Universidad Nacional Autónoma de México

Facultad de Ciencias

La propiedad de especificación en los subshifts y sus propiedades ergódicas.

T E S I S

para obtener el título de:

Licenciado en Matemáticas

presenta:

Héctor Daniel Alonso Orozco

Tutor:

Dr. Ricardo Gómez Aíza

Ciudad Universitaria, 2024

Universidad Nacional Autónoma de México
UNAM
Biblioteca Central
Dirección General de Bibliotecas de la UNAM

UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

1.-Datos del alumno
Alonso
Orozco
Héctor Daniel
5548526294
Universidad Nacional Autónoma de México
Facultad de Ciencias
Matemáticas
314340290

2.-Datos del tutor
Dr.
Ricardo
Gómez
Aíza

3.- Datos sinodal 1
Dra.
Laura Clementina
Eslava
Fernández

4.- Datos sinodal 2
Dr.
Renato
Leriche
Vázquez

5.- Datos sinodal 3
Dr.
Darío
Alatorre
Guzmán

6.- Datos sinodal 4
Dra.
Alma Saraí
Hernández
Torres

7.-Datos de trabajo escrito
La propiedad de especificación en los subshifts y sus propiedades ergódicas.
108 pp
2024

"Deeds will not be less valiant because they are unpraised."
J.R.R. Tolkien

Agradecimientos

No tengo idea de como empezar, es imposible plasmar en solo unas lineas lo importante que han sido estas personas en cada aspecto de mi vida y sin ellas no hay manera en que yo pudiera estar escribiendo esto.

Primero a mis padres, Ricardo y Nena. Gracias por apoyarme desde el primer día que dije que me gustaba la ciencia y apoyarme en cada parte del camino, creyendo más en mi de lo que yo mismo lo hacía. Gracias por todo el trabajo que hicieron para que pudiera seguir mis sueños, hoy soy quien soy por ustedes.

A mis hermanos, Pau, Xime, Faby, Adriana y Luis. Gracias por escucharme, por guiarme en todos estos años e inspirarme. Cada uno de ustedes me ha enseñado a ver la vida de una forma diferente, a vivir por lo que uno ama y seguir adelante sin importar qué tan difícil puede ser.

A Jael y Meli, gracias por seguir conmigo a pesar de todos los años, aún cuando todo era complicado y nos separábamos, siempre estuvieron ahí y no dejé de sentir su cariño ni su apoyo, y aunque siga pasando el tiempo, se que siempre seguirá estando ahí todo su cariño.

A Beni y Vale, ustedes saben lo importantes que son en mi vida, todo lo que representan y cómo me apoyaron en los momentos más complicados que he tenido, nunca podré terminar de agradecer que los encontré y que me dieron su amistad en el momento en que más la necesitaba.

A mi asesor Ricardo, por aceptarme a pesar de que no sabia nada del tema, enseñarme y tenerme paciencia en lo que aprendía y guiarme en esta nueva etapa.

A mis sinodales, Laura Eslava, Darío Alatorre, Saraí Hernández y Renato Leriche, gracias por leer mi trabajo y darme la guía para poder completarlo.

A todos ellos espero algún día volverlos tan orgullosos como yo me siento de ellos y regresarles aunque sea una pequeña parte de todo lo que hicieron por mi. Los quiero mucho.

A la DGAPA-UNAM, ya que la tesis fue realizada también gracias al apoyo de una beca del Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) de la UNAM, proyecto DGAPA-PAPIIT-IN110221.

Prefacio

El trabajo de esta tesis parte de mi interés en buscar entender algunas de las propiedades ergódicas que tienen los sistemas dinámicos discretos, en especial los simbólicos. A lo largo de mis estudios en la licenciatura pude observar lo importantes que son los sistemas dinámicos discretos, no solo como herramientas para solucionar problemas en distintas ramas de las matemáticas, si no como modelos de distintos problemas importantes en física, biología, economía, etc.

Una de las aplicaciones más importantes es la teoría de la información, donde se busca entender la "información"que tiene un sistema y la manera de cuantificarla, a partir de eso surgen conceptos como la entropía y se vuelve realmente importante el poder encontrar la forma de maximizar la información del sistema.

El resolver este problema para cualquier sistema dinámico es un problema bastante complicado,así se han desarrollado propiedades para los sistemas las cuales nos ayudan a buscar la manera de maximizar la entropía y encontrar la medida que lo cumpla. Una de estas propiedades es la especificación.

Esta tesis está enfocada en explicar el concepto de especificación y su relación con la entropía máxima analizando varios artículos que se han escrito en los últimos años, desde el primero que propuso la especificación y su relación con la entropía maximal, hasta algunos que buscan debilitar la especificación y estudian sus propiedades ergódicas. Antes de analizar estos textos se dará una introducción a los conceptos más importantes que utilizaremos. En la primera parte buscaremos dar un panorama básico de los sistemas simbólicos, la entropía y la teoría ergódica, por lo tanto nos limitaremos a explicar y enunciar conceptos y teoremas importantes sin la necesidad de demostrarlos.

En el primer capítulo daremos algunos conceptos importantes en los sistemas simbólicos, los cuales serán ampliamente usados en el trabajo al ser los más simples de analizar con

la propiedad de especificación. En el segundo capítulo daremos las bases de la dinámica topológica y la teoría ergódica fundamentales en el estudio de la información.

Finalmente el tercer capítulo haremos el análisis detallado de la propiedad de especificación enfocándonos en los sistemas simbólicos, daremos ciertas proposiciones que nos ayudarán a demostrar que un sistema dinámico expansivo con especificación tiene una medida de entropía maximal única y finalmente veremos dos debilitaciones de la especificación y cómo este teorema ya no se cumple en general.

Índice general

Introducción **1**

1. Espacios shift **3**

1.1. Espacios shift completos 4

1.2. Códigos de bloques corredizos 6

1.3. Espacios shift de tipo finito 8

1.4. Espacios shift sóficos 13

1.5. Entropía 18

1.6. Notas adicionales al capítulo 25

2. Sistemas Dinámicos **27**

2.1. Dinámica topológica 28

2.2. Shifts como sistemas dinámicos 33

2.2.1. La función zeta de los espacios shift sóficos 37

2.2.2. Representaciones simbólicas de sistemas dinámicos 41

2.3. Teoría Ergódica 43

2.3.1. Cadenas de Markov y automorfismos de Bernoulli. 46

2.3.2. Particiones y entropía métrica 49

3. Especificación **55**

3.1. Propiedades básicas 55

3.2. Especificación y ergodicidad 60

3.3. Casi especificación y especificación no uniforme 78

3.4. Ergodicidad de casi especificación y especificación no uniforme 87

Bibliografía **106**

Introducción

El estudio de los sistemas dinámicos discretos ha sido de suma importancia a lo largo de la historia, ya que cuando tenemos un problema que debemos de modelar no siempre se puede realizar un experimento de forma continua. Por ejemplo si quisiéramos estudiar el crecimiento de población de alguna especie animal, no necesariamente podemos ver su cambio en cada momento, al contrario analizaríamos su cambio en cada cierto tiempo, por ejemplo cada mes.

Esta idea de tener experimentos que solo podemos realizar en tiempos específicos nos lleva a tener que estudiar el comportamiento de manera discreta y por eso estudiamos los sistemas dinámicos discretos. El tipo que más usaremos de un sistema dinámico discreto son los sistemas dinámicos simbólicos, los cuales son sistemas formados con base en espacios de sucesiones infinitas de símbolos. Estas sucesiones son muy útiles en múltiples problemas, desde modelos en física cuyo estudio es más simple desde el punto simbólico, como en problemas de información, donde la información se da por largas cadenas de símbolos, especialmente 0s y 1s, como en el caso binario.

Modelar problemas en física por medio de los sistemas simbólicos es muy interesante porque puede facilitar mucho el estudio del sistema, e.g. cuando la función que modela el sistema, por ejemplo en $\mathbb{R}^n$, puede ser bastante complicada. El problema es entonces cómo podemos elegir un sistema simbólico que sea "similar" a nuestra función original, es decir que sean "conjugados" (equivalentes como sistemas dinámicos). Una manera de intentar resolver este problema es con la entropía. La entropía topológica de un sistema es un número que tiene propiedades muy importantes, una de ellas es que dos sistemas dinámicos equivalentes tienen que tener la misma entropía topológica .

El estudiar la entropía topológica es una parte vital de este trabajo, ya que no sólo nos ayuda a saber si dos sistemas dinámicos pueden ser equivalentes, si no que nos da información realmente importante del sistema como qué tan complejo es, o qué tanta información carga el sistema. Aparte de la entropía topológica podemos definir la entropía métrica (o de teoría de la medida), la cual es una forma de definir la entropía pero en términos de la medida del sistema. La entropía métrica nos da información de la complejidad del sistema con respecto a cómo medimos su conjuntos, y se vuelve interesante el saber si existe una medida que nos pueda llevar a relacionar la entropía métrica y la entropía topológica, lo cual será el problema más importante de nuestro trabajo.

Por definición, la entropía métrica siempre es menor o igual a la entropía topológica, y nos interesa el problema de encontrar medidas de entropía máxima, es decir medidas cuya entropía métrica coincide con la entropía topológica del sistema. Para poder encontrar una forma de resolver este problema, se define una propiedad llamada especificación. La especificación nos permite encontrar elementos en el sistema que se "aproximan" a pedazos de órbitas de cierto tipo. Muchos autores han encontrado la forma de relacionar esta propiedad con el hecho de tener una única medida de entropía máxima, es decir que sea igual a la entropía topológica, y analizaremos sus correspondientes trabajos para tratar de entender esta relación. En especial encontraremos que siempre que un sistema tenga especificación y sea expansivo implica la existencia de una medida de entropía máxima.

El llegar a esa conclusión es bastante complicado y requiere un amplio conocimiento de teoría de la medida, teoría ergódica y sistemas dinámicos. Nuestro propósito es también presentar una monografía con todo el material necesarios para así poder tener una referencia completa que presente las propiedades más importantes de la especificación.

