

INSTITUTO POTOSINO DE INVESTIGACIÓN CIENTÍFICA Y TECNOLÓGICA

DIVISIÓN DE MATEMÁTICAS APLICADAS



TÍTULO DE LA TESIS:

**Cotas en la rapidez de convergencia en el
Teorema Ergódico de Birkhoff en sistemas
dinámicos simbólicos con medida de Gibbs**

QUE PARA OBTENER EL GRADO DE:

Maestro en Control y Sistemas Dinámicos

PRESENTA:

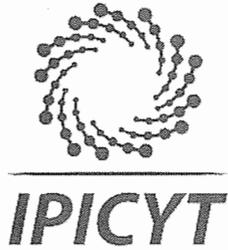
Hugo Alberto Nieto Loredo

DIRECTOR DE LA TESIS:

Dr. César Maldonado Ahumada

San Luis Potosí, S.L.P.

31 de julio de 2019



Constancia de aprobación de la tesis

La tesis "***Cotas en la rapidez de convergencia en el Teorema Ergódico de Birkhoff en sistemas dinámicos simbólicos con medida de Gibbs***" presentada para obtener el Grado de Maestro en Control y Sistemas Dinámicos, fue elaborada por **Hugo Alberto Nieto Loredo** y aprobada el **treinta y uno de julio del dos mil diecinueve** por los suscritos, designados por el Colegio de Profesores de la División de Matemáticas Aplicadas del Instituto Potosino de Investigación Científica y Tecnológica, A.C.

Dr. Cesar Octavio Maldonado Ahumada
Director de la tesis

Dr. Hugo Cabrera Ibarra
Jurado en el Examen

Dr. David Antonio Lizárraga Navarro
Jurado en el Examen

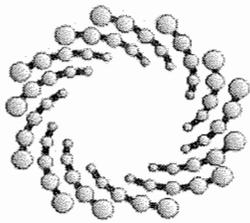
Dr. Haret-Codratian Rosu Barbus
Jurado en el Examen



Créditos institucionales

Esta tesis fue elaborada en la División de Matemáticas Aplicadas del Instituto Potosino de Investigación Científica y Tecnológica, A.C., bajo la dirección del Dr. César Octavio Maldonado Ahumada.

Durante la realización del trabajo el autor recibió una beca académica del Consejo Nacional de Ciencia y Tecnología (862351) y del Instituto Potosino de Investigación Científica y Tecnológica, A. C.



IPICYT

Instituto Potosino de Investigación Científica y Tecnológica, A.C.

Acta de Examen de Grado

El Secretario Académico del Instituto Potosino de Investigación Científica y Tecnológica, A.C., certifica que en el Acta 039 del Libro Primero de Actas de Exámenes de Grado del Programa de Maestría en Control y Sistemas Dinámicos está asentado lo siguiente:

En la ciudad de San Luis Potosí a los 31 días del mes de julio del año 2019, se reunió a las 11:00 horas en las instalaciones del Instituto Potosino de Investigación Científica y Tecnológica, A.C., el Jurado integrado por:

Dr. Haret-Codratian Rosu Barbus	Presidente	IPICYT
Dr. Cesar Octavio Maldonado Ahumada	Secretario	IPICYT
Dr. David Antonio Lizárraga Navarro	Sinodal	IPICYT
Dr. Hugo Cabrera Ibarra	Sinodal	IPICYT

a fin de efectuar el examen, que para obtener el Grado de:

MAESTRO EN CONTROL Y SISTEMAS DINÁMICOS

sustentó el C.

Hugo Alberto Nieto Loredo

sobre la Tesis intitulada:

Cotas en la rapidez de convergencia en el Teorema Ergódico de Birkhoff en sistemas dinámicos simbólicos con medida de Gibbs

que se desarrolló bajo la dirección de

Dr. Cesar Octavio Maldonado Ahumada

El Jurado, después de deliberar, determinó

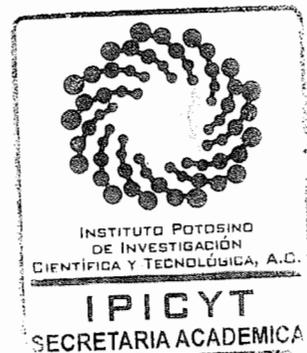
APROBARLO

Dándose por terminado el acto a las 12:40 horas, procediendo a la firma del Acta los integrantes del Jurado. Dando fe el Secretario Académico del Instituto.

A petición del interesado y para los fines que al mismo convengan, se extiende el presente documento en la ciudad de San Luis Potosí, S.L.P., México, a los 31 días del mes de julio de 2019.

Dr. Marcial Bonilla Marín
Secretario Académico

Mtra. Ivonne Lizette Cuevas Vélez
Jefa del Departamento del Posgrado



Contenido

Constancia de aprobación de la tesis	Página I
Créditos institucionales	Página II
Acta de examen	Página III
Contenido	Página IV
Resumen	Página V
Abstract	Página VI
Introducción	Página 1
Estimaciones de la probabilidad de desvío	Página 13
Estimaciones del desvío casi cierto	Página 34
Contribuciones	Página 48
Apéndices	Página 53

Resumen

El teorema ergódico de Birkhoff es uno de los principales resultados de la teoría ergódica y en general de los sistemas dinámicos. El teorema establece condiciones suficientes para la existencia del límite del promedio temporal de funciones medibles. Además, bajo la hipótesis de ergodicidad de la medida, el teorema establece también que el promedio temporal converge casi ciertamente al promedio espacial. Sin embargo, el teorema no dice nada acerca de la rapidez a la que se produce dicha convergencia. En esta tesis estamos interesados en obtener estimaciones de esta rapidez de convergencia en sistemas dinámicos simbólicos cuya medida invariante es una medida de Gibbs. Buscamos dar estas estimaciones en forma de cotas superiores para la probabilidad de desvío y para el desvío casi cierto.

Con respecto a la rapidez de convergencia de la probabilidad de desvío, tomamos como base los resultados de [1], [2] y [3]. Con respecto a la rapidez de convergencia del desvío casi cierto, tomamos como base los resultados de [11] y [12]. Utilizando estos resultados, se obtienen estimaciones aplicables a funciones que no son necesariamente Hölder continuas pero que cumplen la condición de Hölder localmente.

Abstract

The Birkhoff Ergodic Theorem is one of the main results in Ergodic Theory and in general, in dynamical systems. The theorem establishes sufficient conditions for the existence of the limit of time average of measurable functions. Moreover, under the assumption of ergodicity of the measure, the theorem also states that time average converges almost surely to the space average. However, it does not say anything about the rate at which this convergence takes place. In this thesis we are interested in finding estimates of this rate of convergence for symbolic dynamical systems whose invariant measure is a Gibbs measure. We give these estimates in form of upper bounds for the probability of deviation and also for the almost sure deviation.

Regarding the rate of convergence of the probability of deviation, we take as starting point the results in [1], [2] and [3]. Regarding the rate of convergence of the almost sure deviation, we take as starting point the results in [11] and [12]. By using these results, we obtain estimates applicable to functions that are not necessarily Hölder continuous but that satisfy Hölder condition locally.

Índice general

1..	<i>Introducción</i>	1
1.1.	Ejemplo numérico del Teorema Ergódico de Birkhoff	2
1.2.	Definiciones y resultados preliminares para sistemas dinámicos en general	4
1.3.	Enunciado formal del Teorema Ergódico de Birkhoff	5
1.4.	Descripción del problema	7
1.5.	Comentarios acerca de las estimaciones uniformes	7
1.6.	Definiciones y resultados preliminares para sistemas dinámicos simbólicos	9
2..	<i>Estimaciones de la probabilidad de desvío</i>	13
2.1.	Estimaciones de $P_n^\epsilon(f)$ a partir de estimaciones de $p_n^\epsilon(f)$	13
2.2.	Estimación de p_n^ϵ para funciones acotadas y continuas casi en todas partes	19
2.3.	Teoría de grandes desvíos	25
2.4.	Estimación de p_n^ϵ para funciones Lipschitz continuas	31
3..	<i>Estimación de desvío casi cierto</i>	35
3.1.	Estimación de la función de autocorrelación	35
3.2.	Estimación del desvío casi cierto basada en la función de autocorrelación	47
4..	<i>Contribuciones</i>	50
4.1.	Estimación de probabilidad de desvío: p_n^ϵ	50
4.2.	Estimación de probabilidad de desvío: P_n^ϵ	51
4.3.	Estimación de desvío casi cierto	52
4.4.	Discusión de los resultados y conclusiones	53
	<i>Appendices</i>	56
A..	<i>Apéndice</i>	57
A.1.	Teorema de Perron-Frobenius	57
A.2.	Comentarios sobre la notación de orden de magnitud	57
A.3.	Integral con respecto a una medida	58
A.4.	Lemas necesarios en algunas de las demostraciones presentadas	59

1. Introducción

La Teoría Ergódica es el estudio matemático del comportamiento promedio a largo plazo de diferentes clases de sistemas dinámicos [16]. La colección de todos los estados de un sistema forma un *espacio de estados* X y la evolución del sistema es representada por una *transformación* del espacio de estados en sí mismo $T : X \rightarrow X$. Dado un $x \in X$, la secuencia $x, T(x), T^2(x), \dots$ representa la trayectoria del sistema bajo la condición inicial x . El estado del sistema es observado mediante una función $f : X \rightarrow \mathbb{R}$ a la que llamamos *observable*. En este contexto, una pregunta básica de la Teoría Ergódica se refiere a la convergencia del promedio temporal: ¿bajo qué condiciones en X , en f y en T el siguiente límite existe?

$$f^*(x) := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x)).$$

El Teorema ergódico de Birkhoff, uno de los principales resultados de la Teoría Ergódica, establece condiciones para la existencia de f^* para casi todo $x \in X$ cuando f es integrable con respecto a una medida de probabilidad λ . El mismo teorema establece que, si además, la medida es *ergódica*, entonces:

$$f^*(x) = \int_X f d\lambda.$$

Este resultado fue enunciado por primera vez, como hipótesis, por el físico austriaco Ludwig Boltzmann para explicar algunos fenómenos experimentales de la Mecánica Estadística (una introducción a las ideas de Boltzmann puede encontrarse en [17]). Se le llamó entonces *hipótesis ergódica* y fue demostrada en su forma matemática por George David Birkhoff en 1931. En la sección 1.3 presentamos una versión moderna de este teorema.

En esta tesis estamos interesados en obtener estimaciones de la rapidez de convergencia del Teorema ergódico de Birkhoff en sistemas dinámicos en los que la medida es una medida de Gibbs asociada a un potencial ϕ . Aunque todas las definiciones precisas se darán más adelante, es necesario comentar en este punto que, como se explica en la definición 1.6.5, una medida de Gibbs tiene necesariamente un *potencial* asociado. Por lo mismo, en los resultados que se presentan en esta tesis aplicables a sistemas dinámicos con medida de Gibbs, se incluyen las condiciones sobre el observable f y el potencial ϕ bajo las cuales el resultado es aplicable.

El contenido de esta tesis está organizado de la siguiente forma. En lo que resta de este capítulo se presenta el enunciado formal del Teorema ergódico de Birkhoff y

se describe con detalle el problema que se estará estudiando en esta tesis. También se presenta un ejemplo numérico y algunas definiciones y resultados preliminares. En los capítulos 2 y 3 se presentan algunos resultados en la literatura relacionados con la rapidez de convergencia del Teorema de Birkhoff. En particular, en el capítulo 2 se presentan los resultados relacionados con estimaciones de la probabilidad de desvío y en el capítulo 3 se presentan resultados relacionados con la convergencia puntual *casi cierta*. En ambos capítulos se incluyen demostraciones detalladas de los resultados más relevantes. En el capítulo 4 se presentan las contribuciones de este trabajo de tesis y algunos comentarios sobre los resultados obtenidos y sobre el trabajo futuro. Finalmente, el último capítulo es un apéndice en el que se han incluido teoremas, lemas y proposiciones necesarias para algunas de las demostraciones de los capítulos 2 y 3.

1.1. Ejemplo numérico del Teorema Ergódico de Birkhoff

En esta sección se presenta un ejemplo numérico que ilustra la convergencia del Teorema Ergódico de Birkhoff. En la sección 1.1 se presenta el enunciado formal del teorema. Considere los siguientes valores para X , T y f :

$$X = [-1, 1] \tag{1.1}$$

$$T(x) = 1 - 2x^2 \tag{1.2}$$

$$f(x) = \chi_{[0.5, 0.7]}(x) = \begin{cases} 1 & x \in [0.5, 0.7] \\ 0 & x \notin [0.5, 0.7] \end{cases} \tag{1.3}$$

Por conveniencia, denotamos $S_n f(x) = \sum_{i=0}^{n-1} f(T^i(x))$. Estamos, interesados en estudiar el comportamiento límite de la órbita de $\frac{1}{n} S_n f(x_0)$ con una condición inicial x_0 , es decir, el comportamiento de la secuencia $S_1 f(x_0), \frac{1}{2} S_2 f(x_0), \frac{1}{3} S_3 f(x_0), \dots, \frac{1}{n} S_n f(x_0)$ para un número grande de iteraciones n . La figura 1.1 muestra el comportamiento de $\frac{1}{n} S_n f(x_0)$ desde $n = 1$ hasta $n = 100000$ para cuatro condiciones iniciales diferentes. Como puede verse en las figuras, en los cuatro casos, el valor de $\frac{1}{n} S_n f(x_0)$ se estabiliza alrededor de 0.079. Naturalmente esta aproximación numérica no es matemáticamente rigurosa, pero da una idea visual de la convergencia para este caso particular.

Esta convergencia puede ser explicada analíticamente con el Teorema de Birkhoff. Este teorema es un resultado que puede ser aplicado a mapeos más generales que el presentado aquí. En particular, puede ser aplicado a sistemas dinámicos simbólicos con transformaciones definidas en un espacio de símbolos cuya medida invariante es una medida de Gibbs. La rapidez de convergencia del Teorema de Birkhoff en este tipo de sistemas dinámicos, es el tema principal de esta tesis y se aborda a partir de la sección 1.6, después de presentar algunos resultados de la literatura aplicables a sistemas dinámicos en general.

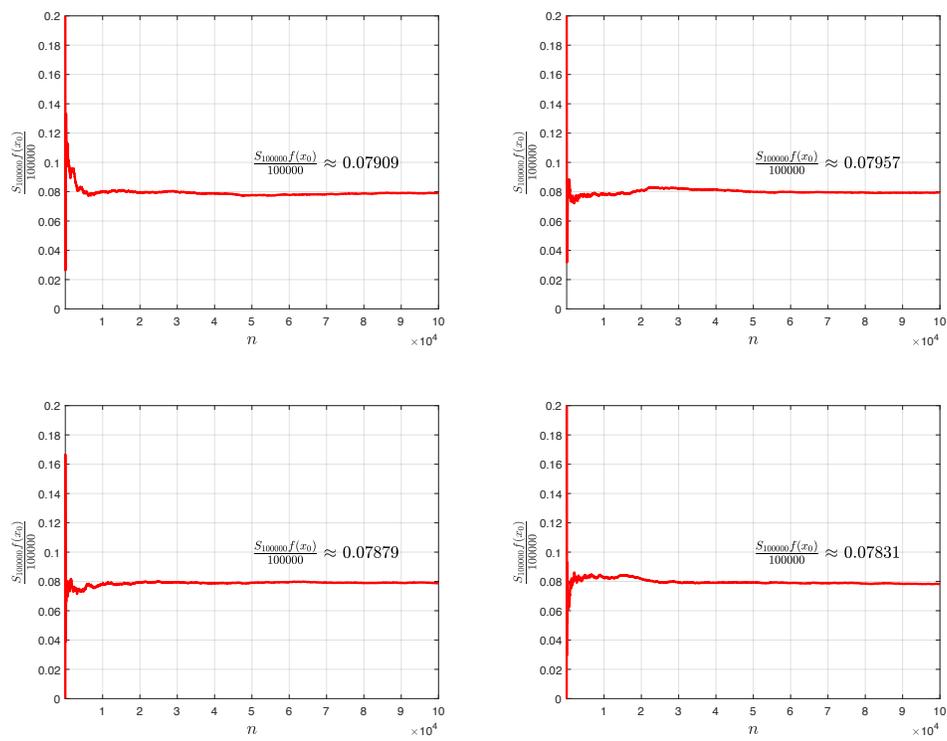


Fig. 1.1: Órbita de $\frac{1}{n}S_n f(x_0)$ con cuatro valores diferentes de x_0 .

1.2. Definiciones y resultados preliminares para sistemas dinámicos en general

En esta sección se dan algunas definiciones y resultados relacionados con teoría de la medida. Estas definiciones y resultados serán útiles en la siguiente sección para presentar el enunciado formal del Teorema de Birkhoff. En el apéndice A.3 se presentan las definiciones relacionadas con integral con respecto a una medida.

Definición 1.2.1. *Sea Ω un conjunto. Una familia de subconjuntos de Ω , \mathcal{B} , es una σ -álgebra sobre Ω si cumple las siguientes propiedades:*

1. $\emptyset \in \mathcal{B}$ y $\Omega \in \mathcal{B}$.
2. Si $E \in \mathcal{B}$ entonces $E^c \in \mathcal{B}$.
3. Si $E_i \in \mathcal{B}, i = 1, 2, \dots$ entonces $\bigcup_{i=1}^{\infty} E_i \in \mathcal{B}$.

Adicionalmente, si Ω es un espacio métrico compacto, llamamos σ -álgebra de Borel a la σ -álgebra más pequeña que contiene a todos los subconjuntos abiertos de Ω .

Definición 1.2.2. *Sea Ω un conjunto y \mathcal{B} una σ -álgebra sobre Ω . Una función $\lambda : \mathcal{B} \rightarrow \mathbb{R} \cup \{\infty\}$ es llamada medida si satisface las siguientes propiedades:*

1. $\lambda(E) \geq 0$, para todo $E \in \mathcal{B}$.
2. $\lambda(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \lambda(E_i)$ siempre que $E_i \in \mathcal{B}$ y $E_i \cap E_j = \emptyset$ para $i \neq j$.
3. $\lambda(\emptyset) = 0$.

Adicionalmente, si $\lambda(\Omega) = 1$, decimos que λ es una medida de probabilidad.

Las siguientes propiedades de la medida pueden consultarse en [6, Sección 9]. Estas propiedades se utilizarán más adelante para las demostraciones de los teoremas presentados en el Capítulo 2.

Proposición 1.2.1 ([6] Sección 9). *Sean $A, B, A_i \in \mathcal{B}, i = 1, \dots, n$. La medida λ satisface las siguientes propiedades:*

1. *Monotonicidad no decreciente: si $A \subseteq B$ entonces $\lambda(A) \leq \lambda(B)$.*
2. *Desigualdad de Boole: $\lambda(\bigcup_{i=1}^n A_i) \leq \sum_{i=1}^n \lambda(A_i)$, donde los conjuntos A_i no son necesariamente disjuntos.*
3. *Diferencia de medidas: Si $A \subseteq B$ entonces $\lambda(B \setminus A) = \lambda(B) - \lambda(A)$.*

Demostración. Ver [6, Sección 9]. □

Definición 1.2.3. *Sea Ω un conjunto, \mathcal{B} una σ -álgebra sobre Ω y λ una medida de probabilidad. Llamamos espacio de probabilidad, a la terna ordenada $(\Omega, \mathcal{B}, \lambda)$.*

Definición 1.2.4. *Sea Ω un conjunto, \mathcal{B} una σ -álgebra sobre Ω y T una transformación $T : \Omega \rightarrow \Omega$. Decimos que una medida λ es T -invariante, o bien que λ es invariante bajo la transformación T , si para todo $E \in \mathcal{B}$:*

$$\lambda(T^{-1}(E)) = \lambda(E). \quad (1.4)$$

Proposición 1.2.2 ([25]). Sea $(\Omega, \mathcal{B}, \lambda)$ un espacio de medida donde Ω es un espacio métrico. Entonces λ es una medida T -invariante si y solo si:

$$\int_{\Omega} (f \circ T) d\lambda = \int_{\Omega} f d\lambda$$

para toda función continua f .

Demostración. Ver por ejemplo [25]. □

Definición 1.2.5. Sea $(\Omega, \mathcal{B}, \lambda)$ un espacio de probabilidad y T una transformación $T : \Omega \rightarrow \Omega$. Decimos que una medida λ es ergódica si es T -invariante y además, para todo conjunto $E \in \mathcal{B}$ que cumpla $T^{-1}(E) = E$, se tiene que $\lambda(E) = 0$ o bien $\lambda(E) = 1$.

Definición 1.2.6. Decimos que un mapeo $T : \Omega \rightarrow \Omega$ es topológicamente mezclante si para cualesquiera conjuntos medibles, no vacíos, $A, B \subset \Omega$ existe N tal que $T^{-n}A \cap B \neq \emptyset$, para todo $n \geq N$.

Definición 1.2.7. Decimos que un mapeo $T : \Omega \rightarrow \Omega$ que preserva la medida λ , es mezclante si para cualesquiera conjuntos medibles, no vacíos, $A, B \subset \Omega$, se cumple:

$$\lim_{n \rightarrow \infty} \lambda(T^{-n}A \cap B) = \lambda(A)\lambda(B).$$

Definición 1.2.8. Decimos que un mapeo $T : \Omega \rightarrow \Omega$ es topológicamente transitivo si para cualesquiera conjuntos medibles, no vacíos, $A, B \subset \Omega$ se cumple $T^{-n}A \cap B \neq \emptyset$, para algún $n \geq 1$.

Es importante notar que, aunque están relacionadas, las definiciones de mapeo mezclante, topológicamente mezclante y topológicamente transitivo, no son equivalentes. Si un mapeo es topológicamente mezclante, entonces es topológicamente transitivo, pero ninguna de estas dos propiedades nos permiten concluir que el mapeo es mezclante.

Definición 1.2.9. Decimos que una propiedad p se cumple para casi todo $x \in \Omega$, o bien que se cumple λ -casi en todas partes, o bien que se cumple casi ciertamente, si el conjunto de las $x \in \Omega$ para las cuales no se cumple p , tiene medida cero, es decir:

$$\lambda(\{x \in \Omega : p \text{ no se cumple para } x\}) = 0. \quad (1.5)$$

Definición 1.2.10. Sea \mathcal{B} la σ -álgebra de Borel sobre Ω y $\mathcal{B}_{\mathbb{R}}$ la σ -álgebra de Borel sobre \mathbb{R} . Una función $f : \Omega \rightarrow \mathbb{R}$ es medible si para todo $B \in \mathcal{B}_{\mathbb{R}}$:

$$f^{-1}(B) \in \mathcal{B}. \quad (1.6)$$

1.3. Enunciado formal del Teorema Ergódico de Birkhoff

En esta sección se presenta el enunciado formal del teorema de Birkhoff. Este resultado puede encontrarse en textos estándar de Teoría Ergódica o sistemas dinámicos en general. Ver por ejemplo en [16], [18].

Teorema 1.3.1 (Teorema Ergódico de Birkhoff). *Sea $(\Omega, \mathcal{B}, \lambda)$ un espacio de probabilidad. Sea $T : \Omega \rightarrow \Omega$ y suponga que λ es una medida T -invariante. Sea $f : \Omega \rightarrow \mathbb{R}$ una función integrable. Definimos:*

$$S_n f(x) = \sum_{i=0}^{n-1} f(T^i(x)). \quad (1.7)$$

Entonces, el siguiente límite existe para casi todo $x \in \Omega$:

$$f^*(x) = \lim_{n \rightarrow \infty} \frac{1}{n} S_n f(x). \quad (1.8)$$

También:

$$\int_{\Omega} f^* d\lambda = \int_{\Omega} f d\lambda. \quad (1.9)$$

Adicionalmente, si λ es una medida ergódica, entonces:

$$f^*(x) = \int_{\Omega} f d\lambda \quad (1.10)$$

Demostración. Ver [16, Teorema 2.3]. □

El lado derecho de (1.8) es el límite del promedio temporal $\frac{1}{n} S_n f(x)$ y el lado derecho de (1.10) se conoce como promedio espacial. Con esta nomenclatura, el Teorema de Birkhoff establece que, en el límite y bajo la condición de ergodicidad de la medida, el promedio temporal converge casi ciertamente al promedio espacial. A este tipo de convergencia le llamamos convergencia *casi segura* o bien convergencia *puntual*. Es importante notar que la igualdad entre el promedio temporal y el promedio espacial ocurre en el límite, pero para valores finitos de n puede haber una diferencia entre estas dos cantidades. A esta diferencia le llamamos *desvío*.

Por otro lado, es un resultado conocido que convergencia puntual implica convergencia en medida (ver por ejemplo [6, Sección 23]), por lo tanto, para todo $\epsilon > 0$, las siguientes cantidades tienden a cero cuando n tiende a infinito:

$$p_n^\epsilon(f) = \lambda \left(\left\{ x \in \Omega : \left| \frac{1}{n} S_n f(x) - f^*(x) \right| \geq \epsilon \right\} \right) \quad (1.11)$$

$$P_n^\epsilon(f) = \lambda \left(\left\{ x \in \Omega : \sup_{k \geq n} \left| \frac{1}{k} S_k f(x) - f^*(x) \right| \geq \epsilon \right\} \right) \quad (1.12)$$

$$(1.13)$$

La ecuación (1.11) representa la probabilidad de que la distancia entre el promedio temporal hasta la iteración n y el promedio espacial, sea mayor o igual que un $\epsilon > 0$ dado. La ecuación (1.12) representa la probabilidad de que la mayor distancia entre los promedios temporal y espacial a partir de la iteración n , sea mayor que un $\epsilon > 0$ dado. A estas dos expresiones en conjunto, les llamamos *probabilidades de desvío*.

La diferencia entre el promedio temporal y el promedio espacial, también tiende a cero cuando n tiende a infinito. Al valor absoluto de esta diferencia le llamamos *desvío casi cierto*:

$$D_n(f) = \left| \frac{1}{n} S_n f(x) - f^*(x) \right| \quad (1.14)$$

A la tasa a la que estas cantidades tienden a cero para valores grandes de n le llamamos *rapidez de convergencia*. Con estas ideas en mente, presentamos a continuación la descripción del problema que estamos interesados en estudiar en esta tesis.

