}^{d_u})}^2 = K,$$

con  $K > 0$  independiente de  $T$ . Finalmente aplicando el Lema 4.2.1 concluimos que existe  $C > 0$  independiente de  $T$  tal que

$$\|u^T\|_{H^k(0, T; \mathbb{R}^{d_u})} + \left\| \mathbb{E}(x^T) - \bar{x} \right\|_{L^2(0, T; \mathbb{R}^{d_x})} + \left\| \mathbb{E}(x^T(t)) - \bar{x} \right\|_{\mathbb{R}^{d_x}} \leq C.$$

**Caso**  $T \leq T_0$ . Nuevamente como  $\Phi > 0$  y  $J^T(u^T) \leq J^T(0)$  tenemos que

$$\left\| \mathbb{E}(x^T) - \bar{x} \right\|_{L^2(0, T; \mathbb{R}^{d_u})}^2 + \|u^T\|_{H^k(0, T; \mathbb{R}^{d_u})}^2 \leq \Phi(\mathbb{E}(x_0)) + T_0 \|\mathbb{E}(x_0) - \bar{x}\|_{\mathbb{R}^{d_x}}^2 = K,$$

concluyendo nuevamente por el Lema 4.2.1. □

Sean  $0 \leq \tau_1 \leq \tau_2 \leq T$ , y sean  $x^{\tau_1}, x^{\tau_2} \in L^2(\Omega; \mathbb{R}^{d_x})$ , así consideremos el funcional auxiliar

$$J_{\tau_1, \tau_2}^{aux}(u) = \int_{\tau_1}^{\tau_2} \|\mathbb{E}(x(t)) - \bar{x}\|_{\mathbb{R}^{d_x}}^2 dt + \|u(t)\|_{H^k(\tau_1, \tau_2; \mathbb{R}^{d_u})}^2, \quad (4.31)$$

para cada  $u = [w, b]^\top \in U_{ad}$ , donde

$$U_{ad} = \left\{ u \in H^k(0, T; \mathbb{R}^{d_x}) : \mathbb{E}(x(\tau_2)) = \mathbb{E}(x^{\tau_2}) \right\},$$

y donde  $x \in C([\tau_0, \tau_1]; L^2(\Omega; \mathbb{R}^{d_x}))$  denota la única solución de

$$\begin{cases} x_t = f(u, x) & \text{en } (\tau_1, \tau_2), \\ x(\tau_1) = x^{\tau_1}. \end{cases} \quad (4.32)$$

Siguiendo los pasos realizados en [13] como en [17], el siguiente lema es de gran importancia para probar nuestro teorema principal.

**Lema 4.2.3.** *Supongamos que tenemos **Hipótesis 1** para  $T_0 > 0$ . Sea  $r > 0$  como en la **Hipótesis 1**, y  $x^{\tau_1}, x^{\tau_2} \in L^2(\Omega; \mathbb{R}^{d_x})$  tales que*

$$\|\mathbb{E}(x^{\tau_i}) - \bar{x}\|_{\mathbb{R}^{d_x}} \leq r,$$

para  $i = 1, 2$ , además  $0 \leq \tau_1 \leq \tau_2 \leq T$  son tales que  $\tau_2 - \tau_1 \geq 2T_0$ . Sea  $u^T \in U_{ad}$  el mínimo global de  $J_{\tau_1, \tau_2}^{aux}$  definido en (4.31) con  $x^T$  denota la solución de (4.32) con  $\mathbb{E}(x^T(\tau_2)) = \mathbb{E}(x^{\tau_2})$ . Entonces existe una constante  $C > 0$  independiente del horizonte tiempo  $T$  y de  $\tau^1, \tau_2 > 0$  tal que

$$\|u^T\|_{H^k(\tau_1, \tau_2; \mathbb{R}^{d_u})} + \|\mathbb{E}(x^T) - \bar{x}\|_{L^2(0, T; \mathbb{R}^{d_x})} + \|\mathbb{E}(x^T(t)) - \bar{x}\| \leq C \left( \|\mathbb{E}(x^{\tau_1}) - \bar{x}\| + \|\mathbb{E}(x^{\tau_2}) - \bar{x}\| \right),$$

para cada  $t \in [\tau_1, \tau_2]$ . Es más, la función  $r \mapsto C$  es no decreciente.

*Demostración.* Dividiremos la demostración en tres pasos.

**Paso 1.** Desde nuestra **Hipótesis 1**, vía una traslación del tiempo en  $\tau_1$ , sabemos que existe un control  $u^a = [w^a, b^a]^\top \in H^k(\tau_1, \tau_1 + T_0; \mathbb{R}^{d_u})$  tal que

$$\|u^a\|_{H^k(\tau_1, \tau_1 + T_0; \mathbb{R}^d)}^2 \leq K \|\mathbb{E}(x^{\tau_1}) - \bar{x}\|^2. \quad (4.33)$$

Además, la solución del sistema

$$\begin{cases} x_t^a(t, \omega) = f(u^a, x^a) & \text{en } (\tau_1, \tau_1 + T_0), \\ x^a(\tau_1) = x^{\tau_1}(\omega), \end{cases}$$

satisface que  $\mathbb{E}(x^a(\tau_1 + T_0)) = \bar{x}$ . Observemos que

$$x^a(t, \omega) = x^{\tau_1}(\omega) + \int_{\tau_1}^t f(w(s), b(s), x(s, \omega)) ds.$$

Tomando norma e integrando en  $\Omega$  tenemos

$$\|x^a(t)\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 \leq \|x^{\tau_1}\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 + \mathbb{E} \left( \left\| \int_{\tau_1}^t f(w(s), b(s), x(s, \omega)) ds \right\|^2 \right). \quad (4.34)$$

Así utilizando el supuesto de que  $f \in Lip_0(\mathbb{R}^{d_x})$ , junto a la linealidad del operador  $\mathbb{E}$  tenemos que

$$\|x^a(t)\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 \leq \|x^{\tau_1}\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 + C_\sigma \mathbb{E} \left( \left( \int_{\tau_1}^t \|w^a(s)\| \|x^a(s)\| ds \right)^2 \right) + \|b^a\|_{L^2(\tau_1, t; \mathbb{R}^d)}^2.$$

Además, utilizando la desigualdad de Hölder obtenemos

$$\begin{aligned} \|x^a(t)\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 &\leq \|x^{\tau_1}\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 + C_\sigma \|w\|_{L^2(\tau_1, \tau_1+T_0; \mathbb{R}^{d \times d})}^2 \int_{\tau_1}^t \|x^a(s)\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 ds \\ &\quad + \|b^a\|_{L^2(\tau_1, \tau_1+T_0; \mathbb{R}^d)}^2. \end{aligned}$$

Aplicando la desigualdad de Gronwall y utilizando la estimación (4.33) tenemos

$$\begin{aligned} \|x^a(t)\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 &\leq (\|x^{\tau_1}\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 + \|b^a\|_{L^2(\tau_1, \tau_1+T_0; \mathbb{R}^n)}^2) \exp(C_\sigma \|w^a\|_{L^2(\tau_1, \tau_1+T_0; \mathbb{R}^{d \times d})}^2) \\ &\leq (\|x^{\tau_1}\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 + C \|\mathbb{E}(x^{\tau_1}) - \bar{x}\|^2) \exp(C \|\mathbb{E}(x^{\tau_1}) - \bar{x}\|^2) \\ &\leq (\|x^{\tau_1}\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 + Cr^2) \exp(Cr^2). \end{aligned} \quad (4.35)$$

Entonces tenemos la cota uniforme en  $T$

$$\|\mathbb{E}(x^a(t))\|^2 \leq \|x^a(t)\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 \leq (\|x^{\tau_1}\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 + Cr^2) \exp(Cr^2) = K. \quad (4.36)$$

Ahora estimamos la diferencia  $\mathbb{E}(x^a(t)) - \bar{x}$  tenemos

$$\|\mathbb{E}(x^a(t)) - \bar{x}\|^2 \leq C(\|\mathbb{E}(x^{\tau_1}) - \bar{x}\|^2 + \int_{\tau_1}^t \|w^a(s)\|^2 \|x^a(s)\|_{L^2(\Omega; \mathbb{R}^{d_x})}^2 + \|b^a(s)\|^2 ds).$$

Así por la estimación (4.36) se sigue que

$$\|\mathbb{E}(x^a(t)) - \bar{x}\|^2 \leq C(\|\mathbb{E}(x^{\tau_1}) - \bar{x}\|^2 + K \int_{\tau_1}^t \|w^a(s)\|^2 + \|b^a(s)\|^2 ds).$$

Finalmente utilizando (4.33) nuevamente tenemos que

$$\begin{aligned} \|\mathbb{E}(x^a(t)) - \bar{x}\|^2 &\leq C(\|\mathbb{E}(x^{\tau_1}) - \bar{x}\|^2 + \|u^a\|_{L^2(\tau_1, \tau_1 + T_0; \mathbb{R}^{d_u})}^2) \\ &\leq C(\|\mathbb{E}(x^{\tau_1}) - \bar{x}\|^2 + \|\mathbb{E}(x^{\tau_1}) - \bar{x}\|^2) \\ &\leq C \|\mathbb{E}(x^{\tau_1}) - \bar{x}\|^2, \end{aligned} \quad (4.37)$$

para todo  $t \in [\tau_1, \tau_1 + T_0]$ .

**Paso 2.** Denotemos por  $x_a = x^a(T_0 + \tau_1, \omega) \in L^2(\Omega; \mathbb{R}^{d_x})$  la trayectoria asociada a  $u^a$  en tiempo  $T_0 + \tau_1$ . Ahora desde nuestra hipótesis de controlabilidad como  $\mathbb{E}(x_a) = \bar{x}$  tenemos que existe  $u^b \in H^k(\tau_1, T_0 + \tau_1; \mathbb{R}^{d_u})$  y  $K > 0$  tal que

$$\|u^b\|_{H^k(0, T; \mathbb{R}^d)}^2 \leq K \|\mathbb{E}(x^{\tau_2}) - \bar{x}\|^2. \quad (4.38)$$

Además, la solución del sistema

$$\begin{cases} x_t^b = f(u^b, x^b) & \text{en } (\tau_1, T_0 + \tau_1), \\ x^b(\tau_1) = x_a, \end{cases}$$

satisface que  $\mathbb{E}(x^b(T_0 + \tau_1)) = \mathbb{E}(x^{\tau_2})$ . Observemos que por la desigualdad de Gronwall junto a (4.36) podemos ver que

$$\|\mathbb{E}(x^b(t))\|^2 \leq C(\|\mathbb{E}(x_a)\|^2 + Cr^2) \exp(Cr^2) \quad (4.39)$$

$$= C(1 + r^2) \exp(Cr^2) = K. \quad (4.40)$$

Por tanto, de forma análoga al paso anterior tenemos

$$\|\mathbb{E}(x^b(t)) - \bar{x}\|^2 \leq C(\|\mathbb{E}(x_a) - \bar{x}\|^2 + \|u^b\|_{L^2(0, T_0; \mathbb{R}^d)}^2).$$

Así utilizando (4.37) y (4.38) concluimos que

$$\|\mathbb{E}(x^b(t)) - \bar{x}\|^2 \leq C(\|\mathbb{E}(x^{\tau_1}) - \bar{x}\|^2 + \|\mathbb{E}(x^{\tau_2}) - \bar{x}\|^2), \quad (4.41)$$

para cada  $t \in [\tau_1, \tau_1 + T_0]$ .