Finalmente estudiaremos un par de propiedades parecidas a nuestra especificación, que son muchos más débiles. Estas propiedades nos dan mucha menos información acerca de las medidas de entropía pero también dan una gran cantidad de problemas abiertos que se pueden estudiar y trabajar más adelante.

Capítulo 1

Espacios shift

En este capítulo la principal área que abordaremos es la dinámica simbólica, la cual se encarga de estudiar los espacios shift. A lo largo de este capítulo estudiaremos los espacios shift, empezando por los shifts completos, los cuales son los espacios shifts más básicos, y sus principales propiedades. Después veremos los espacios shift de tipo finito, los cuales son los principales espacios con los que trabajaremos, al igual que los espacios shift sóficos.

Analizaremos el concepto de entropía en el contexto de los espacios shift, el cual es vital en nuestro estudio de ergodicidad y es uno de los conceptos más importantes en sistemas dinámicos, pues nos ayuda a clasificar los sistemas dinámicos en términos de su complejidad. Por lo tanto explicaremos brevemente lo que es un sistema dinámico y cómo relacionarlos con los espacios shift para formar sistemas simbólicos.

Finalmente estudiaremos el concepto de conjugación entre sistemas dinámicos y su relación con la entropía y veremos por qué es un concepto tan importante para analizar sistemas simbólicos.

Muchos de los conceptos e ideas que veremos en este capítulo serán generalizados y extendidos en los siguientes capítulos, pero estudiarlos primero en particular en el contexto de dinámica simbólica nos da una buena base para continuar. Los siguiente conceptos fueron tomados de [12] y [3]

1.1. Espacios shift completos

Tomemos un conjunto *finito* y discreto (con la topología discreta) de símbolos $\mathcal{A}$ al cual llamaremos *alfabeto*. Los símbolos (o elementos) de un alfabeto los llamaremos *letras*. Estos símbolos pueden ser de cualquier tipo, ya sean las letras del abecedario español, así como los dígitos del 0 al 9, etc. Dado un alfabeto podemos definir el shift completo el cual es la base de la dinámica simbólica.

Definición 1.1. El $\mathcal{A}$-*shift completo*, denotado por $X_{\mathcal{A}}$, se define como

$$X_{\mathcal{A}} = \mathcal{A}^{\mathbb{Z}} = \{x = (x_i)_{i\in\mathbb{Z}} : x_i \in \mathcal{A} \quad \forall\, i \in \mathbb{Z}\}. \tag{1.1}$$

Entonces $X_{\mathcal{A}}$ consiste de todas las sucesiones bi-infinitas de símbolos en el alfabeto $\mathcal{A}$. Si tomamos el alfabeto $\mathcal{A} = \{0, 1, 2, \ldots, r-1\}$, entonces $X_{\mathcal{A}}$ se llama el r-*shift*. Una elemento (i.e. una sucesión) en el espacio shift completo se llama *punto*.

Un concepto que se usará bastante es el de *palabra* o *bloque* w sobre el alfabeto $\mathcal{A}$, esto es una sucesión finita de símbolos del alfabeto y definimos su tamaño $|w|$ como el número de símbolos que compone a la palabra . Por ejemplo si tomamos el alfabeto $\mathcal{A} = \{0, 1\}$ una palabra sería 01000101101 y su tamaño es 11. Si tomamos $|w| = 0$ entonces hablamos de la palabra vacía. Así, las palabras de longitud n sobre el alfabeto $\mathcal{A}$ es precisamente $\mathcal{A}^n$. El conjunto de todas las palabras sobre el alfabeto $\mathcal{A}$ (de cualquier tamaño, incluso la palabra vacía), lo denotaremos por $\mathcal{A}^* = \cup_{n\geq 0}\mathcal{A}^n$.

Dada una palabra $w = a_1 \ldots a_k \in \mathcal{A}^k$ de tamaño k, una *subpalabra* de w es una palabra de la forma $v = a_i \ldots a_j$ donde $1 \leq i \leq j \leq k$. Por ejemplo si tomamos la palabra $w = 0110101010$ entonces un ejemplo de subpalabra es $v = 10101$. Dadas dos palabras u, v podemos concatenarlas para formar una palabra mas larga $w = uv$ de tamaño $|w| = |u| + |v|$ evidentemente la concatenación no es conmutativa pues uv no es igual a vu pues si $u = 1001$ y $v = 011$ entonces $uv = 1001011$ y $vu = 0111001$.

Otro concepto fundamental en la dinámica simbólica es el *mapeo shift* σ, que es precisamente la función que convierte un espacio shift en un sistema dinámico discreto, es decir un espacio que tiene una función en si misma. El mapeo shift σ sobre el espacio shift completo en el alfabeto $\mathcal{A}$ es un mapeo que lleva un punto $x \in X_{\mathcal{A}}$ del shift completo a $y = \sigma(x) \in X_{\mathcal{A}}$ donde el término y_i, es decir el símbolo que se encuentra en la posición i de la sucesión y, está dado por $y_i = x_{i+1}$. Por ejemplo tomemos el punto

$$x = \ldots 01110{,}10101001 \ldots$$

sobre el alfabeto binario $\mathcal{A} = \{0, 1\}$. Entonces

$$\sigma(x) = \ldots 011101{,}0101001 \ldots$$

(el punto decimal indica la posición de la coordenada 0).

Podemos componer el mapeo shift n veces, esto quiere decir que al tomar un punto x movemos todos sus símbolos en él n veces a la izquierda. Definimos un punto de *periodo* n como un punto $x \in \mathcal{A}^{\mathbb{Z}}$ con la propiedad de que que $\sigma^n(x) = x$ y $\sigma^m(x) \neq x$ para $m < n$. En este caso, si en particular $n = 1$, entonces x es un *punto fijo*.

Supongamos que tomamos un espacio shift completo $X_{\mathcal{A}}$ y definimos una lista de palabras o bloques prohibidos $\mathcal{F} \subseteq \mathcal{A}^*$ sobre el espacio shift completo. Esta lista de bloques prohibidos puede ser finita o infinita, pero siempre es numerable porque $\mathcal{A}^*$ siempre es numerable (ya que la cardinalidad de $\mathcal{A}$ es finita).

El conjunto $X_{\mathcal{F}}$ es un subconjunto del shift completo, y lo llamaremos *espacio shift* definido por $\mathcal{F}$. El conjunto $\mathcal{F}$ es un conjunto de bloques prohibidos, por lo tanto el espacio shift es un conjunto de puntos del espacio shift completo en los que no ocurren bloques prohibidos.

Uno de los shifts mas conocidos es el shift *áureo* llamado así por que su entropía topológica (un concepto que se vera mas adelante) es el logaritmo de la razón áurea, el cual está definido como el subconjunto de todas las sucesiones bi-infinitas sobre el alfabeto binario $\mathcal{A} = \{0, 1\}$ con la propiedad de que no pueden haber dos 1s seguidos. Así, si definimos el subconjunto $\mathcal{F} = \{11\}$ como el conjunto de bloques prohibidos, entonces $X_{\mathcal{F}}$ sería el espacio shift áureo.

Si podemos definir un espacio shift a través de un conjunto de bloques prohibidos, entonces, de igual manera, podemos intentar definirlos a través de los bloques que *no* están prohibidos. Sea X un subconjunto del espacio shift completo y $n \geq 1$ un entero positivo. Definimos $\mathcal{B}_n(X)$ el conjunto de todos los n-bloques que ocurren en los puntos de X, es decir,

$$\mathcal{B}_n(X) = \{w = a_1 \dots a_n \in \mathcal{A}^n \ : \ \exists\ x = (x_i)_{i \in \mathbb{Z}} \in X \ \ \& \ \ \exists \ \ k \in \mathbb{Z}, \ \ x_k \dots x_{k+n-1} = w\}.$$

Entonces el *lenguaje* de X está definido como la unión de todos los bloques permitidos, es decir,

$$\mathcal{B}(X) = \bigcup_{n=0}^{\infty} \mathcal{B}_n(X). \tag{1.2}$$

Es importante notar que no toda colección arbitraria de bloques sobre un alfabeto es el lenguaje de un espacio shift. Por lo tanto necesitamos caracterizar cuándo sí lo es.

Proposición 1.1. *1. Sean $X \subseteq \mathcal{A}^{\mathbb{Z}}$ un espacio shift, y sea $\mathcal{L} = \mathcal{B}(X)$ su lenguaje. Si $w \in \mathcal{L}$, entonces*

a) *(propiedad de factor) cualquier subbloque de w pertenece a $\mathcal{L}$, y*

b) *(propiedad de extensión) hay bloques no vacíos u y v en $\mathcal{L}$ tal que $uwv \in \mathcal{L}$.*

2. *La condición (1) caracteriza completamente a los lenguajes de los espacio shift, es decir, $\mathcal{L} \subseteq \mathcal{A}^*$ es el lenguaje de un espacio shift si y solo si se cumple la condición (1).*

3. *El lenguaje de un espacio shift determina el espacio shift, es decir, dos espacios shift $X, Y \subseteq \mathcal{A}^{\mathbb{Z}}$ son iguales si y solo si sus lenguajes son iguales. En otras palabras, $X = Y$ si y solo si $\mathcal{B}(X) = \mathcal{B}(Y)$.*

Si tomamos el complemento (en $\mathcal{A}^*$) del lenguaje de X, entonces obtendremos el conjunto prohibido más grande que caracteriza a X, es decir, siempre tenemos que $X = X_{\mathcal{F}}$ donde $\mathcal{F} = \mathcal{A}^* \setminus \mathcal{B}(X)$, y si $\mathcal{F}' \subseteq \mathcal{A}^*$ es tal que $X = X_{\mathcal{F}'}$, entonces $\mathcal{F}' \subseteq \mathcal{F}$.