1.4. Descripción del problema

Aunque el Teorema de Birkhoff implica que p_n^ϵ , P_n^ϵ y D_n tienden a cero cuando n tiende a infinito, el teorema no dice nada acerca de la rapidez a la que se produce dicha convergencia. En otras palabras, el teorema no brinda información acerca del valor que toman p_n^ϵ , P_n^ϵ y D_n con valores finitos de n . Tomando en cuenta esta situación, tenemos dos objetivos específicos en esta tesis que enumeramos a continuación. En ambos objetivos nos restringimos al caso de sistemas dinámicos simbólicos cuya medida invariante es una medida de Gibbs.

1. El objetivo principal de esta tesis es encontrar estimaciones de (1.11) y (1.12). Buscamos dar estas estimaciones en forma de cotas superiores para p_n^ϵ y P_n^ϵ y buscamos obtenerlas para observables f que son acotadas y continuas casi en todas partes y para potenciales ϕ que son Hölder continuos.
2. Adicionalmente a las estimaciones de p_n^ϵ y P_n^ϵ , también estamos interesados en encontrar una estimación del desvío casi cierto D_n . Al igual que en el caso anterior, buscamos dar esta estimación como una cota superior aplicable a observables f que son acotadas y continuas casi en todas partes. Extendiendo el caso para el punto anterior, realizamos nuestras estimaciones para potenciales ϕ , cuya variación decae en forma polinomial.

1.5. Comentarios acerca de las estimaciones uniformes

Es bien conocido que existen casos en los que la convergencia del Teorema de Birkhoff puede ser arbitrariamente lenta o arbitrariamente rápida. El siguiente Teorema ilustra este punto.

Teorema 1.5.1 ([23]). *Sea $\Omega = [0, 1]$, \mathcal{B} la σ -álgebra generada por todos los intervalos abiertos de Ω y λ la medida de Lebesgue. Sea T una transformación invertible que preserva la medida de Lebesgue. Entonces:*

1. *Para cualquier secuencia monótonicamente convergente a infinito $\{b_n\}$, existe un conjunto medible $A \subset \Omega$ de medida arbitraria $\lambda(A)$ tal que:*

$$\lim_{n \rightarrow \infty} \left| b_n \left(\frac{1}{n} \sum_{i=0}^{n-1} \chi_A(T^i(x)) - \lambda(A) \right) \right| = \infty, \quad (1.15)$$

para casi todo $x \in \Omega$.

2. Para cualquier secuencia monotónicamente convergente a infinito $\{a_n\}$ con $a_1 \geq 2$, existe un conjunto medible $B \subset \Omega$ de medida arbitraria $\lambda(B)$ tal que:

$$\left| \frac{1}{n} \sum_{i=0}^{n-1} \chi_B (T^i(x)) - \lambda(B) \right| \leq \frac{a_n}{n}, \quad (1.16)$$

para todo $n \geq 1$ y para casi todo $x \in \Omega$.

Demostración. Ver [23]. □

Con respecto a la primera parte del Teorema 1.5.1, el Teorema de Birkhoff garantiza que $\frac{1}{n} \sum_{i=0}^{n-1} \chi_A (T^i(x)) - \lambda(A)$ tiende a cero cuando n tiende a infinito. Sin embargo, dado que la secuencia $\{b_n\}$ converge a infinito, la desigualdad (1.15) implica que esta secuencia, cualquiera que sea, converge a infinito más rápido de lo que $\frac{1}{n} \sum_{i=0}^{n-1} \chi_A (T^i(x)) - \lambda(A)$ converge a cero. El Teorema es válido para cualquier secuencia monotónicamente convergente a infinito; en particular, si la secuencia converge arbitrariamente lento, entonces la convergencia del promedio temporal al promedio espacial también es arbitrariamente lenta. En este caso no tiene mucho sentido preguntarse por la rapidez de convergencia del Teorema de Birkhoff: siempre habrá un conjunto A cuya función característica converja más lento que cualquier rapidez de convergencia prescrita.

Con respecto a la segunda parte del Teorema 1.5.1, la situación es análoga. Dado que la convergencia de $\{a_n\}$ a infinito puede ser arbitrariamente lenta y $a_1 \geq 2$, se tiene que $\frac{a_n}{n}$ puede ser arbitrariamente cercano a $\frac{2}{n}$. Esto quiere decir cualquier rapidez de convergencia mayor a $\frac{2}{n}$ es posible. Para lograr una rapidez de convergencia más cercana a $\frac{2}{n}$ solo se tiene que seleccionar una secuencia $\{a_n\}$ que converja más lento. Por ejemplo, considere la secuencia dada por $\{a_n\} = 2\alpha^{n-1}$ con $\alpha > 1$. Entre más cercano a 1 sea el valor de la constante α , más lenta es la rapidez de convergencia de $\{a_n\}$ a infinito y por lo tanto, más cercana a $\frac{2}{n}$ es la rapidez de convergencia del promedio temporal al promedio espacial.

Estos resultados muestran que, en ciertos casos particulares, no es posible obtener estimaciones que dependan solo de la transformación T y que sean independientes del observable f . Aún más, en [4] se presenta el siguiente ejemplo que muestra un caso en el que tampoco es posible obtener estimaciones que dependan solo del observable f y que sean independientes de la transformación T .

Ejemplo 1.5.1. *Considere el espacio de medida dado por el círculo unitario y la medida de Lebesgue. Sea el observable χ_R la función característica del lado derecho del círculo. Para cada $N \in \mathbb{N}$ dividimos el círculo unitario en $2N$ arcos iguales y consideramos una transformación T_N que mapea cada arco en el arco adyacente en sentido antihorario. La figura 1.2 ilustra el ejemplo. Observamos que en este caso:*

$$\int \chi_R d\lambda = \lambda(R) = \frac{1}{2}.$$

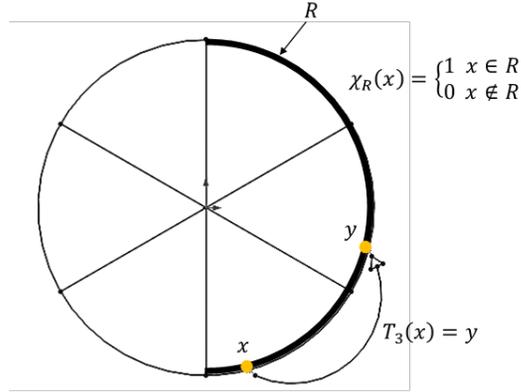


Fig. 1.2: Ilustración del ejemplo 1.5.1 con $N = 3$.

Pero en estas condiciones, la rapidez de convergencia del promedio temporal a $\frac{1}{2}$ depende del valor de N . Entre mayor sea el valor de N , menor es la rapidez de convergencia. Esto significa que no se pueden obtener estimaciones independientes de la transformación T_N . La figura 1.3 muestra la convergencia para algunos valores de N .

A pesar de estos resultados, el problema de determinar la rapidez de convergencia del Teorema de Birkhoff sí tiene interés si se delimita de forma adecuada. Esto implica considerar restricciones tanto en el observable f como en la transformación T . En esta tesis nos enfocaremos en el caso en el que el observable cumple con la condición de regularidad de ser acotada y continua casi en todas partes. La transformación en la que estaremos interesados será el mapeo σ y la medida de Gibbs que definiremos más adelante.

1.6. Definiciones y resultados preliminares para sistemas dinámicos simbólicos

En esta sección presentamos los conceptos básicos de Sistemas Dinámicos Simbólicos y medidas de Gibbs. Para ello definiremos la siguiente notación que utilizaremos a lo largo de esta sección para el contexto simbólico. Sea S un conjunto finito y $\Omega = S^{\mathbb{N}}$. Para $\underline{x}, \underline{y} \in \Omega$, definimos una distancia en Ω mediante $d_\theta(\underline{x}, \underline{y}) = \theta^N$, donde $\theta \in (0, 1)$ y N es el máximo entero para el que $x_i = y_i$ para todo $0 \leq i < N$. Por convención, si $\underline{x} = \underline{y}$, entonces $d_\theta(\underline{x}, \underline{y}) = 0$ mientras que si $x_0 \neq y_0$, entonces $d_\theta(\underline{x}, \underline{y}) = 1$.

Definición 1.6.1. Sea S un conjunto finito y $\Omega = S^{\mathbb{N}}$. Al mapeo $\sigma : \Omega \rightarrow \Omega$ dado por $(\sigma \underline{x})_i = x_{i+1}$, le llamamos mapeo shift.

Definición 1.6.2. Sea S un conjunto finito y $\Omega = S^{\mathbb{N}}$. Al sistema dinámico (Ω, σ) le llamamos sistema dinámico simbólico.

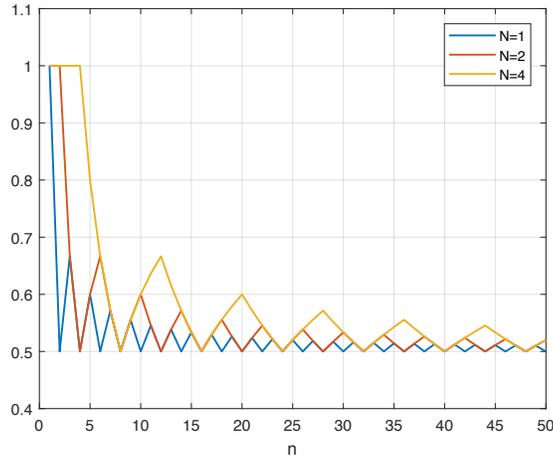


Fig. 1.3: Ilustración del ejemplo 1.5.1 con algunos valores de N .

Definición 1.6.3. Dada una secuencia z_0, \dots, z_{n-1} con $z_i \in S$, al conjunto dado por:

$$[z_0, \dots, z_{n-1}] = \{x \in \Omega_A^+ : x_i = z_i, i = 0, \dots, n-1\},$$

le llamamos cilindro de longitud n .

Definición 1.6.4. Sea $\phi : \Omega \rightarrow \mathbb{R}$ una función continua. Para cada $m \geq 1$, definimos la m -variación de ϕ como:

$$\text{var}_m(\phi) := \sup \{|\phi(x) - \phi(y)| : x_i = y_i, i = 0, \dots, m-1\}.$$

Notamos que la variación mide cuánto puede variar la función ϕ en cilindros de longitud m . También notamos que, dado que ϕ es Hölder continua, $\text{var}_m(\phi)$ tiende a cero cuando m tiende a infinito.

El siguiente lema permite caracterizar tanto a las funciones Hölder continuas como a las funciones Lipschitz continuas con respecto a la distancia definida al principio de esta sección.

Lema 1.6.1. Sea $f : \Omega \rightarrow \mathbb{R}$. Las siguientes proposiciones son equivalentes:

1. f es Hölder continua con exponente α , con respecto a la distancia d_θ .
2. f es Lipschitz continua con respecto a la distancia d_{θ^α} .
3. $\text{var}_m(f) \leq C\theta^{\alpha m}$ para todo $m \geq 1$.

Demostración.

$1 \Rightarrow 2$ Dado que f es Hölder continua con exponente α , se tiene que existe $h \in \mathbb{R}$, tal que, para todo $x, y \in \Omega$:

$$|f(x) - f(y)| \leq h(\theta^N)^\alpha = h(\theta^\alpha)^N = h d_{\theta^\alpha}(x, y).$$

Con lo que se concluye que f es Lipschitz continua con respecto a la distancia d_{θ^α} .

2 \Rightarrow 3 Dado que f es Lipschitz continua con respecto a la distancia d_{θ^α} , se tiene que para todo $\underline{x}, \underline{y} \in \Omega$:

$$|f(\underline{x}) - f(\underline{y})| \leq C(\theta^\alpha)^N.$$

En particular, la desigualdad se cumple para $\underline{x}, \underline{y}$ tales que $x_i = y_i$ con $i = 0, \dots, m-1$. Es decir:

$$\text{var}_m(f) = \sup \{|f(\underline{x}) - f(\underline{y})| : x_i = y_i, i = 0, \dots, m-1\} \leq C(\theta^\alpha)^N.$$

Dado que $x_i = y_i$ para $i = 0, \dots, m-1$, se tiene que $N \geq m$, por lo que $(\theta^\alpha)^N \leq (\theta^\alpha)^m$ (esto es porque $\theta^\alpha \in (0, 1]$) con lo que:

$$\text{var}_m(f) \leq C(\theta^\alpha)^N \leq C(\theta^\alpha)^m = C\theta^{\alpha m}.$$

3 \Rightarrow 1 Sean $\underline{x}, \underline{y} \in \Omega$ y sea M el máximo entero para el que $x_i = y_i$. Entonces:

$$|f(\underline{x}) - f(\underline{y})| \leq \sup \{|f(\underline{x}) - f(\underline{y})| : x_i = y_i, i = 0, \dots, M-1\} = \text{var}_M(f).$$

Pero por hipótesis $\text{var}_M(f) \leq C\theta^{\alpha M}$ por lo que:

$$|f(\underline{x}) - f(\underline{y})| \leq \text{var}_M(f) \leq C\theta^{\alpha M} = C(\theta^M)^\alpha.$$

Con lo que se tiene que f es Hölder continua con exponente α , con respecto a la distancia d_θ . \square

A continuación definimos una medida de Gibbs que es el tipo de medida en el que estaremos interesados de aquí en adelante.

Definición 1.6.5. Decimos que μ_ϕ es una medida de Gibbs con potencial ϕ si existen constantes $C > 1$ y $P > 0$ tales que:

$$C^{-1} \leq \frac{\mu_\phi \{\underline{y} \in \Omega : y_i = x_i, \forall i \in [0, m]\}}{\exp\left(-Pm + \sum_{k=0}^{m-1} \phi(\sigma^k(\underline{x}))\right)} \leq C. \quad (1.17)$$

para todo $\underline{x} \in \Omega$ y para todo $m \geq 1$.

Note que, salvo por una constante, la medida de Gibbs del cilindro $[x_0, \dots, x_{m-1}]$ está dada por $\exp\left(\sum_{k=0}^{m-1} \phi(\sigma^k(\underline{x}))\right)$. Esto quiere decir que el potencial es la función que asigna los pesos de las entradas de la secuencia \underline{x} en el valor de la medida μ_ϕ .

La medida de Gibbs generaliza el concepto de distribución de Gibbs utilizado en Mecánica Estadística. Suponga que un sistema físico tiene n estados posibles denotados $x^{(i)}$, con $i = 1, \dots, n$ y que cada estado tiene una energía asociada denotada $E(x^{(i)})$. Bajo ciertas condiciones, la probabilidad de que el sistema se encuentre en el estado $x^{(j)}$ está dada por la distribución de Gibbs:

$$p(x^{(j)}) = \frac{e^{-\beta E(x^{(j)})}}{\sum_{i=1}^n e^{-\beta E(x^{(i)})}},$$

donde $\beta = \frac{1}{k_B T}$, k_B es la constante de Boltzmann y T es la temperatura en Kelvin. El estado del sistema completo está determinado por el estado de sus partículas individuales. La medida de Gibbs surge entonces como el comportamiento límite de la distribución de Gibbs cuando el sistema está constituido por un gran número de partículas.

Aunque su origen está en Mecánica Estadística, la medida de Gibbs se encuentra implícita en muchos mapeos $f : X \rightarrow X$. Para ilustrar este punto, suponga que hacemos una partición finita del conjunto X y asignamos a cada elemento de la partición un símbolo. A continuación tomamos una condición inicial en X y comenzamos a iterar el mapeo f , pero en lugar de estar interesados en el valor de cada iteración, estamos interesados el símbolo que corresponde al elemento de la partición en la que se encuentre el valor de la iteración. En esta situación, la órbita del mapeo f puede ser vista como una secuencia de símbolos. El conjunto de todas las secuencias posibles forma un espacio de símbolos y el mapeo σ es visto como el mapeo que devuelve el símbolo correspondiente a la siguiente iteración del mapeo.

2. Estimaciones de la probabilidad de desvío

En este capítulo se presentan algunos de los resultados existentes en la literatura relacionados con estimaciones de p_n^ϵ y P_n^ϵ definidas en las ecuaciones (1.11) y (1.12). Brevemente, los resultados que se presentan son como sigue.

En [3] los autores muestran que para funciones esencialmente acotadas, un decaimiento exponencial de p_n^ϵ implica un decaimiento exponencial de P_n^ϵ . Esto quiere decir que, si conocemos una estimación de p_n^ϵ cuyo decaimiento es exponencial, podemos utilizarla para obtener una estimación de P_n^ϵ (la cual tendrá también un decaimiento exponencial).

En [1], los autores muestran que, a partir de una estimación de p_n^ϵ para funciones Hölder continuas, puede obtenerse una estimación de p_n^ϵ para funciones que no son necesariamente Hölder continuas pero que cumplen la condición de Hölder localmente (es decir, dado un $x \in \Omega$, la condición de Hölder se cumple para todo $y \in N_\delta(x)$, donde $N_\delta(x)$ es la vecindad de x de radio δ).

En [8], los autores toman como base una desigualdad de momentos presentada en [9] y la utilizan para obtener una estimación de p_n^ϵ aplicable en el contexto de sistemas dinámicos simbólicos con medida de Gibbs. Esta desigualdad aplica solamente a medidas de Gibbs con potenciales que son Hölder continuos.

En este capítulo se incluye también, una breve introducción a la Teoría de grandes desvíos y se presentan desigualdades de p_n^ϵ obtenidas mediante la aplicación de esta teoría. El contenido de esta introducción está basado principalmente en [7] y [5].

2.1. Estimaciones de $P_n^\epsilon(f)$ a partir de estimaciones de $p_n^\epsilon(f)$

En esta sección se presentan los resultados de [3]. Para el propósito de esta tesis, el resultado principal de este artículo es que, para funciones esencialmente acotadas, un decaimiento exponencial de p_n^ϵ implica un decaimiento exponencial de P_n^ϵ . Este resultado se presenta en el Teorema 2.1.6 y su demostración se basa en el Teorema 2.1.3 y en algunos otros lemas adicionales que también se presentan en esta sección. Comenzamos presentando los lemas necesarios para la demostración del Teorema 2.1.3.

A continuación haremos algunas precisiones importantes que serán aplicables en lo que resta de esta tesis. Observamos que $\|f - f^*\|_\infty = 0$ implica que $f(x) = f^*(x)$ para todo $x \in \Omega$. En este caso, el promedio temporal y el promedio espacial son

iguales para todo $n \geq 0$ y por lo tanto no hay rapidez de convergencia qué estimar. Por ello, de aquí en adelante asumimos que $\|f - f^*\|_\infty \neq 0$. También, asumimos que la medida λ es ergódica. Por el teorema 1.3.1, esto significa que la función $f^*(x) = \lim_{n \rightarrow \infty} A_n f(x)$ es una constante dada por $\int_\Omega f d\lambda$. Por ello, en lo que resta de esta tesis, escribimos f^* en lugar de $f^*(x)$. Por último, para simplificar la notación, de aquí en adelante denotamos $A_n f(x) = \frac{1}{n} S_n f(x)$.

Lema 2.1.1. *Dado $\epsilon > 0$, suponga que se cumple $|A_n f(x) - f^*| < \frac{\epsilon}{2}$, para algún $n \geq 1$ y para todo $x \in \Omega$. Sea M un entero tal que $M \in \left[0, \frac{n\epsilon}{2\|f - f^*\|_\infty}\right]$. Entonces se cumple la siguiente desigualdad:*

$$|A_{n+M} f(x) - f^*| \leq \frac{1}{n+M} \left[\frac{n\epsilon}{2} + M\|f - f^*\|_\infty \right] < \epsilon. \quad (2.1)$$

Demostración. Se observa que:

$$\begin{aligned} A_{n+1} f(x) &= \frac{1}{n+1} \sum_{i=0}^n f(T^i(x)) \\ &= \frac{1}{n+1} \left[\sum_{i=0}^{n-1} f(T^i(x)) + f(T^n(x)) \right] \\ &= \frac{1}{n+1} [nA_n f(x) + f(T^n(x))]. \end{aligned}$$

Luego:

$$(2.2)$$

$$(n+1)A_{n+1} f(x) = nA_n f(x) + f(T^n(x)). \quad (2.3)$$

También, dado que $\|f - f^*\|_\infty = \sup_{x \in \Omega} |f(x) - f^*|$, se tiene:

$$-\|f - f^*\|_\infty \leq f(T^n(x)) - f^* \leq \|f - f^*\|_\infty. \quad (2.4)$$

Por hipótesis, $|A_n f(x) - f^*| < \frac{\epsilon}{2}$, por lo que:

$$\begin{aligned} -\frac{\epsilon}{2} &\leq A_n f(x) - f^* \leq \frac{\epsilon}{2} \\ -\frac{n\epsilon}{2} &\leq nA_n f(x) - nf^* \leq \frac{n\epsilon}{2}. \end{aligned} \quad (2.5)$$

Sumando (2.5) y (2.4) se obtiene:

$$-\frac{n\epsilon}{2} - \|f - f^*\|_\infty \leq nA_n f(x) - nf^* + f(T^n(x)) - f^* \leq \frac{n\epsilon}{2} + \|f - f^*\|_\infty \quad (2.6)$$

Sustituyendo (2.3) en (2.6) se obtiene:

$$-\frac{n\epsilon}{2} - \|f - f^*\|_\infty \leq (n+1)A_{n+1} f(x) - nf^* - f^* \leq \frac{n\epsilon}{2} + \|f - f^*\|_\infty \quad (2.7)$$

Dividiendo entre $n+1$:

$$-\frac{1}{n+1} \left[\frac{n\epsilon}{2} + \|f - f^*\|_\infty \right] \leq A_{n+1} f(x) - f^* \leq \frac{1}{n+1} \left[\frac{n\epsilon}{2} + \|f - f^*\|_\infty \right] \quad (2.8)$$

Observamos que:

$$-\frac{1}{n+M} \left[\frac{n\epsilon}{2} + M\|f - f^*\|_\infty \right] \leq A_{n+M}f(x) - f^* \leq \frac{1}{n+M} \left[\frac{n\epsilon}{2} + M\|f - f^*\|_\infty \right], \quad (2.9)$$

siempre que $M \in \left[0, \frac{n\epsilon}{2\|f - f^*\|_\infty} \right]$. La desigualdad (2.9) es equivalente al extremo izquierdo de (2.1). Por otro lado, como $M \leq \frac{n\epsilon}{2\|f - f^*\|_\infty}$, se puede ver que:

$$\begin{aligned} M \leq \frac{n\epsilon}{2\|f - f^*\|_\infty} &\Rightarrow M < \frac{n\epsilon}{2(\|f - f^*\|_\infty - \epsilon)} \\ &\Rightarrow M(\|f - f^*\|_\infty - \epsilon) < \frac{n\epsilon}{2} \\ &\Rightarrow M\|f - f^*\|_\infty + \frac{n\epsilon}{2} < \epsilon(n+M) \\ &\Rightarrow \frac{1}{n+M} \left[\frac{n\epsilon}{2} + M\|f - f^*\|_\infty \right] < \epsilon. \end{aligned} \quad (2.10)$$

Uniendo (2.10) y (2.9) se llega a (2.1). Esto completa la demostración del Lema 2.1.1. \square

Lema 2.1.2. Sean $r = 1 + \frac{\epsilon}{2\|f - f^*\|_\infty}$ y $n \geq 1$. Para $N \in (0, \infty)$, sea n_N el mínimo entero no menor que nr^N . Entonces se cumple la siguiente desigualdad.

$$P_n^{2\epsilon} \leq \sum_{N=0}^{\infty} \lambda \left\{ x \in \Omega : \sup_{nr^N \leq k < nr^{N+1}} |A_k f(x) - f^*| \geq 2\epsilon \right\} \leq \sum_{N=0}^{\infty} p_{n_N}^\epsilon. \quad (2.11)$$

Demostración. El extremo izquierdo de (2.11) se sigue de los siguientes cálculos:

$$\begin{aligned} P_n^{2\epsilon}(f) &= \lambda \left(\left\{ x \in \Omega : \sup_{k \geq n} |A_k f(x) - f^*| \geq 2\epsilon \right\} \right) \\ &= \lambda \left(\left\{ x \in \Omega : \sup_{n \leq k < nr} |A_k f(x) - f^*| \geq 2\epsilon \right\} \cup \right. \\ &\quad \left. \left\{ x \in \Omega : \sup_{nr \leq k < nr^2} |A_k f(x) - f^*| \geq 2\epsilon \right\} \cup \right. \\ &\quad \left. \left\{ x \in \Omega : \sup_{nr^2 \leq k < nr^3} |A_k f(x) - f^*| \geq 2\epsilon \right\} \cup \dots \right) \\ &\leq \sum_{N=0}^{\infty} \lambda \left\{ x \in \Omega : \sup_{nr^N \leq k < nr^{N+1}} |A_k f(x) - f^*| \geq 2\epsilon \right\}. \end{aligned} \quad (2.12)$$

Donde, en el último paso, se utilizó el inciso 2 de la proposición 1.2.1 enunciada en la sección 1.2. Para demostrar el extremo derecho de (2.11), es suficiente mostrar que la siguiente desigualdad es válida para todo $N \in (0, \infty)$:

$$\lambda \left\{ x \in \Omega : \sup_{nr^N \leq k < nr^{N+1}} |A_k f(x) - f^*| \geq 2\epsilon \right\} \leq \lambda \{ x \in \Omega : |A_{n_N} f(x) - f^*| \geq \epsilon \}. \quad (2.13)$$

Por el inciso 1 de la proposición 1.2.1, para mostrar (2.13), es suficiente mostrar:

$$\left\{ x \in \Omega : \sup_{nr^N \leq k < nr^{N+1}} |A_k f(x) - f^*| \geq 2\epsilon \right\} \subseteq \{x \in \Omega : |A_{n_N} f(x) - f^*| \geq \epsilon\}. \quad (2.14)$$

Por ello, nos enfocaremos en mostrar (2.14) y lo haremos mediante contrapositiva. Sea $y \notin \{x \in \Omega : |A_{n_N} f(x) - f^*| \geq \epsilon\}$. Entonces $|A_{n_N} f(y) - f^*| < \epsilon$. Por el Lema 2.1.1, se tiene que $|A_{n_N+M} f(y) - f^*| < 2\epsilon$, para todo $M \in \left[0, \frac{n_N \epsilon}{2\|f - f^*\|_\infty}\right]$, o bien:

$$|A_m f(y) - f^*| < 2\epsilon, \quad \forall m \in \left[n_N, n_N + \frac{n_N \epsilon}{2\|f - f^*\|_\infty}\right]. \quad (2.15)$$

Dado que n_N es el mínimo entero no menor que nr^N , se tiene que el menor entero del intervalo $[nr^N, nr^{N+1}]$ es n_N . Además:

$$nr^{N+1} = nr^N r \leq n_N r = n_N \left(1 + \frac{\epsilon}{2\|f - f^*\|_\infty}\right).$$

Esto quiere decir que los enteros en el intervalo $[nr^N, nr^{N+1}]$ son un subconjunto de los enteros del intervalo $\left[n_N, n_N + \frac{n_N \epsilon}{2\|f - f^*\|_\infty}\right]$. Por (2.15), se tiene que:

$$|A_m f(y) - f^*| < 2\epsilon, \quad \forall m \in [nr^N, nr^{N+1}]. \quad (2.16)$$

Esto quiere decir que $\sup_{nr^N \leq k < nr^{N+1}} |A_k f(y) - f^*| < 2\epsilon$, y por lo tanto $y \notin \{x \in \Omega : \sup_{nr^N \leq k < nr^{N+1}} |A_k f(x) - f^*| \geq 2\epsilon\}$. Por contrapositiva y monotonicidad de la medida, se concluye que (2.14), y por lo tanto (2.13), se cumplen para todo $N \in (0, \infty)$, lo cual significa que:

$$\sum_{N=0}^{\infty} \lambda \left\{ x \in \Omega : \sup_{nr^N \leq k < nr^{N+1}} |A_k f(x) - f^*| \geq 2\epsilon \right\} \leq \sum_{N=0}^{\infty} p_{n_N}^\epsilon. \quad (2.17)$$

Con (2.17) se completa la demostración del extremo derecho de (2.11) y también la demostración del Lema 2.1.2. \square

El siguiente Teorema, establece que, bajo ciertas condiciones, es posible obtener una desigualdad de P_n^ϵ a partir de una desigualdad de p_n^ϵ .¹

Teorema 2.1.3 ([3]). *Sea f una función esencialmente acotada y sea la función $\varphi(x) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ una función monótona decreciente a cero para $x > x_0 \geq 1$. Suponga que existen $M(\epsilon) > 0$ y n_0 tales que, para todo $n > n_0 \geq 1$:*

$$\int_n^\infty \frac{\varphi(x)}{x} dx \leq M(\epsilon) \varphi(n). \quad (2.18)$$

Entonces, las siguientes proposiciones son equivalentes:

¹ El Teorema en [3] se presenta con la notación de \circ y \mathcal{O} . Vea la sección A.2 del apéndice para revisar los comentarios acerca de esta notación.