**Paso 3.** Ahora definamos

$$u^{aux}(t) := \begin{cases} u^a(t) & \text{en } (\tau_1, T_0 + \tau_1), \\ 0 & \text{en } (\tau_1 + T_0, \tau_2 - T_0), \\ u^b(t - (\tau_2 - \tau_1 - T_0)) & \text{en } (\tau_2 - T_0, \tau_2), \end{cases}$$



y sea  $x^{aux}$  la trayectoria asociada a  $u^{aux}$ . Observemos que por construcción se satisface que

$$x^{aux}(t) := \begin{cases} x^a(t) & \text{en } (\tau_1, T_0 + \tau_1), \\ x^a(T_0 + \tau_1) & \text{en } (\tau_1 + T_0, \tau_2 - T_0), \\ x^b(t - (\tau_2 - \tau_1 - T_0)) & \text{en } (\tau_2 - T_0, \tau_2). \end{cases}$$

Así, la trayectoria  $x^{aux}$  puede ser representada por la Figura 4.2

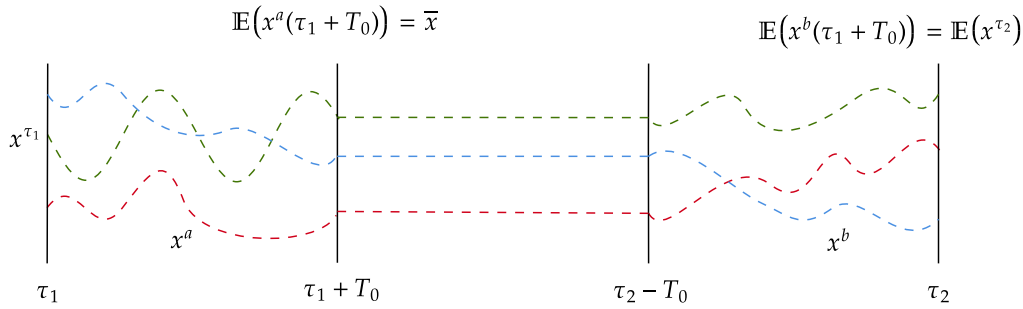


FIGURA 4.2: Representación trayectoria  $x^{aux}(t, \omega)$ .

Además, tenemos que

$$\mathbb{E}(x^{aux}(t)) := \begin{cases} \mathbb{E}(x^a(t)) & \text{en } (\tau_1, T_0 + \tau_1), \\ \bar{x} & \text{en } (\tau_1 + T_0, \tau_2 - T_0), \\ \mathbb{E}(x^b(t - (\tau_2 - \tau_1 - T_0))) & \text{en } (\tau_2 - T_0, \tau_2). \end{cases}$$

En particular tenemos que  $\mathbb{E}(x^{aux}(\tau_2)) = \mathbb{E}(x^b(T_0 + \tau_1)) = \mathbb{E}(x^{\tau_2})$ , por tanto tenemos que  $u^{aux} \in U_{ad}$ . Ahora observemos que si evaluamos  $u^{aux}$  en  $J_{\tau_1, \tau_2}^{aux}$  el funcional (4.31), utilizamos una traslación del intervalo de integración, y utilizando las estimaciones (4.37) y (4.41) tenemos que

$$\begin{aligned} J_{\tau_1, \tau_2}^{aux}(u^{aux}) &= \|u^a\|_{H^k(\tau_1, T_0 + \tau_1; \mathbb{R}^{d_u})}^2 + \|u^b\|_{H^k(\tau_1, T_0 + \tau_1; \mathbb{R}^{d_u})}^2 \\ &\quad + \int_{\tau_1}^{\tau_1 + T_0} \|\mathbb{E}(x^a(t)) - \bar{x}\|^2 dt + \int_{\tau_1}^{\tau_1 + T_0} \|\mathbb{E}(x^b(t)) - \bar{x}\|^2 dt \\ &\leq C \left( \|\mathbb{E}(x^{\tau_1}) - \bar{x}\|^2 + \|\mathbb{E}(x^{\tau_2}) - \bar{x}\|^2 \right). \end{aligned}$$

Donde  $C > 0$  es independiente de  $T, \tau_1, \tau_2 > 0$ . Además, no difícil ver desde (4.36) y (4.39) que  $C$  no es decreciente con respecto a  $r$ . Así, tenemos

$$\begin{aligned} \left\| \mathbb{E}(x^T) - \bar{x} \right\|_{L^2(\tau_1, \tau_2; \mathbb{R}^{d_x})}^2 + \left\| u^T \right\|_{H^k(\tau_1, \tau_2; \mathbb{R}^{d_x})}^2 &\leq J_{\tau_1, \tau_2}^{aux}(u^T) \\ &\leq J_{\tau_1, \tau_2}^{aux}(u^{aux}) \leq C \left( \left\| \mathbb{E}(x^{\tau_1}) - \bar{x} \right\|^2 + \left\| \mathbb{E}(x^{\tau_2}) - \bar{x} \right\|^2 \right). \end{aligned}$$

Finalmente aplicando el Lema 4.2.1 junto al hecho que  $\left\| u^T \right\|_{L^2(\tau_1, \tau_2; \mathbb{R}^{d_u})}$  acotado uniformemente con respecto a  $\tau_1, \tau_2, T$ , se sigue el resultado.  $\square$

A continuación, mostramos el ultimo lema de este capítulo.

**Lema 4.2.4.** *Sea  $X$  un espacio de Banach real tal que  $T > 0$  y  $g \in L^2(0, T; X)$ . Para cada  $\tau \leq \frac{T}{2}$ , existe  $t_1 \in [0, \tau]$  y  $t_2 \in (T - \tau, T]$  tal que*

$$\|g(t_i)\|_X \leq \varepsilon(\tau) \quad \text{para } i = 1, 2,$$

$$\text{donde } \varepsilon(\tau) = \frac{\|g\|_{L^2(0, T; X)}}{\sqrt{\tau}}.$$

*Demostración.* Procedemos por contradicción. Supongamos que

$$\|g(t)\|_X > \varepsilon(\tau) \quad \forall t \in [0, \tau],$$

o

$$\|g(t)\|_X > \varepsilon(\tau) \quad \forall t \in (T - \tau, T].$$

Entonces tenemos que

$$\int_0^T \|g(t)\|_X^2 dt \geq \int_0^\tau \|g(t)\|_X^2 dt + \int_{T-\tau}^T \|g(t)\|_X^2 dt \geq \tau \varepsilon^2(\tau).$$

Por tanto

$$\varepsilon^2(\tau) < \frac{1}{\tau} \int_0^T \|g(t)\|_X^2 dt = \varepsilon^2(\tau).$$

Lo cual es una contradicción, probando así el lema.  $\square$

**Observación 4.2.4.** *Como ya comentamos anteriormente, no probaremos el teorema principal en este documento, aun así, no es difícil notar que la construcción de los lemas previos se realizó de forma completamente análoga a los trabajos [13, 17]. Por lo anterior creemos que es posible probar el teorema principal moldeando las herramientas utilizadas en [13, Teorema 2.1], en relación con las usadas en el Lema 4.2.3*

---

*para probar la propiedad de Turnpike en este contexto. No obstante, esto requiere de una mayor rigurosidad.*



## Capítulo 5

# Soluciones en Promedio

### 5.1. Introducción

En este capítulo analizaremos el comportamiento de la dinámica promedio de ciertas ecuaciones.

Como hemos visto en los Capítulos 1,2,3 y 4, obtuvimos resultados para la dinámica promedio de un problema con aleatoriedad sobre los coeficientes, sin embargo, no sabemos con claridad que sistema satisface el promedio, o cuál es la dinámica que lo subyace. Es natural analizar el promedio de las soluciones cuando se tiene aleatoriedad, e intentar dar respuesta a que sistema modela la dinámica promedio. Un ejemplo de esto es cuando se tiene un río, en el cual la velocidad de su caudal varía aleatoriamente dentro de algún intervalo, en tal caso lo más natural e interesante sería analizar el promedio de velocidades, más que analizar el caudal para cada velocidad.

Dar respuesta a que dinámica soluciona el sistema promedio, no es una respuesta trivial. Es más, como bien se comenta en [31], la solución promedio de una ecuación hiperbólica puede satisfacer una dinámica parabólica. En efecto, consideremos el espacio de probabilidad  $(\Omega, \mathcal{F}, \mu)$  y la ecuación de onda

$$\begin{cases} u_{tt} - \alpha^2 u_{xx} = 0 & (x, t) \in \mathbb{R} \times (0, \infty), \\ u(x, 0) = g(x) & x \in \mathbb{R}, \\ \partial_t u(x, 0) = 0 & x \in \mathbb{R}. \end{cases} \quad (5.1)$$

En donde  $g \in C^2(\mathbb{R})$ ,  $\alpha : \Omega = \mathbb{R} \rightarrow \mathbb{R}$  una variable aleatoria que se distribuye normal de parámetros  $(0, 1)$ , en tal caso su función de densidad  $\rho$  tiene la forma

$$\rho(\alpha) = \frac{1}{\sqrt{2\pi}} e^{-\frac{|\alpha|^2}{2}},$$

y además,  $d\mu = \rho(\alpha)d\alpha$ . Por otro lado, sabemos desde la fórmula de D’Alambert [14, Capítulo 2.4] que la solución de (5.1) viene dada por

$$u(x, t, \omega) = \frac{1}{2} \left( g(x + \alpha t) + g(x - \alpha t) \right).$$

Por lo tanto, si promediamos  $u(t, x, \omega)$  con respecto a  $\omega$  tenemos que

$$\bar{u}(x, t) = \int_{\Omega} u(t, x, \omega) d\mu = \int_{\mathbb{R}} \frac{1}{2} \left( g(x + \alpha t) + g(x - \alpha t) \right) \rho(\alpha) d\alpha. \quad (5.2)$$

Así tomado  $z_1 = x + \alpha t$  y  $z_2 = x - \alpha t$  en (5.2) tenemos

$$\bar{u}(x, t) = \frac{1}{2t} \left( \int_{-\infty}^{\infty} g(z_1) \rho\left(\frac{z_1 - x}{t}\right) dz_1 - \int_{\infty}^{-\infty} g(z_2) \rho\left(\frac{-(z_2 - x)}{t}\right) dz_2 \right).$$

Como la función  $\rho$  es simétrica con respecto al origen, obtenemos la fórmula para el promedio

$$\bar{u}(x, t) = \frac{1}{t} \int_{\mathbb{R}} g(z) \rho\left(\frac{z - x}{t}\right) dz = \frac{1}{t\sqrt{2\pi}} \int_{\mathbb{R}} g(z) e^{\frac{-(z-x)^2}{2t^2}} dz. \quad (5.3)$$

Además, observemos que

$$\bar{u}(x, 0) = \int_{\mathbb{R}} g(x) \rho(\alpha) d\alpha = g(x) \int_{\Omega} d\mu = g(x),$$

Por tanto, como la expresión (5.3) corresponde a una convolución entre la condición inicial y el núcleo de la ecuación del calor, tenemos que  $\bar{u}$  soluciona la ecuación

$$\begin{cases} \bar{u}_t - \frac{1}{t} \Delta \bar{u} = 0 & (x, t) \in \mathbb{R} \times (0, \infty), \\ \bar{u}(x, 0) = g(x) & x \in \mathbb{R}. \end{cases}$$

Finalmente si definimos  $\varphi(x, t) = \bar{u}(x, \sqrt{2t})$  tenemos que

$$\begin{cases} \varphi_t - \Delta \varphi = 0 & (x, t) \in \mathbb{R} \times (0, \infty), \\ \varphi(x, 0) = g(x) & x \in \mathbb{R}. \end{cases}$$

Desde lo anterior podemos observar que (al menos en este caso) existe un efecto regularizante tras tomar promedio, efecto regularizante heredado desde la ecuación del calor. Por lo tanto, promediar las soluciones puede generar un impacto no trivial sobre la dinámica inicial.