La proposición anterior nos dice que cualquier bloque en el lenguaje de un espacio shift se puede extender con otros dos bloques de cada lado y obtener un nuevo bloque que pertenece al lenguaje del espacio shift. El inverso no siempre es cierto, es decir, dados dos bloques $u, v \in \mathcal{B}(X)$, no necesariamente existe $w \in \mathcal{B}(X)$ tal que $uwv \in \mathcal{B}(X)$. Por ejemplo si tomamos el alfabeto binario $\mathcal{A}(X) = \{0, 1\}$ y el espacio shift de sucesiones que cuentan con todas sus entradas 0 a excepción de una entrada que es 1. En este espacio shift las palabra 01 y 10 son palabras válidas pero es imposible unirlas por que implicaría la existencia de dos entradas diferentes de cero. Si se cumple que para cualesquiera dos bloques $u, v \in \mathcal{B}(X)$, podemos encontrar un bloque $w \in \mathcal{B}(X)$ tal que $uwv \in \mathcal{B}(X)$, entonces decimos que el espacio shift X es *irreducible*.

1.2. Códigos de bloques corredizos

Nos gustaría encontrar funciones que transformen un espacio shift en otro, no necesariamente sobre el alfabeto. Dichas transformaciones deberán preservar la estructura topológica y dinámica de los espacios shift. Los mapeos de bloques van a ser las funciones que vamos a considerar ahora. Supongamos que tenemos una sucesión $x = \ldots x_{-1}x_0x_1 \ldots$ en un espacio shift X sobre el alfabeto $\mathcal{A}$, y queremos transformarla a una sucesión $y = \ldots y_{-1}y_0y_1 \ldots$ de un espacio shift Y sobre un alfabeto $\mathcal{U}$. Fijemos dos enteros no negativos $m, n \geq 0$ (lo que implica que $-m \leq n$). Para determinar la coordenada y_i usamos una función $\Phi\colon \mathcal{A}^{m+n+1} \to \mathcal{U}$, de forma que y_i es la imagen bajo Φ de la ventana de coordenadas de x que va de $i - m$ a $i + n$, es decir, $y_i = \Phi(x_{[i-m,i+n]})$. Esta regla determina una función $\phi\colon X \to Y$ que denomi-

namos un *mapeo bloques*. La dependencia de ϕ en términos de Φ la denotaremos escribiendo $\phi = \Phi_\infty$.

Si tenemos un espacio shift X sobre $\mathcal{A}$ y un mapeo bloques $\Phi\colon \mathcal{B}_{m+n+1}(X) \to \mathcal{U}$, entonces el mapeo $\phi : X \to \mathcal{U}^{\mathbb{Z}}$ definido por $y = \phi(x)$ con $y_i = \Phi(x_{[i-m,i+n]})$ es llamado el *código de bloques corredizo* con *memoria* m y *anticipación* n inducido por Φ (o simplemente *código de bloques*). Es decir es una función que va a tomar bloques de tamaño $m+n$ de un punto $x \in A$ y los va transformar en símbolos de y. A los números m y n los llamamos memoria y anticipación por que dado un elemento x_i el bloque corredizo solo necesita saber que paso m pasos antes, por eso es su memoria, y necesita saber que pasara n pasos después por eso es la anticipación. De la definición de código de bloques corredizo se desprende inmediatamente el siguiente resultado.

Proposición 1.2. *Sean X y Y dos espacios shift. Si $\phi : X \to Y$ es un código de bloques corredizo, entonces $\phi \circ \sigma_X = \sigma_Y \circ \phi$.*

Es importante notar que la conmutatividad de ϕ con σ no es suficiente para decir que ϕ es un código de bloques corredizo. Para eso debemos de pedir, aparte de la conmutatividad, que exista $N \geq 0$ tal que $\phi(x)_0$ es una función de $x_{[-N,N]}$.

Si un código de bloques corredizo $\phi : X \to Y$ es suprayectivo (o simplemente sobre), entonces ϕ se llama *código factor* de X a Y, y un espacio shift Y es *factor* de X si existe un código factor de X a Y. Si el código de bloques corredizo es inyectivo entonces se le llama una *incrustación* (o *encaje*) de X en Y. Si ϕ es biyectivo y la inversa $\phi^{-1}\colon Y \to X$ es un código de bloques corredizo, entonces ϕ es una *conjugación* entre X y Y, y estos dos espacios shift se dicen que son *conjugados*. Si dos espacios shift son conjugados, entonces podemos pensarlos como dos espacios *esencialmente* iguales, los cuales comparten sus propiedades, y lo único que cambia es el alfabeto que lo conforma (y por lo tanto su lenguaje, pero ambos lenguajes son esencialmente el mismo). Una de las propiedades que conserva la conjugación es el número de puntos de periodo n, es decir, $|\{x \in X \ : \ \sigma^n(x) = x\}|$. Por lo tanto este número es un invariante de conjugación. En general, los invariantes forman una parte muy importante del estudio en dinámica topológica y teoría ergódica, ya que nos permiten saber cuándo dos sistemas comparten ciertas propiedades, en particular si son conjugados.

Algo importante que debemos de saber es que la imagen de un espacio shift bajo un código de bloques deslizante en un espacio shift, y que si el código de bloques deslizante es inyectivo y sobre entonces la inversa es automáticamente un código de bloques deslizante (por lo tanto, en el párrafo anterior, hemos sido redundantes al definir la conjugación).

1.3. Espacios shift de tipo finito

Diremos que un espacio shift $X \subseteq \mathcal{A}^{\mathbb{Z}}$ es *de tipo finito* si como el espacio shift que puede ser descrito bajo un conjunto de bloques prohibidos $\mathcal{F} \subseteq \mathcal{A}^{\mathbb{Z}}$ de tamaño finito. Por ejemplo, el espacio shift áureo es de tipo finito ya que está descrito por un conjunto finito $\mathcal{F} = \{11\}$, es decir, $X = X_{\mathcal{F}}$. Observemos que dos conjuntos de palabras prohibidos $\mathcal{F}, \mathcal{F}' \subseteq \mathcal{A}^*$ puede ser distintos y aún así determinar el mismo espacio shift, es decir, es posible que $X_{\mathcal{F}} = X_{\mathcal{F}'}$. Por ejemplo, en el caso de espacio shift áureo, definamos $\mathcal{F}' = \{11, 111, 1111, 11111, \ldots\} = \{1\}^* \setminus \{\epsilon, 1\}$, donde ϵ denota a la palabra vacía. Entonces $\mathcal{F}'$ es un conjunto infinito y claramente tenemos que $X_{\mathcal{F}} = X_{\mathcal{F}'}$, donde $\mathcal{F} = \{11\}$.

Si tomamos un espacio shift de tipo finito tal que pueda ser descrito por un conjunto de bloques prohibidos todos de tamaño $M + 1$ entonces decimos que tiene *memoria* M. Siempre podemos tomar un espacio shift de tipo finito y hacer que todos los bloques de $\mathcal{F}$ sean del mismo tamaño. Para eso, primero tomamos los bloques del tamaño más grande en un conjunto finito de bloques prohibidos que generan al espacio shift. Supongamos que dicho tamaño máximo es $M+1$ con $M \geq 0$. Luego construimos todos los bloques de tamaño $M+1$ que contengan al menos un bloque de $\mathcal{F}$ de longitud no máxima. Esto hace que el nuevo conjunto de bloques prohibidos $\mathcal{F}'$ sea más grande, pero ahora todos los elementos de $\mathcal{F}'$ son del mismo tamaño, es decir, $\mathcal{F}' \subseteq \mathcal{A}^{M+1}$.

Por ejemplo tomemos el alfabeto $\mathcal{A} = \{a, b, c, d, \}$ y el conjunto $\mathcal{F} = \{ab, bcd\}$, la palabra mas grande es de tamaño 3, así que podemos crear las palabras $\{aba, abb, abc, abd, aab, bab, cab, dab\}$ todas estas son prohibidas pues contienen ab y salvo la palabra bcd no hay mas palabras de tamaño tres prohibidas, por lo tanto el nuevo conjunto $\mathcal{F}' = \{aba, abb, abc, abd, aab, bab, cab, dab, bcd\}$ forma al mismo shift que $\mathcal{F}$. Esto nos permite tener un criterio sencillo para ver si un elemento $x \in \mathcal{A}^{\mathbb{Z}}$ pertenece a X, a saber, basta que todos los bloques de longitud $M + 1$ de x no pertenezcan a $\mathcal{F}'$. Por lo tanto para cualquier espacio shift de tipo finito, siempre podemos encontrar una *memoria* $M \geq 0$ tal que, X es un espacio shift de tipo finito de memoria M.

Teorema 1.1. *Un espacio shift X es un shift finito con memoria M si y solo si para cualquier uv, $vw \in \mathcal{B}(X)$ y $|v| \leq M$ entonces $uvw \in \mathcal{B}(X)$.*

Teorema 1.2. *Si un espacio shift X es conjugado a un espacio shift de tipo finito Y, entonces X es de tipo finito.*

Estos dos teoremas son muy importantes para poder caracterizar los espacios shift de tipo finito. Un resultado importante es el hecho de que los espacios shift de tipo finito,

son esencialmente los espacios shifts definidos por gráficas dirigidas. Iremos ahora en esta dirección.

Definición 1.2. Una *gráfica dirigida* (o *digráfica*, o simplemente una *gráfica*, ya que en este trabaja todas las gráficas van a ser dirigidas), denotada por $G = (\mathcal{V}, \mathcal{E})$, consiste de un conjunto finito $\mathcal{V} = \mathcal{V}(G)$ de *vértices* (o *nodos*, o incluso también *estados*), junto con un conjunto finito $\mathcal{E} = \mathcal{E}(G)$ de *aristas*, cada arista $e \in \mathcal{E}(G)$ empieza en un vértice denotado por $i(e) \in \mathcal{V}(G)$ y termina en un vértice $t(e) \in \mathcal{V}(G)$.

Definición 1.3. Sean G y H dos gráficas. Un *homomorfismo de gráficas* de G a H consiste de un par de mapeos $\partial\phi : \mathcal{V}(G) \to \mathcal{V}(H)$ y $\phi : \mathcal{E}(G) \to \mathcal{E}(H)$ tales que $i(\phi(e)) = \partial\phi(i(e))$ y $t(\phi(e)) = \partial\phi(t(e))$ para todas las aristas de G y se denota por $(\partial\phi, \phi) : G \to H$. Si los dos mapeos son inyectivos, entonces se dice que es una *incrustación* (o *encaje*) de gráficas, y si los dos son inyectivos y sobre (es decir, biyectivos), entonces se dice que es un *isomorfismo de gráficas*. Si dos gráficas tienen un isomorfismo de gráficas se dice que son isomorfas.