1. Existen constantes $C_1(\epsilon) > 0$ y $n_1 \geq 1$ tales que $p_n^\epsilon(f) \leq C_1(\epsilon)\varphi(n), \forall n \geq n_1$.
2. Existen constantes $C_2(\epsilon) > 0$ y $n_2 \geq 1$ tales que $P_n^\epsilon(f) \leq C_2(\epsilon)\varphi(n), \forall n \geq n_2$

Demostración.

1 \Rightarrow 2 Suponga que se cumple la proposición 1 y defina r y n_N igual que en el Lema 2.1.2. Se tiene que $n_N \geq nr^N > n \geq n_1$. Utilizando el Lema 2.1.2 se tiene:

$$P_n^{2\epsilon} \leq \sum_{N=0}^{\infty} p_{n_N}^\epsilon \leq C_1(\epsilon) \sum_{N=0}^{\infty} \varphi(n_N), \quad (2.19)$$

para todo $n \geq n_1$. Dado que φ es monótona decreciente para todo $x > x_0$, se tiene que $\varphi(n_N) \leq \varphi(nr^N)$, para todo $n > x_0$ y para todo $N \in (0, \infty)$. Por lo tanto:

$$\begin{aligned} \sum_{N=0}^{\infty} \varphi(n_N) &\leq \sum_{N=0}^{\infty} \varphi(nr^N) = \varphi(n) + \sum_{N=1}^{\infty} \varphi(nr^N) \\ &\leq \varphi(n) + \int_0^{\infty} \varphi(nr^x) dx. \end{aligned}$$

Donde hemos utilizado el Lema A.4.2 presentado en el apéndice de la tesis. Ahora hacemos el cambio de variable $y = nr^x$, con $r = 1 + \frac{\epsilon}{2\|f-f^*\|_\infty}$. Dado que r no depende de y , se tiene que $dx = \frac{dy}{nr^x \ln r} = \frac{dy}{y \ln r}$ y por lo tanto:

$$\sum_{N=0}^{\infty} \varphi(n_N) \leq \varphi(n) + \frac{1}{\ln r} \int_n^{\infty} \frac{\varphi(y)}{y} dy \quad (2.20)$$

$$\leq \varphi(n) + \frac{M(\epsilon)}{\ln r} \varphi(n) = \left(1 + \frac{M(\epsilon)}{\ln r}\right) \varphi(n). \quad (2.21)$$

Que es válida siempre que (2.18) sea válida, es decir, para todo $n > n_0$. Uniendo esta última desigualdad con 2.19 se obtiene $P_n^{2\epsilon} \leq C_1(\epsilon) \left(1 + \frac{M(\epsilon)}{\ln r}\right) \varphi(n)$, para todo $n > \max(n_0, n_1, x_0)$. Haciendo $C_2(\epsilon) = C_1(\epsilon) \left(1 + \frac{M(\epsilon)}{\ln r}\right)$ y $n_2 = \max(n_0, n_1, x_0)$ se completa la demostración de la primera implicación.

2 \Rightarrow 1 Para demostrar esta implicación, primero mostraremos, mediante contrapositiva, que se cumple $p_n^\epsilon \leq P_n^\epsilon$ para todo $\epsilon > 0$ y $n \geq 1$.

Sea $y \notin \{x \in \Omega : \sup_{k \geq n} |A_k f(x) - f^*| \geq \epsilon\}$; esto quiere decir que:

$$\sup_{k \geq n} |A_k f(y) - f^*| < \epsilon.$$

En particular $|A_n f(y) - f^*| < \epsilon$, por lo que se tiene además que

$$y \notin \{x \in \Omega : |A_n f(x) - f^*| \geq \epsilon\}.$$

Por contrapositiva se tiene que si $y \in \{x \in \Omega : |A_n f(x) - f^*| \geq \epsilon\}$, entonces $y \in \{x \in \Omega : \sup_{k \geq n} |A_k f(x) - f^*| \geq \epsilon\}$. Esto significa que:

$$\{x \in \Omega : |A_n f(x) - f^*| \geq \epsilon\} \subseteq \left\{x \in \Omega : \sup_{k \geq n} |A_k f(x) - f^*| \geq \epsilon\right\}. \quad (2.22)$$

Por la proposición 1.2.1, se tiene que (2.22) implica para todo $n \geq 1$:

$$\lambda(\{x \in \Omega : |A_n f(x) - f^*| \geq \epsilon\}) \leq \lambda\left(\left\{x \in \Omega : \sup_{k \geq n} |A_k f(x) - f^*| \geq \epsilon\right\}\right). \quad (2.23)$$

Sustituyendo (1.11) y (1.12) en (2.23) se obtiene $p_n^\epsilon \leq P_n^\epsilon$.

Ahora suponga que se cumple la proposición 2. Dado que $p_n^\epsilon \leq P_n^\epsilon$, se tiene que $p_n^\epsilon \leq P_n^\epsilon \leq C_2(\epsilon)\varphi(n)$ para todo $n > n_2$. Haciendo $C_1(\epsilon) = C_2(\epsilon)$ y $n_1 = n_2$ se completa la demostración del Teorema. \square

A partir del Teorema 2.1.3, los autores de [3] obtienen estimaciones de P_n^ϵ aplicables para diferentes formas de decaimiento de p_n^ϵ . En esta tesis revisaremos solamente el caso de decaimiento exponencial, que es el caso en el que estamos interesados para aplicarlo más adelante. Los dos lemas y el teorema siguientes muestran en conjunto, la forma en que se obtienen estas estimaciones de P_n^ϵ a partir del decaimiento exponencial de p_n^ϵ . Antes de presentar los lemas, recordamos la definición de la función gamma incompleta $\Gamma(a, x) = \int_x^\infty e^{-t} t^{a-1} dt$ con $a, x \in \mathbb{R}$.

Lema 2.1.4. *Para todo $\gamma(\epsilon) > 0$ y $n \geq 1$, la función dada por $\varphi(x) = e^{-\gamma(\epsilon)x}$ cumple la desigualdad 2.18 con $M(\epsilon) \geq e^{\gamma(\epsilon)}\Gamma(0, \gamma(\epsilon))$.*

Demostración.

$$\frac{1}{\varphi(n)} \int_n^\infty \frac{\varphi(x)}{x} dx = e^{\gamma(\epsilon)n} \int_n^\infty \frac{e^{-\gamma(\epsilon)x}}{x} dx = e^{\gamma(\epsilon)n} \Gamma(0, \gamma(\epsilon)n) \leq e^{\gamma(\epsilon)} \Gamma(0, \gamma(\epsilon)).$$

\square

Lema 2.1.5. *Sea $\psi(x) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ una función que satisface 2.18. Entonces, para todo $\delta(\epsilon) > 0$, la función $\varphi(x) = \psi(x^{\delta(\epsilon)})$ también satisface 2.18.*

Demostración.

$$\int_n^\infty \frac{\varphi(x)}{x} dx = \int_n^\infty \frac{\psi(x^{\delta(\epsilon)})}{x} dx = \frac{1}{\delta(\epsilon)} \int_{n^{\delta(\epsilon)}}^\infty \frac{\psi(y)}{y} dy \leq \frac{M(\epsilon)}{\delta(\epsilon)} \psi(n^{\delta(\epsilon)}) = \frac{M(\epsilon)}{\delta(\epsilon)} \varphi(n).$$

Donde hemos hecho el cambio de variable $y = x^\delta$ con lo que $dx = \frac{xdy}{y\delta}$. Esto completa la demostración del Lema 2.1.5. \square

Teorema 2.1.6 ([3]). *Suponga que f es una función esencialmente acotada. Entonces para cada $\epsilon > 0$ fijo, si $p_n^\epsilon \leq C(\epsilon)e^{-\gamma(\epsilon)n^{\delta(\epsilon)}}$ para algunas constantes $\gamma(\epsilon) > 0$, $\delta(\epsilon) > 0$ para todo $n > n_0(\epsilon)$, $n_0(\epsilon) \in \mathbb{Z}^+$, entonces:*

$$P_n^{2\epsilon} < C(\epsilon) \left(1 + \frac{\ln(1 + 1/\gamma(\epsilon))}{\delta(\epsilon) \ln r(\epsilon)}\right) e^{-\gamma(\epsilon)n^{\delta(\epsilon)}}. \quad (2.24)$$

Donde $r = 1 + \frac{\epsilon}{2\|f - f^*\|_\infty}$.

Demostración. Utilizando los lemas 2.1.4 y 2.1.5 se tiene que $\varphi(x) = e^{-\gamma(\epsilon)x^{\delta(\epsilon)}}$ satisface 2.18 con $M(\epsilon) \geq \frac{e^{\gamma(\epsilon)}\Gamma(0, \gamma(\epsilon))}{\delta(\epsilon)}$. Notamos que se satisfacen las condiciones de la proposición 1) del Teorema 2.1.3 con $C_1(\epsilon) = C(\epsilon)$ y $n_1 = n_0(\epsilon)$. Por lo tanto, utilizando la proposición 2) del mismo teorema, se satisface:

$$\begin{aligned} P_n^{2\epsilon} &\leq C(\epsilon) \left(1 + \frac{e^{\gamma(\epsilon)}\Gamma(0, \gamma(\epsilon))}{\delta(\epsilon) \ln r(\epsilon)} \right) e^{-\gamma(\epsilon)n^{\delta(\epsilon)}} \\ &< C(\epsilon) \left(1 + \frac{\ln(1 + 1/\gamma(\epsilon))}{\delta(\epsilon) \ln r(\epsilon)} \right) e^{-\gamma(\epsilon)n^{\delta(\epsilon)}}. \end{aligned}$$

Donde, en la última desigualdad hemos utilizado la siguiente propiedad de la función gamma incompleta [10, desigualdad 5.1.20], válida para todo $x > 0$:

$$e^x \Gamma(0, x) < \ln \left(1 + \frac{1}{x} \right).$$

Esto completa la demostración del Teorema 2.1.6. \square

2.2. Estimación de p_n^ϵ para funciones acotadas y continuas casi en todas partes

En esta sección se presentan con todo detalle, los resultados de [1]. Estos resultados permiten encontrar estimaciones de p_n^ϵ para funciones acotadas y continuas casi en todas partes, utilizando como base estimaciones para funciones Hölder continuas. A lo largo de esta sección, usaremos la siguiente notación.

- Sea $f : \Omega \rightarrow \mathbb{R}$ una función acotada y continua λ casi en todas partes. Denotamos la *norma infinito de f* como:

$$\|f\|_\infty := \sup_{y \in \Omega} |f(y)|. \quad (2.25)$$

- Para $x \in \Omega$ y $\delta > 0$, definimos los siguientes conjuntos:

$$\begin{aligned} S_1^\delta(x) &:= \{f(y) + 2\|f\|_\infty \delta^{-\alpha} d^\alpha(x, y) : y \in \Omega\}, \\ S_2^\delta(x) &:= \{f(y) - 2\|f\|_\infty \delta^{-\alpha} d^\alpha(x, y) : y \in \Omega\}. \end{aligned}$$

- Para $\delta > 0$, denotamos la *vecindad de x de radio δ* como:

$$N_\delta(x) := \{y \in \Omega : d(x, y) \leq \delta\}. \quad (2.26)$$

- Con $f : \Omega \rightarrow \mathbb{R}$ una función medible, denotamos la *medida de la función f* como:

$$\lambda(f) = \int_\Omega f d\lambda.$$

Presentamos ahora algunas definiciones y lemas, y posteriormente presentamos los teoremas de [1] y su demostración.

Definición 2.2.1. Sea Ω un conjunto equipado con una métrica d . Decimos que una función $f : \Omega \rightarrow \mathbb{R}$ es Hölder continua con exponente $\alpha \in (0, 1]$ y lo denotamos $f \in H_\alpha(\Omega)$, si existe una constante $h \geq 0$ tal que, para todo $x, y \in \Omega$:

$$|f(x) - f(y)| \leq h d^\alpha(x, y). \quad (2.27)$$

Denotamos h_f al mínimo valor de h para el que se cumple (2.27). Adicionalmente, si la desigualdad (2.27) se cumple con $\alpha = 1$, decimos que f es una función Lipschitz continua.

Definición 2.2.2. Decimos que una función $f : \Omega \rightarrow \mathbb{R}$ pertenece a la familia \mathcal{F}_α si es λ integrable y además, para todo $\delta > 0$, existen funciones Hölder continuas $g_1^\delta, g_2^\delta \in H_\alpha(\Omega)$ para las cuales:

$$g_1^\delta \leq f \leq g_2^\delta \quad (2.28)$$

$$\lambda(g_1^\delta) + l(\delta) \geq \lambda(f) \geq \lambda(g_2^\delta) - l(\delta), \quad (2.29)$$

donde $l(\delta) \rightarrow 0$ cuando $\delta \rightarrow 0$.

El siguiente teorema permite caracterizar a las funciones acotadas que pertenecen a la familia \mathcal{F}_α . Esta caracterización será útil posteriormente para obtener una estimación de p_n^ϵ para una función $f \in \mathcal{F}_\alpha$. Aunque la prueba del Teorema se encuentra originalmente en [1], en esta tesis la presentamos con todos los detalles que no aparecen desglosados en el artículo.

Teorema 2.2.1 ([1]). Sea $\alpha \in (0, 1]$. Una función acotada f pertenece a la familia \mathcal{F}_α si y solo si f es continua λ casi en todas partes.

La demostración del teorema se basa en una serie de lemas que presentamos a continuación. Posteriormente presentamos la demostración del teorema, la cual consistirá solamente de algunas líneas referenciando cada uno de los lemas.

Lema 2.2.2. La desigualdad (2.28) se cumple con las funciones g_1^δ y g_2^δ dadas por:

$$g_1^\delta(x) = \inf_{y \in \Omega} S_1^\delta(x) \quad (2.30)$$

$$g_2^\delta(x) = \sup_{y \in \Omega} S_2^\delta(x). \quad (2.31)$$

Demostración. Observamos que, dado que $x \in \Omega$, se tiene que:

$$\begin{aligned} f(x) + 2\|f\|_\infty \delta^{-\alpha} d^\alpha(x, x) &\in S_1^\delta(x) \\ f(x) - 2\|f\|_\infty \delta^{-\alpha} d^\alpha(x, x) &\in S_2^\delta(x). \end{aligned}$$

Por propiedades de la métrica, $d(x, x) = 0$ por lo que se tiene que $f(x) \in S_1^\delta(x)$ y $f(x) \in S_2^\delta(x)$. Por definición de \inf y \sup se tiene que:

$$\begin{aligned} f(x) &\geq \inf_{y \in \Omega} S_1^\delta(x) = g_1^\delta(x) \\ f(x) &\leq \sup_{y \in \Omega} S_2^\delta(x) = g_2^\delta(x). \end{aligned}$$

Estas dos desigualdades son equivalentes a (2.28). □

Lema 2.2.3. Las funciones g_1^δ y g_2^δ definidas mediante (2.30) y (2.31) son Hölder continuas con respecto a la métrica definida en Ω .

Demostración. Utilizando el Lema A.4.4 presentado en el Apéndice, se tiene que para todo $x, z \in \Omega$:

$$\begin{aligned} \left| g_1^\delta(x) - g_1^\delta(z) \right| &= \left| \inf_{y \in \Omega} \{f(y) + 2\|f\|_\infty \delta^{-\alpha} d^\alpha(x, y)\} - \inf_{y \in \Omega} \{f(y) + 2\|f\|_\infty \delta^{-\alpha} d^\alpha(z, y)\} \right| \\ &\leq \sup_{y \in \Omega} \{ |f(y) + 2\|f\|_\infty \delta^{-\alpha} d^\alpha(x, y) - f(y) - 2\|f\|_\infty \delta^{-\alpha} d^\alpha(z, y)| \} \\ &= \sup_{y \in \Omega} \{ 2\|f\|_\infty \delta^{-\alpha} |d^\alpha(x, y) - d^\alpha(z, y)| \} \\ &\leq 2\|f\|_\infty \delta^{-\alpha} \sup_{y \in \Omega} \{ |d^\alpha(x, z)| \} \\ &= 2\|f\|_\infty \delta^{-\alpha} d^\alpha(x, z). \end{aligned}$$

Con lo que g_1^δ es Hölder continua con mínima constante Hölder $h_f \leq 2\|f\|_\infty \delta^{-\alpha}$. Similarmente para g_2^δ :

$$\begin{aligned} \left| g_2^\delta(x) - g_2^\delta(z) \right| &= \left| \sup_{y \in \Omega} \{f(y) - 2\|f\|_\infty \delta^{-\alpha} d^\alpha(x, y)\} - \sup_{y \in \Omega} \{f(y) - 2\|f\|_\infty \delta^{-\alpha} d^\alpha(z, y)\} \right| \\ &\leq \sup_{y \in \Omega} \{ |f(y) - 2\|f\|_\infty \delta^{-\alpha} d^\alpha(x, y) - f(y) + 2\|f\|_\infty \delta^{-\alpha} d^\alpha(z, y)| \} \\ &\leq \sup_{y \in \Omega} \{ 2\|f\|_\infty \delta^{-\alpha} |d^\alpha(z, y) - d^\alpha(x, y)| \} \\ &\leq 2\|f\|_\infty \delta^{-\alpha} \sup_{y \in \Omega} \{ |d^\alpha(x, z)| \} \\ &= 2\|f\|_\infty \delta^{-\alpha} d^\alpha(x, z). \end{aligned}$$

Con lo que g_2^δ también es Hölder continua con mínima constante Hölder $h_f \leq 2\|f\|_\infty \delta^{-\alpha}$. Esto completa la demostración del Lema 2.2.3. \square

Lema 2.2.4.

- 1) Sea $\bar{y} \in \Omega$ tal que $g_1^\delta(x) = f(\bar{y}) + 2\|f\|_\infty \delta^{-\alpha} d^\alpha(x, \bar{y})$. Entonces $\bar{y} \in N_\delta(x)$.
- 2) Sea $\tilde{y} \in \Omega$ tal que $g_2^\delta(x) = f(\tilde{y}) - 2\|f\|_\infty \delta^{-\alpha} d^\alpha(x, \tilde{y})$. Entonces $\tilde{y} \in N_\delta(x)$.

Demostración. Demostraremos 1) y 2) por contradicción. Sea $\bar{y} \in \Omega$ tal que $g_1^\delta(x) = f(\bar{y}) + 2\|f\|_\infty \delta^{-\alpha} d^\alpha(x, \bar{y})$ y supongamos que $\bar{y} \notin N_\delta(x)$. Por (2.26) se tiene la siguiente serie de implicaciones:

$$\begin{aligned} d(x, \bar{y}) > \delta &\Rightarrow d^\alpha(x, \bar{y}) > \delta^\alpha \\ &\Rightarrow f(\bar{y}) + 2\|f\|_\infty \delta^{-\alpha} d^\alpha(x, \bar{y}) > f(\bar{y}) + 2\|f\|_\infty \delta^{-\alpha} \delta^\alpha \\ &\Rightarrow g_1^\delta(x) > f(\bar{y}) + 2\|f\|_\infty \\ &\Rightarrow g_1^\delta(x) > \|f\|_\infty. \end{aligned} \tag{2.32}$$

Pero por (2.25), se tiene que $\|f\|_\infty \geq f(x), \forall x \in \Omega$, lo cual conduce a que $g_1^\delta(x) > f(x)$. Pero esto contradice el resultado del lema 2.2.2. Por contradicción concluimos

que $\bar{y} \in N_\delta(x)$. Similarmente, sea $\tilde{y} \in \Omega$ tal que $g_2^\delta(x) = f(\tilde{y}) - 2\|f\|_\infty \delta^{-\alpha} d^\alpha(x, \tilde{y})$ y supongamos que $\tilde{y} \notin N_\delta(x)$. Por (2.26) se tiene la siguiente serie de implicaciones:

$$\begin{aligned} d(x, \tilde{y}) > \delta &\Rightarrow d^\alpha(x, \tilde{y}) > \delta^\alpha \\ &\Rightarrow f(\tilde{y}) - 2\|f\|_\infty \delta^{-\alpha} d^\alpha(x, \tilde{y}) < f(\tilde{y}) - 2\|f\|_\infty \delta^{-\alpha} \delta^\alpha \\ &\Rightarrow g_2^\delta(x) < f(\tilde{y}) - 2\|f\|_\infty \\ &\Rightarrow g_2^\delta(x) < -\|f\|_\infty. \end{aligned} \tag{2.33}$$

Nuevamente (2.33) y (2.25) conducen a una contradicción con el resultado del lema 2.2.2, por lo que se concluye que $\tilde{y} \in N_\delta(x)$. Esto completa la demostración del lema 2.2.4. \square

Lema 2.2.5. *Defina:*

$$\omega_f(\delta, x) := \sup_{y \in N_\delta(x)} |f(x) - f(y)|. \tag{2.34}$$

Entonces, con g_1^δ y g_2^δ dadas por (2.30) y (2.31) se cumple:

$$f(x) - \omega_f(\delta, x) \leq g_1^\delta(x) \leq f(x) \leq g_2^\delta(x) \leq f(x) + \omega_f(\delta, x). \tag{2.35}$$

Demostración. Dado que ya hemos probado (2.28), para probar (2.35) sólo debemos probar sus extremos. Es conveniente observar que (2.34) implica lo siguiente para todo $y \in N_\delta(x)$:

$$\omega_f(\delta, x) \geq |f(x) - f(y)| \geq f(x) - f(y) \tag{2.36}$$

$$\omega_f(\delta, x) \geq |f(x) - f(y)| = |f(y) - f(x)| \geq f(y) - f(x). \tag{2.37}$$

Teniendo esto en mente, sea \bar{y} definido igual que en el lema 2.2.4 y considere los siguientes cálculos:

$$\begin{aligned} f(x) - g_1^\delta(x) &= f(x) - f(\bar{y}) - 2\|f\|_\infty \delta^{-\alpha} d^\alpha(x, \bar{y}) \\ &\leq f(x) - f(\bar{y}) \\ &\leq \omega_f(\delta, x). \end{aligned} \tag{2.38}$$

Donde hemos utilizado las definiciones de $g_1^\delta(x)$ y $\omega_f(\delta, x)$, así como el hecho de que $\bar{y} \in N_\delta(x)$. La desigualdad (2.38) es equivalente al extremo izquierdo de (2.35). Similarmente:

$$\begin{aligned} g_2^\delta(x) - f(x) &= f(\tilde{y}) - 2\|f\|_\infty \delta^{-\alpha} d^\alpha(x, \tilde{y}) - f(x) \\ &\leq f(\tilde{y}) - f(x) \\ &\leq \omega_f(\delta, x). \end{aligned} \tag{2.39}$$

La desigualdad (2.39) es equivalente al extremo derecho de (2.35). Esto completa la demostración del Lema 2.2.5. \square

Lema 2.2.6. *La desigualdad (2.29) se cumple con $l(\delta) = \omega_f(\delta, x)$.*

Demostración. Este resultado se sigue de aplicar la medida λ a la desigualdad (2.35):

$$\begin{aligned} f(x) - \omega_f(\delta, x) \leq g_1^\delta(x) &\Rightarrow \lambda(f) - \lambda(\omega_f) \leq \lambda(g_1^\delta) \\ &\Rightarrow \lambda(g_1^\delta) + \lambda(\omega_f) \geq \lambda(f) \end{aligned} \quad (2.40)$$

$$\begin{aligned} g_2^\delta(x) \leq f(x) + \omega_f(\delta, x) &\Rightarrow \lambda(g_2^\delta) \leq \lambda(f) + \lambda(\omega_f) \\ &\Rightarrow \lambda(g_2^\delta) - \lambda(\omega_f) \leq \lambda(f). \end{aligned} \quad (2.41)$$

Las desigualdades (2.40) y (2.41) son equivalentes a (2.29) si definimos $l(\delta) = \omega_f(\delta, x)$. Notamos también que $\omega_f(\delta, x) \rightarrow 0$ cuando $\delta \rightarrow 0$. Esto completa la demostración del Lema 2.2.6. \square

A partir de estos lemas, presentamos a continuación la demostración del Teorema 2.2.1.