Por otro lado, como bien se presenta en el Capítulo 1 para analizar la controlabilidad de un sistema lineal, podemos hacerlo mediante un sistema dual. En [31] se prueba en particular que,

para la ecuación del calor con control interior se verifica la controlabilidad aproximada en promedio, y la controlabilidad a cero en promedio. Las técnicas usadas para tener este resultado son basadas en la descomposición del operador Laplaciano Dirichlet mediante sus autovalores, representado la solución mediante una serie de Fourier. Lo anterior permite al autor obtener una expresión explícita para el promedio del sistema dual asociado a la ecuación del calor, para posteriormente probar la desigualdad de observabilidad. En tal demostración tener una expresión explícita de la solución del sistema adjunto, adquiere un rol primordial. En el caso general, esto se puede ilustrar si consideramos el sistema (1.31) con  $B(\omega) = j_{U,V}$  el operador inclusión entre  $U$  y  $V$ , es decir

$$\begin{cases} u_t(t) + A(\omega)u(t) = \chi_E j_{U,V} f(t) & \text{en } (0, T], \\ u(0) = u_0, \end{cases} \quad (5.4)$$

donde  $E \subset [0, T]$  y  $f \in L^2(E; U)$  es el control <sup>1</sup>.

Supongamos que queremos probar la controlabilidad a cero del sistema (5.4). En tal caso podríamos querer obtener la desigualdad de observabilidad a cero (Definición 1.5.10), es decir, encontrar  $K > 0$  tal que

$$\left\| \int_{\Omega} v(0, \omega) d\mu \right\|_{V'}^2 \leq K \int_0^T \chi_E \left\| \int_{\Omega} v(t, \omega) d\mu \right\|_U^2 dt, \quad (5.5)$$

con  $v$  solución del sistema dual

$$\begin{cases} -v_t(t) + A^*(\omega)v(t) = 0 & \text{en } [0, T) \\ v(T) = v_0 \end{cases} \quad (5.6)$$

Entonces si denotamos como  $\bar{v}$  al promedio de la solución de (5.6), la desigualdad (5.5) se traduce en

$$\|\bar{v}(0)\|_{V'}^2 \leq K \int_0^T \chi_E \left\| \int_{\Omega} \bar{v}(t) \right\|_U^2 dt,$$

Por lo anterior, tener una expresión explícita para el promedio del sistema dual, puede llegar a tener un rol fundamental al momento de querer obtener alguna propiedad de controlabilidad del sistema (5.4).

---

<sup>1</sup>La notación puede ser vista con más detalle en la Sección 1.5.3.

Motivados por entender la dinámica que subyace el promedio de un sistema, es que consideremos la siguiente ecuación

$$\begin{cases} u_t + Au + \alpha Bu = 0, \\ u(0) = u_0, \end{cases} \quad (5.7)$$

donde  $\alpha$  una variable aleatoria en el intervalo  $[a, b]$  y supondremos que  $A, B \in \mathcal{L}(\mathbb{R}^n)$ . Como señalamos, estamos interesados en analizar la solución promedio ecuación, es decir

$$\bar{u}(t) = \int_a^b u(t, \alpha) d\mu. \quad (5.8)$$

## 5.2. Operadores conmutativos

En esta sección nos enfocaremos en el caso que  $A, B$  conmutan, es decir, que  $AB = BA$ . Utilizando esto junto a las propiedades de la matriz exponencial, es que podremos obtener una expresión explícita para el promedio  $\bar{u}$ . Además, representaremos la dinámica promedio como una combinación de soluciones asociadas a los valores extremos relacionados a la variable aleatoria a considerar.

### 5.2.1. Estimaciones Principales

En lo que sigue supondremos que  $\alpha$  se distribuye uniforme en  $[a, b]$  ( $\alpha \sim U(a, b)$ ), es decir,  $\alpha$  toma valores en el intervalo  $[a, b]$ , y su función de densidad corresponde a  $\frac{1}{b-a}$ . Así comencemos observando que la solución de la ecuación (5.7) viene dada por

$$u(t, \alpha) = e^{-(A+\alpha B)t} u_0. \quad (5.9)$$

Entonces expresando  $u(t, \omega)$  como serie de potencias, y utilizando el hecho que  $A, B$  conmutan tenemos que

$$u(t, \alpha) = e^{-(A+\alpha B)t} u_0 = e^{-At} e^{-\alpha Bt} u_0 = e^{-At} \sum_{n=0}^{\infty} \frac{(-1)^n \alpha^n (Bt)^n}{n!} u_0. \quad (5.10)$$

Ahora para calcular la solución, podemos utilizar el teorema de la convergencia dominada por la convergencia de la matriz exponencial, para obtener

$$\begin{aligned} \bar{u}(t) &= \frac{e^{-At}}{b-a} \int_a^b \sum_{n=0}^{\infty} \frac{(-1)^n (Bt)^n \alpha^n}{n!} u_0 d\alpha = \frac{e^{-At}}{b-a} \sum_{n=0}^{\infty} \int_a^b \frac{(-1)^n (Bt)^n \alpha^n}{n!} u_0 d\alpha \\ &= \frac{e^{-At}}{b-a} \sum_{n=0}^{\infty} \frac{(-1)^n (Bt)^n (b^{n+1} - a^{n+1})}{(n+1)!} u_0 \end{aligned}$$



En particular si  $B$  es invertible, entonces

$$\begin{aligned}
 \bar{u}(t) &= \frac{e^{-At}}{b-a} \sum_{n=0}^{\infty} \frac{(-1)^n (Bt)^n (b^{n+1} - a^{n+1})}{(n+1)!} u_0 \\
 &= \frac{e^{-At} B^{-1}}{t(b-a)} \left( - \sum_{n=0}^{\infty} \frac{(-1)^{n+1} (Bt)^{n+1} b^{n+1}}{(n+1)!} u_0 + \sum_{n=0}^{\infty} \frac{(-1)^{n+1} (Bt)^{n+1} a^{n+1}}{(n+1)!} u_0 \right) \\
 &= \frac{e^{-At} B^{-1}}{t(b-a)} \left( \left[ I - I - \sum_{n=1}^{\infty} \frac{(-1)^n (Bt)^n b^n}{n!} u_0 \right] - \left[ I - I - \sum_{n=1}^{\infty} \frac{(-1)^n (Bt)^n a^n}{n!} u_0 \right] \right) \\
 &= \frac{e^{-At} B^{-1}}{t(b-a)} \left( \left[ I - \sum_{n=0}^{\infty} \frac{(-1)^n (Bt)^n b^n}{n!} u_0 \right] - \left[ I - \sum_{n=0}^{\infty} \frac{(-1)^n (Bt)^n a^n}{n!} u_0 \right] \right) \\
 &= \frac{e^{-At} B^{-1}}{t(b-a)} \left( e^{-aBt} - e^{-bBt} \right) u_0 = \frac{B^{-1}}{t(b-a)} \left( e^{-(A+aB)t} - e^{-(A+bB)t} \right) u_0,
 \end{aligned}$$

donde  $I$  denota la matriz identidad. Por tanto, tenemos que el promedio

$$\bar{u}(t) = \frac{B^{-1}}{t(b-a)} \left( e^{-(A+aB)t} - e^{-(A+bB)t} \right) u_0. \quad (5.11)$$

**Observación 5.2.1.** La convergencia de la representación para el promedio (5.10), es heredada desde la convergencia de la solución (5.9).

**Observación 5.2.2.** Como bien muestra las estimaciones anteriores, cuando  $\alpha \sim U([a, b])$  y  $B$  es invertible, podemos expresar el promedio  $\bar{u}$  como una combinación de las soluciones de los sistemas asociados a los valores extremos del parámetro  $\alpha$ . En efecto, para ello basta notar que

$$\bar{u}(t) = \frac{B^{-1}}{t(b-a)} \left( e^{-(A+aB)t} - e^{-(A+bB)t} \right) u_0 = \frac{B^{-1}}{t(b-a)} (u^a - u^b), \quad (5.12)$$

donde  $u^a$  y  $u^b$  son soluciones de

$$\begin{cases} u_t^a + (A + aB)u^a = 0, \\ u^a(0) = u_0, \end{cases} \quad y \quad \begin{cases} u_t^b + (A + bB)u^b = 0, \\ u^b(0) = u_0. \end{cases}$$

**Observación 5.2.3.** 1. Desde las estimaciones anterior, surgen naturalmente la pregunta, si el sistema promedio siempre es posible expresarlo como una combinación de las soluciones de los sistemas asociados a los valores extremos del parámetro aleatorio, y si no es así, en qué casos no. Esto quizás puede llegar a depender de la variable aleatoria que sea considerada. Una posible forma de analizar este problema cuando  $\alpha$  es arbitrario, es utilizando las herramientas desarrolladas en [11], donde

se prueba la controlabilidad de la ecuación del calor aleatoria con coeficiente de difusión distribuido arbitrariamente.

2. Observemos que en el ejemplo anterior consideramos un lado derecho nulo. Otro posible problema abierto es considerar el lado derecho no nulo e intentar obtener una expresión para el promedio.
3. Es posible pensar las relaciones anteriores en el caso en donde (5.7) es una ecuación lineal con  $A, B$  dos operadores que conmutan, e intentar análoga las estimaciones. No obstante, para ello se necesita mayor rigurosidad, posiblemente utilizando la regularizada de Yoshida, y basándose en la teoría de Semigrupos.

### 5.2.2. Ejemplos operadores conmutativos

Considerando  $G \subset \mathbb{R}^n$ . Como bien se señala en la observación 5.2.3, si fuese posible obtener una expresión como (5.11) para el caso en que  $A, B$  son operadores que conmutan, entonces sería posible obtener una forma explícita para la solución promedio de los siguientes sistemas

- Ejemplo 1: Ecuación del calor.

$$\begin{cases} u_t - \Delta u + \alpha p u = 0 & \text{en } G \times (0, 1), \\ u = 0 & \text{sobre } \partial G \times (0, 1), \\ u(0) = u_0, \end{cases}$$

donde  $p \in \mathbb{R}$ . De forma inmediata tenemos que el operador  $\Delta$  conmuta con la constante  $p \in \mathbb{R}$ .

- Ejemplo 2: Ecuación de Kuramoto Sivashinsky lineal.

$$\begin{cases} u_t + \Delta^2 u - \alpha \Delta u = 0 & \text{en } G \times (0, 1), \\ u = \Delta u = 0 & \text{sobre } \partial G \times (0, 1), \\ u(0) = u_0, \end{cases}$$

donde  $\Delta^2$  denota el bi-laplaciano.

- Ejemplo 3: Ecuación convección difusión en dominio periódico.

$$u_t - \Delta u + \alpha \mathbf{v} \cdot \nabla u = 0,$$

donde  $G = [0, 1]^d$ ,  $\mathbf{v}$  es un campo de velocidades y  $u$  es  $[0, 1]^d$ -periódica.

**Observación 5.2.4.** Además, de los ejemplos anteriores, podría ser considerado cualquier ejemplo donde el operador  $A \equiv 0$  o  $A \equiv I$ , con  $I$  la identidad del espacio.