Definición 1.4. Sea $G = (\mathcal{V}, \mathcal{E})$. Para cada par de vértices $I, J \in \mathcal{V}$, sea $A_{I,J}$ el número de aristas que tiene como estado inicial I y como estado final J. Entonces la *matriz de adyacencia* de G es la matriz cuadrada $A = A(G) = [A_{I,J}]_{I,J\in\mathcal{V}}$.

Esto nos permite convertir un gráfica en una matriz, la cual es mucho mas fácil de manipular y estudiar, pero también podemos tener la operación inversa y de una matriz cuadrada obtener una gráfica:

Definición 1.5. Sea $A = [A_{I,J}]$ una matriz cuadrada de dimensiones $n \times n$ con entradas enteras no-negativas. La gráfica de A es la gráfica $G = G(A)$ con vértices $\mathcal{V}(G) = \{1, 2, \ldots, n\}$ y con $A_{I,J}$ aristas con estado inicial I y estado final J, con $I, J \in \mathcal{V}(G)$.

Definición 1.6. Sea G una gráfica con un conjunto de aristas $\mathcal{E}$ y matriz de adyacencia A. El *shift de aristas* X_G es el espacio shift sobre el alfabeto $A = \mathcal{E}$ definido por

$$X_G = \{\xi = (\xi_i)_{i\in\mathbb{Z}} \in \mathcal{E}^{\mathbb{Z}} : t(\xi_i) = i(\xi_{i+1}) \text{ para todo } i \in \mathbb{Z}\}.$$

Una *caminata* en una gráfica, es una sucesión de aristas donde el estado terminal de la primera arista es el mismo estado inicial que la segunda arista, y así sucesivamente, por lo tanto un elemento en el shift de aristas es una caminata (doblemente) infinita.

Proposición 1.3. *Si G es una gráfica con matriz de adyacencia A, entonces el shift de aristas correspondiente $X_G = X_A$ es un shift de tipo finito con memoria 1.*

Las gráficas pueden tener aristas que no aparecen en el shift de aristas, esto debido a que algunas aristas pueden tener estados terminales que no sean el inicio de ninguna otra arista,

o viceversa. A este tipo de vértices se les llama *varados*. Una gráfica es *esencial* si ninguno de sus vértices está varado. Generalmente usaremos gráficas esenciales, así que si tomamos un gráfica H tal que su conjunto de vértices es subconjunto del conjunto de vértices de G, y lo mismo con las aristas la llamamos *subgráfica* y buscamos aplicar la siguiente proposición.

Proposición 1.4. *Si G es una gráfica entonces existe una única subgráfica H de G tal que H es esencial y su shift de aristas es el mismo.*

Definición 1.7. Un *camino* de de longitud n (o simplemente un *camino*) en una gráfica G es una sucesión finita de aristas $\pi = e_1e_2\dots e_n$ las cuales cumplen que $t(e_i) = i(e_{i+1})$ para todo $i = 1, \dots, n-1$. El tamaño de un camino está dado por el número de aristas que lo componen. Dado el camino $\pi = e_1e_2\dots e_n$ entonces el vértice inicial del camino es $i(\pi) = i(e_1)$ y el vértice terminal del camino es $t(\pi) = t(e_n)$. Si el vértice terminal de un camino es el mismo que el inicial, entonces el camino lo llamamos un *ciclo*. Si un ciclo no interseca ninguno de sus vértices excepto el inicial y el terminal, entonces es un *ciclo simple*.

Proposición 1.5. *Sea $G = (V, E)$ una gráfica con matriz de adyacencia A y sea $m \geq 0$ un entero no negativo.*

1. *El número de caminos posibles de longitud m de $I \in V$ a $J \in V$ está dado por $(A^m)_{I,J}$.*
2. *El número de ciclos de longitud m en G esta dado por $tr(A^m)$, donde $tr(A)$ es la traza de A.*

Definición 1.8. Una gráfica G se dice *irreducible* si para cada par de nodos I y J existe un camino que comience en I y termine en J.

Proposición 1.6. *Una gráfica esencial es irreducible si y solo si su shift de aristas asociado es irreducible.*

Esta proposición es muy útil ya que en varios casos puede ser más fácil observar la irreducibilidad para gráficas que para un shift. Toda esta información de gráficas nos ayudará mucho pues veremos que cualquier shift de tipo finito puede ser convertido por medio de un código en un shift de aristas. De manera general llamaremos *código* a cualquier función que transforme un espacio shift en otro. Es importante notar que ésto no implica que todos los shifts finitos sean shifts de aristas, tomemos por el ejemplo el shift áureo. Sabemos que este espacio shift solo tiene un par de aristas 0 y 1, por el Teorema 1.4 existe una subgráfica H esencial cuyo shift de aristas es el mismo. Entonces al tener dos aristas y ser esencial solo puede tener 1 o 2 vértices. Si tuviera un solo vértice, ambas aristas tendrían el mismo nodo inicial y terminal, por lo tanto tendríamos el shift completo en dos símbolos, lo cual es una contradicción. Por otro lado, si tuviera dos vértices, entonces tendríamos caminos infinitos

solo formados por la palabra 01 o 10 pues el inicial de una arista sería el terminal de la otra. Ninguno de estos espacios shift es el shift áureo y por lo tanto no es un shift de aristas. Por esto hay que buscar una forma tal que, siempre que sea posible, dado un shift podamos construir otro que sí sea un shift de aristas. Esto lo podremos hacer precisamente cuando el espacio shift es de tipo finito, como veremos más adelante.

Definición 1.9. Sea X un espacio shift sobre el alfabeto $\mathcal{A}$. Si pensamos al conjunto $\mathcal{B}_n(X)$ de todos los bloques permitidos de tamaño $n \geq 1$ como conjunto de símbolos, podemos construir un nuevo alfabeto $\mathcal{A}_X^{[n]}$ donde cada letra es una de las palabras permitidas de tamaño n, es decir $\mathcal{A}_X^{[n]} = \mathcal{B}_n(X)$. Al ser un alfabeto, existe el shift completo sobre ese alfabeto y definimos el *código del bloques de altura* n como $\beta_n : X \to (\mathcal{A}_X^{[n]})^{\mathbb{Z}}$ tal que $(\beta_n(x))_{[i]} = x_{[i,i+n-1]}$.

Lo que hace este código es convertir la i-ésima coordenada de una sucesión x en la palabra de tamaño n en x que empieza en i. Un elemento en $(\mathcal{A}_X^{[n]})^{\mathbb{Z}}$ es una sucesión de palabras que se sobreponen pues tanto $(\beta_n(x))_{[i]} = x_{[i,i+n-1]} = u$ como $(\beta_n(x))_{[i+1]} = x_{[i+1,i+n]} = v$ contienen a la subpalabra $x_{[i+1,1+n-1]}$, tal que si $u = u_1u_2 \dots u_n$ y $v = v_1v_2 \dots v_n$ entonces $u_2 \dots u_n = v_1 \dots v_{n-1} = x_{[i+1,i+n-1]}$.

Definición 1.10. Sea X un espacio shift. Entonces el *espacio shift de bloques de altura* N *de* X, denotado por $X^{[N]}$, es la imagen del código $\beta_N(X)$ en el shift completo $\mathcal{A}_X^{[N]}$.

Proposición 1.7. *Los espacios shift de bloques de altura N de un espacio shift son espacios shift para cualquier N.*

Una observación importante es que el mapeo β_n es claramente invertible. Para ver esto, sea $y \in X^{[N]}$, es decir, sea $x \in X$ tal que $y = \beta_n(x)$. Entonces cada uno de los símbolos de y es un n-bloque de x, y los símbolos consecutivos de y, vistos como n-bloques de x, se sobreponen en $n-1$ de sus coordenadas. Por construcción entonces basta considerar el primer símbolo de cada uno de estos n-bloques que conforman a y para recuperar x. En virtud de la proposición anterior, podemos concluir que X y $X^{[N]}$ son *esencialmente el mismo* (es decir, son *conjugados*, concepto que definiremos con precisión más adelante).

Teorema 1.3. *Si X es un espacio shift de tipo finito de memoria $M \geq 0$, entonces existe una gráfica G tal que $X^{[M+1]} = X_G$.*

Idea de la demostración. Primero notemos que si $M = 0$, entonces estamos hablando de que X es un espacio shift completo. En este caso, definimos la gráfica G con un solo vértice y el mismo número de aristas que los símbolos en el espacio shift completo, y ya. Ahora supongamos que $M \leq 1$. Definimos el conjunto de vértices de G como $\mathcal{V} = \mathcal{B}_M(X)$, esto es, todos los M-bloques permitidos del espacio shift X. Ahora supongamos que tomamos

$I = a_1a_2\dots a_M$ y $J = b_1b_2\dots b_M$ dos vértices en G, los cuales son bloques de tamaño M de X. Si se cumple que $a_2a_3\dots a_m = b_1b_2\dots b_{M-1}$, entonces $a_1b_1b_2\dots b_M = a_1\dots a_Mb_M$ y si esta sucesión pertenece a $\mathcal{B}(X)$, entonces se crea una arista de I a J, de otra manera no se unen los nodos. De esta manera una caminata bi-infinita a lo largo de G es una sucesión de $[M+1]$-bloques pues cada arista está formada por bloques válidos de tamaño $M+1$ que se sobreponen progresivamente. □

Utilizando el proceso descrito en la idea de demostración, podemos transformar cualquier espacio shift de tipo finito en una espacio shift por aristas. Notemos entonces que al considerar el espacio shift de bloques de altura N asociado a un espacio shift de tipo finito X, de memoria N, estamos esencialmente reduciendo la memoria. Reducir la memoria de un espacio shift de tipo finito es una técnica muy útil y estándar dentro del estudio de los espacios shift de tipo finito.