Demostración del Teorema 2.2.1.

\Rightarrow Esta implicación se demuestra mediante contrapositiva utilizando el Teorema ???. La doble implicación de ese Teorema nos permite concluir que si f no es acotada y continua λ casi en todas partes, entonces no puede ser λ -integrable. Dado que no puede ser λ -integrable, tampoco puede pertenecer a la familia \mathcal{F}_α .

\Leftarrow Dado que f es acotada y continua λ casi en todas partes, se tiene por el Teorema ??? que también es λ -integrable. El Lema 2.2.3 permite concluir que g_1^δ y g_2^δ son Hölder continuas y los lemas 2.2.2 y 2.2.6 muestran que además cumplen las ecuaciones (2.28) y (2.29) donde $l(\delta) = \omega_f(\delta, x)$ tiende a cero cuando δ tiende a cero. Estos resultados en conjunto, muestran f satisface la definición 2.2.2 y por lo tanto pertenece a la familia \mathcal{F}_α . Esto completa la demostración del Teorema 2.2.1. \square

El siguiente Teorema permite obtener una estimación de $p_n^\epsilon(f)$ para funciones f acotadas y continuas λ casi en todas partes. Esta estimación se obtiene a partir de estimaciones de p_n^ϵ para funciones Hölder continuas. Su demostración se basa en el Teorema 2.2.1 y en la aplicación del Teorema de Birkhoff a las funciones λ -integrables g_1^δ y g_2^δ . Aunque esta demostración se encuentra originalmente en [1], en esta tesis la presentamos con todos los detalles no desglosados en el artículo.

Teorema 2.2.7 ([1]). *Sea Ω un espacio métrico y sea $(\Omega, \mathcal{B}, \lambda)$ un espacio de probabilidad. Sea $T : \Omega \rightarrow \Omega$ y suponga que λ es una medida ergódica. Entonces, para toda función $f : \Omega \rightarrow \mathbb{R}$ acotada y continua λ casi en todas partes, para todo $\delta > 0$ y para todo $\alpha \in (0, 1]$, existen funciones Hölder continuas $g_1^\delta, g_2^\delta \in H_\alpha(\Omega)$ y $l(\delta) > 0$ tales que, para cualquier $\epsilon > 0$ y para todo $n \geq 1$:*

$$p_n^{\epsilon+l(\delta)}(f) \leq p_n^\epsilon(g_1^\delta) + p_n^\epsilon(g_2^\delta), \quad (2.42)$$

donde $l(\delta) \rightarrow 0$ cuando $\delta \rightarrow 0$.

Demostración. Dado que f es acotada y continua λ casi en todas partes, por el teorema 2.2.1 se tiene que $f \in \mathcal{F}_\alpha$. Es decir, existen funciones Hölder continuas $g_1^\delta, g_2^\delta \in H_\alpha(\Omega)$ tales que, para todo $\delta > 0$ se cumplen (2.28) y (2.29). La desigualdad (2.28) implica:

$$A_n g_1^\delta(x) \leq A_n f(x) \leq A_n g_2^\delta(x). \quad (2.43)$$

También, las condiciones del Teorema de Birkhoff se cumplen con las funciones λ -integrables f , g_1^δ y g_2^δ con lo que se tiene que:

$$\lambda(f) = \int f d\lambda = \int f^* d\lambda = f^*, \quad (2.44)$$

$$\lambda(g_1^\delta) = \int g_1^\delta d\lambda = \int g_1^{\delta*} d\lambda = g_1^{\delta*}, \quad (2.45)$$

$$\lambda(g_2^\delta) = \int g_2^\delta d\lambda = \int g_2^{\delta*} d\lambda = g_2^{\delta*}. \quad (2.46)$$

Utilizando las ecuaciones (2.44), (2.45) y (2.46), definimos los siguientes conjuntos:

$$F_n^\delta = \{x \in \Omega : |A_n f(x) - f^*| \geq \epsilon + l(\delta)\}, \quad (2.47)$$

$$G_{1n}^\delta = \left\{x \in \Omega : |A_n g_1^\delta(x) - g_1^{\delta*}| \geq \epsilon\right\}, \quad (2.48)$$

$$G_{2n}^\delta = \left\{x \in \Omega : |A_n g_2^\delta(x) - g_2^{\delta*}| \geq \epsilon\right\}. \quad (2.49)$$

Usando estas definiciones y la ecuación (1.11), podemos escribir:

$$p_n^{\epsilon+l(\delta)}(f) = \lambda(F_n^\delta), \quad (2.50)$$

$$p_n^\epsilon(g_1^\delta) = \lambda(G_{1n}^\delta), \quad (2.51)$$

$$p_n^\epsilon(g_2^\delta) = \lambda(G_{2n}^\delta). \quad (2.52)$$

Utilizando (2.50), (2.51) y (4.14) podemos reescribir (2.42) como:

$$\lambda(F_n^\delta) \leq \lambda(G_{1n}^\delta) + \lambda(G_{2n}^\delta). \quad (2.53)$$

Por otro lado, utilizando la proposición 1.2.1, sabemos que:

$$\lambda(G_{1n}^\delta \cup G_{2n}^\delta) \leq \lambda(G_{1n}^\delta) + \lambda(G_{2n}^\delta). \quad (2.54)$$

Por lo tanto, para mostrar (2.42), es suficiente mostrar que:

$$\lambda(F_n^\delta) \leq \lambda(G_{1n}^\delta \cup G_{2n}^\delta). \quad (2.55)$$

Por la monotonicidad de la medida (proposición 1.2.1), se tiene que si $F_n^\delta \subseteq G_{1n}^\delta \cup G_{2n}^\delta$, entonces se cumple (2.55). Nos enfocaremos entonces, en mostrar que:

$$x \in F_n^\delta \Rightarrow x \in G_{1n}^\delta \cup G_{2n}^\delta. \quad (2.56)$$

Una forma de mostrar esto es mediante la proposición contrapositiva. Sea $x \notin G_{1n}^\delta \cup G_{2n}^\delta$. Esto significa que $x \notin G_{1n}^\delta$ y que $x \notin G_{2n}^\delta$. Dado que $x \notin G_{1n}^\delta$, se tiene de

(2.48) que $|A_n g_1^\delta(x) - \lambda(g_1^\delta)| < \epsilon$. Esta desigualdad, usando (2.29) y (2.43), tiene la siguiente serie de implicaciones:

$$\begin{aligned}
-\epsilon < A_n g_1^\delta(x) - \lambda(g_1^\delta) < \epsilon &\Rightarrow -\epsilon - l(\delta) < A_n g_1^\delta(x) - \lambda(g_1^\delta) - l(\delta) \\
&\Rightarrow \epsilon + l(\delta) > \lambda(g_1^\delta) + l(\delta) - A_n g_1^\delta(x) \\
&\Rightarrow \epsilon + l(\delta) > \lambda(f) - A_n g_1^\delta(x) \\
&\Rightarrow \epsilon + l(\delta) > \lambda(f) - A_n f(x). \tag{2.57}
\end{aligned}$$

Análogamente, dado que $x \notin G_{2n}^\delta$, se tiene de (2.49) que $|A_n g_2^\delta(x) - \lambda(g_2^\delta)| < \epsilon$. Esta desigualdad, usando (2.29) y (2.43), tiene la siguiente serie de implicaciones:

$$\begin{aligned}
-\epsilon < A_n g_2^\delta(x) - \lambda(g_2^\delta) < \epsilon &\Rightarrow A_n f(x) - \lambda(g_2^\delta) < \epsilon \\
&\Rightarrow \lambda(g_2^\delta) - A_n f(x) > -\epsilon \\
&\Rightarrow \lambda(g_2^\delta) - l(\delta) - A_n f(x) > -\epsilon - l(\delta) \\
&\Rightarrow \lambda(f) - A_n f(x) > -\epsilon - l(\delta). \tag{2.58}
\end{aligned}$$

Las desigualdades (2.57) y (2.58) son equivalentes a $|A_n f(x) - \lambda(f)| < \epsilon + l(\delta)$. Recordamos que $F_n^\delta = \{x \in \Omega : |A_n f(x) - \lambda(f)| \geq \epsilon + l(\delta)\}$ por lo que se tiene que $x \notin F_n^\delta$. Por contrapositiva se concluye que (2.56) es verdadera y por lo tanto $F_n^\delta \subseteq G_{1n}^\delta \cup G_{2n}^\delta$.

Por la monotonicidad de la medida, esto significa que $\lambda(F_n^\delta) \leq \lambda(G_{1n}^\delta \cup G_{2n}^\delta)$ y por lo tanto $\lambda(F_n^\delta) \leq \lambda(G_{1n}^\delta) + \lambda(G_{2n}^\delta)$. Utilizando las definiciones de p_n^ϵ y P_n^ϵ se tiene que $p_n^{\epsilon+l(\delta)}(f) \leq p_n^\epsilon(g_1^\delta) + p_n^\epsilon(g_2^\delta)$. Esto completa la demostración del teorema 2.2.7. \square

Es importante notar que las hipótesis del Teorema 2.2.7 son muy generales en el sentido de que no establecen restricciones sobre el espacio de probabilidad más que Ω sea un espacio métrico y λ sea una medida ergódica con respecto a una transformación $T : \Omega \rightarrow \Omega$. Esto es importante ya que en la sección 4.1, estaremos interesados en aplicar el teorema en espacios de probabilidad en los que Ω es un espacio de secuencias y λ es una medida de Gibbs.

2.3. Teoría de grandes desvíos

En esta sección se presenta un breve resumen de la teoría de grandes desvíos. Esta teoría estudia la probabilidad de que la media empírica de n variables aleatorias, tome valores alejados ϵ de su valor esperado. A partir de esta teoría, es posible encontrar estimaciones de p_n^ϵ . Para comenzar, consideramos la suma de variables aleatorias reales X_1, X_2, \dots, X_n :

$$S_n(X) = \sum_{i=1}^n X_i \tag{2.59}$$

Estamos interesados en estudiar el comportamiento estadístico de $\frac{1}{n}S_n$, en particular, nos interesa saber si sigue alguna distribución de probabilidad y qué forma tiene esta distribución. Si asumimos que las variables aleatorias X_i son independientes e

idénticamente distribuidas, por el Teorema del Límite Central, $\frac{1}{n}S_n$ tiende a una variable aleatoria distribuida de forma Normal. Sin embargo, si las variables aleatorias no son independientes o idénticamente distribuidas, no podemos aplicar el Teorema del Límite Central directamente en todos los casos. Además, es importante notar que este teorema es válido en el límite y no dice nada acerca de la distribución de $\frac{1}{n}S_n$ para valores finitos de n . La Teoría de Grandes Desvíos se enfoca en encontrar la distribución de probabilidad de la variable aleatoria $\frac{1}{n}S_n$ denotada $P(\frac{1}{n}S_n = s)$. Note que P es la probabilidad con la que se distribuyen las variables aleatorias X_1, X_2, \dots, X_n . En este contexto, presentamos la siguiente definición:

Definición 2.3.1. *Decimos que una variable aleatoria A_n cumple un principio de grandes desvíos con función tasa $I(a)$ si el siguiente límite existe:*

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \ln P(A_n = a) = I(a). \quad (2.60)$$

A partir de esta definición, se sigue que, si podemos escribir $P(A_n = a)$ como $P(A_n = a) = e^{-nI(a) + \mathcal{O}(n)}$, donde $\frac{\mathcal{O}(n)}{n} \rightarrow 0$ cuando $n \rightarrow \infty$, entonces A_n cumple un principio de grandes desvíos. Esto es así ya que:

$$\begin{aligned} P(A_n = a) = e^{-nI(a) + \mathcal{O}(n)} &\Rightarrow \ln P(A_n = a) = -nI(a) + \mathcal{O}(n) \\ &\Rightarrow -\frac{1}{n} \ln P(A_n = a) = I(a) - \frac{\mathcal{O}(n)}{n} \\ &\Rightarrow \lim_{n \rightarrow \infty} -\frac{1}{n} \ln P(A_n = a) = I(a). \end{aligned}$$

Con valores de n muy grandes, $\frac{\mathcal{O}(n)}{n} \approx 0$ y podemos aproximar $P(A_n = s)$ mediante:

$$P(A_n = a) \approx e^{-nI(a)} \quad (2.61)$$

El siguiente ejemplo ilustra el caso de una variable aleatoria, para la cual se puede demostrar de forma directa, que cumple un principio de grandes desvíos y se puede encontrar la función tasa.

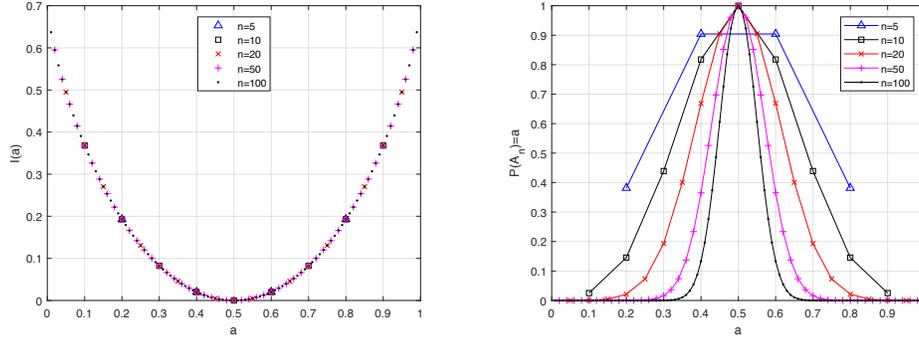
Ejemplo 2.3.1. *Considere la variable aleatoria $A_n(X) = \frac{1}{n}S_n(X)$ donde las X_i son variables aleatorias independientes que toman valores de 0 o 1 con igual probabilidad. Dados valores x_1, \dots, x_n , con $x_i \in \{0, 1\}$, por la independencia de las X_i , se tiene que:*

$$P(X = x) = P(X_1 = x_1, \dots, X_n = x_n) = \frac{1}{2^n}.$$

Para obtener $P(A_n(X) = a)$, tenemos que sumar sobre todos los valores de x_1, \dots, x_n tales que $A_n(x) = a$. Dado que la probabilidad $P(X = x) = \frac{1}{2^n}$ es la misma para toda secuencia x_1, \dots, x_n , con $x_i \in \{0, 1\}$, se tiene:

$$P(A_n(X) = a) = \sum_{x: A_n(x)=a} P(X = x) = \frac{1}{2^n} \# \{x : A_n(x) = a\}.$$

Donde $\#$ denota la cardinalidad del conjunto. Ahora hacemos notar que, dado que $A_n(x)$ sólo puede tomar los valores $0, 1/n, 2/n, \dots, 1$, $A_n(x) = a$ es equivalente a que



$$(a) I(a) = \ln 2 + a \ln a + (1-a) \ln(1-a)$$

$$(b) P(A_n = a) \approx e^{-nI(a)}$$

Fig. 2.1: Función tasa (a) y distribución de probabilidad de A_n (b) para el ejemplo 2.3.1

el número de 1's en x es an . Esto significa que $\#\{x : A_n(x) = a\}$ está dada por el número de formas diferentes en las que x puede tener an veces el valor de 1. A partir de esta observación y utilizando la fórmula combinatoria, la distribución de A_n puede escribirse:

$$P(A_n = a) = \frac{1}{2^n} \frac{n!}{(an)![(1-a)n]}.$$

Tomando n suficientemente grande, podemos utilizar la fórmula de Stirling, $n! \approx n^n e^{-n}$, con lo que la ecuación anterior puede escribirse como:

$$\begin{aligned} P(A_n = a) &\approx \frac{1}{2^n} \frac{n^n e^{-n}}{(an)^{an} e^{-an} ((1-a)n)^{(1-a)n} e^{-(1-a)n}} \\ &= \frac{1}{2^n} \frac{1}{(an)^{an} e^{-an} ((1-a)n)^{(1-a)n} e^{-(1-a)n} n^{-n} e^n} \\ &= \frac{1}{2^n} \frac{1}{a^{an} (1-a)^{(1-a)n}}. \end{aligned}$$

Si definimos $I(a) := \ln 2 + a \ln a + (1-a) \ln(1-a)$ se tiene que:

$$-nI(a) = \ln(2^{-n}) + \ln(a^{-na}) + \ln\left((1-a)^{-n(1-a)}\right)$$

Luego:

$$P(A_n = a) \approx e^{-nI(a)}.$$

Aplicando directamente la definición, vemos que A_n cumple un principio de grandes desvíos con función tasa dada por $I(a) = \ln 2 + a \ln a + (1-a) \ln(1-a)$. La figura 2.1 muestra las gráficas de $I(a)$ y $P(A_n = a)$ para diferentes valores de n .

No siempre es posible concluir directamente que una variable aleatoria cumple un principio de grandes desvíos ni encontrar la función tasa. En estos casos, el

Teorema de Gärtner-Ellis permite obtener las mismas conclusiones para una clase más general de variables aleatorias. Una versión simplificada de este teorema se presenta a continuación. La versión completa de este Teorema y su demostración, pueden encontrarse en [7].

Teorema 2.3.1. *Sea S_n definida como en la ecuación (2.59) y sea A_n la variable aleatoria dada por $A_n = \frac{1}{n}S_n$. Defina:*

$$\lambda(k) := \lim_{n \rightarrow \infty} \frac{1}{n} \ln \mathbb{E} \left[e^{nkA_n} \right], \quad (2.62)$$

si el límite existe. Entonces, si $\lambda(k)$ es diferenciable, A_n cumple un principio de grandes desvíos con función tasa dada por:

$$I(a) = \sup_{k \in \mathbb{R}} \{ka - \lambda(k)\}. \quad (2.63)$$

Demostración. Ver Capítulo 2, Sección 3 de [7]. □

A continuación se presentan los siguientes ejemplos para ilustrar la aplicación del Teorema de Gärtner-Ellis.

Ejemplo 2.3.2. *Mostraremos que la media empírica de variables aleatorias independientes e idénticamente distribuidas de forma Normal con media μ y varianza σ^2 , cumple un principio de grandes desvíos y obtendremos la función tasa. Por el Teorema de Gärtner-Ellis, para mostrar que A_n cumple un principio de grandes desvíos debemos mostrar que $\lambda(k)$ es diferenciable. En este caso podemos escribir:*

$$\begin{aligned} \lambda(k) &= \lim_{n \rightarrow \infty} \frac{1}{n} \ln \mathbb{E} \left[e^{nkA_n} \right] \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \ln \mathbb{E} \left[e^{k \sum_{i=1}^n X_i} \right] \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \ln \mathbb{E} \left[e^{kX_1} e^{kX_2} \dots e^{kX_n} \right]. \end{aligned}$$

Dado que X_1, X_2, \dots, X_n son independientes:

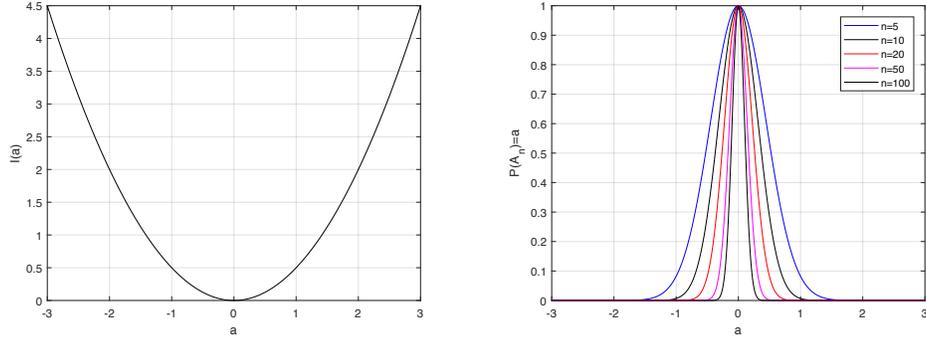
$$\lambda(k) = \lim_{n \rightarrow \infty} \frac{1}{n} \ln \left(\mathbb{E} \left[e^{kX_1} \right] \mathbb{E} \left[e^{kX_2} \right] \dots \mathbb{E} \left[e^{kX_n} \right] \right).$$

Y dado que X_1, X_2, \dots, X_n son idénticamente distribuidas:

$$\begin{aligned} \lambda(k) &= \lim_{n \rightarrow \infty} \frac{1}{n} \ln \left(\mathbb{E} \left[e^{kX} \right] \right)^n \\ &= \lim_{n \rightarrow \infty} \ln \left(\mathbb{E} \left[e^{kX} \right] \right) \\ &= \ln \left(\mathbb{E} \left[e^{kX} \right] \right). \end{aligned}$$

Donde X es cualquiera de las variables X_1, X_2, \dots, X_n . Observamos que $\mathbb{E} \left[e^{kX} \right] = e^{\mu k + \frac{1}{2}\sigma^2 k^2}$ para una variable distribuida de forma normal, por lo que:

$$\begin{aligned} \lambda(k) &= \ln \left(e^{\mu k + \frac{1}{2}\sigma^2 k^2} \right) \\ &= \mu k + \frac{1}{2}\sigma^2 k^2. \end{aligned}$$



$$(a) I(a) = \frac{(a - \mu)^2}{2\sigma^2}$$

$$(b) P(A_n = a) \approx e^{-\frac{n(a-\mu)^2}{2\sigma^2}}$$

Fig. 2.2: Función tasa (a) y distribución de probabilidad de A_n (b) para el ejemplo 2.3.2

Observamos que $\lambda(k)$ es diferenciable por lo que concluimos que A_n cumple un principio de grandes desvíos con:

$$\begin{aligned} I(a) &= \sup_{k \in \mathbb{R}} \{ka - \lambda(k)\} \\ &= \sup_{k \in \mathbb{R}} \left\{ ka - \mu k - \frac{1}{2}\sigma^2 k^2 \right\}. \end{aligned}$$

Dado que:

$$\frac{d}{dk} \left(ka - \mu k - \frac{1}{2}\sigma^2 k^2 \right) = a - \mu - \sigma^2 k = 0,$$

Luego, $k = \frac{a - \mu}{\sigma^2}$ y por lo tanto:

$$\begin{aligned} I(a) &= \left(\frac{a - \mu}{\sigma^2} \right) a - \left(\frac{a - \mu}{\sigma^2} \right) \mu + \frac{1}{2}\sigma^2 \left(\frac{a - \mu}{\sigma^2} \right)^2 \\ &= \frac{a^2 - 2\mu a + \mu^2}{\sigma^2} - \frac{(a - \mu)^2}{2\sigma^2} = \frac{(a - \mu)^2}{2\sigma^2}. \end{aligned}$$

Esto quiere decir que:

$$P(A_n = a) \approx e^{-\frac{n(a-\mu)^2}{2\sigma^2}}.$$

La figura 2.2 muestra las gráficas de $I(a)$ y de $P(A_n = a)$ para este ejemplo.

Ejemplo 2.3.3. Sea una cadena de Markov formada por las variables aleatorias X_1, X_2, \dots, X_n que toman valores en un conjunto finito $\Sigma = \{1, \dots, m\}$. Suponemos que la matriz de transición de estados, $\mathbf{\Pi}$, es irreducible y denotamos sus elementos como $\pi_{i,j}$. Estamos interesados en obtener un principio de grandes desvíos para la variable aleatoria:

$$S_n = \frac{1}{n} \sum_{i=1}^n f(X_i).$$

donde $f : \Sigma \rightarrow \mathbb{R}$. Para ello considere la secuencia $x_1, \dots, x_n \in \Sigma$. Dado un valor inicial $X_0 = x_0 \in \Sigma$, la probabilidad de que la cadena de Markov X_1, X_2, \dots, X_n tome los valores x_1, \dots, x_n es:

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n | X_0 = x_0) = \pi_{x_0, x_1} \prod_{i=1}^{n-1} \pi_{x_i, x_{i+1}}.$$

En los cálculos siguientes, $\mathbb{E}_{x_0} [e^{nkS_n}]$ denota el valor esperado de e^{nkS_n} dado el estado inicial x_0 . Para comenzar, es conveniente notar que:

$$\begin{aligned} \mathbb{E}_{x_0} [e^{nkS_n}] &= \mathbb{E}_{x_0} [e^{k \sum_{i=1}^n f(X_i)}] = \mathbb{E}_{x_0} [e^{kf(X_1)} \dots e^{kf(X_n)}] \\ &= \sum_{x_1, \dots, x_n \in \Sigma} e^{kf(x_1)} \dots e^{kf(x_n)} \pi_{x_0, x_1} \prod_{i=1}^{n-1} \pi_{x_i, x_{i+1}} \\ &= \sum_{x_1, \dots, x_n \in \Sigma} \pi_{x_0, x_1} e^{kf(x_1)} \dots \pi_{x_{n-1}, x_n} e^{kf(x_n)} \end{aligned} \quad (2.64)$$

$$= \sum_{x_1, \dots, x_n \in \Sigma} \left(\prod_{i=0}^{n-1} \pi_{x_i, x_{i+1}} e^{kf(x_{i+1})} \right). \quad (2.65)$$

Donde se ha utilizado la definición de valor esperado para variables aleatorias discretas $\mathbb{E}[X] = \sum_{i=1}^n x_i p(x_i)$. Es importante notar que la sumatoria en (2.64) es sobre todos los valores de $x_i \in \Sigma$. Para ejemplificar, si $n = 3$, $m = 2$, la sumatoria en (2.64) está compuesta de $2^3 = 8$ términos. Definimos la matriz $\mathbf{\Pi}_k$ como $(\mathbf{\Pi}_k)_{i,j} = \pi_{i,j} e^{kf(x_j)}$ y utilizando el lema A.4.1 podemos escribir:

$$\mathbb{E}_{x_0} [e^{nkS_n}] = \sum_{x_n=1}^m (\mathbf{\Pi}_k^n)_{x_0, x_n}.$$

Calculamos ahora $\lambda(k)$:

$$\begin{aligned} \lambda(k) &= \lim_{n \rightarrow \infty} \frac{1}{n} \ln \mathbb{E}_{x_0} [e^{nkS_n}] \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \ln \left(\sum_{x_n=1}^m (\mathbf{\Pi}_k^n)_{x_0, x_n} \right). \end{aligned}$$

Dado que $(\mathbf{\Pi}_k)_{i,j} = \pi_{i,j} e^{kf(x_j)} \geq 0$, se tiene que $\mathbf{\Pi}_k$ es una matriz irreducible. Por el inciso 3 del Teorema A.1, se tiene que:

$$\lambda(k) = \lim_{n \rightarrow \infty} \frac{1}{n} \ln \left(\sum_{x_n=1}^m (\mathbf{\Pi}_k^n)_{x_0, x_n} \right) = \ln \rho(\mathbf{\Pi}_k).$$

Donde $\rho(\mathbf{\Pi}_k)$ es el eigenvalor dominante de $\mathbf{\Pi}_k$. De acuerdo al Teorema A.1, esta última expresión es válida para todo $x_0 \in \Sigma$. En nuestro caso, esto significa que la expresión es válida para cualquier condición inicial. La función $\lambda(k)$ es diferenciable por lo que podemos concluir que S_n cumple un principio de grandes desvíos con función tasa:

$$I(s) = \sup_{k \in \mathbb{R}} \{ks - \ln \rho(\mathbf{\Pi}_k)\}.$$

2.4. Estimación de p_n^ϵ para funciones Lipschitz continuas

En esta subsección se presenta un resultado obtenido en [8] con respecto a sistemas dinámicos simbólicos con medida de Gibbs. En este artículo, los autores toman como base una desigualdad exponencial obtenida en [9] y obtienen una desigualdad de concentración para sistemas dinámicos simbólicos con medida de Gibbs aplicable a funciones Lipschitz continuas. En esta subsección se detallan los resultados de [8] relevantes para esta tesis. En particular, el resultado que nos interesa demostrar es el Teorema 2.4.4. Para comenzar presentamos la siguiente definición.