### 5.3. Operadores no conmutativos

En este caso nos centraremos en analizar la dinámica promedio cuando  $A, B$  son matrices no conmutativas en (5.7). En este caso el problema de conocer la dinámica promedio de (5.7) se vuelve mucha más difícil, pues en este caso es importante el orden del producto entre las potencias de  $A$  y  $B$ . Además, en este caso, las estimaciones realizadas en (5.10) ya no son válidas, pues no siempre se verifica que  $e^{-(A+\lambda B)t} = e^{-At}e^{-\lambda Bt}$ .

#### 5.3.1. Estimaciones Principales

Por lo señalado anteriormente, es que para tener una expresión para el promedio con respecto a  $\lambda$  de la solución (5.8), nos basaremos en el siguiente resultado para álgebras no conmutativas. La demostración será omitida, pero puede ser encontrada en [53].

En lo que sigue definimos  $d_B(A) = [B, A]$ , donde  $[\cdot, \cdot]$  denota el corchete de Lie.

**Teorema 5.3.1** (Teorema del binomio para álgebras no conmutativas). *Sea  $A, B$  matrices (u operadores), no necesariamente conmutativos. Entonces se verifica que*

$$(A + B)^n = \sum_{k=0}^n \binom{n}{k} \{ (A + d_B)^k I \} B^{n-k}, \quad (5.13)$$

donde  $I$  corresponde a la matriz identidad.

Si bien el teorema anterior nos entrega una expresión para  $(A + B)^n$  cuando  $A, B$  no necesariamente conmutan, no podemos saber directamente que forma toma la expresión  $\int_0^1 (A + \lambda B)^n d\lambda$ , pues no es posible factorizar el término  $\lambda$  directamente. Para ello tenemos el siguiente teorema.

**Teorema 5.3.2.** *Dado  $k \geq 0$  y consideremos las funciones  $\theta_k^i$  definidas por las siguientes relaciones*

$$\theta_k^0(A, B) = A^k \quad 0 \leq k, \quad (5.14a)$$

$$\theta_i^i(A, B) = 0 \quad 1 \leq i \leq k, \quad (5.14b)$$

$$\theta_{k+1}^i(A, B) = A\theta_k^i(A, B) + d_B\theta_k^{i-1}(A, B) \quad 1 \leq i \leq k. \quad (5.14c)$$

Entonces se satisface que

$$(A + d_B)^k I = \sum_{i=0}^k \theta_k^i(A, B). \quad (5.15)$$

Además, las funciones  $\theta_k^i(A, B)$  son homogéneas de grado  $i$  para cada  $0 \leq i \leq k$  con respecto a su segunda entrada, es decir, para  $\lambda$  una constante real,  $\theta_k^i$  verifica que

$$\theta_k^i(A, \lambda B) = \lambda^i \theta_k^i(A, B). \quad (5.16)$$

*Demostración.* Primero probaremos que se tiene la igualdad (5.15), para ello procedemos por inducción. Consideremos  $k = 0$ , así utilizando la propiedad (5.14a) tenemos que

$$(A + d_B)^0 I = I = A^0 = \theta_0^0 = \sum_{i=0}^0 \theta_0^i(A, B).$$

Luego asumimos que se tiene la igualdad (5.15) para  $k > 0$ . Así por las propiedades (5.14) tenemos que

$$\begin{aligned} (A + d_B)^{k+1} I &= (A + d_B)(A + d_B)^k I = (A + d_B) \left( \sum_{i=0}^k \theta_k^i(A, B) \right) \\ &= \sum_{i=0}^k \left\{ A\theta_k^i(A, B) + d_B\theta_k^i(A, B) \right\} \\ &= \sum_{i=0}^{k-1} \left\{ A\theta_k^i(A, B) + d_B\theta_k^i(A, B) \right\} \quad (\text{Por la propiedad (5.14b)}) \\ &= A\theta_k^0(A, B) + \sum_{i=1}^{k-1} A\theta_k^i(A, B) + \sum_{i=0}^{k-2} d_B\theta_k^i(A, B) + d_B\theta_k^{k-1}(A, B) \\ &= A\theta_k^0(A, B) + \sum_{i=1}^{k-1} \left\{ A\theta_k^i(A, B) + d_B\theta_k^{i-1}(A, B) \right\} + d_B\theta_k^{k-1}(A, B). \end{aligned} \quad (5.17)$$

Ahora observemos que si tomamos  $i = k$  en (5.14c) y utilizando la propiedad (5.14b) tenemos que

$$\theta_{k+1}^k(A, B) = A\theta_k^k(A, B) + d_B\theta_k^{k-1}(A, B) = d_B\theta_k^{k-1}(A, B).$$

Además, desde (5.14a) es claro que  $A\theta_k^0(A, B) = \theta_{k+1}^0(A, B)$ . Así juntando lo anterior y utilizando nuevamente la propiedad (5.14c) en (5.17) tenemos que

$$\begin{aligned} (A + d_B)^{k+1} I &= A\theta_k^0(A, B) + \sum_{i=1}^{k-1} \left\{ A\theta_k^i(A, B) + d_B\theta_k^{i-1}(A, B) \right\} + d_B\theta_k^{k-1}(A, B) \\ &= \theta_{k+1}^0(A, B) + \sum_{i=1}^{k-1} \theta_{k+1}^i(A, B) + \theta_{k+1}^k(A, B) \\ &= \sum_{i=0}^k \theta_{k+1}^i(A, B) + \theta_{k+1}^{k+1}(A, B) = \sum_{i=0}^{k+1} \theta_{k+1}^i(A, B). \end{aligned}$$

Probando así la igualdad (5.17). Ahora para probar la homogeneidad de  $\theta_k^i$  procedemos también por inducción. Primero observamos que para  $k \geq 0$  fijo pero arbitrario, si  $i = 0$ , entonces  $\theta_k^0(A, \lambda B) = A^k = \lambda^0 \theta_k^0(A, B)$ . Antes de continuar observemos que la función  $d_B$  satisface que

$d_{\lambda B} = \lambda d_B$ . En efecto, sea  $M$  una matriz arbitraria, así desde la definición tenemos que

$$d_{\lambda B}M = [\lambda B, M] = \lambda BM - M(\lambda B) = \lambda(BM - MB) = \lambda d_B M.$$

Ahora suponemos la propiedad para  $i$ , es decir

$$\theta_k^i(A, \lambda B) = \lambda^i \theta_k^i(A, B).$$

Así para  $i + 1 \leq k$ , primero observemos que si  $i + 1 = k$  entonces por la propiedad (5.14b) tenemos que

$$\theta_k^k(A, \lambda B) = 0 = \lambda^k \theta_k^k(A, B).$$

Por otro lado, si  $i + 1 \leq k$ , entonces existe  $m > 0$  tal que  $i + 1 = k - m$ . Así aplicando la propiedad (5.14c) recursivamente  $m$  veces tenemos que

$$\begin{aligned} \theta_k^{i+1}(A, \lambda B) &= A\theta_{k-1}^{i+1}(A, \lambda B) + d_{\lambda B}\theta_{k-1}^i(A, \lambda B) \\ &= A\left(A\theta_{k-2}^{i+1}(A, \lambda B) + d_B\theta_{k-2}^i(A, \lambda B)\right) + d_{\lambda B}\theta_{k-1}^i(A, \lambda B) \\ &= A^2\theta_{k-2}^{i+1}(A, \lambda B) + Ad_B\theta_{k-2}^i(A, \lambda B) + d_{\lambda B}\theta_{k-1}^i(A, \lambda B) \\ &= A^2\theta_{k-2}^{i+1}(A, \lambda B) + \sum_{j=1}^2 A^{j-1}d_{\lambda B}\theta_{k-j}^i(A, \lambda B) \\ &= A^2\left(A\theta_{k-3}^{i+1}(A, \lambda B) + d_{\lambda B}\theta_{k-3}^i(A, \lambda B)\right) + \sum_{j=1}^2 A^{j-1}d_{\lambda B}\theta_{k-j}^i(A, \lambda B) \\ &= A^3\theta_{k-3}^{i+1}(A, \lambda B) + A^2d_{\lambda B}\theta_{k-3}^i(A, \lambda B) + \sum_{j=1}^2 A^{j-1}d_{\lambda B}\theta_{k-j}^i(A, \lambda B) \\ &= A^3\theta_{k-3}^{i+1}(A, \lambda B) + \sum_{j=1}^3 A^{j-1}d_{\lambda B}\theta_{k-j}^i(A, \lambda B) \\ &\quad \vdots \\ &= A^m\theta_{k-m}^{i+1}(A, \lambda B) + \sum_{j=1}^m A^{j-1}d_{\lambda B}\theta_{k-j}^i(A, \lambda B). \end{aligned} \tag{5.18}$$

Como  $i + 1 = k - m$ , usando la propiedad (5.14b) tenemos que  $\theta_{k-m}^{i+1}(A, \lambda B) = 0$ . Por tanto, volviendo a (5.18) y utilizando la hipótesis de inducción tenemos que

$$\begin{aligned}
 \theta_k^{i+1}(A, \lambda B) &= A^m \theta_{k-m}^{i+1}(A, \lambda B) + \sum_{j=1}^m A^{j-1} d_{\lambda B} \theta_{k-j}^i(A, \lambda B) \\
 &= \sum_{j=1}^m A^{j-1} d_{\lambda B} \theta_{k-j}^i(A, \lambda B) \\
 &= \sum_{j=1}^m A^{j-1} \lambda d_B \lambda^i \theta_{k-j}^i(A, B) \\
 &= \lambda^{i+1} \sum_{j=1}^m A^{j-1} d_B \theta_{k-j}^i(A, B) \\
 &= \lambda^{i+1} \left( A^m \theta_{k-m}^{i+1}(A, B) + \sum_{j=1}^m A^{j-1} d_B \theta_{k-j}^i(A, B) \right).
 \end{aligned}$$

Finalmente utilizando el argumento recursivo que utilizamos en (5.18) tenemos que

$$\begin{aligned}
 \theta_k^{i+1}(A, \lambda B) &= \lambda^{i+1} \left( A^m \theta_{k-m}^{i+1}(A, B) + \sum_{j=1}^m A^{j-1} d_B \theta_{k-j}^i(A, B) \right) \\
 &= \lambda^{i+1} \left( A^{m-1} \theta_{k-m-1}^{i+1}(A, B) + \sum_{j=1}^{m-1} A^{j-1} d_B \theta_{k-j}^i(A, B) \right) \\
 &\quad \vdots \\
 &= \lambda^{i+1} \left( A \left( A \theta_{k-2}^{i+1}(A, B) + d_B \theta_{k-2}^i(A, B) \right) + d_B \theta_{k-1}^i(A, B) \right) \\
 &= \lambda^{i+1} \left( A \theta_{k-1}^{i+1}(A, B) + d_B \theta_{k-1}^i(A, B) \right) \\
 &= \lambda^{i+1} \theta_k^{i+1}(A, B).
 \end{aligned}$$

Concluyendo la homogeneidad de grado  $i$  de  $\theta_k^i(A, B)$ . □

Por tanto, si remplazamos (5.15), en (5.13) tenemos que

$$(A + B)^n = \sum_{k=0}^n \sum_{i=0}^k \binom{n}{k} \theta_k^i(A, B) B^{n-k}.$$

En particular utilizando (5.16) se verifica que

$$(A + \lambda B)^n = \sum_{k=0}^n \sum_{i=0}^k \binom{n}{k} \theta_k^i(A, \lambda B) (\lambda B)^{n-k} = \sum_{k=0}^n \sum_{i=0}^k \binom{n}{k} \lambda^{n-k+i} \theta_k^i(A, B) B^{n-k}.$$