Por ejemplo en el espacio shift áureo, sabemos que es un shift de memoria 1, sin embargo no es un shift por aristas, como vimos anteriormente. Consideremos el espacio shift de bloques de altura 2. Los 1-bloques válidos, es decir, el alfabeto, son 1 y 0. Los 2-bloques válidos son 01, 10 y 00 (el único 2-bloque inválido es 11). Por lo tanto creamos tres vértices con las etiquetas de estos tres 2-bloques válidos. Luego, considerando las superposiciones entre estos tres 2-bloques válidos, colocamos una arista del vértice 01 al vértice 10, una arista del vértice 10 al 01, y una arista del vértice 00 al 00 (es decir, un ciclo de longitud 1 en 00). Hemos recodificado el espacio shift áureo en un nuevo espacio shift por aristas.

Así como un espacio shift puede ser recodificado en una nueva presentación en bloques, lo mismo podemos hacer cuando un espacio shift es un espacio shift de aristas:

Definición 1.11. Sea G un gráfica. Para $N \geq 2$ definimos la N-ésima gráfica de aristas más grande $G^{[N]}$ de G como la gráfica que tiene como conjunto de vértices a todos los caminos de tamaño N en G, y para cada par de vértices $e_1e_2\dots e_N$, $f_1f_2\dots f_N$, si $e_2\dots e_N = f_1\dots f_{N-1}$, entonces colocamos una arista entre ellos.

Observemos que las entradas de la matriz de adyacencia de $G^{[N]}$ son 0s y 1s. El lector familiar con teoría de gráficas encontrará semejanzas entre la definición anterior y el concepto de *gráfica de líneas*.

Proposición 1.8. *Si G es una gráfica entonces $(X_G)^{[N]} = X_{G^{[N]}}$.*

Definición 1.12. Sea B una matriz de $r \times r$ de ceros y unos, esto es una matriz de adyacencia de una gráfica G donde solo hay a lo más una arista entre dos vértices. Entonces el espacio shift de vértices $\hat{X}_B = \hat{X}_G$ es el espacio shift con alfabeto $A = \{1, 2\dots r\}$ definido por

$$\hat{X}_B = \{x = (x_i)_{i\in\mathbb{Z}} \in A^{\mathbb{Z}} \text{ tal que } B_{x_i,x_{i+1}} = 1 \text{ para todo } i \in \mathbb{Z}\}.$$

Proposición 1.9.

- *Si renombramos los símbolos, los espacios shift de tipo finito de memoria* 1 *son los mismos que los espacios shift por vértices.*
- *Renombrando los símbolos, cada espacio shift de aristas es un espacio shift de vértices (de otra gráfica).*
- *Si X es un espacio shift de tipo finito de memoria M, entonces $X^{[M]}$ es un espacio shift de memoria* 1*, es decir, es un espacio shift de vértices. De hecho, existe una gráfica G tal que $X^{[M]} = \hat{X}_G$ y $X^{[M+1]} = X_G$.*

Esta proposición nos da una relación entre espacios shift de tipo finito, espacios shift de aristas y shifts de vértices. Nos da la capacidad de poder cambiar entre ellos y trabajar en el que nos sea más cómodo. Por lo tanto, los espacios shift relacionados con gráficas nos ayudarán mucho para trabajar con muchos tipos diferentes de espacios shift. De hecho, hay todavía una clase más amplia de espacios shift definidos por gráficas "etiquetadas", tema que a continuación abordaremos.

1.4. Espacios shift sóficos

Definición 1.13. Una *gráfica etiquetada* $\mathcal{G}$ está formada por el par (G, L), donde $G = (\mathcal{V}, \mathcal{E})$ es una gráfica y $L\colon \mathcal{E} \to A$ es un "etiquetado" de las aristas con un alfabeto A. De esta forma, para cada arista $e \in \mathcal{E}$, la función L la asocia un símbolo del alfabeto A. La *gráfica subyacente* de $\mathcal{G}$ es G.

La función L no tiene por que ser inyectiva, de hecho en general no lo va a ser, y mandará muchas aristas al mismo símbolo del alfabeto. Para las gráficas etiquetadas también podemos definir una matriz de adyacencia $A_{\mathcal{G}}$ donde la entrada I, J esta dada por la suma formal de las etiquetas de las aristas del vértice I al vértice J, si no hay aristas entre dos vértices, la entrada es formalmente *el vacío*. Una gráfica etiquetada es *irreducible* si su gráfica subyacente es irreducible.

Definición 1.14. Sea $\mathcal{G} = (G, L_G)$ y $\mathcal{H} = (H, L_H)$ dos gráficas etiquetadas. Un *homomorfismo entre dos gráficas etiquetadas* $\mathcal{G}$ y $\mathcal{H}$ es un homomorfismo de gráficas $(\partial\phi, \phi)\colon G \to H$ tal que $L_H(\phi(e)) = L_G(e)$ para todas las aristas $e \in \mathcal{E}(G)$. Si tanto $\partial\phi$ como ϕ son biyectivas, entonces decimos que el homomorfismo es de hecho un *isomorfismo de gráficas etiquetadas.*

Sea $\mathcal{G} = (G, L)$ una gráfica etiquetada. Si tomamos un camino $\pi = e_1 e_2 \dots e_n$ en G, entonces podemos considerar "etiquetar" el camino, considerando la etiqueta de cada arista

que lo conforma, es decir, la imagen $L(\pi) = L(e_1)L(e_2)\dots L(e_n)$. Esto es un n-bloque en el alfabeto A. Esto puede ser trasladado a una caminata bi infinita en G, donde etiquetamos cada uno de las aristas. El conjunto de todas las caminatas bi infinitas etiquetadas está denotado como $X_{\mathcal{G}}$. Formalmente,

$$X_{\mathcal{G}} = \{(L(x_i))_{i\in\mathbb{Z}} \ : \ (x_i)_{i\in\mathbb{Z}} \in X_G\}.$$

Definición 1.15. Un subconjunto X de un shift completo es un *shift sófico* si $X = X_{\mathcal{G}}$ para alguna gráfica etiquetada $\mathcal{G}$. La presentación de un shift sófico X es una gráfica etiquetada $\mathcal{G}$ donde $X_{\mathcal{G}} = X$.

Es importante observar que la presentación de un shift sófico no es única (por ejemplo, cualquier gráfica etiquetada con el alfabeto $\{0\}$ da lugar al espacio shift completo en un símbolo, siempre y cuando la gráfica contenga al menos un camino bi-infinito). De hecho, dos gráficas etiquetadas no isomorfas pueden dar lugar al mismo espacio shift sófico. Por lo tanto, siempre es importante observar qué presentación estamos tomando en cuenta y lo que podemos obtener de ella.

Sea un bloque w en $\mathcal{B}(X)$ donde X es un espacio shift sófico. Decimos que w tiene una presentación π, donde π es un camino en la gráfica etiquetada, si $L(\pi) = w$. La presentación de w no es única, no solo por que hay diferentes formas de presentar el shift, si no que una misma presentación podemos encontrar diferentes presentaciones para el mismo bloque.

Teorema 1.4. *Los espacios shift sóficos son espacios shifts.*

Teorema 1.5. *Cada espacio shift de tipo finito es un espacio shift sófico.*

Estos dos teoremas nos muestran dos cualidades importantes de los shifts sóficos. Primero nos aseguran que son espacios shift, y segundo que siempre que tengamos un espacio shift de tipo finito lo podemos ver como un espacio shift sófico, aunque el inverso no se cumple, es decir, no todos los espacios shift sóficos son de tipo finito.

Proposición 1.10. *Un espacio shift sófico es un espacio shift de tipo finito si y solo si tiene una presentación (G, L) tal que L_∞ es un conjugación.*

Dado un shift X cualquiera queremos saber si ese shift es sófico, para eso podemos usar el siguiente teorema.

Teorema 1.6. *Un espacio shift es sófico si y solo si es un factor de un shift de tipo finito.*

Corolario 1.1. *El factor de un espacio shift sófico es sófico.*

Corolario 1.2. *Un espacio shift que sea conjugado a un shift sófico es sófico.*

El teorema anterior y sus corolarios nos dicen que la colección de espacios shift sóficos es la colección más pequeña que contiene a los shifts de tipo finito y sus factores.

Pero obtener factores de un shift de tipo finito puede ser complicado, o más aún, comprobar si un espacio shift es factor de un espacio shift sófico. Por eso mostraremos otra forma de probar que un espacio shift es sófico.

Definición 1.16. Sea X un espacio shift y w una palabra en $\mathcal{B}(X)$. El *conjunto sucesor* $F_X(w)$ de w en X es el conjunto que contiene todas las palabras en $\mathcal{B}(X)$ que pueden seguir a w en X, es decir, $\mathcal{F}_X(w) = \{v \in \mathcal{B}(X) : wv \in \mathcal{B}(X)\}$. La colección de todos los conjuntos sucesores de X está denotada por $\mathcal{C}_X$.

Supongamos que X es un espacio shift sobre un alfabeto $\mathcal{A}$ tal que $\mathcal{C}_X$ es finito. Entonces podemos construir la gráfica etiquetada $\mathcal{G} = (G, L)$ llamada la *gráfica de conjuntos sucesores*, donde el conjunto de vértices es precisamente el conjunto sucesor $\mathcal{C}_X$, y donde para todo bloque $w \in \mathcal{B}(X)$ y cualquier símbolo $a \in A$, si se cumple que $wa \in \mathcal{B}(X)$, entonces creamos una arista etiquetada con a del vértice $\mathcal{F}_X(w)$ al vértice $\mathcal{F}_X(wa)$.

Proposición 1.11. *Si X es un espacio shift con un número finito de conjuntos sucesores y $\mathcal{G}$ es la gráfica de conjuntos sucesores, entonces G es una presentación de X, en particular X es sófico.*

La proposición anterior nos dice que cuando tengamos un espacio shift y podamos obtener sus conjuntos sucesores, si resulta que hay un número finito de ellos, entonces podemos asegurar que nuestro espacio shift es sófico. Esta propiedad de hecho caracteriza a los espacios shift sóficos:

Teorema 1.7. *Un espacio shift es sófico si y solo si tiene un número finito de conjunto sucesores.*

Definición 1.17. Una gráfica etiquetada $\mathcal{G} = (G, L)$ se dice que es *derecha resolvente* si para cada vértice I en G, las aristas que empiezan en I tienen diferentes etiquetas. Una presentación derecha resolvente de un espacio shift sófico es una presentación de X que es precisamente derecha resolvente.