Definición 2.4.1. Sea $K : \Omega^n \rightarrow \mathbb{R}$ una función de n variables denotadas $\underline{x}^{(0)}, \dots, \underline{x}^{(n-1)}$. Para cada $j = 0, \dots, n-1$ defina:

$$\text{Lip}_j(K) := \sup_{\underline{x}^{(0)}, \dots, \underline{x}^{(n-1)}} \sup_{\underline{x}^{(j)} \neq \underline{y}^{(j)}} \frac{|K(\underline{x}^{(0)}, \dots, \underline{x}^{(j-1)}, \underline{x}^{(j)}, \underline{x}^{(j+1)}, \dots, \underline{x}^{(n-1)}) - K(\underline{x}^{(0)}, \dots, \underline{x}^{(j-1)}, \underline{y}^{(j)}, \underline{x}^{(j+1)}, \dots, \underline{x}^{(n-1)})|}{d_\theta(\underline{x}^{(j)}, \underline{y}^{(j)})} \quad (2.66)$$

Decimos que K es una función Lipschitz separable de n variables si, para $j = 0, \dots, n-1$:

$$\text{Lip}_j K < \infty. \quad (2.67)$$

A continuación se presenta la desigualdad exponencial obtenida en [9].

Teorema 2.4.1 ([9]). Sea μ_ϕ una medida de Gibbs con potencial Hölder continuo ϕ . Entonces existe una constante $D = D(\phi) > 0$ tal que, para todo entero $n \geq 1$ y para toda función K Lipschitz separable de n variables, se tiene:

$$\int e^{K(\underline{x}, \dots, \sigma^{n-1}(\underline{x}))} d\mu_\phi(\underline{x}) \leq e^{\int K(\underline{y}, \dots, \sigma^{n-1}(\underline{y})) d\mu_\phi(\underline{y})} d\mu_\phi(\underline{x}) e^{D \sum_{i=0}^{n-1} \text{Lip}_i^2(K)}. \quad (2.68)$$

Demostración. Ver [9]. □

Los siguientes lemas serán la base para demostrar el Teorema 2.4.4:

Lema 2.4.2 (Desigualdad de Markov). Sea $\epsilon > 0$, λ una medida de probabilidad y X una variable aleatoria no negativa, con valor esperado finito y que toma valores en Ω . Entonces $\lambda\{X \geq \epsilon\} \leq \frac{\mathbb{E}[X]}{\epsilon}$.

Demostración. Para todo $\epsilon > 0$, defina:

$$s(X) = \begin{cases} \epsilon & X \geq \epsilon \\ 0 & X < \epsilon \end{cases}$$

Entonces, se tiene que: $0 \leq s(X) \leq X$. Dado que X y $s(X)$ son no negativas:

$$\mathbb{E}[X] = \int_{\Omega} X d\lambda \geq \int_{\Omega} s(X) d\lambda = \int_{X \geq \epsilon} \epsilon d\lambda = \epsilon \int_{X \geq \epsilon} d\lambda = \epsilon \lambda\{X \geq \epsilon\}. \quad (2.69)$$

□

Lema 2.4.3 (Desigualdad de Markov para funciones crecientes). *Sea $f(X)$ una función creciente, con valor esperado finito. Entonces $\lambda\{X \geq \epsilon\} \leq \frac{\mathbb{E}[f(X)]}{f(\epsilon)}$.*

Demostración. Como f es creciente, se tiene que $X \geq \epsilon$ si y solo si $f(X) \geq f(\epsilon)$. Por lo tanto $\lambda\{X \geq \epsilon\} = \lambda\{f(X) \geq f(\epsilon)\}$. Usando el Lema 2.4.2, se tiene:

$$\lambda\{X \geq \epsilon\} = \lambda\{f(X) \geq f(\epsilon)\} \leq \frac{\mathbb{E}[f(X)]}{f(\epsilon)}. \quad (2.70)$$

□

Observación 2.4.1. *Con $\kappa > 0$ y $f(X) = e^{\kappa X}$, la desigualdad 2.70 se convierte en:*

$$\lambda\{X \geq \epsilon\} = e^{-\kappa\epsilon} \mathbb{E}[e^{\kappa X}] = e^{-\kappa\epsilon} \int_{\Omega} e^{\kappa X} d\lambda. \quad (2.71)$$

El siguiente resultado obtenido en [8] proporciona una desigualdad de p_n^ϵ para sistemas dinámicos simbólicos con medida de Gibbs aplicable a funciones Lipschitz continuas.

Teorema 2.4.4 ([8]). *Sea S un conjunto finito y $\Omega = S^{\mathbb{N}}$. Sea f una función Lipschitz continua cuya mínima constante de Lipschitz con respecto a la métrica d_θ , es $|f|_\theta$. Sea σ el mapeo shift y μ_ϕ una medida de Gibbs con potencial Hölder continuo ϕ . Entonces existe una constante $D = D(\phi) > 0$ tal que se cumple la siguiente desigualdad:*

$$\mu_\phi \left\{ \underline{x} \in \Omega : \left| \frac{1}{n} (f(\underline{x}) + \dots + f(\sigma^{n-1}\underline{x})) - \int_{\Omega} f d\mu_\phi \right| \geq \epsilon \right\} \leq 2 \exp \left(-\frac{n\epsilon^2}{4D|f|_\theta^2} \right) \quad (2.72)$$

Demostración. Sean μ_ϕ una medida de Gibbs con potencial ϕ y X una variable aleatoria dada por $X = K(\underline{x}, \dots, \sigma^{n-1}(\underline{x})) - \int_{\Omega} K(\underline{y}, \dots, \sigma^{n-1}(\underline{y})) d\mu_\phi(\underline{y})$, donde K una función Lipschitz separable de n variables. La desigualdad (2.71) se convierte en:

$$\begin{aligned} & \mu_\phi \left\{ \underline{x} \in \Omega : K(\underline{x}, \dots, \sigma^{n-1}(\underline{x})) - \int_{\Omega} K(\underline{y}, \dots, \sigma^{n-1}(\underline{y})) d\mu_\phi(\underline{y}) \geq \epsilon \right\} \\ & \leq e^{-\kappa\epsilon} \int_{\Omega} e^{\kappa[K(\underline{x}, \dots, \sigma^{n-1}(\underline{x})) - \int_{\Omega} K(\underline{y}, \dots, \sigma^{n-1}(\underline{y})) d\mu_\phi(\underline{y})]} d\mu_\phi(\underline{x}) \\ & = e^{-\kappa\epsilon} \int_{\Omega} e^{\kappa K(\underline{x}, \dots, \sigma^{n-1}(\underline{x}))} e^{-\int_{\Omega} \kappa K(\underline{y}, \dots, \sigma^{n-1}(\underline{y})) d\mu_\phi(\underline{y})} d\mu_\phi(\underline{x}) \\ & = e^{-\kappa\epsilon} e^{-\int_{\Omega} \kappa K(\underline{y}, \dots, \sigma^{n-1}(\underline{y})) d\mu_\phi(\underline{y})} \int_{\Omega} e^{\kappa K(\underline{x}, \dots, \sigma^{n-1}(\underline{x}))} d\mu_\phi(\underline{x}) \\ & \leq e^{-\kappa\epsilon} e^{D \sum_{i=0}^{n-1} \text{Lip}_i^2(\kappa K)} = \exp \left(-\kappa\epsilon + \kappa^2 D \sum_{i=0}^{n-1} \text{Lip}_i^2(K) \right). \end{aligned}$$

Donde en la última desigualdad, hemos utilizado el Teorema 2.4.1. Esta desigualdad es válida para todo $\kappa > 0$. Podemos obtener una desigualdad óptima, al minimizar

sobre κ :

$$\begin{aligned} \frac{d}{d\kappa} \left(-\kappa\epsilon + \kappa^2 D \sum_{i=0}^{n-1} \text{Lip}_i^2(K) \right) &= -\epsilon + 2\kappa D \sum_{i=0}^{n-1} \text{Lip}_i^2(K) = 0 \\ \kappa &= \frac{\epsilon}{2D \sum_{i=0}^{n-1} \text{Lip}_i^2(K)}. \end{aligned}$$

Con lo que se llega a:

$$\begin{aligned} \mu_\phi \left\{ \underline{x} \in \Omega : K(\underline{x}, \dots, \sigma^{n-1}(\underline{x})) - \int_{\Omega} K(\underline{y}, \dots, \sigma^{n-1}(\underline{y})) d\mu_\phi(\underline{y}) \geq \epsilon \right\} \\ \leq \exp \left(- \left(\frac{\epsilon}{2D \sum_{i=0}^{n-1} \text{Lip}_i^2(K)} \right) \epsilon + \left(\frac{\epsilon}{2D \sum_{i=0}^{n-1} \text{Lip}_i^2(K)} \right)^2 D \sum_{i=0}^{n-1} \text{Lip}_i^2(K) \right) \\ = \exp \left(- \frac{\epsilon^2}{4D \sum_{i=0}^{n-1} \text{Lip}_i^2(K)} \right). \end{aligned} \quad (2.73)$$

En particular, defina:

$$K(\underline{x}, \dots, \sigma^{n-1}(\underline{x})) := f(\underline{x}) + f(\sigma \underline{x}) + \dots + f(\sigma^{n-1} \underline{x}), \quad (2.74)$$

donde f es una función Lipschitz continua. En este caso se tiene que:

$$\begin{aligned} \mu_\phi \left\{ \underline{x} \in \Omega : f(\underline{x}) + \dots + f(\sigma^{n-1} \underline{x}) - \int_{\Omega} (f(\underline{y}) + \dots + f(\sigma^{n-1} \underline{y})) d\mu_\phi(\underline{y}) \geq n\epsilon \right\} \\ \leq \exp \left(- \frac{n^2 \epsilon^2}{4D \sum_{i=0}^{n-1} \text{Lip}_i^2(f(\underline{x}) + \dots + f(\sigma^{n-1} \underline{x}))} \right) \\ = \exp \left(- \frac{n^2 \epsilon^2}{4Dn \text{Lip}_i^2(f)} \right) \\ \leq \exp \left(- \frac{n\epsilon^2}{4D|f|_\theta^2} \right). \end{aligned} \quad (2.75)$$

Ahora hacemos notar que:

$$\begin{aligned} \mu_\phi \left\{ \underline{x} \in \Omega : f(\underline{x}) + \dots + f(\sigma^{n-1} \underline{x}) - \int_{\Omega} (f(\underline{y}) + \dots + f(\sigma^{n-1} \underline{y})) d\mu_\phi(\underline{y}) \geq n\epsilon \right\} \\ = \mu_\phi \left\{ \underline{x} \in \Omega : \frac{1}{n} \sum_{i=0}^{n-1} f(\sigma^i \underline{x}) - \int_{\Omega} f(\underline{y}) d\mu_\phi(\underline{y}) \geq \epsilon \right\}. \end{aligned} \quad (2.76)$$

Utilizando (2.75) y (2.76) se llega a:

$$\mu_\phi \left\{ \underline{x} \in \Omega : \frac{1}{n} \sum_{i=0}^{n-1} f(\sigma^i \underline{x}) - \int_{\Omega} f(\underline{y}) d\mu_\phi(\underline{y}) \geq \epsilon \right\} \leq \exp \left(- \frac{n\epsilon^2}{4D|f|_\theta^2} \right). \quad (2.77)$$

Aplicando el mismo procedimiento con $K(\underline{x}, \dots, \sigma^{n-1}(\underline{x})) = -f(\underline{x}) - f(\sigma \underline{x}) - \dots - f(\sigma^{n-1} \underline{x})$, se obtiene la desigualdad análoga:

$$\mu_\phi \left\{ \underline{x} \in \Omega : \int_{\Omega} f(\underline{y}) d\mu_\phi(\underline{y}) - \frac{1}{n} \sum_{i=0}^{n-1} f(\sigma^i \underline{x}) \geq \epsilon \right\} \leq \exp \left(- \frac{n\epsilon^2}{4D|f|_\theta^2} \right). \quad (2.78)$$

Uniendo (2.77) y (2.78) se obtiene (2.72). Esto completa la demostración del Teorema 2.4.4. \square

3. Estimación de desvío casi cierto

En este capítulo presentamos algunos de los resultados de [11] y [12] relevantes para esta tesis. Brevemente los resultados son como sigue. En [11], el autor muestra que, para funciones f que son Hölder continuas y potenciales ϕ cuya variación presenta un decaimiento polinomial, es posible obtener una estimación de la función de autocorrelación $\rho_f(N)$ cuyo decaimiento también será polinomial.

En [12], el autor muestra que, si la función de autocorrelación presenta un decaimiento polinomial, entonces es posible obtener una estimación del desvío de $A_n f$ con respecto a su valor esperado. Uniendo estos resultados, el autor de [11] obtiene una desigualdad de desvío de $A_n f$ para potenciales ϕ cuya variación decae de forma polinomial y para funciones f Hölder continuas. Este resultado se presenta en esta tesis como el Corolario 3.2.1 y en la sección 4.3 lo retomamos para ampliar su aplicabilidad a funciones f acotadas y continuas μ_ϕ casi en todas partes.

3.1. Estimación de la función de autocorrelación

El objetivo de esta sección es demostrar detalladamente el Teorema 3.1.1. La demostración requerirá un buen número de resultados preliminares por lo que primero se enuncian algunas definiciones, después se enuncian los lemas necesarios y posteriormente se presenta la demostración del teorema. Para comenzar, definimos la función de autocorrelación de orden N , de una función cuadrado integrable $f : \Omega \rightarrow \mathbb{R}$ como:

$$\rho_f(N) = \int_{\Omega} (f \circ \sigma^N) f d\mu_\phi - \left(\int_{\Omega} f d\mu_\phi \right)^2.$$

Si σ es mezclante y preserva la medida μ_ϕ , entonces la función de autocorrelación tiende a cero cuando n tiende a infinito. La tasa del decaimiento de la función de autocorrelación mide la velocidad con la que la dinámica del sistema se vuelve independiente de las condiciones iniciales.

En la sección 1.6, definimos el conjunto Ω a partir de un conjunto finito S . El conjunto Ω fue definido entonces como $\Omega = S^{\mathbb{N}}$. De aquí en adelante en esta tesis, estaremos interesados en el conjunto Ω_A^+ definido a continuación.

Definición 3.1.1. *Sea $S = \{1, 2, \dots, k\}$ donde k es un entero que pertenece al intervalo $(0, \infty)$. Sea A una matriz cuadrada de $k \times k$ cuyos elementos son 0 o 1.*

Llamamos subshift de tipo finito al conjunto:

$$\Omega_A = \left\{ \underline{x} \in S^{\mathbb{Z}} : A_{x_i, x_{i+1}} = 1, i \in \mathbb{Z} \right\}.$$

Llamamos subshift de tipo finito de un solo lado, al conjunto:

$$\Omega_A^+ = \left\{ \underline{x} \in S^{\mathbb{N}_0} : A_{x_i, x_{i+1}} = 1, i \geq 0 \right\}.$$

Al conjunto S le llamamos alfabeto y a la matriz A le llamamos matriz de transición.

Aunque los conjuntos Ω_A y Ω_A^+ están definidos a partir de $S = \{1, 2, \dots, k\}$, es importante notar que, si estamos interesados en cualquier otro alfabeto finito S' , podemos definir una biyección entre S y S' , $\pi : S' \rightarrow S$, y las definiciones de Ω_A y Ω_A^+ pueden modificarse con $A_{\pi(x_i), \pi(x_{i+1})}$ en lugar de $A_{x_i, x_{i+1}}$.

Un subshift de tipo finito puede ser visto como un espacio de símbolos restringido, ya que una secuencia \underline{x} con dos símbolos consecutivos cualesquiera $x_i, x_{i+1} \in S$, solo puede pertenecer al conjunto Ω_A si el elemento $A_{x_i, x_{i+1}}$ es igual a 1. Pese a esta restricción, es importante notar que cualquier función definida en Ω , también estará definida en Ω_A .

Si el decaimiento de la variación es polinomial, entonces el decaimiento de la función de autocorrelación también es polinomial. Esto se establece en el Teorema 3.1.1 y el objetivo principal de esta sección es detallar la demostración del Teorema presentada en [11]. Antes de enunciar el teorema presentamos dos definiciones más y algunas ideas relacionadas con ellas. Denotamos $\mathcal{M}(\Omega)$ al conjunto de medidas en Ω y $\mathcal{M}_T(\Omega)$ al conjunto de las medidas en Ω que son T -invariantes. También, denotamos $C(\Omega)$ al conjunto de funciones continuas $f : \Omega \rightarrow \mathbb{R}$. Finalmente, denotamos $L_1(\Omega, \lambda)$ al conjunto de funciones integrables con respecto a la medida λ y $\|f\|_1 = \int_{\Omega} |f| d\mu$.

Sea $\mathcal{C} = \{C_1, \dots, C_k\}$ una partición del conjunto Ω_A^+ . Si \mathcal{C} y \mathcal{D} son dos particiones de Ω_A^+ , podemos definir una nueva partición \mathcal{E} mediante:

$$\mathcal{E} = \mathcal{C} \vee \mathcal{D} = \{C_i \cap D_j : C_i \in \mathcal{C}, D_j \in \mathcal{D}\}.$$

A una partición definida de esta forma le llamamos refinación. Notamos que podemos definir una partición mediante los cilindros de longitud uno, es decir:

$$\mathcal{C} = \{[z_0] : z_0 \in S\} = \{\underline{x} \in \Omega_A : x_0 = z_0\} : z_0 \in S\}.$$

También notamos que $\sigma^{-1}\mathcal{C} = \{\sigma^{-1}C_i : C_i \in \mathcal{C}\}$ forma una nueva partición del conjunto Ω_A^+ . Aunque σ no es un mapeo invertible, podemos definir $\sigma^{-1}C_i$ como $\sigma^{-1}C_i = \{\underline{x} \in \Omega_A^+ : \sigma \underline{x} \in C_i\}$.

Definición 3.1.2. *Considere el espacio de probabilidad $(\Omega_A^+, \mathcal{B}, \lambda)$ donde \mathcal{B} es la σ -álgebra generada por los cilindros y λ es una medida de probabilidad en Ω_A^+ . Sea $\mathcal{C} = \{C_1, \dots, C_k\}$ una partición del conjunto Ω_A^+ . Entonces:*

1. Definimos la entropía de la partición \mathcal{C} mediante:

$$H_\lambda(\mathcal{C}) = - \sum_{i=1}^k \lambda(C_i) \log(\lambda(C_i)).$$

2. Definimos la entropía de la medida λ asociada al mapeo σ y a la partición \mathcal{C} como:

$$h_\lambda(\sigma, \mathcal{C}) = \lim_{m \rightarrow \infty} \frac{1}{m} H_\lambda(\mathcal{C} \vee \sigma^{-1}\mathcal{C} \vee \dots \vee \sigma^{-(m-1)}\mathcal{C}).$$

3. Definimos la entropía de la medida λ como:

$$h_\lambda(\sigma) = \sup_{\mathcal{C}} \{h_\lambda(\sigma, \mathcal{C})\}.$$

Definición 3.1.3. Definimos el estado de equilibrio asociado a un potencial ϕ como la medida η tal que:

$$P(\phi) = \sup \left\{ h_\lambda(\sigma) + \int_{\Omega_A^+} \phi d\lambda : \lambda \in \mathcal{M}_\sigma(\Omega_A^+) \right\} = h_\eta(\sigma) + \int_{\Omega_A^+} \phi d\eta.$$

Proposición 3.1.1 ([15] Lema 1.5). Sean $\phi, \psi \in C(\Omega)$. Si $\phi = \psi + f \circ \sigma - f + c$, para alguna $f \in C(\Omega)$ y $c \in \mathbb{R}$, entonces ϕ y ψ tienen los mismos estados de equilibrio.

Demostración. Para facilitar la notación, solo en esta demostración, la integral sobre el conjunto Ω_A^+ la denotamos simplemente como \int_Ω . Suponga que μ es un estado de equilibrio de ϕ y $\bar{\mu}$ es un estado de equilibrio de ψ donde $\phi = \psi + f \circ \sigma - f + c$. Entonces $P(\phi) = h_\mu(\sigma) + \int_\Omega \phi d\mu$ y $P(\psi) = h_{\bar{\mu}}(\sigma) + \int_\Omega \psi d\bar{\mu}$. Por propiedades del supremo, se tiene que:

$$h_{\bar{\mu}}(\sigma) + \int_\Omega \psi d\bar{\mu} \geq h_\mu(\sigma) + \int_\Omega \psi d\mu \quad (3.1)$$

$$h_\mu(\sigma) + \int_\Omega \phi d\mu \geq h_{\bar{\mu}}(\sigma) + \int_\Omega \phi d\bar{\mu}. \quad (3.2)$$

Dado que μ y $\bar{\mu}$ son medidas σ -invariantes, por la proposición 1.2.2 se tiene que:

$$\begin{aligned} \int_\Omega f \circ \sigma d\mu &= \int_\Omega f d\mu \\ \int_\Omega f \circ \sigma d\bar{\mu} &= \int_\Omega f d\bar{\mu}. \end{aligned}$$

Usando estas ecuaciones y la ecuación (3.2), se tiene:

$$h_\mu(\sigma) + \int_\Omega \psi d\mu + \int_\Omega c d\mu \geq h_{\bar{\mu}}(\sigma) + \int_\Omega \psi d\bar{\mu} + \int_\Omega c d\bar{\mu}.$$

Dado que $\int_\Omega c d\mu = \int_\Omega c d\bar{\mu} = c$, se tiene que:

$$h_\mu(\sigma) + \int_\Omega \psi d\mu \geq h_{\bar{\mu}}(\sigma) + \int_\Omega \psi d\bar{\mu}. \quad (3.3)$$

Uniendo (3.1) y (3.3) se llega a:

$$P(\psi) = h_{\bar{\mu}}(\sigma) + \int_{\Omega} \psi d\bar{\mu} = h_{\mu}(\sigma) + \int_{\Omega} \psi d\mu,$$

con lo que se concluye que μ también es un estado de equilibrio de ψ . Los mismos cálculos permiten concluir que un estado de equilibrio de ψ también es un estado de equilibrio de ϕ . Esto completa la demostración de la proposición 3.1.1. \square

A continuación presentamos el teorema principal de esta sección. Este teorema proporciona una estimación de la función de autocorrelación para funciones f que son Hölder continuas y para potenciales ϕ cuya variación decae de forma polinomial. Aunque en [11] el teorema se presenta para el conjunto Ω_A , en esta tesis nos restringimos al caso menos general Ω_A^+ .

Teorema 3.1.1 ([11]). *Sea Ω_A^+ un subshift de tipo finito, $\sigma : \Omega_A^+ \rightarrow \Omega_A^+$ el mapeo shift transitivo y sea $\phi : \Omega_A^+ \rightarrow \mathbb{R}$ el potencial asociado al estado de equilibrio μ_{ϕ} . Suponga además que $f : \Omega_A^+ \rightarrow \mathbb{R}$ es una función Hölder continua. Si existen constantes $r > 2$ y $D > 0$ tales que $\text{var}_n(\phi) \leq \frac{D}{n^r}$ para todo $n \geq 1$, entonces $\rho_f(N) = \mathcal{O}\left(\frac{1}{N^{r-2-\epsilon}}\right)$ para todo $\epsilon > 0$.*

Es importante notar que, por la definición del orden de magnitud $\mathcal{O}(\cdot)$ presentada en el apéndice, el Teorema equivale a decir que existen constantes C y N_0 tales que $|\rho_f(N)| \leq \frac{C}{N^{r-2-\epsilon}}$ para todo $N \geq N_0$. Al final de este capítulo presentamos algunas estimaciones de N_0 obtenidas con un método numérico. Por facilidad en la exposición, pospondremos la demostración del Teorema hasta el final del capítulo, una vez que dispongamos de los lemas y definiciones necesarias. Comenzamos con dos definiciones centrales para la demostración del Teorema: el operador de Ruelle y la esperanza condicional.