Finalmente, la solución de (5.7) viene dada por

$$u(t, \lambda) = e^{-(A+\lambda B)t} x_0 = \sum_{n=0}^{\infty} \frac{(-t)^n}{n!} (A + \lambda B)^n x_0 = \sum_{n=0}^{\infty} \sum_{k=0}^n \sum_{i=0}^k \frac{(-t)^n \lambda^{n-k+i} \theta_k^i(A, B) B^{n-k}}{(n-k)!k!} x_0.$$

Por tanto si asumimos que  $\lambda \sim U([a, b])$  entonces el promedio denotado como  $\bar{u}$  verifica que

$$\bar{u}(t) = \int_a^b u(t, \lambda) d\lambda = \frac{1}{b-a} \sum_{n=0}^{\infty} \sum_{k=0}^n \sum_{i=0}^k \frac{(-t)^n (b^{n-k+i+1} - a^{n-k+i+1}) \theta_k^i(A, B) B^{n-k}}{(n-k)!k!(n-k+i+1)} x_0.$$

Es claro de la expresión anterior podemos separar los términos que se restan en la serie, pues cada serie separada converge por si sola. Además, definiendo la función

$$\Phi(x) = \sum_{n=0}^{\infty} \sum_{k=0}^n \sum_{i=0}^k \frac{(-t)^n x^{n-k+i+1} \theta_k^i(A, B) B^{n-k}}{(n-k)!k!(n-k+i+1)} x_0,$$

tenemos que el promedio puede ser expresado como

$$\bar{u}(t) = \frac{\Phi(b) - \Phi(a)}{b-a}. \quad (5.19)$$

En particular si  $B$  es invertible podemos definir

$$\Phi^*(x) = \sum_{n=0}^{\infty} \sum_{k=0}^n \sum_{i=0}^k \frac{(-t)^{n+1} x^{n-k+i+1} \theta_k^i(A, B) B^{n-k+1}}{(n-k)!k!(n-k+i+1)} x_0,$$

y obtener una expresión similar a la expresión (5.12), como sigue

$$\bar{u}(t) = \frac{B^{-1}}{t(b-a)} \left( \Phi^*(a) - \Phi^*(b) \right). \quad (5.20)$$

**Observación 5.3.1.** 1. Tanto desde la expresión (5.19) como desde la expresión (5.20), tenemos que la solución del sistema se puede expresar como una combinación de funciones asociadas a los valores extremos del parámetro  $\lambda$ . Por tanto, es natural preguntarse si  $\lambda$  tiene una distribución arbitraria es posible expresar el promedio  $\bar{u}$  como combinación de las funciones asociadas a los valores extremos del parámetro  $\lambda$ .

2. Si bien al igual como en caso conmutativo es posible escribir la solución como una combinación

de los valores extremos, no sabemos con certeza que ecuación satisfacen las funciones  $\Phi$  o  $\Phi^*$ . Por lo anterior un interesante problema es intentar averiguar qué dinámica subyace tanto a  $\Phi$  como a  $\Phi^*$ . No obstante, intentar abordar esta pregunta utilizando las herramientas anteriormente desarrolladas, podría implicar una gran dificultad desde el punto de vista del algebraico.

3. Si bien lo desarrollado en esta sección nos entrega una formula pseudo explicita para el promedio, no señalamos que sea la única forma de resolver este problema. Otra posible herramienta para este problema podría ser utilizando la fórmula de Lie (o formula de Trotter). Esta fórmula no señala que

$$e^{A+B} = \left( \lim_{n \rightarrow \infty} e^{\frac{A}{n}} e^{\frac{B}{n}} \right)^n,$$

cuando  $A, B$  son matrices no necesariamente conmutativas. Es más, el Teorema de TrotterKato ([23]) nos señala que es posible aplicar esta fórmula para  $A, B$  operadores no acotados. En particular, lo anterior es aplicando cuando se quiere aproximar soluciones de ecuaciones diferenciales parciales, teniendo aplicaciones numéricas.

4. Observemos que en el ejemplo anterior consideramos un lado derecho nulo. un posible problema a futuro es considerara el lado derecho no nulo e intentar obtener una expresión para el promedio, posiblemente vía la fórmula de Duhamel.
5. De forma análoga como se comentó en el caso de matrices conmutativas, es posible reformular las estimaciones realizadas en esta sección en el caso que  $A, B$  sean operador no acotado, posiblemente utilizando la regularizada de Yoshida. No obstante, como primer paso es necesario reconstruir formalmente las estimaciones anteriores.

### 5.3.2. Ejemplos operadores no conmutativos

En relación al último punto de la observación 5.3.1, de ser posible obtener una expresión como (5.19) o como (5.20) en el caso de operador (posiblemente no acotados) no conmutativos, se podría tener una forma explícita para el promedio de los siguientes sistemas.

1. Ejemplo 1: Ecuación convección difusión.

$$\begin{cases} u_t - \Delta u + \lambda \mathbf{v} \cdot \nabla u = 0 & \text{en } G \times (0, 1), \\ u = f & \text{sobre } \partial G \times (0, 1), \\ u(0) = u_0, \end{cases}$$

donde  $G \subset \mathbb{R}^n$ ,  $\mathbf{v}$  es un campo de velocidades.



2. Ecuación con dos tipos de operadores de difusión.

$$\begin{cases} u_t - \Delta u - \lambda \operatorname{Div}(\sigma(x) \cdot \nabla u) = 0 & \text{en } G \times (0, 1), \\ u = f & \text{sobre } \partial G \times (0, 1), \\ u(0) = u_0. \end{cases}$$

donde  $G \subset \mathbb{R}^n$  y  $\sigma$  una función que depende del espacio.

3. Ejemplo 2: Ecuación de Kuramoto Sivashinsky lineal.

$$\begin{cases} u_t + \Delta^2 u - \lambda \Delta u = 0 & \text{en } G \times (0, 1), \\ u = \Delta u = f & \text{sobre } \partial G \times (0, 1), \\ u(0) = u_0, \end{cases}$$

donde  $\Delta^2$  denota el bi-laplaciano y  $f$  una función no nula.

Como podemos ver en este último ejemplo en comparación a al caso conmutativo, en general cuando las condiciones de frontera son no nulas, los operadores no conmutan.

## 5.4. Ecuación de Burger's

En relación con ecuaciones diferenciales no lineales, en esta sección analizaremos la ecuación de Burger's sin viscosidad para distintos tipos de condiciones iniciales. La ecuación de Burger's corresponde a una ecuación relacionada con las leyes de conservación, usualmente utilizada para modelar fenómenos de transporte en mecánica de fluidos, en acústica no lineal, en dinámica de gases, etc. Al ser una ecuación relacionada con las leyes de conservación en medios que fluyen, es de esperar que usualmente se tiene una familia de flujos, en donde los miembros de esta familia pueden tener velocidades distintas, y solo se conoce el rango en donde se mueve esta.

Para introducir el problema consideremos  $\lambda \sim U([0, 1])$  y consideramos  $u_\lambda$  la solución de la siguiente ecuación de Burger's

$$\begin{cases} u_t + \lambda u u_x = 0 & (t, x) \in (0, T) \times \mathbb{R} \\ u(0, x) = u^0(x) & \text{en } x \in \mathbb{R}. \end{cases} \quad (5.21)$$

Además, definimos su promedio como

$$y(t, x) = \int_0^1 u_\lambda(t, x) d\lambda.$$

En lo que sigue nos enfocaremos en encontrar una expresión para el promedio.

### 5.4.1. Promedio en presencia de Choque

Consideremos el problema (5.21) con condición inicial

$$u^0(x) = \begin{cases} u_L & x \leq 0, \\ u_R & x > 0. \end{cases}$$

En este caso asumimos que  $u_L > u_R \geq 0$ , es decir, estamos en presencia de choque ([14, Capítulo 3]). Definimos la velocidad de choque como  $s = \frac{u_R + u_L}{2}$ , así tenemos

$$u_\lambda(t, x) = \begin{cases} u_L & x \leq \lambda st, \\ u_R & x > \lambda st. \end{cases}$$

Por tanto, el promedio corresponde a

$$y(t, x) = \begin{cases} u_L & x \leq 0, \\ f(t, x) & x \in (0, st], \\ u_R & x > st, \end{cases}$$

donde

$$f(t, x) = \int_0^1 \left( u_L \chi_{\left\{ \frac{x}{st} \leq \lambda \right\}} + u_R \chi_{\left\{ \frac{x}{st} > \lambda \right\}} \right) d\lambda = \int_0^{\frac{x}{st}} u_L d\lambda + \int_{\frac{x}{st}}^1 u_R d\lambda = u_L \left( 1 - \frac{x}{st} \right) + u_R \frac{x}{st}.$$

Finalmente juntando todo lo anterior para este caso tenemos que

$$y(t, x) = \begin{cases} u_L & x \leq 0, \\ u_L \left( 1 - \frac{x}{st} \right) + u_R \frac{x}{st} & x \in (0, st], \\ u_R & x > st. \end{cases}$$

**Observación 5.4.1.** En este caso observamos que la solución promedio tiene una estructura semejante a la solución de la ecuación de Burger's para cada  $\lambda$ , aun así, en el sector que ocurre el choque la solución promedio pondera las dos ondas, haciendo que el frente de choque en este caso tenga una estructura distinta. Es más, tras tomar promedio la solución promedio adquiere una mayor regularidad siendo en este caso continua.

### 5.4.2. Promedio en presencia de refracción

Ahora consideremos el caso en que ocurre refracción, es decir, suponemos que  $0 < u_L < u_R$ , donde la condición inicial  $u^0$  es dada por

$$u^0(x) = \begin{cases} u_L & x \leq 0, \\ u_R & x > 0. \end{cases}$$

Así tenemos que la solución corresponde a

$$u_\lambda(t, x) = \begin{cases} u_L & x \leq \lambda u_L t, \\ \frac{x}{\lambda t} & x \in (\lambda u_L t, \lambda u_R t], \\ u_R & x > \lambda u_R t, \end{cases} = \begin{cases} u_L & \frac{x}{u_L t} \leq \lambda, \\ \frac{x}{\lambda t}, & \lambda \in \left[ \frac{x}{u_R t}, \frac{x}{u_L t} \right), \\ u_R & \frac{x}{u_R t} > \lambda. \end{cases}$$

Por tanto, el promedio viene dado por

$$y(t, x) = \int_0^1 u_L \chi_{\left\{ \frac{x}{u_L} \leq \lambda \right\}} d\lambda + \int_0^1 \frac{x}{\lambda t} \chi_{\left\{ \lambda \in \left[ \frac{x}{u_R}, \frac{x}{u_L} \right) \right\}} d\lambda + \int_0^1 u_R \chi_{\left\{ \frac{x}{u_R} > \lambda \right\}} d\lambda.$$

Así calculando las integrales anteriores tenemos

$$\int_0^1 u_L \chi_{\left\{ \frac{x}{u_L} \leq \lambda \right\}} d\lambda = \begin{cases} u_L & x \leq 0, \\ u_L \left( 1 - \frac{x}{tu_L} \right) & x \in (0, tu_L], \\ 0, & x > tu_L. \end{cases}$$

$$\int_0^1 \frac{x}{\lambda t} \chi_{\left\{ \lambda \in \left[ \frac{x}{u_R}, \frac{x}{u_L} \right) \right\}} d\lambda = \begin{cases} 0 & x \leq 0, \\ \frac{x}{t} \ln \left( \frac{u_R}{u_L} \right) & x \in (0, tu_L], \\ \frac{x}{t} \ln \left( \frac{tu_R}{x} \right) & x \in (tu_L, tu_R], \\ 0 & x > tu_R. \end{cases}$$

$$\int_0^1 u_R \chi_{\left\{ \frac{x}{u_R} > \lambda \right\}} d\lambda = \begin{cases} 0 & x \leq 0, \\ \frac{x}{t} & x \in (0, tu_R], \\ u_R & x > tu_R. \end{cases}$$

Finalmente, juntando lo anterior obtenemos

$$y(t, x) = \begin{cases} u_L & x \leq 0, \\ u_L + \frac{x}{t} \ln \left( \frac{u_R}{u_L} \right) & x \in (0, tu_L], \\ \frac{x}{t} + \frac{x}{t} \ln \left( \frac{tu_R}{x} \right) & x \in (tu_L, tu_R], \\ u_R & x > tu_R. \end{cases} \quad (5.22)$$

Así realizando una simulación numérica, obtenemos que

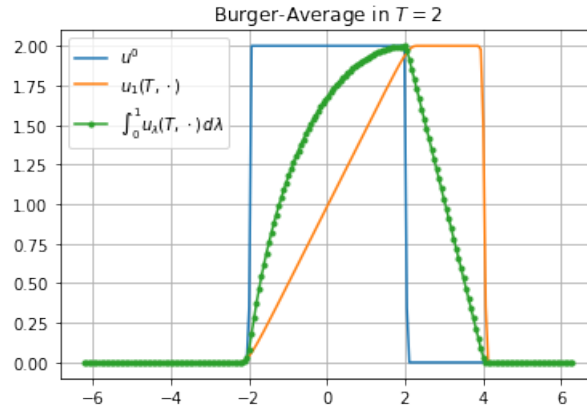


FIGURA 5.1: Refracción de la ecuación de Burger's promedio, con condición inicial  $u^0$ , y  $u_1$  la solución de la ecuación de Burger's con  $\lambda = 1$ .