Teorema 1.8. *Todo espacio shift sófico tiene una presentación derecha resolvente.*

Por ejemplo, ya vimos que la gráfica de conjuntos sucesores de un espacio shift sófico X es una presentación del mismo, y por construcción, es siempre derecha resolvente.

Generalmente el precio a pagar por tener una representación derecha resolvente es que el número de vértices de la presentación es mucho mayor que una que no es derecha resolvente, así que nos gustaría encontrar la representación derecha resolvente más simple.

Definición 1.18. Una *presentación derecha resolvente mínima* de un espacio shift sófico X es una representación derecha resolvente con el mínimo número de vértices posibles.

Definición 1.19. Sea $\mathcal{G} = (G, L)$ una gráfica etiquetada, y sea $I \in V$ es un vértice de G. El *conjunto sucesor* $F_{\mathcal{G}}(I)$ de I en $\mathcal{G}$ es la colección de etiquetas de los caminos que comienzan en I. Decimos que $\mathcal{G}$ es *sucesivamente separada* si estados diferentes tienen conjuntos sucesores distintos.

Sea $\mathcal{G} = (G, L)$ una gráfica etiquetada con el alfabeto $\mathcal{A}$. Dados dos vértices I, J, decimos que son *equivalentes* si tienen el mismo conjunto sucesor. El ser equivalente es claramente una relación de equivalencia. Podemos entonces construir las clases de equivalencia $\{\mathcal{I}_1, \mathcal{I}_2, \ldots\}$ de los vértices que sean equivalentes. Con ellas, podemos seguidamente construir la gráfica etiquetada $\mathcal{H}$ tal que los nodos son las clases de equivalencia, y colocamos una arista etiquetada con $a \in \mathcal{A}$ entre dos nodos $\mathcal{I}, \mathcal{J}$ siempre que existan dos nodos $I \in \mathcal{I}$ y $J \in \mathcal{J}$ en la gráfica original unidos por un arista de I a J etiquetada con a. Llamaremos a esta gráfica la *gráfica unida* de $\mathcal{G}$.

Lema 1.1. *Sea $\mathcal{G}$ una gráfica etiquetada y $\mathcal{H}$ la gráfica unida de $\mathcal{G}$. Entonces $\mathcal{H}$ es sucesivamente separada y $X_{\mathcal{H}} = X_{\mathcal{G}}$. Más aún, si $\mathcal{G}$ es irreducible, entonces también lo es $\mathcal{H}$ y si $\mathcal{G}$ es derecha resolvente, entonces también lo es $\mathcal{H}$.*

Proposición 1.12. *Una presentación derecha resolvente mínima de un espacio shift sófico es sucesivamente separada.*

Si $\mathcal{G}$ es una gráfica etiquetada irreducible, entonces el $X_{\mathcal{G}}$ es un espacio shift sófico irreducible, pues si tomamos dos palabras w, v en $\mathcal{B}(X_{\mathcal{G}})$, entonces existen caminos en la gráfica subyacente π, τ tales que la etiqueta de estos caminos son las palabras w, v. Al ser la gráfica subyacente irreducible, podemos encontrar un camino σ con etiqueta u tal que el nodo inicial de σ es el nodo terminal de π, y el nodo terminal de σ es el nodo inicial de τ, por lo tanto la etiqueta del camino $\pi\sigma\tau$ es igual a la unión de las etiquetas de cada uno y entonces tenemos que se cumple que wuv está en el lenguaje del espacio shift representado por $\mathcal{G}$ y por lo tanto es irreducible. El inverso no siempre se cumple, es decir que un espacio shift sófico irreducible puede tener representaciones que no son irreducibles (considérense dos copias idénticas y ajenas de una gráfica etiquetada).

Lema 1.2. *Supongamos que X es un espacio shift sófico irreducible, y que $\mathcal{G}$ es una presentación derecha resolvente mínima de X. Entonces $\mathcal{G}$ es irreducible.*

Proposición 1.13. *Un espacio shift sófico es irreducible si y solo si tiene una representación irreducible.*

Definición 1.20. Sean $\mathcal{G} = (G, L)$ una gráfica etiquetada. Una palabra $w \in \mathcal{B}(X_{\mathcal{G}})$ es una palabra *sincronizante* para $\mathcal{G}$ si todo los caminos en G que sean presentaciones de w terminan en el mismo vértice I, en cuyo caso diremos que w se *enfoca* en I.

Lema 1.3. *Supongamos que $\mathcal{G}$ es una gráfica etiquetada derecha resolvente y que w es una palabra sincronizante para $\mathcal{G}$. Entonces cualquier palabra de la forma wu en el lenguaje de $X_{\mathcal{G}}$ es también sincronizante en $\mathcal{G}$, y si w se enfoca en I, entonces el conjunto sucesor de w es el mismo que el de I.*

Proposición 1.14. *Sea $\mathcal{G}$ una gráfica etiquetada derecha resolvente que es sucesivamente separada. Entonces cada palabra $u \in \mathcal{B}(X_{\mathcal{G}})$ puede ser extendida a la derecha a una palabra sincronizante uw.*

Proposición 1.15. *Si $\mathcal{G}$ y $\mathcal{H}$ son representaciones derecha resolventes, irreducibles y sucesivamente separadas de un mismo espacio shift sófico, entonces $\mathcal{G}$ y $\mathcal{H}$ son isomorfas como gráficas etiquetadas.*

Teorema 1.9. *Cualquier par de presentaciones mínimas derechas resolventes de un espacio shift sófico son isomorfas entre ellas.*

Corolario 1.3. *Sea X un espacio shift sófico irreducible. Entonces una gráfica derecha resolvente $\mathcal{G}$ es la representación derecha resolvente mínima de X si y solo si es irreducible y sucesivamente separada.*

Corolario 1.4. *Sea X un espacio shift sófico irreducible, y sea $\mathcal{G}$ una representación irreducible derecha resolvente de X. Entonces la gráfica unida de $\mathcal{G}$ es la representación derecha resolvente mínima de X.*

Todas estas proposiciones y teoremas nos sirven para mostrar que cuando tengamos un espacio shift sófico, la "mejor" presentación que podemos tomar es la presentación derecha resolvente mínima. Ésto es debido a que todas las representaciones son isomorfas y no solo estamos tomando la que tiene menos nodos, aparte es derecha resolvente y sucesivamente separada, es decir que no tendrá diferentes presentaciones de la misma palabra a partir de un nodo. De cierta manera la presentación derecha resolvente mínima actúa como una forma canónica de todas las presentaciones.

1.5. Entropía

Hemos visto que los espacios shift se pueden caracterizar por los conjuntos de palabras que están permitidas, o equivalentemente, por las que no. Dados dos espacios shifts sobre el mismo alfabeto, podemos pensar que entre más palabras prohibidas tenga uno de ellos, y por lo tanto menos palabras permitidas, es menos complicado (o complejo). Es decir, si tomamos un símbolo x en el alfabeto, la cantidad de elementos que le pueda seguir para que sea una palabra permitida será, en general, mayor mientras más palabras de tamaño 2 tenga. Esto se puede extender a palabras de mayor tamaño. Entonces entre más grande sea el conjunto de las palabras permitidas, hay más opciones de símbolos para extender una palabra, por eso decimos que es más complicado.

Dado que el conjunto de las palabras permitidas es un conjunto infinito, nos gustaría encontrar una forma de cuantificar esta complejidad de los espacios shift tomando todas las palabras permitidas.

Definición 1.21. Sea X un espacio shift. La *entropía* de X está definida como

$$h(X) = \lim_{n\to\infty} \frac{1}{n} \log |B_n(X)|. \tag{1.3}$$

donde $\log = \log_2$

La entropía mide el crecimiento exponencial de la sucesión de la cantidad de bloques permitidos de tamaño n, es decir la complejidad, pues entre más bloques permitidos tenga, más complejo es el sistema. Se utiliza el logaritmo en base dos, por conveniencia pues el creciminto de las palabras suele ser 2^k.

Si el alfabeto de X está denotado por A, entonces $|B_n(X)| \leq |A|^n$, por lo tanto $\log(|B_n(X)|) \leq \log(|A|^n)$ entonces $h(X) = \frac{1}{n}\log(|B_n(X)|) \leq \log|A|$ lo cual nos dice que la entropía de un espacio shift no diverge al infinito, pero ¿cómo podemos saber que si converge a un límite fijo?

Lema 1.4. *Sea $a_1, a_2, \ldots$ una sucesión no-negativa de números tal que $a_{m+n} \leq a_m + a_n$ para todo $m, n \geq 1$. Entonces $\lim_{n\to\infty}(a_n/n)$ existe y es igual a $\inf_{n\geq 1} a_n/n$.*

Proposición 1.16. *Si X es un espacio shift, entonces $\lim_{n\to\infty} \frac{1}{n}\log|B_n(X)|$ existe y es igual a $\inf_{n\geq 1} \frac{1}{n}\log|B_n(X)|$.*

Demostración. Para dos enteros $m, n \geq 1$, un bloque $B_{n+m}(X)$ está determinado por el bloque de tamaño m inicial y el bloque sucesor de tamaño n. Entonces la cantidad de bloques

de tamaño $m+n$ debe de ser a lo más las combinaciones posibles de bloques de tamaño m y n, por lo tanto $|B_{m+n}(X)| \leq |B_m(X)| \cdot |B_n(X)|$, entonces aplicando el logaritmo obtenemos que $\log|B_{m+n}(X)| \leq \log|B_m(X)| + \log|B_n(X)|$ y podemos aplicar el lema anterior a una sucesión $a_n = \log|B_n(X)|$. □

Proposición 1.17. *Si Y es un factor de X, entonces $h(Y) \leq h(X)$.*

Ahora sabemos que si dos espacios shift son conjugados, entonces cada uno es factor del otro, por lo tanto la entropía es la misma. Por lo tanto, la entropía es invariante bajo conjugación.