Definición 3.1.4. *Para $\phi \in C(\Omega_A^+)$ y $\underline{x} \in \Omega_A^+$ definimos el operador de Ruelle $\mathcal{L}_{\phi} : C(\Omega_A^+) \rightarrow C(\Omega_A^+)$ mediante:*

$$\mathcal{L}_{\phi}(f(\underline{x})) = \sum_{\underline{y} \in \sigma^{-1}\underline{x}} e^{\phi(\underline{y})} f(\underline{y})$$

Observación 3.1.1. *Es importante notar que las iteraciones del operador de Ruelle están dadas por:*

$$\mathcal{L}_{\phi}^n(f(\underline{x})) = \sum_{\underline{y} \in \sigma^{-n}\underline{x}} \exp\left(\sum_{k=0}^{n-1} \phi(\sigma^k \underline{y})\right) f(\underline{y}) \quad (3.4)$$

La siguiente proposición forma parte de la demostración del Teorema 3.3 y el Corolario 3.3 (i) de [14].

Proposición 3.1.2 (Teorema 3.3 y Corolario 3.3(i) [14]). *Sea Ω_A^+ un subshift de tipo finito de un solo lado y $\sigma : \Omega_A^+ \rightarrow \Omega_A^+$ un mapeo topológicamente mezclante. Entonces se cumple lo siguiente:*

- a) Si $\sum_{n=0}^{\infty} \text{var}_n(\phi) < \infty$, entonces existe una función $r \in C(\Omega_A^+)$ y una constante $\lambda > 0$ tales que $\mathcal{L}_\phi r(\underline{x}) = \lambda r(\underline{x})$.
- b) Defina $\phi_0(\underline{x}) := \phi(\underline{x}) + \log r(\underline{x}) - \log r(\sigma \underline{x})$. Entonces ϕ y ϕ_0 tienen un único estado de equilibrio denotado μ_ϕ .
- c) μ_ϕ es la única medida en Ω_A^+ tal que $\int \mathcal{L}_{\phi_0}(f) d\mu_\phi = \int f d\mu_\phi$ para toda $f \in C(\Omega_A^+)$.

Demostración. La demostración detallada se presenta en [14, Teorema 3.3 y Corolario 3.3(i)]. En esta tesis solo esbozamos la demostración presentada en el artículo. Para el inciso a), primero se demuestra que el conjunto $\mathcal{M}(\Omega_A^+)$ es compacto, acotado y no vacío. Después se aplica el Teorema de Schauder-Tychonoff (Ver [26], pág. 456), para concluir que el mapeo $G : \mathcal{M}(\Omega_A^+) \rightarrow \mathcal{M}(\Omega_A^+)$ dado por $G(\nu(f)) = \frac{\nu(\mathcal{L}_\phi f)}{\nu(\mathcal{L}_\phi 1)}$ tiene un punto fijo. Es decir existe una medida de probabilidad $\mu_\phi \in \mathcal{M}(\Omega_A^+)$ tal que $\mu_\phi(\mathcal{L}_\phi f) = \mu_\phi(\mathcal{L}_\phi 1)\mu_\phi(f)$. Posteriormente se muestra que el conjunto:

$$\Lambda = \left\{ f \in C(\Omega_A^+) : f \geq 0, \mu_\phi(f) = 1, f(\underline{x}) \leq f(\underline{y}) \exp \left(\sum_{k=m+1}^{\infty} \text{var}_m(\phi) \right) \right. \\ \left. \forall k \geq 1, \text{ siempre que } x_i = y_i, i = 0, \dots, m \right\},$$

es un conjunto compacto, acotado y no vacío y se concluye, utilizando nuevamente el Teorema de Schauder-Tychonoff, que el mapeo $E(f) : \Lambda \rightarrow \Lambda$ dado por $E(f) = \frac{\mathcal{L}_\phi f}{\mu_\phi(\mathcal{L}_\phi 1)}$ tiene un punto fijo. Es decir, existe $r \in \Lambda$ tal que $\frac{\mathcal{L}_\phi r}{\mu_\phi(\mathcal{L}_\phi 1)} = r$, o bien $\mathcal{L}_\phi r = \lambda r$ con $\lambda = \mu_\phi(\mathcal{L}_\phi 1)$.

Para el inciso b), la proposición 3.1.1 permite concluir directamente que ϕ y ϕ_0 tienen los mismos estados de equilibrio. Para la unicidad, se observa que ϕ_0 se puede escribir como $\phi_0 = \log \left(\frac{e^{\phi r}}{r \circ \sigma} \right)$. Después se utiliza el inciso a) de la proposición 3.1.2 y el hecho de que $\mathcal{L}_{\phi_0} 1 = 1$ para concluir que $\frac{e^{\phi r}}{r \circ \sigma}$ pertenece al conjunto:

$$G = \left\{ g \in C(\Omega_A^+) : g > 0, \sum_{y \in \sigma^{-1}x} g(y) = 1, \text{ para todo } x \in \Omega_A^+ \right\}.$$

Con este resultado, el autor utiliza el Teorema 3.1 del mismo artículo para concluir que el estado de equilibrio de ϕ_0 es único.

Para el inciso c), se utiliza el hecho de que $\mathcal{L}_{\phi_0} 1 = 1$ con lo que el resultado del inciso a) es $\mu_\phi(\mathcal{L}_{\phi_0} f) = \mu_\phi(f)$ o bien $\int \mathcal{L}_{\phi_0}(f) d\mu_\phi = \int f d\mu_\phi$. \square

El siguiente lema permite concluir que, además de ser un estado de equilibrio, μ_ϕ es de hecho una medida de Gibbs (definición 2.4.1).

Lema 3.1.2. Si μ es un estado de equilibrio de ϕ_0 , entonces existe una constante $0 < \gamma < 1$ tal que:

$$\gamma \leq \frac{\mu[x_0, \dots, x_{n-1}]}{\exp \left(\sum_{i=0}^{n-1} \phi_0(\sigma^i(\underline{x})) \right)} \leq \gamma^{-1} \quad (3.5)$$

Demostración. Ver [15, Teorema 1.16]. La demostración del lema se basa en la convergencia de las iteraciones del operador de Ruelle a la medida μ_ϕ y en las propiedades de la función r y la constante λ de la proposición 3.1.2. \square

El siguiente Lema nos permite relacionar la función de autocorrelación $\rho_f(N)$ con las iteraciones del operador de Ruelle $\mathcal{L}_{\phi_0}^N(f)$.

Lema 3.1.3. *Sea $\rho_f(N)$ la función de autocorrelación de orden N de una función $f : \Omega_A^+ \rightarrow \mathbb{R}$ y sea $\mathcal{L}_{\phi_0}^N f$ como en la ecuación (3.4). Entonces, para todo $N \geq 1$:*

$$\rho_f(N) \leq \|\mathcal{L}_{\phi_0}^N(f)\|_1.$$

Demostración. A partir de la definición se tiene:

$$\begin{aligned} \rho_f(N) &= \int_{\Omega} (f \circ \sigma^N) f d\mu_\phi - \left(\int_{\Omega} f d\mu_\phi \right)^2 \leq \int_{\Omega} (f \circ \sigma^N) f d\mu_\phi \\ &= \int_{\Omega} \mathcal{L}_{\phi_0}^N((f \circ \sigma^N) f) d\mu_\phi. \end{aligned}$$

Utilizando la ecuación (3.4), se tiene que:

$$\mathcal{L}_{\phi_0}^N((f \circ \sigma^N) f) = \sum_{\underline{y} \in \sigma^{-N} \underline{x}} \exp\left(\sum_{k=0}^{N-1} \phi_0(\sigma^k \underline{y})\right) (f(\sigma^N(\underline{y}))) f(\underline{y}).$$

Dado que $\sigma^N(\underline{y}) = \underline{x}$, se tiene:

$$\begin{aligned} \mathcal{L}_{\phi_0}^N((f \circ \sigma^N) f) &= \sum_{\underline{y} \in \sigma^{-N} \underline{x}} \exp\left(\sum_{k=0}^{N-1} \phi_0(\sigma^k \underline{y})\right) f(\underline{x}) f(\underline{y}) \\ &= f(\underline{x}) \sum_{\underline{y} \in \sigma^{-N} \underline{x}} \exp\left(\sum_{k=0}^{N-1} \phi_0(\sigma^k \underline{y})\right) f(\underline{y}) = f(\underline{x}) \mathcal{L}_{\phi_0}^N f(\underline{x}), \end{aligned}$$

y por lo tanto:

$$\rho_f(N) \leq \int_{\Omega} \mathcal{L}_{\phi_0}^N((f \circ \sigma^N) f) d\mu_\phi = \int_{\Omega} f \mathcal{L}_{\phi_0}^N f d\mu_\phi \leq \|\mathcal{L}_{\phi_0}^N(f)\|_1.$$

\square

Definición 3.1.5. *Sea $(\Omega, \mathcal{B}, \lambda)$ un espacio de medida, $f : \Omega \rightarrow \mathbb{R}$ una función λ -integrable y $\mathcal{C} \subset \mathcal{B}$ una sub sigma álgebra de \mathcal{B} . Definimos la esperanza condicional de f con respecto de \mathcal{C} , como el operador $\mathbb{E}(f|\mathcal{C}) : L_1(\Omega, \lambda) \rightarrow L_1(\Omega, \lambda)$ tal que:*

$$\int_C f d\lambda = \int_C \mathbb{E}(f|\mathcal{C}) d\lambda,$$

para todo $C \in \mathcal{C}$.

Observación 3.1.2. *Cuando \mathcal{C} es una partición del conjunto Ω , la siguiente desigualdad también es válida:*

$$\int_{\Omega} f d\lambda = \int_{\Omega} \mathbb{E}(f|\mathcal{C}) d\lambda. \quad (3.6)$$

Una partición finita de Ω forma una sub sigma álgebra sobre Ω . Dada la definición de esperanza condicional, podemos preguntarnos cuál es la forma que toma la esperanza condicional cuando la sub sigma álgebra \mathcal{C} es la partición del conjunto Ω dada por:

$$\mathcal{C}_n = \bigvee_{i=0}^{n-1} \sigma^{-i}(\mathcal{C}).$$

A continuación mostraremos que la esperanza condicional con respecto a esta partición está dada por:

$$\mathbb{E}(f|\mathcal{C}_n)(\underline{x}) = \sum_{C \in \mathcal{C}_n} \left(\frac{\int_C f d\lambda}{\lambda(C)} \right) \chi_C(\underline{x}). \quad (3.7)$$

Para mostrar esto, es suficiente mostrar que (3.7) cumple con la definición de esperanza condicional dada por (3.6). Los siguientes cálculos muestran precisamente esta afirmación:

$$\begin{aligned} \int_C \mathbb{E}(f|\mathcal{C}_n)(\underline{x}) d\lambda(\underline{x}) &= \int_C \sum_{C \in \mathcal{C}_n} \left(\frac{\int_C f d\lambda}{\lambda(C)} \right) \chi_C(\underline{x}) d\lambda(\underline{x}) = \int_C \left(\frac{\int_C f d\lambda}{\lambda(C)} \right) d\lambda(\underline{x}) \\ &= \frac{\int_C f d\lambda}{\lambda(C)} \int_C d\lambda(\underline{x}) = \frac{\int_C f d\lambda}{\lambda(C)} \lambda(C) = \int_C f d\lambda \end{aligned} \quad (3.8)$$

Lema 3.1.4. *Defina:*

$$K(\underline{x}, \underline{z}) = \sum_{C \in \mathcal{C}_n} \left(\frac{e^{S_n(\phi_0(x_C))}}{\mu(C)} \right) \chi_C(\underline{z}),$$

donde $S_n(\phi_0(x_C)) = \exp\left(\sum_{i=0}^{n-1} \phi_0(\sigma^i(x_C))\right)$ y $x_C = i_0, \dots, i_{n-1}, x_0, x_1, \dots$. Entonces $\mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\mathcal{C}_n)(\underline{x})) = \int_{\Omega} K(\underline{x}, \underline{z}) f(\underline{z}) d\mu(\underline{z})$ para toda $f \in L_1(\Omega, \mu)$.

Demostración. Primero observamos que:

$$\begin{aligned} \int_{\Omega} K(\underline{x}, \underline{z}) f(\underline{z}) d\mu(\underline{z}) &= \int_{\Omega} \sum_{C \in \mathcal{C}_n} \left(\frac{e^{S_n(\phi_0(x_C))}}{\mu(C)} \right) \chi_C(\underline{z}) f(\underline{z}) d\mu(\underline{z}) \\ &= \sum_{C \in \mathcal{C}_n} \int_{\Omega} \left(\frac{e^{S_n(\phi_0(x_C))}}{\mu(C)} \right) \chi_C(\underline{z}) f(\underline{z}) d\mu(\underline{z}) \\ &= \sum_{C \in \mathcal{C}_n} \frac{e^{S_n(\phi_0(x_C))}}{\mu(C)} \int_{\Omega} \chi_C(\underline{z}) f(\underline{z}) d\mu(\underline{z}) \\ &= \sum_{C \in \mathcal{C}_n} \frac{e^{S_n(\phi_0(x_C))}}{\mu(C)} \int_C f(\underline{z}) d\mu(\underline{z}) \\ &= \sum_{\underline{y} \in \sigma^{-n}\underline{x}} \frac{e^{S_n(\phi_0(\underline{y}))}}{\mu(C_{\underline{y}})} \int_{C_{\underline{y}}} f(\underline{z}) d\mu(\underline{z}). \end{aligned} \quad (3.9)$$

Donde $C_y := [y_0, \dots, y_{n-1}]$. Por otro lado:

$$\begin{aligned}
\mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\mathcal{C}_n)(\underline{x})) &= \sum_{y \in \sigma^{-n}\underline{x}} \left(\exp \left(\sum_{k=0}^{n-1} \phi_0(\sigma^k y) \right) \sum_{C \in \mathcal{C}_n} \left(\frac{\int_C f d\mu}{\mu(C)} \right) \chi_C(y) \right) \\
&= \sum_{y \in \sigma^{-n}\underline{x}} \left(e^{S_n(\phi_0(y))} \sum_{C \in \mathcal{C}_n} \left(\frac{\int_C f d\mu}{\mu(C)} \right) \chi_C(y) \right) \\
&= \sum_{y \in \sigma^{-n}\underline{x}} \frac{e^{S_n(\phi_0(y))}}{\mu(C_y)} \int_{C_y} f(z) d\mu(z). \tag{3.10}
\end{aligned}$$

Donde hemos utilizado el hecho de que cada $y \in \sigma^{-n}\underline{x}$ pertenece a un solo cilindro denotado C_y . Comparando (3.9) y (3.10) vemos que:

$$\mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\mathcal{C}_n)(\underline{x})) = \int_{\Omega} K(\underline{x}, z) f(z) d\mu(z).$$

Esto completa la demostración del Lema 3.1.4. \square

El siguiente lema nos permite obtener una cota para $\left\| \mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\mathcal{C}_n)) \right\|_1$.

Lema 3.1.5. *Sea γ la constante de la ecuación (3.5). Para todo $n \geq 1$ se cumple:*

$$\left\| \mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\mathcal{C}_n)) \right\|_1 \leq (1 - \gamma) \|f\|_1.$$

Demostración. Por el inciso c) de la Proposición 3.1.2, se tiene que:

$$\begin{aligned}
\left\| \mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\mathcal{C}_n)) \right\|_1 &= \int_{\Omega} |\mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\mathcal{C}_n))(\underline{x})| d\mu(\underline{x}) \\
&= \int_{\Omega} |\mathbb{E}(f|\mathcal{C}_n)(\underline{x})| d\mu(\underline{x}).
\end{aligned}$$

Sin pérdida de generalidad, consideramos $\int_{\Omega} f(z) d\mu(z) = 0$. Utilizando (3.8) notamos que esto implica que $\int_{\Omega} \mathbb{E}(f|\mathcal{C}_n)(\underline{x}) d\mu(\underline{x}) = 0$. Defina:

$$\Omega_1 := \{x \in \Omega : \mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\mathcal{C}_n))(\underline{x}) \geq 0\}.$$

Por el lema A.4.3 presentado en el apéndice se tiene que:

$$\int_{\Omega} |\mathbb{E}(f|\mathcal{C}_n)(\underline{x})| d\mu(\underline{x}) = 2 \int_{\Omega_1} \mathbb{E}(f|\mathcal{C}_n)(\underline{x}) d\mu(\underline{x}).$$

También, por el Lema 3.1.4 se tiene:

$$\begin{aligned}
2 \int_{\Omega_1} \mathbb{E}(f|\mathcal{C}_n)(\underline{x}) d\mu(\underline{x}) &= 2 \int_{\Omega_1} \int_{\Omega} K(\underline{x}, z) f(z) d\mu(z) d\mu(\underline{x}) \\
&= 2 \int_{\Omega_1} \int_{\Omega} K(\underline{x}, z) f(z) d\mu(z) d\mu(\underline{x}) - 2\mu(\Omega_1\gamma) \int_{\Omega} f(z) d\mu(z).
\end{aligned}$$

Donde el segundo término del lado derecho no afecta la igualdad ya que hemos asumido que $\int_{\Omega} f(z) d\mu(z) = 0$. La expresión anterior se puede reescribir como:

$$\begin{aligned}
& 2 \int_{\Omega_1} \int_{\Omega} K(x, z) f(z) d\mu(z) d\mu(x) - 2\mu(\Omega_1) \gamma \int_{\Omega} f(z) d\mu(z) \\
&= 2 \int_{\Omega} \int_{\Omega_1} K(x, z) f(z) d\mu(x) d\mu(z) - 2\gamma \int_{\Omega} \mu(\Omega_1) f(z) d\mu(z) \\
&= 2 \int_{\Omega} \int_{\Omega_1} K(x, z) f(z) d\mu(x) d\mu(z) - 2\gamma \int_{\Omega} \int_{\Omega_1} f(z) d\mu(x) d\mu(z) \\
&= 2 \int_{\Omega} \left(\int_{\Omega_1} f(z) (K(x, z) - \gamma) d\mu(x) \right) d\mu(z) \\
&= 2 \int_{\Omega} f(z) \left(\int_{\Omega_1} (K(x, z) - \gamma) d\mu(x) \right) d\mu(z). \tag{3.11}
\end{aligned}$$

Ahora, dado que z pertenece al cilindro $[z_0, \dots, z_{n-1}]$ y utilizando el Lema 3.1.2 se tiene que:

$$K(x, z) - \gamma = \sum_{C \in \mathcal{C}_n} \left(\frac{e^{S_n(\phi_0(x_C))}}{\mu(C)} \right) \chi_C(z) - \gamma = \frac{e^{S_n(\phi_0(z_0, \dots, z_{n-1}, x_0, \dots))}}{\mu[z_0, \dots, z_{n-1}]} - \gamma \geq 0.$$

Esto quiere decir que $K(x, z) - \gamma \geq 0$ y por lo tanto:

$$\begin{aligned}
2 \int_{\Omega} f(z) \left(\int_{\Omega_1} (K(x, z) - \gamma) d\mu(x) \right) d\mu(z) &\leq 2 \int_{\Omega} f(z) \left(\int_{\Omega} (K(x, z) - \gamma) d\mu(x) \right) d\mu(z) \\
&\leq 2 \int_{\Omega_0} f(z) \left(\int_{\Omega} (K(x, z) - \gamma) d\mu(x) \right) d\mu(z).
\end{aligned}$$

Donde $\Omega_0 = \{x \in \Omega : f(x) \geq 0\}$. También, tomando en cuenta que:

$$\begin{aligned}
\int_{\Omega} K(x, z) d\mu(x) &= \int_{\Omega} \sum_{C \in \mathcal{C}_n} \left(\frac{e^{S_n(\phi_0(x_C))}}{\mu(C)} \right) \chi_C(z) d\mu(x) \\
&= \sum_{C \in \mathcal{C}_n} \left(\int_{\Omega} e^{S_n(\phi_0(x_C))} d\mu(x) \right) \frac{\chi_C(z)}{\mu(C)} \\
&= \sum_{C \in \mathcal{C}_n} \left(\int_{\Omega} \mathcal{L}_{\phi_0}^n \chi_C(x) d\mu(x) \right) \frac{\chi_C(z)}{\mu(C)} \\
&= \sum_{C \in \mathcal{C}_n} \left(\int_{\Omega} \chi_C(x) d\mu(x) \right) \frac{\chi_C(z)}{\mu(C)} \\
&= \sum_{C \in \mathcal{C}_n} \chi_C(z) = 1.
\end{aligned}$$

Llegamos finalmente a:

$$\begin{aligned}
\|\mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\mathcal{C}_n))\|_1 &\leq 2 \int_{\Omega_0} f(z) (1 - \gamma) d\mu(z) \\
&= (1 - \gamma) \int_{\Omega} |f(z)| d\mu(z) = (1 - \gamma) \|f\|_1.
\end{aligned}$$

Esto completa la demostración del Lema 3.1.5. \square

Observación 3.1.3. Podemos aplicar el lema 3.1.5 a la función $\mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\mathcal{C}_n)(\underline{x}))$ con lo que obtenemos:

$$\left\| (\mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\mathcal{C}_n)))^2 \right\|_1 \leq (1-\gamma) \left\| \mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\mathcal{C}_n)) \right\|_1 \leq (1-\gamma)^2 \|f\|_1.$$

Inductivamente podemos ver que:

$$\left\| (\mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\mathcal{C}_n)))^m \right\|_1 \leq (1-\gamma)^m \|f\|_1. \quad (3.12)$$

Lema 3.1.6. Para todo $n \geq 0$ se tiene que $\|f - \mathbb{E}(f|\mathcal{C}_n)\|_1 \leq \text{var}_n(f)$.

Demostración.

$$\|f - \mathbb{E}(f|\mathcal{C}_n)\|_1 \leq \|f - \mathbb{E}(f|\mathcal{C}_n)\|_\infty \leq \text{var}_n(f).$$

□

Lema 3.1.7. Si $\text{var}_n(f) < \infty$ para todo $l, n \geq 0$, entonces:

$$\text{var}_n(\mathcal{L}_{\phi_0}^l f) \leq C_n(\phi) \|f\|_\infty + \text{var}_{n+l}(f), \quad (3.13)$$

donde $C_n(\phi) = \sum_{k=n}^{\infty} \text{var}_k(\phi)$.

Demostración. Ver [14, Teorema 3.1]. La demostración se basa en utilizar la ecuación (3.4) en $\left| \mathcal{L}_{\phi_0}^l f(x) - \mathcal{L}_{\phi_0}^l f(y) \right|$, acotar mediante la desigualdad del triángulo y usar el hecho de que $\sum_{y \in \sigma^{-1}x} \frac{e^{\phi(y)h(y)}}{h(\sigma y)} = 1$. □

Lema 3.1.8. Con $n, m \geq 0$ y $N = nm$ se cumple:

$$\left\| \mathcal{L}_{\phi_0}^N f - (\mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\mathcal{C}_n)))^m \right\|_1 = \mathcal{O}(C_n(\phi)).$$

Demostración. A fin de aligerar la notación, definimos:

$$T(\cdot) = \mathcal{L}_{\phi_0}^n(\mathbb{E}(\cdot|\mathcal{C}_n)) \quad (3.14)$$

$$U(\cdot) = \mathcal{L}_{\phi_0}^n(\cdot). \quad (3.15)$$

Hacemos notar que $T^{k-1}(T(U^{m-k}f)) = T^k(U(U^{m-k-1}f))$ para todo $k \geq 0$, con lo que se tiene que:

$$\begin{aligned} |U^m f(\underline{x}) - T^m f(\underline{x})| &= |U(U^{m-1}f(\underline{x})) - T(U^{m-1}f(\underline{x}))| \\ &\quad + T[U(U^{m-2}f(\underline{x})) - T(U^{m-2}f(\underline{x}))] + \dots + T^{m-1}[Uf(\underline{x}) - Tf(\underline{x})] \\ &\leq \sum_{k=0}^{m-1} \left| T^k[U(U^{m-k-1}f(\underline{x})) - T(U^{m-k-1}f(\underline{x}))] \right|. \end{aligned}$$

Por lo tanto:

$$\begin{aligned}
\left\| \mathcal{L}_{\phi_0}^N f - \left(\mathcal{L}_{\phi_0}^n \left(\mathbb{E} \left(f | \alpha^{(n)} \right) \right) \right)^m \right\|_1 &= \left\| (\mathcal{L}_{\phi_0}^n)^m f - \left(\mathcal{L}_{\phi_0}^n \left(\mathbb{E} \left(f | \alpha^{(n)} \right) \right) \right)^m \right\|_1 \\
&\leq \sum_{k=0}^{m-1} \left\| T^k [U(U^{m-k-1} f) - T(U^{m-k-1} f)] \right\|_1 \\
&\leq \sum_{k=0}^{m-1} \left\| T^k [U(U^{m-k-1} f) - U(\mathbb{E}(U^{m-k-1} f | \mathcal{C}_n))] \right\|_1 \\
&\leq \sum_{k=0}^{m-1} \left\| T^k [U[U^{m-k-1} f - \mathbb{E}(U^{m-k-1} f | \mathcal{C}_n)]] \right\|_1 \\
&\leq \sum_{k=0}^{m-1} (1 - \gamma)^k \left\| U[U^{m-k-1} f - \mathbb{E}(U^{m-k-1} f | \mathcal{C}_n)] \right\|_1 \\
&\leq \sum_{k=0}^{m-1} (1 - \gamma)^k \left\| U^{m-k-1} f - \mathbb{E}(U^{m-k-1} f | \mathcal{C}_n) \right\|_1 \\
&\leq \sum_{k=0}^{m-1} (1 - \gamma)^k \text{var}_n(U^{m-k-1} f) \\
&\leq \sum_{k=0}^{m-1} (1 - \gamma)^k (C_n(\phi) \|f\|_\infty + \text{var}_{n(m-k)}(f)) \\
&\leq \sum_{k=0}^{m-1} (1 - \gamma)^k C_n(\phi) \|f\|_\infty + \sum_{k=0}^{m-1} (1 - \gamma)^k \text{var}_{n(m-k)}(f) \\
&\leq C_n(\phi) \sum_{k=0}^{m-1} \left((1 - \gamma)^k \|f\|_\infty + 1 \right) \\
&= \mathcal{O}(C_n(\phi)). \tag{3.16}
\end{aligned}$$

Donde hemos utilizado el inciso (c) de la proposición 3.1.2, la observación 3.1.3 y los lemas 3.1.6 y 3.1.7. Note que la última igualdad se cumple ya que, para todo $n, m \geq 1$ y $k \in [0, m - 1]$:

$$\text{var}_{n(m-k)}(f) \leq C_n(\phi)$$

Esto completa la demostración del lema 3.1.8. \square

Finalmente tenemos todos los resultados necesarios para demostrar el Teorema 3.1.1. A continuación se presenta la demostración.