**Observación 5.4.2.** Observemos que nuevamente el promedio (5.22) es expresado como una combinación de las soluciones asociadas a los valores extremos de  $\lambda$ , siendo este caso a diferencia del anterior, una combinación no lineal. Lo anterior puede ser reafirmado analizando la Figura 5.1.

### 5.4.3. Promedio de una solución suave

Supongamos que  $u^0 \in C^1(0, L)$  y  $t \in (0, T)$  tal que la solución de la ecuación de Burger's no tiene choque, es decir

$$T < \frac{1}{\sup_{x \in \mathbb{R}} \left| \frac{d}{dx} u^0(x) \right|}.$$

Ahora usando el método de las características [14, Capítulo 3] tenemos que la solución de la ecuación viene dada por

$$u_\lambda(t, x) = u^0(y_\lambda), \quad \text{donde } x = y_\lambda + \lambda t u^0(y_\lambda).$$

Por tanto, el promedio corresponde a

$$y(t, x) = \int_0^1 u_\lambda(t, x) d\lambda = \int_0^1 u^0(y_\lambda) d\lambda. \quad (5.23)$$

Además, si derivamos con respecto a  $\lambda$  la relación  $x = y_\lambda + \lambda t u^0(y_\lambda)$  tenemos que

$$\frac{d}{d\lambda} y_\lambda = -\frac{t u^0(y_\lambda)}{1 + \lambda t \frac{d}{dx} u^0(y_\lambda)} = \frac{-t |u^0(y_\lambda)|^2}{u^0(y_\lambda) + (x - y_\lambda) \frac{d}{dx} u^0(y_\lambda)}.$$

Ahora si suponemos que  $u^0 > 0$  y recordemos que  $y_1$  es caracterizado por  $x - y_1 = tu^0(y_1)$ , además, tenemos que  $u_1(t, x) = u^0(y_1)$ . Así haciendo el cambio de variable  $z = y_\lambda$  en (5.23) obtenemos

$$\begin{aligned} y(t, x) &= \int_0^1 u^0(z) \left( \frac{dy_\lambda}{d\lambda} \right)^{-1} dz = \int_{y_1}^x \frac{\left( u^0(z) + (x - z) \frac{d}{dx} u^0(z) \right)}{tu^0(z)} dz \\ &= \frac{x - y_1}{t} + \frac{1}{t} \int_{y_1}^x (x - z) \frac{d}{dz} (\ln(u^0(z))) dz \end{aligned} \quad (5.24)$$

$$= u_1(t, x) + \frac{1}{t} (x - z) (\ln(u^0(z))) \Big|_{y_1}^x + \frac{1}{t} \int_{y_1}^x (\ln(u^0(z))) dz. \quad (5.25)$$

Entonces definiendo

$$F(z) = \int \ln(u^0(z)) dz,$$

tenemos que

$$\begin{aligned} y(t, x) &= u_1(t, x) - \frac{1}{t} (x - y_1) (\ln(u^0(y_1))) + \frac{1}{t} \int_{y_1}^x \frac{d}{dz} (F(z)) dz \\ &= u_1(t, x) - u_1(t, x) \ln(u_1(t, x)) + \frac{F(x) - F(x - tu_1(t, x))}{t}. \end{aligned}$$

Luego si consideramos el operador

$$G(u^0, u_1)(x, t) := u_1(t, x) - u_1(t, x) \ln(u_1(t, x)) + \frac{F(x) - F(x - tu_1(t, x))}{t}, \quad (5.26)$$

obtenemos la siguiente expresión para el promedio

$$y(t, x) = \int_0^1 u_\lambda(t, x) d\lambda = G(u^0, u_1)(t, x). \quad (5.27)$$

**Observación 5.4.3.** Al igual como en los ejemplos anteriores, podemos observar que el promedio corresponde a una combinación de soluciones asociadas a los valores extremos del parámetro  $\lambda$ . Al igual que en caso de refracción, la ponderación de las soluciones para obtener la solución promedio es no lineal.

### Ejemplo (Promedio Solución Suave)

Con intenciones de ejemplificar la efectividad de la relación (5.27), es que presentamos el siguiente breve ejemplo.

Consideremos la condición inicial suave

$$u^0(x) = e^{-\frac{x^2}{\sigma^2}}.$$

En particular se satisface

$$F(x) := \int \ln(u^0(x)) dx = -\frac{x^3}{3\sigma^2}.$$

Obteniendo así, en virtud de la relación (5.27) tenemos que

$$y(t) = G(u^0, u_1)(x, t) := u_1(t, x) - u_1(t, x) \ln(u(t, x)) - \frac{(x^3 - (x - tu_1(t, x))^3)}{3t\sigma^2}.$$

Supongamos que  $\sigma = 2$ , en tal caso no existen choques, es decir, se satisface la condición

$$\frac{1}{\sup_{x \in \mathbb{R}} \left| \frac{d}{dx} u^0(x) \right|} = \frac{e^{\frac{1}{2}}}{\sqrt{2}} \sigma > 2.$$

En consecuencia, podemos considerar  $T = 2$ . Así tenemos la siguiente simulación numérica.

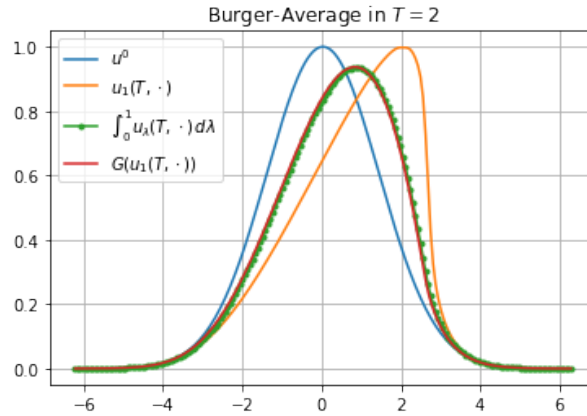


FIGURA 5.2: Reconstrucción del promedio de la ecuación de Burger's con condición inicial  $u^0$ , y  $u_1$  la solución de la ecuación de Burger's con  $\lambda = 1$ .

**Observación 5.4.4.** En la simulación anterior (Figura 5.2), se puede observar que efectivamente la función  $G$  definida en (5.26) se acerca bastante al promedio numérico. Por otro lado, observemos que desde (5.26), la solución promedio puede ser vista con una combinación de las curvas  $u_1$  y  $u^0$ . Lo anterior puede observarse en la simulación numérica de la Figura 5.2.

#### 5.4.4. Otra representación del promedio

Observemos que en el caso general de la ecuación de Burger's (5.21), es posible absorber el parámetro aleatorio a través de un re-escalamiento del tiempo. Para ello podemos consideremos el cambio de variable  $u(t, x) = v(\lambda t, x)$ , en tal caso

$$\begin{cases} u_t = \lambda v_t, \\ u_x = v_x, \\ v(0, x) = u(0, x) = u^0(x). \end{cases}$$

Así  $v(t, x)$  satisface que

$$\begin{cases} v_t + vv_x = 0 & (t, x) \in (0, \lambda T) \times \mathbb{R}, \\ v(0, x) = u^0(x) & \text{en } x \in \mathbb{R}. \end{cases} \quad (5.28)$$

Por otro lado, el promedio verifica que

$$y(t, x) = \int_0^1 u(t, x) d\lambda = \int_0^1 v(\lambda t, x) d\lambda.$$

Así utilizando el cambio de variable  $\lambda t = z$  tenemos que

$$y(t, x) = \frac{1}{t} \int_0^t v(z, x) dz. \quad (5.29)$$

Ahora si derivamos la expresión anterior con respecto a  $x$  tenemos

$$y_x(t, x) = \frac{1}{t} \int_0^t v_x(z, x) dz. \quad (5.30)$$

Luego si  $v(z, x) = 0$  desde (5.29) tenemos que  $y(t, x) = 0$ . En caso contrario, utilizando la ecuación (5.28) tenemos

$$\begin{aligned} y_x(t, x) &= \frac{1}{t} \int_0^t v_x(z, x) dz = \frac{1}{t} \int_0^t \frac{v_t(z, x)}{v(z, x)} dz = \frac{1}{t} \int_0^t \partial_t (\ln(v(z, x))) dz \\ &= \frac{1}{t} \ln \left( \frac{v(0, x)}{v(t, x)} \right). \end{aligned} \quad (5.31)$$

Luego si suponemos que  $v$  tiene soporte compacto, entonces integrando entre  $(-\infty, x)$  la expresión (5.31) tenemos

$$y(t, x) = \frac{1}{t} \int_{-\infty}^x \ln \left( \frac{v(0, z)}{v(t, z)} \right) dz.$$

**Observación 5.4.5.** 1. Notemos que hasta la igualdad (5.29), no hemos utilizado la totalidad estructura de la ecuación que se resuelve  $u$ . En general podemos llegar a la representación (5.29) para el promedio, siempre y cuando el cambio de variable  $u(t, x) = v(\lambda t, x)$  permita quitar el parámetro  $\lambda$  de la ecuación. Esto es posible cuando tenemos una ecuación de la forma

$$\begin{cases} u_t + \lambda F(u, Du) = 0, \\ u(0) = u^0(x). \end{cases}$$

En tal caso, el promedio puede ser representado por (5.29).