Corolario 1.5. *Si X y Y son conjugados, entonces $h(X) = h(Y)$.*

Proposición 1.18. *Si Y está encajado en X, entonces $h(Y) \leq h(X)$.*

Proposición 1.19. *Sea $\mathcal{G} = (G, L)$ una gráfica etiquetada derecha resolvente. Entonces $h(X_{\mathcal{G}}) = h(X_G)$.*

Definición 1.22. Para una espacio shift X denotemos como $p_n(X)$ el numero de puntos en X que tienen periodo n.

Proposición 1.20. *Sea X un espacio shift. Entonces* $\limsup_{n\to\infty} \frac{1}{n}\log(p_n(X)) \leq h(X)$.

Proposición 1.21. *Sea $A \neq 0$ una matriz cuadrada de $r \times r$, con entradas reales no-negativas, con un eigenvector positivo V (es decir que todas sus entradas son positivas). Entonces el eigenvalor λ correspondiente es real positivo y existen constantes positivas c_0, d_0 tales que $c_0\lambda^n \leq \Sigma_{I,J=1}^r (A^n)_{IJ} \leq d_0\lambda^n$. En particular, si A es la matriz de adyacencia de una gráfica G, entonces $h(X_G) = \log\lambda$.*

Definición 1.23. Una matriz A no-negativa es *irreducible* si para cada par de índices ordenados I, J, existen algún $n = n(I, J) \geq 0$ tal que $A^n_{IJ} > 0$. Adoptaremos al convención de que para cualquier matriz A, tenemos que $A^0 = Id$. Una matriz no-negativa A es *esencial* si ninguna de sus columnas o filas es cero.

Claramente, A es esencial si y sólo si $G(A)$ es esencial.

Teorema 1.10 (Perron-Frobenius)**.** *Sea $A \neq 0$ una matriz irreducible. Entonces A tiene un eigenvector positivo v_A con un eigenvalor correspondiente $\lambda_A > 0$ que es geométricamente simple (su eigenespacio correspondiente es unidimensional) y algebraicamente simple (es una raíz simple del polinomio característico). Si μ es otro eigenvalor para A, entonces $|\mu| \leq \lambda_A$. Cualquier eigenvector positivo de A es un múltiplo positivo de v_A.*

Teorema 1.11.

- *Si G es una gráfica irreducible, entonces $h(X_G) = \log(\lambda_{A(G)})$.*

- *Si X es un espacio shift de tipo finito irreducible de memoria M y G es una gráfica esencial para la cual $X^{[M+1]} = X_G$, entonces $h(X) = \log(\lambda_{A(G)})$.*

Teorema 1.12. *Sea X un shift sófico irreducible y sea $\mathcal{G} = (G, L)$ una presentación irreducible derecha resolvente de X. Entonces $h(X) = \log(\lambda_{A(G)})$.*

Estos teoremas nos muestran una forma fácil de calcular la entropía de un espacio shift sófico (en particular un shift de tipo finito), conociendo la matriz de adyacencia A de la gráfica subyacente, aplicando el teorema de Perron-Frobenius y encontrando un mayor eigenvector positivo.

Teorema 1.13. *Si X es un espacio shift sófico irreducible, entonces*

$$\limsup_{n\to\infty} \frac{1}{n} \log(p_n(X)) = h(X).$$

Definición 1.24. Para un espacio shift X, $q_n(X)$ denota el número de puntos en X que tienen periódo mínimo n, es decir,

$$q_n(X) = |\{x \in X \ : \ \sigma^n(x) = x \text{ y } \sigma^k(x) \neq x \text{ si } k \in \{1, \ldots, n-1\}\}|.$$

Corolario 1.6. *Si X es un espacio shift sófico irreducible, entonces*

$$\limsup_{n\to\infty} \frac{1}{n} \log(q_n(X)) = h(X).$$

Sea A una matriz cuadrada de $r \times r$ con entradas no-negativas (no necesariamente enteras). Podemos asociarle a A una gráfica simple (es decir, sin aristas múltiples), de la siguiente manera. Definimos $A^{\#}$ como una matriz cuadrada de $r \times r$ con entradas en $\{0, 1\}$, donde para cualesquiera $I, J \in \{1, \ldots, r\}$, tenemos $A^{\#}_{I,J} = 1$ si y sólo si $A_{I,J} \neq 0$. La gráfica simple asociada a A es la gráfica G con matriz de adyacencia $A^{\#}$. Si $(A^n)_{IJ} > 0$ para algún n, es decir, si existe un camino en G de I a J, entonces decimos que $I \rightsquigarrow J$, y si además se cumple que $J \rightsquigarrow I$, entonces diremos que I y J se *comunican*. La comunicación entre vértices es un relación de equivalencia, por lo que podemos partir a los vértices en *clases de comunicación*.

Podemos crear una gráfica H donde su conjunto de vértices esté dado por las clases de comunicación de una gráfica G, y donde existe una arista entre clases $E, F \in \mathcal{V}(H)$ si y sólo si las clases son diferentes y existe una arista en G que tenga un estado en E y otro en F. Evidentemente H no puede tener un ciclo, pues de ser así existiría un camino entre nodos de diferentes clases en G. Por lo tanto H es un árbol. En particular, tiene al menos un estado

que no tiene aristas que salgan de él, el cual es llamado *sumidero*, y al menos un estado que no tiene aristas que entren a él, llamado *fuente*.

Si quitamos los sumideros de la gráfica H, podemos ordenar las clases de comunicación de tal forma que si las clases de comunicación ordenadas son $C_1, C_2, \ldots, C_k$ entonces solo puede existir un camino en H de C_j a C_i si y solo si j es mayor que i. Con este orden podemos hacer que A sea una matriz triangular ordenando los bloques dados por las clases de comunicación.

Para cada $i = 1, \ldots k$, denotemos por G_i a la subgráfica inducida de G con conjunto de vértices C_i, de forma que sus aristas son solo las que tienen estados iniciales y terminales en C_i. Evidentemente estas subgráficas son irreducibles pues definimos las clases de comunicación como los vértices que tienen caminos entre ellos. Llamamos a estas gráficas G_i las *componentes irreducibles* de G. De igual forma, dada una matriz no-negativa A, las submatrices A_i que se obtienen a partir de A al considerar únicamente las entradas que corresponden a los vértices de las componentes irreducibles G_i de $G(A^{\#})$ las llamamos componentes irreducibles de A. Análogamente, X_{G_i} son las las componentes irreducibles de X_G.

Por ejemplo tomemos la siguiente matriz:

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{pmatrix}. \tag{1.4}$$

Al enumerar cada columna y línea del 1 al 6, esta matriz representa la cantidad de aristas que hay entre dos nodos, entonces la gráfica $\mathcal{G}$ relacionada con la matriz es la figura 1.1

Como podemos ver tenemos cuatro clases de comunicación $C_1 = 1, C_2 = 2, C_3 = 3, 5$ y $C_4 = 4, 6$, con las clases como nodos podemos dibujar la gráfica $\mathcal{H}$ que aparece en la figura 1.2

Con $\mathcal{H}$ obtenemos la matriz de aristas ordenando los nodos de acuerdo a las clases de

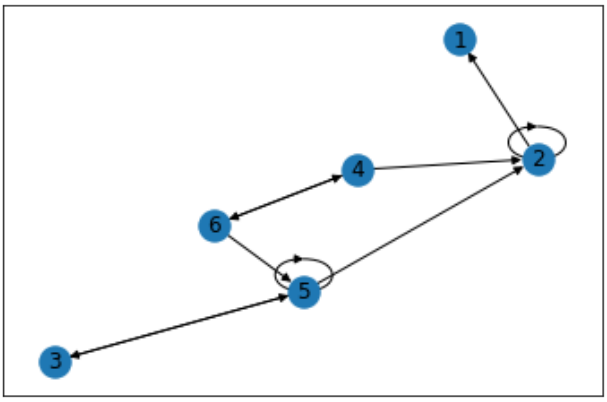

Figura 1.1: Grafica de $\mathcal{G}$.

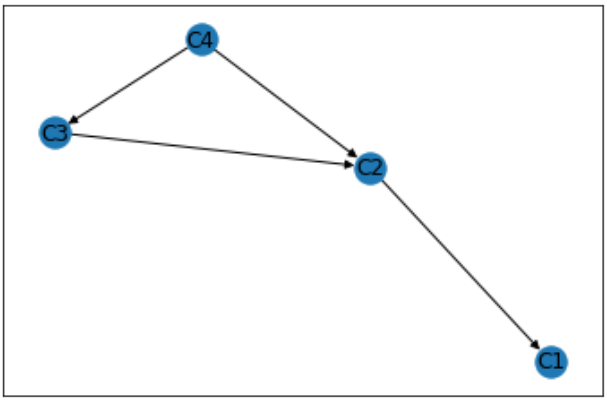

Figura 1.2: Grafica de $\mathcal{H}$.

equivalencia lo cual sería $\{C_1, C_2, C_3, C_4\} = \{1, 2, 3, 5, 4, 6\}$

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{pmatrix}. \tag{1.5}$$

De esta manera obtenemos una matriz casi triangular.

Definición 1.25. Sea A una matriz no-negativa con componentes irreducibles $A_1, \ldots, A_k$. El *eigenvalor de Perron* λ_A de A está dado por $\lambda_A = \text{máx}_{1 \leq i \leq k} \lambda_{A_i}$.

Lema 1.5. *Para una matriz arbitraria no-negativa A, el eigenvalor de Perron es el eigenvalor más grande de A.*

Teorema 1.14. *Sea G una gráfica con matriz de adyacencia A entonces $h(X_G) = \log \lambda_A$.*

Este teorema nos muestra una forma fácil de calcular la entropía de shifts sóficos, pues sabemos que si $\mathcal{G} = (G, L)$ es una presentación derecha resolvente de $X = X_{\mathcal{G}}$, entonces $h(X_{\mathcal{G}}) = h(X_G) = \log \lambda_{A(G)}$. Como vimos, el eigenvalor de Perron es un eigenvalor de una componente irreducible de A y al ser el mayor eigenvalor positivo de una matriz irreducible, implica que la entropía del espacio shift asociado es $\log \lambda_A$, y por lo tanto cada espacio shift sófico contiene un espacio shift sófico irreducible cuya entropía es la entropía del sistema original.