Demostración del Teorema 3.1.1. Notamos que:

$$\begin{aligned}
\rho_f(N) &\leq \|\mathcal{L}_{\phi_0}^N f\|_1 \\
&= \int |\mathcal{L}_{\phi_0}^N f| d\mu - \int |(\mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\alpha^{(n)})))^m| d\mu + \int |(\mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\alpha^{(n)})))^m| d\mu \\
&= \int (|\mathcal{L}_{\phi_0}^N f| - |(\mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\alpha^{(n)})))^m|) d\mu + \int |(\mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\alpha^{(n)})))^m| d\mu \\
&\leq \int |\mathcal{L}_{\phi_0}^N f - (\mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\alpha^{(n)})))^m| d\mu + \int |(\mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\alpha^{(n)})))^m| d\mu \\
&= \|\mathcal{L}_{\phi_0}^N f - (\mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\alpha^{(n)})))^m\|_1 + \|(\mathcal{L}_{\phi_0}^n(\mathbb{E}(f|\alpha^{(n)})))^m\|_1. \quad (3.17)
\end{aligned}$$

Donde hemos utilizado el hecho de que $|a| - |b| \leq |a - b|$ junto con el Teorema C de [6, Sección 23]. Lo que sigue de la demostración se basa en utilizar las desigualdades (3.17), (3.12) y (3.16). A partir de estas desigualdades, con $nm = N$ obtenemos:

$$\rho_f(N) \leq \|\mathcal{L}_{\phi_0}^N f\|_1 = \mathcal{O}(\max(C_n(\phi), (1 - \gamma)^m)).$$

Dado que $\text{var}_n(\phi) \leq D \left(\frac{1}{n^r}\right)$ se tiene:

$$\begin{aligned}
C_n(\phi) &= \sum_{k=n}^{\infty} \text{var}_k(\phi) \leq D \sum_{k=n}^{\infty} \frac{1}{k^r} \leq D \left(\int_{n-1}^{\infty} x^{-r} dx \right) = D \left(\frac{(n-1)^{1-r}}{r-1} \right) \\
&= \frac{D}{r-1} (n-1)^{1-r} = \mathcal{O}\left(\frac{1}{n^{r-1}}\right).
\end{aligned}$$

Elegimos $n = \lfloor N^\beta \rfloor$ y $m = \lfloor N^{1-\beta} \rfloor$ con $0 < \beta < 1$. Entonces:

$$\rho_f(N) = \mathcal{O}\left(\max\left(\frac{1}{N^{\beta(r-1)}}, (1 - \gamma)^{N^{1-\beta}}\right)\right).$$

Podemos evaluar el máximo en el límite utilizando logaritmos:

$$\lim_{N \rightarrow \infty} \frac{\log\left((1 - \gamma)^{N^{1-\beta}}\right)}{\log N^{\beta(1-r)}} = \frac{\log(1 - \gamma)}{\beta(1-r)} \lim_{N \rightarrow \infty} \frac{N^{1-\beta}}{\log N}$$

Utilizando la regla de L'Hôpital para evaluar el límite, se tiene que:

$$\begin{aligned}
\lim_{N \rightarrow \infty} \frac{N^{1-\beta}}{\log N} &= \lim_{N \rightarrow \infty} \frac{\frac{d}{dN} N^{1-\beta}}{\frac{d}{dN} \log N} \\
&= \lim_{N \rightarrow \infty} \frac{(1 - \beta)N^{-\beta}}{\frac{1}{N}} = (1 - \beta) \lim_{N \rightarrow \infty} \frac{N}{N^\beta} = \infty,
\end{aligned}$$

Esto significa que existe N_0 tal que, para todo $N > N_0$:

$$\left| \log(1 - \gamma)N^{1-\beta} \right| > |\beta(1 - r) \log N|$$

Dado que $\log(1 - \gamma) < 0$ y $\beta(1 - r) < 0$, se tiene:

$$\begin{aligned}
\left| \log(1 - \gamma)N^{1-\beta} \right| > |\beta(1 - r) \log N| &\Rightarrow -\log(1 - \gamma)N^{1-\beta} > -\beta(1 - r) \log N \\
&\Rightarrow \log(1 - \gamma)N^{1-\beta} < \beta(1 - r) \log N
\end{aligned}$$

Con lo que $\rho_f(N) = \mathcal{O}\left(\frac{1}{N^{\beta(r-1)}}\right)$. Tomando β suficientemente cercano a 1 se tiene que, para todo $\epsilon > 0$:

$$\rho_f(N) = \mathcal{O}\left(\frac{1}{N^{r-1-\epsilon}}\right)$$

Esto completa la demostración del teorema 3.1.1. \square

Hacemos notar que el valor de N_0 puede aproximarse numéricamente encontrando las raíces de la ecuación:

$$\frac{N_0^{1-\beta}}{\log N_0} - \beta \frac{1-r}{\log(1-\gamma)} = 0$$

La figura 3.1 muestra algunos valores de N_0 obtenidos utilizando el método de Newton-Raphson. Como puede verse en las gráficas, el valor de N_0 puede ser extremadamente grande, especialmente para valores de β cercanos a 1.

3.2. Estimación del desvío casi cierto basada en la función de autocorrelación

El siguiente Teorema proporciona una relación entre el decaimiento de $\rho_f(n)$ y el desvío de $A_n f$ con respecto a su valor esperado. Por simplicidad, el Teorema se presenta solamente para dos casos particulares que retomaremos más adelante. Todos los casos en los que el Teorema es aplicable y su demostración pueden encontrarse en [12].

Teorema 3.2.1 ([12]). *Suponga que la función de autocorrelación $\rho_f(n)$ satisface la siguiente relación con $\tau > 0$:*

$$\rho_f(n) = \mathcal{O}(n^{-\tau}). \quad (3.18)$$

Entonces, para casi todo $\underline{x} \in \Omega$, también se cumple la siguiente relación:

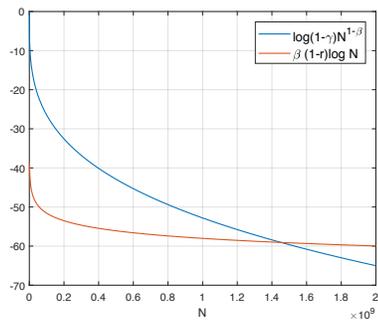
$$A_n f(\underline{x}) - f^* = \mathcal{O}(g_\gamma^\tau(n)). \quad (3.19)$$

Donde, con $\gamma > 0$, $g_\gamma^\tau(n)$ toma las siguientes formas dependiendo del rango en el que se encuentre τ :

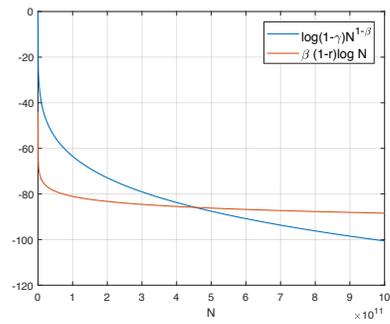
$$g_\gamma^\tau(n) = \begin{cases} \frac{(\log n)^{1/2} (\log \log n)^{(1+\gamma)/2}}{n^{\tau/2}} & 0 < \tau < 1 \\ \frac{(\log n)^{3/2} (\log \log n)^{1/2} (\log \log \log n)^{(1+\gamma)/2}}{n^{1/2}} & \tau > 1 \end{cases} \quad (3.20)$$

Demostración. Ver teoremas 3 y 4 de [12]. \square

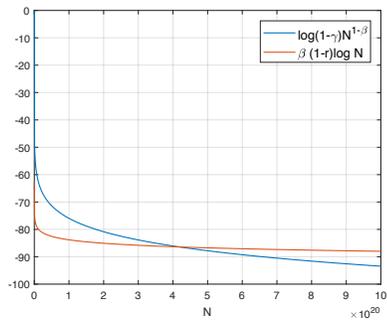
Utilizando los resultados del Teorema 3.1.1 en el Teorema 3.2.1, podemos obtener una estimación del desvío de $A_n f$ con respecto a su valor esperado, aplicable a funciones f que son Hölder continuas y a potenciales cuya variación decae de forma polinomial. El siguiente Corolario muestra cómo obtener esta estimación.



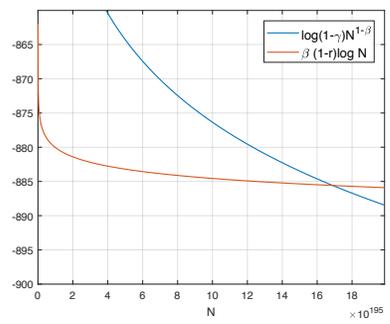
(a) $\beta = 0.7, \gamma = 0.1, r = 5, N_0 \approx 1.4533 \times 10^9$



(b) $\beta = 0.8, \gamma = 0.33, r = 5, N_0 \approx 4.5378 \times 10^{11}$



(c) $\beta = 0.91, \gamma = 0.7, r = 3, N_0 \approx 4.1909 \times 10^{20}$



(d) $\beta = 0.98, \gamma = 0.1, r = 3, N_0 \approx 1.6893 \times 10^{196}$

Fig. 3.1: Diferentes valores de N_0 obtenidos numéricamente.

Corolario 3.2.1 ([11]). *Suponga que se cumplen las condiciones del Teorema 3.1.1. Entonces, existen constantes $C, N \in \mathbb{R}$ tales que la siguiente relación se cumple para todo $n > N$, para todo $0 < \epsilon < r - 2$ y para casi todo $\underline{x} \in \Omega_A^+$:*

$$\left| \frac{1}{n} \sum_{i=0}^{n-1} f(\sigma^i \underline{x}) - \int_{\Omega_A^+} f d\mu_\phi(\underline{y}) \right| \leq C g_\epsilon^\tau(n). \quad (3.21)$$

Donde $\tau = r - 2 - \epsilon$.

Demostración. Dado que se cumplen las condiciones del Teorema 3.1.1, se tiene que $\rho_f(n) = \mathcal{O}\left(\frac{1}{n^{r-2-\epsilon}}\right)$. Esta expresión es equivalente a (3.18) con $\tau = r - 2 - \epsilon$. Cuando $r < 3$ se tiene que $\tau < 1$; cuando $r > 3 + \epsilon$ se tiene que $\tau > 1$.

Por el Teorema 3.2.1 se tiene que se cumple la desigualdad (3.19) la cual implica que $A_n f(\underline{x}) - f^* = \mathcal{O}(g_\epsilon^\tau(n))$. Por definición, de la notación \mathcal{O} , se tiene que existen constantes $C, N \in \mathbb{R}$ tales que la relación (3.21) se cumple para todo $n > N$ y para casi toda $\underline{x} \in \Omega$. Esto completa la demostración del Corolario 3.2.1. \square

4. Contribuciones

En este capítulo se presentan las contribuciones de este trabajo de tesis. Estas contribuciones son en su mayoría, corolarios basados en los teoremas 2.1.6, 2.2.7, 2.4.4 y 3.1.1. También en esta sección se presenta una breve discusión de los resultados obtenidos.

4.1. Estimación de probabilidad de desvío: p_n^ϵ

En esta sección se utilizan los los teoremas 2.2.7 y 2.4.4. El Teorema 2.2.7 permite obtener una estimación de p_n^ϵ para funciones f que son localmente Hölder continuas. Para aplicar este Teorema, se requieren estimaciones de p_n^ϵ para funciones Hölder continuas. En particular, la desigualdad (2.72) del Teorema 2.4.4, válida para sistemas dinámicos simbólicos con medida de Gibbs, puede utilizarse en el Teorema 2.2.7 para obtener una estimación de p_n^ϵ para funciones localmente Hölder continuas en esta clase de sistemas dinámicos. El resultado es el siguiente corolario.

Corolario 4.1.1. *Sea S un conjunto finito y $\Omega = S^{\mathbb{N}}$. Defina una distancia en Ω como en la sección 1.6. Sea $f : \Omega \rightarrow \mathbb{R}$ una función acotada y continua μ_ϕ casi en todas partes, $\delta > 0$ y $\alpha \in (0, 1]$. Sean g_1^δ y g_2^δ como en el Teorema 2.2.7 y ω_f como en el Lema 2.2.5. Sea μ_ϕ una medida de Gibbs con potencial ϕ . Entonces se cumple la siguiente desigualdad para todo $\epsilon > \mu_\phi(\omega_f(\delta, x))$ y $n \geq 1$:*

$$\mu_\phi \left\{ x \in \Omega : \left| \frac{1}{n} \sum_{i=0}^{n-1} f(\sigma^i x) - \int_{\Omega} f d\mu_\phi(y) \right| \geq \epsilon \right\} \leq 4 \exp \left(-\frac{n}{4D} \left(\frac{\epsilon - \mu_\phi(\omega_f(\delta, x))}{\max(|g_1^\delta|_{\theta^\alpha}, |g_2^\delta|_{\theta^\alpha})} \right)^2 \right). \quad (4.1)$$

Demostración. Las funciones g_1^δ y g_2^δ son Hölder continuas con respecto a d_θ y por el Lemma 1.6.1, pueden ser vistas como funciones Lipschitz continuas con respecto a la distancia d_{θ^α} . Esto significa que la desigualdad (2.72) es válida para g_1^δ y g_2^δ y

por lo tanto se puede utilizar en la desigualdad (2.42) del Teorema 2.2.7:

$$\begin{aligned}
& \mu_\phi \left\{ \underline{x} \in \Omega : \left| \frac{1}{n} \sum_{i=0}^{n-1} f(\sigma^i \underline{x}) - \int_{\Omega} f d\mu_\phi(\underline{y}) \right| \geq \epsilon \right\} \\
& \leq \mu_\phi \left\{ \underline{x} \in \Omega : \left| \frac{1}{n} \sum_{i=0}^{n-1} g_1^\delta(\sigma^i \underline{x}) - \int_{\Omega} g_1^\delta d\mu_\phi(\underline{y}) \right| \geq \epsilon - \mu_\phi(\omega_f) \right\} \\
& + \mu_\phi \left\{ \underline{x} \in \Omega : \left| \frac{1}{n} \sum_{i=0}^{n-1} g_2^\delta(\sigma^i \underline{x}) - \int_{\Omega} g_2^\delta d\mu_\phi(\underline{y}) \right| \geq \epsilon - \mu_\phi(\omega_f) \right\} \\
& \leq 2 \exp \left(-\frac{n(\epsilon - \mu_\phi(\omega_f))^2}{4D|g_1^\delta|_{\theta^\alpha}^2} \right) + 2 \exp \left(-\frac{n(\epsilon - \mu_\phi(\omega_f))^2}{4D|g_2^\delta|_{\theta^\alpha}^2} \right) \\
& \leq 4 \exp \left(-\frac{n(\epsilon - \mu_\phi(\omega_f))^2}{4D \max(|g_1^\delta|_{\theta^\alpha}, |g_2^\delta|_{\theta^\alpha})^2} \right) \\
& = 4 \exp \left(-\frac{n}{4D} \left(\frac{\epsilon - \mu_\phi(\omega_f)}{\max(|g_1^\delta|_{\theta^\alpha}, |g_2^\delta|_{\theta^\alpha})} \right)^2 \right).
\end{aligned}$$

Esto completa la demostración del Corolario 4.1.1. \square

4.2. Estimación de probabilidad de desvío: P_n^ϵ

Una primera estimación de P_n^ϵ para funciones Lipschitz continuas puede obtenerse directamente utilizando los Teoremas 2.4.4 y 2.1.6. El resultado puede enunciarse como sigue. Sea f una función Lipschitz continua cuya mínima constante de Lipschitz es $|f|_\theta$, $\sigma : \Omega \rightarrow \Omega$ el mapeo dado por $(\sigma \underline{x})_i = x_{i+1}$, μ_ϕ una medida de Gibbs con potencial ϕ . Por el Teorema 2.4.4 se cumple la desigualdad (2.72). A partir de esta desigualdad, notamos que se cumplen las condiciones del Teorema 2.1.6 con $C(\epsilon) = 2$, $\gamma(\epsilon) = -\frac{\epsilon^2}{4D|f|_\theta^2}$ y $\delta(\epsilon) = 1$. Con estas constantes, la desigualdad (2.24) puede escribirse como:

$$\begin{aligned}
P_n^{2\epsilon}(f) &= \mu_\phi \left(\left\{ \underline{x} \in \Omega : \sup_{k \geq n} \left| \frac{1}{k} S_k f(\underline{x}) - f^* \right| \geq \epsilon \right\} \right) \\
&< 2 \left(1 + \frac{\ln(1 + 4D|f|_\theta^2/\epsilon^2)}{\ln r(\epsilon)} \right) \exp \left(-\frac{n\epsilon^2}{4D|f|_\theta^2} \right). \tag{4.2}
\end{aligned}$$

Para todo $\epsilon > 0$ y $r(\epsilon) = 1 + \frac{\epsilon}{2\|f-f^*\|_\infty}$.

Es importante notar que este resultado es válido para funciones Lipschitz continuas. Alineados con los objetivos presentados en la sección 1.4, buscamos extender esta desigualdad para funciones acotadas y continuas μ_ϕ casi en todas partes. Esta generalización de la desigualdad 4.2 puede obtenerse también de forma directa aplicando el Corolario 4.1.1 y el Teorema 2.1.6. El Corolario 4.1.1 permite encontrar una estimación de p_n^ϵ para funciones acotadas y continuas μ_ϕ casi en todas partes. Esta estimación tiene un comportamiento asintótico exponencial para el cual, puede aplicarse el Teorema 2.1.6. El resultado es una estimación de P_n^ϵ , válida para

funciones acotadas y continuas μ_ϕ casi en todas partes. El corolario se encunacia a continuación.

Corolario 4.2.1. *Suponga que se cumplen las condiciones del Corolario 4.1.1. Entonces también se cumple la siguiente desigualdad para todo $\epsilon > \mu_\phi(\omega_f(\delta, \underline{x}))$ y $n \geq 1$:*

$$P_n^{2\epsilon}(f) < 4 \left(1 + \frac{\ln \left(1 + \frac{4D}{\psi^2(\epsilon)} \right)}{\ln r(\epsilon)} \right) \exp \left(-\frac{n\psi^2(\epsilon)}{4D} \right). \quad (4.3)$$

donde:

$$\psi(\epsilon) = \frac{\epsilon - \mu_\phi(\omega_f(\delta, \underline{x}))}{\max(|g_1^\delta|_{\theta^\alpha}, |g_2^\delta|_{\theta^\alpha})}$$

Demostración. El resultado sigue de utilizar la desigualdad (4.2.1) del Corolario 4.1.1 en la desigualdad (2.24) del Teorema 2.1.6 con:

$$\begin{aligned} C(\epsilon) &= 4 \\ \delta(\epsilon) &= 1 \\ \gamma(\epsilon) &= 4D \left(\frac{\max(|g_1^\delta|_{\theta^\alpha}, |g_2^\delta|_{\theta^\alpha})}{\epsilon - \mu_\phi(\omega_f(\delta, \underline{x}))} \right)^2. \end{aligned}$$

Esto completa la demostración del Corolario 4.2.1. \square

4.3. Estimación de desvío casi cierto

El Teorema 3.2.1 es aplicable a funciones f que son Hölder continuas. Sin embargo, utilizando algunos de los resultados de la sección 2.2, es posible extender este teorema a funciones que son acotadas y continuas μ_ϕ casi en todas partes. El siguiente corolario proporciona esta generalización.

Corolario 4.3.1. *Sea Ω_A^+ un subshift de tipo finito, $\sigma : \Omega_A^+ \rightarrow \Omega_A^+$ un mapeo shift transitivo y sea $\phi : \Omega_A^+ \rightarrow \mathbb{R}$ el potencial asociado a la medida de equilibrio μ_ϕ . Suponga además que $f : \Omega_A^+ \rightarrow \mathbb{R}$ es una función acotada y continua μ_ϕ casi en todas partes. Si existen constantes $r > 2$ y $D > 0$ tales que $\text{var}_n(\phi) \leq \frac{D}{n^r}$ para todo $n \geq 1$, entonces, para todo $0 < \epsilon < r - 2$ y para para casi toda $\underline{x} \in \Omega_A^+$ se cumple la siguiente desigualdad de desvío:*

$$\left| \frac{1}{n} \sum_{i=0}^{n-1} f(\sigma^i \underline{x}) - \int_{\Omega} f d\mu_\phi(\underline{y}) \right| \leq C g_\gamma^\tau(n) + \mu_\phi(\omega_f(\delta, \underline{x})), \quad (4.4)$$

donde $\tau = r - 2 - \epsilon$.

Antes de comenzar la demostración del corolario 4.3.1, reproducimos aquí las ecuaciones (2.35) y (2.43)-(2.46), obtenidas en la sección 2.2 y que se utilizarán en esta sección en el contexto simbólico:

$$f(\underline{x}) - \omega_f(\delta, \underline{x}) \leq g_1^\delta(\underline{x}) \leq f(\underline{x}) \leq g_2^\delta(\underline{x}) \leq f(\underline{x}) + \omega_f(\delta, \underline{x}). \quad (4.5)$$

$$A_n g_1^\delta(\underline{x}) \leq A_n f(\underline{x}) \leq A_n g_2^\delta(\underline{x}). \quad (4.6)$$

$$\mu_\phi(f) = \int f d\mu_\phi = \int f^* d\mu_\phi = f^*, \quad (4.7)$$

$$\mu_\phi(g_1^\delta) = \int g_1^\delta d\mu_\phi = \int g_1^{\delta*} d\mu_\phi = g_1^{\delta*}, \quad (4.8)$$

$$\mu_\phi(g_2^\delta) = \int g_2^\delta d\mu_\phi = \int g_2^{\delta*} d\mu_\phi = g_2^{\delta*}. \quad (4.9)$$

Demostración. Dado que f es acotada y continua μ_ϕ casi en todas partes, se tiene que las funciones g_1^δ y g_2^δ definidas en (2.30) y (2.31) están bien definidas y son Hölder continuas. Por lo tanto, podemos utilizar el Corolario 3.2.1 con estas funciones. Las desigualdades que se obtienen son, respectivamente:

$$\left| A_n g_1^\delta(\underline{x}) - g_1^{\delta*} \right| \leq C g_\gamma^\tau(n) \quad (4.10)$$

$$\left| A_n g_2^\delta(\underline{x}) - g_2^{\delta*} \right| \leq C g_\gamma^\tau(n), \quad (4.11)$$

A partir de (4.10) y (4.11) podemos obtener:

$$A_n g_1^\delta(\underline{x}) - g_1^{\delta*} \geq -C g_\gamma^\tau(n) \quad (4.12)$$

$$A_n g_2^\delta(\underline{x}) - g_2^{\delta*} \leq C g_\gamma^\tau(n) \quad (4.13)$$

Por otro lado, la desigualdad (4.5) implica:

$$f^* - \mu_\phi(\omega_f(\delta, \underline{x})) \leq g_1^{\delta*} \leq f^* \leq g_2^{\delta*} \leq f^* + \mu_\phi(\omega_f(\delta, \underline{x})), \quad (4.14)$$

Por lo tanto, utilizando (4.6) y el extremo derecho de (4.14), se tiene:

$$A_n f(\underline{x}) - f^* \leq A_n g_2^\delta(\underline{x}) - f^* \leq A_n g_2^\delta(\underline{x}) - g_2^{\delta*} + \mu_\phi(\omega_f(\delta, \underline{x})) \leq C g_\gamma^\tau(n) + \mu_\phi(\omega_f(\delta, \underline{x})). \quad (4.15)$$

Similarmente, utilizando (4.6) y el extremo izquierdo de (4.14), se tiene:

$$A_n f(\underline{x}) - f^* \geq A_n g_1^\delta(\underline{x}) - f^* \geq A_n g_1^\delta(\underline{x}) - g_1^{\delta*} - \mu_\phi(\omega_f(\delta, \underline{x})) \geq -C g_\gamma^\tau(n) - \mu_\phi(\omega_f(\delta, \underline{x})). \quad (4.16)$$

Uniendo (4.15) y (4.16) se obtiene:

$$-C g_\gamma^\tau(n) - \mu_\phi(\omega_f(\delta, \underline{x})) \leq A_n f(\underline{x}) - f^* \leq C g_\gamma^\tau(n) + \mu_\phi(\omega_f(\delta, \underline{x})). \quad (4.17)$$

Esta última desigualdad es equivalente a (4.4), por lo tanto la demostración del Corolario 4.3.1 está completa. \square

4.4. Discusión de los resultados y conclusiones

En esta sección se presentan algunos comentarios acerca de los resultados obtenidos en esta tesis así como otros potenciales temas de investigación a futuro relacionados con tales resultados.

Las desigualdades (4.2.1), (4.3) y (4.4) son los resultados más importantes de este trabajo de tesis. Todas ellas son válidas en el contexto de sistemas dinámicos simbólicos con medida de Gibbs, con funciones observables f que son acotadas y continuas μ_ϕ casi en todas partes. Estos resultados extienden las desigualdades (2.72) y (3.19), válidas en el mismo tipo de sistema dinámico, pero con funciones observables f que son Hölder continuas.

Enfatizamos en este punto que las desigualdades (4.2.1), (4.3) y (4.4) son válidas para funciones observables mucho más generales que las funciones Hölder continuas. Para ver esto, note que el espacio de funciones acotadas y continuas casi en todas partes contiene al espacio de funciones continuas, el cual contiene a su vez al espacio de funciones uniformemente continuas, el cual contiene a su vez al espacio de funciones Hölder continuas. Con esto se puede ver que las desigualdades obtenidas en esta tesis abarcan un espacio más amplio de funciones observables que no se tenía en la literatura.

También es importante notar que la obtención de estos resultados es posible gracias a la existencia de las funciones Hölder continuas g_1^δ y g_2^δ , para cada función f acotada y continua λ casi en todas partes (ecuaciones (2.30) y (2.31)). En particular, las desigualdades (2.28) y (2.29), permiten extender a la función f , las cotas aplicables a las funciones g_1^δ y g_2^δ . Con respecto a la desigualdad de P_n^ϵ , esta fue posible obtenerla gracias a la forma exponencial de p_n^ϵ en (4.2.1).