2. Como se ha comentado en las otras secciones, queda abierta la pregunta sobre qué ocurre cuando la variable  $\lambda$  no se distribuye uniforme, es más, como bien podemos ver en el análisis anterior, solo consideremos  $\lambda$  en el intervalo  $[0, 1]$ , por tanto, queda abierta la si estos mismos pasos pueden reproducirse cuando  $\lambda \in [a, b]$ .
3. Como bien hemos comentado, tras obtener una expresión para la dinámica promedio esta depende de solo los valores extremos del parámetro aleatorio. Por lo cual nos pregunta cuando, y cuando no ocurre esto, y por supuesto no interesa saber porque ocurre esto. Lo anterior posiblemente es un problema abierto no trivial.
4. A modo de ejemplo de las estimaciones anterior, si consideremos la ecuación de Hamilton-Jacobi en una dimensión con solución  $w^H$ , es decir

$$\begin{cases} w_t^H + \lambda H(w_x^H) = 0 & (t, x) \in (0, T) \times \mathbb{R} \\ w^H(0, x) = w^0(x) & \text{en } x \in \mathbb{R}, \end{cases} \quad (5.32)$$

donde  $H$  es una función convexa llamada Hamiltoniano. Así si suponemos que existe  $F = H_x(v)$ , entonces si definimos  $w_x^H = u$  y derivando la ecuación (5.32) con respecto a  $x$ , tenemos que  $u$  satisface la ecuación de Burger's

$$\begin{cases} u_t + \bar{\lambda} u u_x = 0 & (t, x) \in (0, T) \times \mathbb{R}. \\ u(0, x) = w_x^0(x) & \text{en } x \in \mathbb{R}. \end{cases} \quad (5.33)$$

Por tanto, utilizando (5.30) y suponiendo que  $w$  tiene soporte compacto, tenemos que el promedio de la ecuación de Burger's (5.33), el cual denotaremos como  $y^B$ , puede ser representado mediante la solución de la ecuación de Hamilton-Jacobi como sigue

$$y^B(t, x) = \frac{1}{t} \int_{-\infty}^x \int_0^t w^H(z, v) dz dv.$$

Es decir, podemos recuperar una representación para el promedio de la ecuación de Burger's, a través de una ponderación no local de la solución de la ecuación de Hamilton-Jacobi.



5. Esta sección motiva el análisis y caracterización de la dinámica promedio de la ecuación de Burger's no viscosa. Otro posible problema podría ser analizar la versión con viscosidad

$$\begin{cases} u_t - \lambda u_{xx} + uu_x = 0 & (t, x) \in (0, T) \times \mathbb{R}, \\ u(0, x) = u^0(x) & \text{en } x \in \mathbb{R}, \end{cases}$$

ecuación cuya solución se puede tener de forma explícita vía la transformación Hopf-Cole [14, Capítulo 4.4]. Para tener una expresión explícita para el promedio de esta ecuación, podría pensarse en tomar promedio a la solución explícita, no obstante, esto tiene una gran dificultad al momento de realizar las estimaciones.

6. Dentro de los distintos intentó para resolver el problema de analizar la dinámica promedio de la ecuación Burger's, no se analizó con mucho detalle la caracterización vía el semigrupo no lineal asociado a esta ecuación. En caso de tener una expresión u herramientas para resolver este problema vía semigrupos se podría generalizar a campos gradientes [14, Capítulo 9.6].



# Bibliografía

- [1] Mouna Abdelli y Abdelhak Hafdallah. «Averaged null controllability for some hyperbolic equations depending on a parameter». En: *J. Math. Anal. Appl.* 495.1 (2021), págs. 124697, 19. ISSN: 0022-247X. DOI: [10.1016/j.jmaa.2020.124697](https://doi.org/10.1016/j.jmaa.2020.124697). URL: <https://doi-org.usm.idm.oclc.org/10.1016/j.jmaa.2020.124697>.
- [2] Francesca Albertini y Eduardo D. Sontag. «For neural networks, function determines form». En: *Neural Networks* 6.7 (1993), págs. 975 -990. ISSN: 0893-6080. DOI: [https://doi.org/10.1016/S0893-6080\(09\)80007-5](https://doi.org/10.1016/S0893-6080(09)80007-5). URL: <http://www.sciencedirect.com/science/article/pii/S0893608009800075>.
- [3] Logan D. R. Beal y col. «GEKKO Optimization Suite». En: *Processes* 6.8 (2018). ISSN: 2227-9717. DOI: [10.3390/pr6080106](https://doi.org/10.3390/pr6080106). URL: <https://www.mdpi.com/2227-9717/6/8/106>.
- [4] Alain Bensoussan y col. *Representation and control of infinite dimensional systems*. Second. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc. Boston. MA., 2007, págs. xxviii+575. ISBN: 978-0-8176-4461-1; 0-8176-4461-X. DOI: [10.1007/978-0-8176-4581-6](https://doi.org/10.1007/978-0-8176-4581-6). URL: <https://doi-org.usm.idm.oclc.org/10.1007/978-0-8176-4581-6>.
- [5] Lorenz T. Biegler. «An overview of simultaneous strategies for dynamic optimization». En: *Chemical Engineering and Processing: Process Intensification* 46.11 (2007). Special Issue on Process Optimization and Control in Chemical Engineering and Processing, págs. 1043 -1053. ISSN: 0255-2701. DOI: <https://doi.org/10.1016/j.cep.2006.06.021>. URL: <http://www.sciencedirect.com/science/article/pii/S0255270107001122>.
- [6] Tobias Breiten y Laurent Pfeiffer. «On the Turnpike Property and the Receding-Horizon Method for Linear-Quadratic Optimal Control Problems». En: *SIAM Journal on Control and Optimization* 58 (ene. de 2020), págs. 1077-1102. DOI: [10.1137/18M1225811](https://doi.org/10.1137/18M1225811).
- [7] Haim Brezis. *Function Analysis, Sobolev Spaces and Partial Differential Equations*. Ene. de 2010. ISBN: 978-0-387-70913-0. DOI: [10.1007/978-0-387-70914-7](https://doi.org/10.1007/978-0-387-70914-7).
- [8] G.F. Carey y Bruce A. Finlayson. «Orthogonal collocation on finite elements». En: *Chemical Engineering Science* 30.5 (1975), págs. 587 -596. ISSN: 0009-2509. DOI: [https://doi.org/10.1016/0009-2509\(75\)80031-5](https://doi.org/10.1016/0009-2509(75)80031-5). URL: <http://www.sciencedirect.com/science/article/pii/0009250975800315>.
- [9] T. Cazenave y col. *An Introduction to Semilinear Evolution Equations*. Oxford lecture series in mathematics and its applications. Clarendon Press, 1998. ISBN: 9780198502777. URL: <https://books.google.cl/books?id=pcBo9WZBHyMC>.

- [10] J.M. Coron y American Mathematical Society. *Control and Nonlinearity*. Mathematical surveys and monographs. American Mathematical Society, 2007. ISBN: 9780821849187. URL: <https://books.google.cl/books?id=R5h9X2yYKqkC>.
- [11] Jeremy Coulson, Bahman Ghahsifard y Abdol-Reza Mansouri. «On average controllability of random heat equations with arbitrarily distributed diffusivity». En: *Automatica* 103 (2019), págs. 46 -52. ISSN: 0005-1098. DOI: <https://doi.org/10.1016/j.automatica.2019.01.014>.
- [12] Robert Dorfman, Paul A. Samuelson y Robert M. Solow. *Linear programming and economic analysis*. A Rand Corporation Research Study. McGraw-Hill Book Co., Inc., New York-Toronto-London, 1958, págs. ix+525.
- [13] Carlos Esteve y col. *Turnpike in Lipschitz-nonlinear optimal control*. 2020. arXiv: [2011.11091 \[math.OC\]](https://arxiv.org/abs/2011.11091).
- [14] Lawrence C. Evans. *Partial differential equations*. Providence, R.I.: American Mathematical Society, 2010. ISBN: 9780821849743 0821849743.
- [15] T. Faulwasser y col. «The interval turnpike property for adjoints». En: *arXiv: Optimization and Control* (2020).
- [16] Timm Faulwasser y col. «On turnpike and dissipativity properties of continuous-time optimal control problems». En: *Automatica* 81 (2017), págs. 297 -304. ISSN: 0005-1098. DOI: <https://doi.org/10.1016/j.automatica.2017.03.012>. URL: <http://www.sciencedirect.com/science/article/pii/S0005109817301358>.
- [17] Borjan Geshkovski y col. «Large-time asymptotics in deep learning». En: (ago. de 2020).
- [18] Ian Goodfellow, Yoshua Bengio y Aaron Courville. *Deep learning*. Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA, 2016, págs. xxii+775. ISBN: 978-0-262-03561-3.
- [19] Lars Grüne, Manuel Schaller y Anton Schiela. «Sensitivity analysis of optimal control for a class of parabolic PDEs motivated by model predictive control». En: *SIAM J. Control Optim.* 57.4 (2019), págs. 2753-2774. ISSN: 0363-0129. DOI: [10.1137/18M1223083](https://doi-org.usm.idm.oclc.org/10.1137/18M1223083). URL: <https://doi-org.usm.idm.oclc.org/10.1137/18M1223083>.
- [20] Lars Grüne, Manuel Schaller y Anton Schiela. «Exponential sensitivity and turnpike analysis for linear quadratic optimal control of general evolution equations». En: *Journal of Differential Equations* 268.12 (2020), págs. 7311 -7341. ISSN: 0022-0396. DOI: <https://doi.org/10.1016/j.jde.2019.11.064>. URL: <http://www.sciencedirect.com/science/article/pii/S0022039619305984>.
- [21] Bernhard Haak, Duc-Trung Hoang y El-Maati Ouhabaz. «Controllability and observability for non-autonomous evolution equations: the averaged Hautus test». En: *Systems Control Lett.* 133 (2019), págs. 104524, 8. ISSN: 0167-6911. DOI: [10.1016/j.sysconle.2019.104524](https://doi-org.usm.idm.oclc.org/10.1016/j.sysconle.2019.104524). URL: <https://doi-org.usm.idm.oclc.org/10.1016/j.sysconle.2019.104524>.
- [22] Aitziber Ibañez. «Optimal control of the Lotka-Volterra system: Turnpike property and numerical simulations». En: *Journal of Biological Dynamics* 11 (dic. de 2017), págs. 25-41. DOI: [10.1080/17513758.2016.1226435](https://doi.org/10.1080/17513758.2016.1226435).

- [23] Kazufumi Ito y Franz Kappel. «The Trotter-Kato theorem and approximation of PDEs». En: *Math. Comp.* 67.221 (1998), págs. 21-44. ISSN: 0025-5718. DOI: [10.1090/S0025-5718-98-00915-6](https://doi-org.usm.idm.oclc.org/10.1090/S0025-5718-98-00915-6). URL: <https://doi-org.usm.idm.oclc.org/10.1090/S0025-5718-98-00915-6>.
- [24] Rafail Khasminskii. *Stochastic stability of differential equations*. second. Vol. 66. Stochastic Modelling and Applied Probability. With contributions by G. N. Milstein and M. B. Nevelson. Springer, Heidelberg, 2012, págs. xviii+339. ISBN: 978-3-642-23279-4. DOI: [10.1007/978-3-642-23280-0](https://doi-org.usm.idm.oclc.org/10.1007/978-3-642-23280-0). URL: <https://doi-org.usm.idm.oclc.org/10.1007/978-3-642-23280-0>.
- [25] Yann Lecun. «A theoretical framework for back-propagation». English (US). En: *Proceedings of the 1988 Connectionist Models Summer School, CMU, Pittsburg, PA*. Ed. por D. Touretzky, G. Hinton y T. Sejnowski. Morgan Kaufmann, 1988, págs. 21-28.
- [26] J. Lions. «Contrôle à moindres regrets des systèmes distribués». En: 1992.
- [27] Jacques Louis Lions. «Optimal Control of Systems Governed by Partial Differential Equations». En: 170 (ene. de 1971), pág. 400. ISSN: 0072-7830.
- [28] Jérôme Lohéac y Enrique Zuazua. «Averaged controllability of parameter dependent conservative semigroups». En: *Journal of Differential Equations* 262 (nov. de 2016). DOI: [10.1016/j.jde.2016.10.017](https://doi-org.usm.idm.oclc.org/10.1016/j.jde.2016.10.017).
- [29] Jérôme Lohéac y Enrique Zuazua. «From averaged to simultaneous controllability of parameter dependent finite-dimensional systems». En: (jun. de 2015).
- [30] Qi Lü y Enrique Zuazua. «Averaged controllability for random evolution Partial Differential Equations». En: *Journal de Mathématiques Pures et Appliquées* 105 (ene. de 2015). DOI: [10.1016/j.matpur.2015.11.004](https://doi-org.usm.idm.oclc.org/10.1016/j.matpur.2015.11.004).
- [31] Qi Lü y Enrique Zuazua. «Averaged controllability for random evolution Partial Differential Equations». En: *Journal de Mathématiques Pures et Appliquées* 105 (ene. de 2015). DOI: [10.1016/j.matpur.2015.11.004](https://doi-org.usm.idm.oclc.org/10.1016/j.matpur.2015.11.004).
- [32] Francisco J. Marín, Jesús Martínez-Frutos y Francisco Periago. «Robust averaged control of vibrations for the Bernoulli-Euler beam equation». En: *J. Optim. Theory Appl.* 174.2 (2017), págs. 428-454. ISSN: 0022-3239. DOI: [10.1007/s10957-017-1128-x](https://doi-org.usm.idm.oclc.org/10.1007/s10957-017-1128-x). URL: <https://doi-org.usm.idm.oclc.org/10.1007/s10957-017-1128-x>.
- [33] J. Martínez-Frutos y col. «Robust optimal Robin boundary control for the transient heat equation with random input data». En: *Internat. J. Numer. Methods Engrg.* 108.2 (2016), págs. 116-135. ISSN: 0029-5981. DOI: [10.1002/nme.5210](https://doi-org.usm.idm.oclc.org/10.1002/nme.5210). URL: <https://doi-org.usm.idm.oclc.org/10.1002/nme.5210>.
- [34] Jesús Martínez-Frutos y Francisco Periago Esparza. *Optimal control of PDEs under uncertainty*. SpringerBriefs in Mathematics. An introduction with application to optimal shape design of structures, BCAM SpringerBriefs. Springer, Cham; BCAM Basque Center for Applied Mathematics, Bilbao, 2018, págs. xix+122. ISBN: 978-3-319-98209-0; 978-3-319-98210-6. DOI: [10.1007/978-3-319-98210-6](https://doi-org.usm.idm.oclc.org/10.1007/978-3-319-98210-6). URL: <https://doi-org.usm.idm.oclc.org/10.1007/978-3-319-98210-6>.

- [35] R. Kress, V. Mazya, y V. Kozlov. «Linear integral equations». En: *Applied Mathematical Sciences*. Springer, New York, third edition. 82 (dic. de 1989), págs. 25-41. DOI: [10.1007/978-1-4614-9593-2](https://doi.org/10.1007/978-1-4614-9593-2).
- [36] Warren S. McCulloch y Walter Pitts. «A logical calculus of the ideas immanent in nervous activity». En: *Bull. Math. Biophys.* 5 (1943), págs. 115-133. ISSN: 0007-4985. DOI: [10.1007/bf02478259](https://doi.org/10.1007/bf02478259). URL: <https://doi-org.usm.idm.oclc.org/10.1007/bf02478259>.
- [37] Anthony N. Michel, Jay A. Farrell y Wolfgang Porod. «Qualitative analysis of neural networks». En: *IEEE Trans. Circuits and Systems* 36.2 (1989), págs. 229-243. ISSN: 0098-4094. DOI: [10.1109/31.20200](https://doi.org/10.1109/31.20200). URL: <https://doi-org.usm.idm.oclc.org/10.1109/31.20200>.
- [38] Darko Mitrovic, Andrej Novak y Tarik Uzunović. «Averaged control for fractional ODEs and fractional diffusion equations». En: *J. Funct. Spaces* (2018), Art. ID 8095728, 8. ISSN: 2314-8896. DOI: [10.1155/2018/8095728](https://doi.org/10.1155/2018/8095728). URL: <https://doi-org.usm.idm.oclc.org/10.1155/2018/8095728>.
- [39] Darko Mitrovi, Andrej Novak y Tarik Uzunovic. «Averaged Control for Fractional ODEs and Fractional Diffusion Equations». En: *Journal of Function Spaces* 2018 (jul. de 2018), págs. 1-8. DOI: [10.1155/2018/8095728](https://doi.org/10.1155/2018/8095728).
- [40] Gisèle Mophou, Romario Gildas Foko Tiomela y Ali Seibou. «Optimal control of averaged state of a parabolic equation with missing boundary condition». En: *Internat. J. Control* 93.10 (2020), págs. 2358-2369. ISSN: 0020-7179. DOI: [10.1080/00207179.2018.1556810](https://doi.org/10.1080/00207179.2018.1556810). URL: <https://doi-org.usm.idm.oclc.org/10.1080/00207179.2018.1556810>.
- [41] J.R. Munkres. *Topology*. Featured Titles for Topology. Prentice Hall, Incorporated, 2000. ISBN: 9780131816299. URL: <https://books.google.cl/books?id=XjoZAQAAIAAJ>.
- [42] J. v. Neumann. «A Model of General Economic Equilibrium 1». En: *The Review of Economic Studies* 13.1 (abr. de 1945), págs. 1-9. ISSN: 0034-6527. DOI: [10.2307/2296111](https://doi.org/10.2307/2296111). eprint: <https://academic.oup.com/restud/article-pdf/13/1/1/6932720/13-1-1.pdf>. URL: <https://doi.org/10.2307/2296111>.
- [43] Fernando J. Pineda. «Generalization of back-propagation to recurrent neural networks». En: *Phys. Rev. Lett.* 59.19 (1987), págs. 2229-2232. ISSN: 0031-9007. DOI: [10.1103/PhysRevLett.59.2229](https://doi.org/10.1103/PhysRevLett.59.2229). URL: <https://doi-org.usm.idm.oclc.org/10.1103/PhysRevLett.59.2229>.
- [44] Alessio Porretta y Enrique Zuazua. «Long time versus steady state optimal control». En: *SIAM J. Control Optim.* 51.6 (2013), págs. 4242-4273. ISSN: 0363-0129. DOI: [10.1137/130907239](https://doi.org/10.1137/130907239). URL: <https://doi-org.usm.idm.oclc.org/10.1137/130907239>.
- [45] Alessio Porretta y Enrique Zuazua. «Remarks on Long Time Versus Steady State Optimal Control». En: vol. 15. Nov. de 2016, págs. 67-89. ISBN: 978-3-319-39091-8. DOI: [10.1007/978-3-319-39092-5\\_5](https://doi.org/10.1007/978-3-319-39092-5_5).
- [46] E. D. Sontag e Y. Qiao. «Further results on controllability of recurrent neural networks». En: *Systems Control Lett.* 36.2 (1999), págs. 121-129. ISSN: 0167-6911. DOI: [10.1016/S0167-6911\(98\)00068-1](https://doi.org/10.1016/S0167-6911(98)00068-1). URL: [https://doi-org.usm.idm.oclc.org/10.1016/S0167-6911\(98\)00068-1](https://doi-org.usm.idm.oclc.org/10.1016/S0167-6911(98)00068-1).
- [47] Eduardo Sontag. *Mathematical Control Theory: Deterministic Finite-Dimensional Systems*. Ene. de 1998. DOI: [10.1007/978-1-4612-0577-7](https://doi.org/10.1007/978-1-4612-0577-7).

- [48] T. T. Soong. *Random differential equations in science and engineering*. Mathematics in Science and Engineering, Vol. 103. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1973, págs. xiii+327.
- [49] Emmanuel Trelat y Can Zhang. «Integral and measure-turnpike properties for infinite-dimensional optimal control systems». En: *Mathematics of Control, Signals, and Systems* 30 (mayo de 2017). DOI: [10.1007/s00498-018-0209-1](https://doi.org/10.1007/s00498-018-0209-1).
- [50] Emmanuel Trelat, Can Zhang y Enrique Zuazua. «Steady-State and Periodic Exponential Turnpike Property for Optimal Control Problems in Hilbert Spaces». En: *SIAM Journal on Control and Optimization* 56 (oct. de 2016). DOI: [10.1137/16M1097638](https://doi.org/10.1137/16M1097638).
- [51] Emmanuel Trélat y Enrique Zuazua. «The turnpike property in finite-dimensional nonlinear optimal control». En: *Journal of Differential Equations* 258.1 (2015), págs. 81 -114. ISSN: 0022-0396. DOI: <https://doi.org/10.1016/j.jde.2014.09.005>. URL: <http://www.sciencedirect.com/science/article/pii/S0022039614003568>.
- [52] Fredi Tröltzsch. «Optimal control of partial differential equations. Theory, methods and applications». En: 112 (ene. de 2010). DOI: [10.1090/gsm/112](https://doi.org/10.1090/gsm/112).
- [53] Walter Wyss. «Quasi-multiplication and \*-algebras». En: *Reports on Mathematical Physics* 18.1 (1980), págs. 87 -95. ISSN: 0034-4877. DOI: [https://doi.org/10.1016/0034-4877\(80\)90041-5](https://doi.org/10.1016/0034-4877(80)90041-5). URL: <http://www.sciencedirect.com/science/article/pii/0034487780900415>.
- [54] Wenhuan Yu, Fengling Zhang y Weisong Xie. «Differentiability of C0-semigroups with respect to parameters and its application». En: *Journal of Mathematical Analysis and Applications* 279.1 (2003), págs. 78 -96. ISSN: 0022-247X. DOI: [https://doi.org/10.1016/S0022-247X\(02\)00440-7](https://doi.org/10.1016/S0022-247X(02)00440-7). URL: <http://www.sciencedirect.com/science/article/pii/S0022247X02004407>.
- [55] Sebastian Zamorano. «Turnpike Property for Two-Dimensional Navier-Stokes Equations». En: *Journal of Mathematical Fluid Mechanics* 20 (jul. de 2018), págs. 869-888.
- [56] A. Zaslavski. *Turnpike Properties in the Calculus of Variations and Optimal Control*. Nonconvex Optimization and Its Applications. Springer US, 2005. ISBN: 9780387281551. URL: <https://books.google.com.jm/books?id=Svaze6jHZFYC>.
- [57] Eberhard Zeidler. *Nonlinear functional analysis and its applications*. II/A. Linear monotone operators, Translated from the German by the author and Leo F. Boron. Springer-Verlag, New York, 1990, págs. xviii+467. ISBN: 0-387-96802-4. DOI: [10.1007/978-1-4612-0985-0](https://doi.org/10.1007/978-1-4612-0985-0). URL: <https://doi-org.usm.idm.oclc.org/10.1007/978-1-4612-0985-0>.
- [58] Enrique Zuazua. «Averaged control». En: *Automatica J. IFAC* 50.12 (2014), págs. 3077-3087. ISSN: 0005-1098. DOI: [10.1016/j.automatica.2014.10.054](https://doi.org/10.1016/j.automatica.2014.10.054). URL: <https://doi-org.usm.idm.oclc.org/10.1016/j.automatica.2014.10.054>.
- [59] Enrique Zuazua. «Large time control and turnpike properties for wave equations». En: *Annual Reviews in Control* 44 (2017), págs. 199 -210. ISSN: 1367-5788. DOI: <https://doi.org/10.1016/j.arcontrol.2017.04.002>. URL: <http://www.sciencedirect.com/science/article/pii/S1367578817300330>.