A continuación, hacemos algunos comentarios acerca del posible trabajo futuro. Las propiedades de las funciones g_1^δ y g_2^δ pueden utilizarse para extender otras desigualdades válidas para funciones Hölder continuas f . Una parte del posible trabajo futuro, consiste en aplicar la misma técnica mostrada en esta tesis, a estas desigualdades y obtener con ello nuevas desigualdades válidas para funciones acotadas y continuas casi en todas partes. Por ejemplo, en [20] y [19], se presentan estimaciones de la función de correlación en otro tipo de sistemas dinámicos, para las cuales podría intentar aplicarse la misma técnica presentada en esta tesis.

Otra posible dirección del trabajo futuro es determinar si las desigualdades aplicables a funciones Hölder continuas, pueden ser obtenidas como un caso particular de las desigualdades aplicables a funciones acotadas y continuas casi en todas partes. Un primer paso en esta dirección consiste en determinar la forma que adquieren las funciones g_1^δ y g_2^δ y el valor que toma la expresión $\mu_\phi(\omega_f)$ cuando la función f es Hölder continua. Si bajo esta condición sobre f , se demuestra que $g_1^\delta = g_2^\delta = f$ y además $\mu_\phi(\omega_f) = 0$, entonces la forma de las desigualdades (4.2.1), (4.3) y (4.4), se asemeja más a la forma de las desigualdades aplicables en el caso Hölder.

Otra extensión importante, aunque más a largo plazo, sería encontrar desigualdades similares para potenciales más generales que los presentados aquí. Recordemos que las desigualdades para p_n^ϵ y P_n^ϵ son válidas con potenciales Hölder continuos y la desigualdad para el desvío casi cierto es válida para potenciales cuya variación decae de forma polinomial. Sin embargo, la existencia de un único estado de equilibrio ha sido establecida formalmente para potenciales de variación sumable ([14]), por lo

que tiene sentido la pregunta acerca de la rapidez de convergencia del Teorema de Birkhoff en sistemas dinámicos simbólicos para este tipo de potenciales. Recientemente, en [21] los autores muestran que la desigualdad de momentos obtenida en [9], continúa siendo válida con potenciales más generales que las funciones Hölder continuas. Este importante resultado podría utilizarse como punto de partida para ampliar la aplicabilidad de las desigualdades presentadas en esta tesis a este tipo de potenciales.

Apéndice

A. Apéndice

A.1. Teorema de Perron-Frobenius

Teorema A.1.1. *Sea \mathbf{B} una matriz irreducible de $m \times m$. Entonces \mathbf{B} tiene un eigenvalor ρ tal que:*

1. $\rho > 0$ es real y es el eigenvalor dominante de \mathbf{B} .
2. Existen eigenvectores izquierdo y derecho (únicos salvo por multiplicación por una constante) correspondientes al eigenvalor ρ que tienen coordenadas estrictamente positivas.
3. Sea $\phi = (\phi_1, \dots, \phi_m)$ un vector con coordenadas estrictamente positivas. Entonces, para todo $i \in \Sigma$:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln \left(\sum_{j=1}^m (\mathbf{B}^n)_{i,j} \phi_j \right) = \ln \rho$$

Demostración. Ver [7]. □

A.2. Comentarios sobre la notación de orden de magnitud

A lo largo de esta tesis se utilizan las siguientes definiciones de los símbolos de Landau \mathfrak{o} y \mathcal{O} :

Definición A.2.1. *Sea Ω un espacio equipado con una métrica y sean $f : \Omega \rightarrow \mathbb{R}$ y $\varphi : \Omega \rightarrow \mathbb{R}$.*

1. *Escribimos $f(n) = \mathfrak{o}(\varphi(n))$ cuando $n \rightarrow \infty$, si para todo $\epsilon > 0$ existe $M \in \mathbb{R}$ tal que:*

$$|f(n)| < \epsilon |\varphi(n)| \quad \text{para todo } n > M$$

2. *Escribimos $f(n) = \mathcal{O}(\varphi(n))$ cuando $n \rightarrow \infty$, si existen constantes $C, N \in \mathbb{R}$ tales que:*

$$|f(n)| \leq C |\varphi(n)| \quad \text{para todo } n > N$$

De acuerdo a estas definiciones, $f(n) = o(\varphi(n))$ implica que $f(n) = \mathcal{O}(\varphi(n))$. También, si $\varphi(n) \neq 0$ y utilizando la definición de límite en infinito, es fácil notar lo siguiente:

$$\begin{aligned} f(n) = o(\varphi(n)) &\Rightarrow |f(n)| < \epsilon |\varphi(n)| \Rightarrow \left| \frac{f(n)}{\varphi(n)} \right| < \epsilon \Rightarrow \left| \frac{f(n)}{\varphi(n)} - 0 \right| < \epsilon \\ &\Rightarrow \lim_{n \rightarrow \infty} \frac{f(n)}{\varphi(n)} = 0 \end{aligned}$$

Análogamente, si $\varphi(n) \neq 0$, entonces:

$$\begin{aligned} f(n) = \mathcal{O}(\varphi(n)) &\Rightarrow |f(n)| \leq C |\varphi(n)| \Rightarrow \left| \frac{f(n)}{\varphi(n)} \right| \leq C \Rightarrow \left| \frac{f(n)}{\varphi(n)} \right| < C + \epsilon \\ &\Rightarrow \left| \frac{f(n)}{\varphi(n)} \right| - C < \epsilon \Rightarrow \lim_{n \rightarrow \infty} \left(\left| \frac{f(n)}{\varphi(n)} \right| - C \right) = 0 \Rightarrow \lim_{n \rightarrow \infty} \left(\left| \frac{f(n)}{\varphi(n)} \right| \right) = C \\ &\Rightarrow \lim_{n \rightarrow \infty} \left| \frac{f(n)}{\varphi(n)} \right| < \infty \end{aligned}$$

Es importante notar que a partir de estas observaciones, las definiciones A.2.1 son equivalentes a las presentadas en [13, Sección 1.6].

A.3. Integral con respecto a una medida

En esta sección presentamos las definiciones básicas de integración con respecto a una medida. Para un tratamiento más detallado, sugerimos las referencias [24], [6]. A lo largo de esta sección, consideramos un espacio de medida $(\Omega, \mathcal{B}, \lambda)$.

Definición A.3.1. Una función $s : \Omega \rightarrow \mathbb{R}$ es una función simple si solamente puede tomar un número finito de valores.

Sean a_i los distintos valores que puede tomar una función simple $s : \Omega \rightarrow \mathbb{R}$ y sea $A_i = \{x \in \Omega : s(x) = a_i\}$. Entonces la función simple s puede escribirse como:

$$s(x) = \sum_i a_i \chi_{A_i}(x) \quad (\text{A.1})$$

donde:

$$\chi_A(x) = \begin{cases} 0 & x \in A \\ 1 & x \notin A \end{cases} \quad (\text{A.2})$$

Definición A.3.2. Sea $s : \Omega \rightarrow [0, \infty)$ una función simple con $a_i > 0$ para todo i y suponga que $A_i \in \mathcal{B}$ para todo i . Con estas condiciones, decimos que s es integrable con respecto a la medida λ (o bien que es λ -integrable) y además, definimos la integral de la función simple s mediante:

$$\int s d\lambda = \sum_i a_i \lambda(A_i) \quad (\text{A.3})$$

Definición A.3.3. Sea $f : \Omega \rightarrow [0, \infty)$ una función medible. Entonces la integral de f con respecto de λ , denotada $\int f d\lambda$, es definida mediante:

$$\int f d\lambda = \sup \left\{ \int s d\lambda : s \text{ es una función simple, } s \text{ es } \lambda \text{ integrable, } s \leq f \right\}$$

Definición A.3.4. Sea $A \in \mathcal{B}$ y $f : A \rightarrow [0, \infty)$ una función medible. Entonces la integral de f con respecto de λ , sobre A , denotada $\int_A f d\lambda$, es definida mediante:

$$\int_A f d\lambda = \int f \chi_A d\lambda$$

Definición A.3.5. Sea $f : \Omega \rightarrow \mathbb{R}$ una función medible. Denotamos $f^+ = \max(f, 0)$ y $f^- = \min(f, 0)$. Si las integrales $\int f^+ d\lambda$ y $\int f^- d\lambda$ son ambas finitas, entonces decimos que f es integrable con respecto a la medida λ o bien que es λ -integrable, y definimos la integral de f con respecto de λ :

$$\int f d\lambda = \int f^+ d\lambda - \int f^- d\lambda$$

Definición A.3.6. Sea $A \in \mathcal{B}$ y $f : A \rightarrow \mathbb{R}$ una función medible. Si las integrales $\int_A f^+ d\lambda$ y $\int_A f^- d\lambda$ son ambas finitas, entonces la integral de f con respecto de λ , sobre A es definida mediante:

$$\int_A f d\lambda = \int_A f^+ d\lambda - \int_A f^- d\lambda$$

A.4. Lemas necesarios en algunas de las demostraciones presentadas

El siguiente lema se utiliza en el ejemplo 2.3.3.

Lema A.4.1. Sea A una matriz cuadrada de $m \times m$ cuyos elementos son $a_{i,j}$, $\Sigma = \{1, \dots, m\}$, $x_1, \dots, x_n \in \Sigma$. Entonces, para todo $x_0 \in \Sigma$:

$$\sum_{x_n=1}^m (A^n)_{x_0, x_n} = \sum_{x_1, \dots, x_{n-1}} \prod_{i=0}^{n-1} a_{x_i, x_{i+1}} \quad (\text{A.4})$$

Demostración. Por definición de producto matricial, $(A^2)_{i,j} = \sum_{k=1}^m a_{i,k} a_{k,j}$. Esta misma definición puede aplicarse al producto $A^2 A$ obteniéndose:

$$(A^3)_{i,j} = \sum_{k=1}^m \left((A^2)_{i,k} \right) a_{k,j} = \sum_{k=1}^m \left(\sum_{l=1}^m a_{i,l} a_{l,j} \right) a_{k,j}$$

Por lo tanto:

$$\sum_{j=1}^m (A^3)_{i,j} = \sum_{j=1}^m \sum_{k=1}^m \sum_{l=1}^m a_{i,l} a_{l,j} a_{k,j}$$

Por conveniencia en la notación, definimos: $i = x_0$, $l = x_1$, $k = x_2$, $j = x_3$ con $x_1, x_2, x_3 \in \Sigma$. Con esta nueva notación podemos escribir:

$$\sum_{x_3=1}^m (A^3)_{x_0, x_3} = \sum_{x_3=1}^m \sum_{x_2=1}^m \sum_{x_1=1}^m a_{x_0, x_1} a_{x_1, x_2} a_{x_2, x_3} = \sum_{x_1, x_2, x_3} \left(\prod_{i=0}^2 a_{x_i, x_{i+1}} \right)$$

Por inducción se llega a (A.4) con lo que se completa la demostración del Lema A.4.1. \square

El siguiente lema se utiliza en las demostraciones de los Teoremas 2.1.3 y 3.1.1.

Lema A.4.2. *Suponga que $f : \mathbb{R} \rightarrow \mathbb{R}$ es positiva y no creciente en $[1, \infty)$. Entonces, si f es Riemann integrable se cumple la siguiente desigualdad:*

$$\int_n^\infty f(x)dx \leq \sum_{k=n}^\infty f(k) \leq \int_{n-1}^\infty f(x)dx$$

Demostración. Dado que f es positiva y no creciente se tiene para todo $k \geq 1$:

$$\int_k^{k+1} f(x)dx \leq f(k) \int_k^{k+1} dx = f(k)(k+1-k) = f(k)$$

Similarmente:

$$\int_{k-1}^k f(x)dx \geq f(k) \int_{k-1}^k dx = f(k)(k-(k-1)) = f(k)$$

Con lo que:

$$\begin{aligned} \int_k^{k+1} f(x)dx &\leq f(k) \leq \int_{k-1}^k f(x)dx \\ \sum_{k=n}^\infty \int_k^{k+1} f(x)dx &\leq \sum_{k=n}^\infty f(k) \leq \sum_{k=n}^\infty \int_{k-1}^k f(x)dx \\ \int_n^\infty f(x)dx &\leq \sum_{k=n}^\infty f(k) \leq \int_{n-1}^\infty f(x)dx \end{aligned}$$

Esto completa la demostración del lema A.4.2. \square

El siguiente lema se utiliza en la demostración del Teorema 3.1.5.

Lema A.4.3. *Sea $f : \Omega \rightarrow \mathbb{R}$ una función integrable con respecto a la medida μ . Defina $\Omega_0 = \{x \in \Omega : f(x) \geq 0\}$. Si $\int f(z)d\mu(z) = 0$ entonces $\int |f(z)|d\mu(z) = 2 \int_{\Omega_0} f(z)d\mu(z)$*

Demostración.

$$\begin{aligned} \int f(z)d\mu(z) &= \int_{\Omega_0} f(z)d\mu(z) + \int_{\Omega_0^c} f(z)d\mu(z) = 0 \\ \int_{\Omega_0} f(z)d\mu(z) &= - \int_{\Omega_0^c} f(z)d\mu(z) \end{aligned}$$

Por otro lado:

$$\begin{aligned} \int |f(z)|d\mu(z) &= \int_{\Omega_0} |f(z)|d\mu(z) + \int_{\Omega_0^c} |f(z)|d\mu(z) \\ &= \int_{\Omega_0} f(z)d\mu(z) - \int_{\Omega_0^c} f(z)d\mu(z) \\ &= \int_{\Omega_0} f(z)d\mu(z) + \int_{\Omega_0} f(z)d\mu(z) = 2 \int_{\Omega_0} f(z)d\mu(z) \end{aligned}$$

Esto completa la demostración del Lema A.4.3. \square

Las siguientes proposiciones y el lema A.4.4 se utilizan en la demostración del lema 2.2.3.

Proposición A.4.1. *Si f y g son funciones acotadas, entonces:*

$$\inf_{y \in \Omega} \{f(y) + g(y)\} \geq \inf_{y \in \Omega} \{f(y)\} + \inf_{y \in \Omega} \{g(y)\} \quad (\text{A.5})$$

$$\sup_{y \in \Omega} \{f(y) + g(y)\} \leq \sup_{y \in \Omega} \{f(y)\} + \sup_{y \in \Omega} \{g(y)\}. \quad (\text{A.6})$$

Demostración. Mostraremos primero que se cumple (A.5). Se tiene que, para todo $y \in \Omega$, $f(y) \geq \inf_{y \in \Omega} \{f(y)\}$ y $g(y) \geq \inf_{y \in \Omega} \{g(y)\}$, por lo que $f(y) + g(y) \geq \inf_{y \in \Omega} \{f(y)\} + \inf_{y \in \Omega} \{g(y)\}$. En particular, $\inf_{y \in \Omega} \{f(y) + g(y)\} \geq \inf_{y \in \Omega} \{f(y)\} + \inf_{y \in \Omega} \{g(y)\}$, con lo que se demuestra (A.5).

Similarmente demostraremos que se cumple (A.6). Se tiene que, para todo $y \in \Omega$, $f(y) \leq \sup_{y \in \Omega} \{f(y)\}$ y $g(y) \leq \sup_{y \in \Omega} \{g(y)\}$, por lo que $f(y) + g(y) \leq \sup_{y \in \Omega} \{f(y)\} + \sup_{y \in \Omega} \{g(y)\}$. En particular, $\sup_{y \in \Omega} \{f(y) + g(y)\} \leq \sup_{y \in \Omega} \{f(y)\} + \sup_{y \in \Omega} \{g(y)\}$, con lo que se demuestra (A.6). Esto completa la demostración de la Proposición A.4.1. \square

Proposición A.4.2. *Si, para todo $y \in \Omega$ se cumple $f(y) \leq g(y)$ y g es acotada por arriba, entonces:*

$$\sup_{y \in \Omega} \{f(y)\} \leq \sup_{y \in \Omega} \{g(y)\}. \quad (\text{A.7})$$

Demostración. Dado que g es acotada por arriba, se tiene que $g(y) \leq \sup_{y \in \Omega} \{g(y)\}$. También, dado que, para todo $y \in \Omega$ se cumple $f(y) \leq g(y)$, se tiene que $f(y) \leq \sup_{y \in \Omega} \{g(y)\}$. En particular, $\sup_{y \in \Omega} \{f(y)\} \leq \sup_{y \in \Omega} \{g(y)\}$. Esto completa la demostración de la Proposición A.4.2. \square

Proposición A.4.3. *Sea f una función acotada por abajo. Entonces:*

$$\sup_{y \in \Omega} \{-f(y)\} = - \inf_{y \in \Omega} \{f(y)\}.$$

Demostración. Sea $\alpha = \inf_{y \in \Omega} \{f(y)\}$. Por definición, $\alpha \leq f(y)$ para todo $y \in \Omega$. Esto quiere decir que $-\alpha \geq -f(y)$ para todo $y \in \Omega$, lo cual a su vez quiere decir que $-\alpha$ es una cota superior de $-f$. Queremos mostrar que $-\alpha$ es la menor de las cotas superiores de $-f$. Suponga que existe otra cota superior de $-f$, que denotamos b , tal que $-\alpha \geq b \geq -f(y)$. Entonces $-b \leq f(y)$ para todo $y \in \Omega$. Es decir, $-b$ es una cota inferior de f . Dado que, por definición, α es la mayor de las cotas inferiores de f , se tiene que $\alpha \geq -b$, o bien $-\alpha \leq b$. Pero por hipótesis $-\alpha \geq b$ por lo que se concluye que $b = -\alpha$ es la menor de las cotas superiores de $-f$. Por definición de supremo, $-\alpha = \sup_{y \in \Omega} \{-f(y)\}$. Dado que $\alpha = \inf_{y \in \Omega} \{f(y)\}$, se tiene que $\sup_{y \in \Omega} \{-f(y)\} = - \inf_{y \in \Omega} \{f(y)\}$. Esto completa la demostración de la proposición A.4.3. \square

Lema A.4.4. *Si f y g son funciones acotadas, entonces:*

$$\left| \inf_{y \in \Omega} \{f(y)\} - \inf_{y \in \Omega} \{g(y)\} \right| \leq \sup_{y \in \Omega} \{|f(y) - g(y)|\} \quad (\text{A.8})$$

$$\left| \sup_{y \in \Omega} \{f(y)\} - \sup_{y \in \Omega} \{g(y)\} \right| \leq \sup_{y \in \Omega} \{|f(y) - g(y)|\}. \quad (\text{A.9})$$

Demostración. Primero demostraremos que se cumple (A.8). Utilizando las proposiciones A.4.1, A.4.2 y A.4.3 se tiene:

$$\begin{aligned} \inf_{y \in \Omega} \{f(y)\} &= \inf_{y \in \Omega} \{f(y) - g(y) + g(y)\} \\ &\geq \inf_{y \in \Omega} \{f(y) - g(y)\} + \inf_{y \in \Omega} \{g(y)\} \\ \inf_{y \in \Omega} \{f(y)\} - \inf_{y \in \Omega} \{g(y)\} &\geq \inf_{y \in \Omega} \{f(y) - g(y)\} \end{aligned} \quad (\text{A.10})$$

$$\begin{aligned} &= -\sup_{y \in \Omega} \{g(y) - f(y)\} \\ &\geq -\sup_{y \in \Omega} \{|g(y) - f(y)|\} \\ &= -\sup_{y \in \Omega} \{|f(y) - g(y)|\}. \end{aligned} \quad (\text{A.11})$$

Similarmente, intercambiando f y g en la desigualdad (A.10) se tiene:

$$\inf_{y \in \Omega} \{g(y)\} - \inf_{y \in \Omega} \{f(y)\} \geq \inf_{y \in \Omega} \{g(y) - f(y)\}.$$

Multiplicando por -1 ambos lados de la desigualdad y aplicando las proposiciones A.4.2 y A.4.3 se obtiene:

$$\begin{aligned} -\inf_{y \in \Omega} \{g(y)\} + \inf_{y \in \Omega} \{f(y)\} &\leq -\inf_{y \in \Omega} \{g(y) - f(y)\} \\ &= \sup_{y \in \Omega} \{f(y) - g(y)\} \\ &\leq \sup_{y \in \Omega} \{|f(y) - g(y)|\}. \end{aligned} \quad (\text{A.12})$$

Uniendo las desigualdades (A.11) y (A.12) se obtiene (A.8). Similarmente demostraremos que se cumple también (A.9).

$$\begin{aligned} \sup_{y \in \Omega} \{f(y)\} &= \sup_{y \in \Omega} \{f(y) - g(y) + g(y)\} \\ &\leq \sup_{y \in \Omega} \{f(y) - g(y)\} + \sup_{y \in \Omega} \{g(y)\} \\ \sup_{y \in \Omega} \{f(y)\} - \sup_{y \in \Omega} \{g(y)\} &\leq \sup_{y \in \Omega} \{f(y) - g(y)\} \end{aligned} \quad (\text{A.13})$$

$$\leq \sup_{y \in \Omega} \{|f(y) - g(y)|\}. \quad (\text{A.14})$$

Intercambiando f y g en (A.13):

$$\sup_{y \in \Omega} \{g(y)\} - \sup_{y \in \Omega} \{f(y)\} \leq \sup_{y \in \Omega} \{g(y) - f(y)\}. \quad (\text{A.15})$$

Multiplicando por -1 ambos lados de la desigualdad:

$$\begin{aligned} -\sup_{y \in \Omega} \{g(y)\} + \sup_{y \in \Omega} \{f(y)\} &\geq -\sup_{y \in \Omega} \{g(y) - f(y)\} \\ &\geq -\sup_{y \in \Omega} \{|g(y) - f(y)|\} \\ &= -\sup_{y \in \Omega} \{|f(y) - g(y)|\}. \end{aligned} \tag{A.16}$$

Uniendo (A.14) y (A.16) se obtiene (A.9). Esto completa la demostración del Lema A.4.4. \square

Bibliografía

- [1] KACHUROVSKII, A.G. Y I.V. PODVIGIN. Large Deviations and Rates of Convergence in the Birkhoff Ergodic Theorem: From Hölder Continuity to Continuity. *Dokl. Akad. Nauk*, 466(1):12-15, 2016.
- [2] KACHUROVSKII, A.G. Y I.V. PODVIGIN. Correlations, Large Deviations, and Rates of Convergence in Ergodic Theorems for Characteristic Functions. *Dokl. Akad. Nauk*, 461(5):509-512, 2015.
- [3] KACHUROVSKII, A.G. Y I.V. PODVIGIN. Large Deviations and the Rate of Convergence in the Birkhoff Ergodic Theorem. *Dokl. Akad. Nauk*, 94(4):569-577, 2013.
- [4] KACHUROVSKII, A.G. The rate of convergence in ergodic theorems. *Russ. Math. Surv.* 51(4) 653-703, 1996.
- [5] TOUCHETTE, H. A basic introduction to large deviations: Theory, applications, simulations. 2011
- [6] HALMOS, P. R. Measure Theory. *Graduate Texts in Mathematics*, Springer New York, 1976
- [7] DEMBO A., ZEITOUNI O. Large Deviations Techniques And Applications. *Applications of Mathematics*, Springer Verlag New York, 2009
- [8] CHAZOTTES R., MALDONADO C. Concentration bounds for entropy estimation of one-dimensional Gibbs measures. *Nonlinearity*, Vol. 24 (2011) 2371-2381
- [9] COLLET P., MARTÍNEZ S., SCHMITT B Exponential inequalities for dynamical measures of expanding maps of the interval. *Probab. Theory Relat. Fields*, Vol 123, 301-22
- [10] ABRAMOWITZ M., AND STEGUN I. Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables, *National Bureau of Standards*, Washington, D. C., 1964; *Nauka*, Moscow, 1979)
- [11] POLLICOTT M. Rates of mixing for potentials of summable variation, *Trans. Amer. Math. Soc.* 352(2) 843-853, 1999
- [12] GAPOSHKIN V. F. On the dependence of the rate of convergence in the strong law of large numbers for stationary processes on the rate of decay of the correlation function, *English transl. in Theory Probab.* Vol. 26 (4) (1981) 706-720

-
- [13] HARDY, G. H. AND WRIGHT, E. M. An Introduction to the Theory of Numbers, 5th ed. Oxford, England: Clarendon Press, 1979
- [14] WALTERS, P. Ruelle's operator theorem and g-measures, *Trans. Amer. Math. Soc.* 214 375-387, 1975
- [15] BOWEN, R. Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms, *Lecture Notes in Mathematics* 470, Springer Verlag Berlin Heidelberg, Second edition 2017
- [16] PETERSEN K. Ergodic Theory, 1st ed., Cambridge University Press, 1989
- [17] CERCIGNANI C. Ludwig Boltzmann: the man who trusted atoms, Oxford University Press, 2006
- [18] AFRAIMOVICH V., HSU SZE-BI Lectures on chaotic dynamical systems *AMS/IP studies in advanced mathematics*, 2003
- [19] CHERNOV, N. ZHANG, HONG-KUN. Improved Estimates for Correlations in Billiards *Communications in Mathematical Physics.* 277. 305-321, 2008
- [20] MELBOURNE, IAN Large and moderate deviations for slowly mixing dynamical systems *Proceedings of The American Mathematical Society* 137. 1735-1741, 2008
- [21] R. CHAZOTTES, J. MOLES, J. UGALDE, E. Gaussian Concentration bound for equilibrium states with potentials of summable variation *arXiv:1902.07146v1*
- [22] HALASZ, G. Remarks on the remainder in Birkhoff's ergodic theorem *Ada Math. Hungar.* 28, 389-395, 1976
- [23] KRENGEL, U. On the speed of convergence in the ergodic theorem *Monatsh. Math.* 86:1, 3-6, 1978
- [24] DE BARRA, G. Measure Theory and Integration *Royal Holloway College, University of London*, 2000
- [25] DING JIU, ZHOU AIHUI Statistical Properties of Deterministic Systems *University Press, Beijing and Springer-Verlag GmbH Berlin Heidelberg*, 2009
- [26] DUNFORD N., SCHWARTZ J. Linear Operators Part I: General Theory *Interscience Publishers Inc, New York*