Calcular la entropía de un espacio shift X puede ser bastante complicado. Una forma para calcularla es aproximar X por espacios shift de tipo finito. Sea $\mathcal{F} = \{w_1, w_2, \ldots\}$ una colección infinita de palabras prohibidas tal que $X = X_{\mathcal{F}}$ y definimos como $\mathcal{F}_n = \{w_1, w_2, \ldots, w_n\}$, entonces podemos tomar los espacios shift de tipo finito $X_{\mathcal{F}_n}$ donde $X_{\mathcal{F}_j} \subseteq X_{\mathcal{F}_i}$ si j es mayor que i, entonces $\cap_{n=1}^{\infty} X_{\mathcal{F}_n} = X_{\cup_{n=1}^{\infty} \mathcal{F}_n} = X_{\mathcal{F}}$. En virtud del siguiente resultado, la idea es entonces aplicar el teorema anterior a cada espacio shift de aristas asociado a cada espacio shift de tipo finito en estas sucesiones anidadas de espacios shift.

Proposición 1.22. *Sean $X_1 \supseteq X_2 \supseteq X_3 \supseteq \ldots$ espacios shift de los cuales su intersección es el espacio shift X. Entonces $h(X_n) \to h(X)$ si $n \to \infty$.*

Con esta proposición podemos ver que con la construcción anterior se puede aproximar la entropía de X. El problema de calcular la entropía de esta manera es que el tamaño de la matriz de adyacencia de los espacios shift de tipo finito crece mucho, haciendo que los cálculos para encontrar el eigenvalor de Perron sean muy tardados. El otro problema es que la proposición (ni su demostración, la cual omitimos) no nos dice qué tan grande tiene que ser n para que la aproximación de la entropía tenga un error pequeño.

Teorema 1.15. *Sea A una matriz irreducible y sea B otra matriz tal que $0 \leq B \leq A$. Si existe un par de índices i, j tales que $B_{i,j} < A_{i,j}$, entonces $\lambda_B < \lambda_A$.*

Corolario 1.7. *Sea X es un espacio shift sófico e irreducible. Si Y es subshift propio de X, entonces $h(Y) < h(X)$.*

Definición 1.26. Sea A una matriz no-negativa. El *periódo* de un estado I, denotado por $\mathrm{per}(I)$, es el máximo común divisor de los enteros $n \geq 1$ para el cual $(A^n)_{I,I} > 0$. Si no existe un entero que lo cumpla definimos $\mathrm{per}(I) = \infty$. El periodo $\mathrm{per}(A)$ de una matriz no-negativa A es el máximo común divisor de los periodos de sus estados. Una matriz es *aperiódica* si tiene periodo 1. El periodo $\mathrm{per}(G)$ de una gráfica G es el periodo de su matriz de adyacencia.

Lema 1.6. *Si A es irreducible, entonces todos los estados tienen el mismo periodo, por lo que per(A) es el periodo de cualquiera de sus estados.*

Entonces podemos definir el periodo de A directamente como el máximo común divisor de los enteros $n \geq 1$ para los cuales $\mathrm{tr}(A^n) > 0$ (de nueva cuenta, si no existe un entero que satisfaga ésto, entonces definimos $\mathrm{per}(A) = \infty$).

Definición 1.27. Sea X un espacio shift. El periodo $\mathrm{per}(X)$ de X es el máximo común divisor de los enteros $n \geq 1$ para los cuales $p_n(X) > 0$ o es ∞ si no existe el entero (es decir, si X no tiene puntos periódicos) donde $p_n(X)$ son los puntos periódicos de periodo n.

Si X y Y son conjugados, entonces $p_n(X) = p_n(Y)$ para todo n. Entonces el periodo de ambos espacios shift es el mismo. Por lo tanto el periodo es invariante bajo conjugación.

Proposición 1.23. *Si G es una gráfica, entonces* $\mathrm{per}(X_G) = \mathrm{per}(G)$.

Sea $A \neq [0]$ una matriz irreducible con periodo finito $p \geq 1$. Diremos que los estados I, J son *periódicamente equivalentes*, denotados por $I \sim J$, si existe un camino en $G(A)$ de I a J cuya longitud sea divisible entre p. Es fácil ver que $\sim$ es una relación de equivalencia.

Proposición 1.24. *Sea $A \neq [0]$ irreducible con periodo p. Entonces existen exactamente p clases periódicas, las cuales pueden ser ordenadas como $D_0, D_1, D_2, \ldots, D_{p-1}$ de tal forma que cada arista que comience en D_i termina en $D_{(i+1) \text{ mód } p}$.*

Definición 1.28. Una matriz es *primitiva* si es irreducible y aperiódica. Una gráfica es primitiva si su matriz de adyacencia es primitiva.

Teorema 1.16. *Sea A una matriz no-negativa. Los siguientes enunciados son equivalentes:*

1. *A es primitiva.*
2. *$A^N > 0$ para algún $N \geq 1$.*
3. *$A^N > 0$ para todo N suficientemente grande.*

Definición 1.29. Un espacio shift X es *mezclante* si para cualquier par ordenado $u, v \in \mathcal{B}(X)$ hay una N tal que para cada $n \geq N$ existe una palabra $w \in \mathcal{B}_n(x)$ tal que $uwv \in \mathcal{B}(X)$.

Proposición 1.25. 1. *La propiedad de mezclado es invariante bajo conjugación (de hecho, un factor de un espacio shift mezclante es también mezclante).*

2. *Si G es una gráfica esencial, entonces el espacio shift de aristas X_G es mezclante si y solo si G es primitiva.*

3. *Un espacio shift de tipo finito es mezclante si y solo si es conjugado a un shift de aristas X_G donde G sea primitiva.*

4. *Un espacio shift de tipo finito es mezclante si y solo si es irreducible y* $\mathrm{per}(X) = 1$.

Teorema 1.17. *Si A es una matriz primitiva y $\mu \neq \lambda_A$ es un eigenvalor de A, entonces $|\mu| < \lambda_A$.*

Definiremos v_A como el eigenvector de Perron de una matriz cuadrada no-negativa A y w_A como el eigenvector de Perron de la matriz transpuesta A^T. Este vector de Perron w_A también es conocido como el *eigenvector izquierdo de Perron*, y v_A es el *eigenvector derecho de Perron.*

Teorema 1.18. *Sea A una matriz primitiva con eigenvalor de Perron λ. Sean v, w los eigenvectores derecho e izquierdo de Perron respectivamente para A, esto es $v, w > 0$ son tales que $Av = \lambda v$, $wA = \lambda w$, y supongamos que están normalizados, es decir, que $w \cdot v = 1$. Entonces para cada par de estados I y J se cumple que $(A^n)_{I,J} = [(v_I w_J) + \rho_{IJ}(n)]\lambda^n$ donde $\rho_{IJ}(n) \to 0$ si $n \to \infty$.*

Corolario 1.8. 1. *Si A es una matriz primitiva con entradas enteras, entonces los siguientes límites existen y son iguales a* $\log \lambda_A$*:*

   a) $\lim_{n\to\infty} \frac{1}{n} \log(p_n(X_A))$.
   b) $\lim_{n\to\infty} \frac{1}{n} \log(q_n(X_A))$.
   c) $\lim_{n\to\infty} \frac{1}{n} \log(A^n_{I,J})$ *para cada I, J.*

2. *Para un espacio shift sófico mezclante X se cumple que*

$$\lim_{n\to\infty} \frac{\log(p_n(X))}{n} = \lim_{n\to\infty} \frac{\log(q_n(X))}{n} = h(X). \tag{1.6}$$

1.6. Notas adicionales al capítulo

Los espacios shift son realmente importantes, porque, como pudimos ver a lo largo del capítulo, pueden ser relacionados con diferentes tipos de sistemas dinámicos y con gráficas. Los sistemas simbólicos tienen un estudio bastante completo y generalizado, por lo tanto es

muy útil trabajar en ellos y es precisamente por eso que gran parte de nuestro estudio será aplicado en ellos.

Todo este capítulo nos muestra la utilidad de la conjugación en los sistemas simbólicos. Una vez que tenemos dos sistemas conjugados entonces son básicamente lo mismo, por eso es que buscar la forma de saber si dos sistemas son conjugados es un problema muy importante.

En los sistemas simbólicos, gracias a su relación con las gráficas, existen varias maneras de probar la conjugación, pero la entropía sigue siendo uno de los invariantes de conjugación más importantes y con el que se suele trabajar más, así que vale la pena estudiar aún más el concepto fuera de los espacios shift.

Capítulo 2

Sistemas Dinámicos

En el capítulo anterior estudiamos brevemente lo que es un sistema dinámico y nos enfocamos en la dinámica simbólica. Existen muchos sistemas dinámicos y la importancia de la dinámica simbólica radica en el hecho de que muchos sistemas dinámicos se pueden modelar a través de un sistema dinámico simbólico. En este capítulo abordaremos a los sistemas dinámicos de forma más general, en la que los sistemas simbólicos son un caso particular.

Primero estudiaremos qué ocurre cuando construimos un sistema dinámico en un espacio topológico cualquiera. En estos sistemas veremos en especial cómo se define la entropía topológica. Para los sistemas simbólicos, la entropía topológica coincidirá con la entropía que ya definimos en los espacios shift, por lo tanto nos será muy útil como invariante topológico y será importante conocer sus propiedades generales.

Finalmente, estudiaremos la teoría ergódica, la cual es la base de todo el trabajo de especificación. La teoría ergódica estudia las características de un sistema dinámico junto con sus diferentes medidas invariantes asociadas. Para cada medida podremos definir la entropía métrica asociada y estudiaremos su relación con la entropía topológica. Esto nos dará mucha información acerca de los sistemas y sus medidas. Estos resultado son obtenidos de [3], [6], [7] y [8].

2.1. Dinámica topológica

Empezaremos dando unas definiciones básicas de topología.

Definición 2.1. Sea X un espacio topológico, X es *localmente compacto* si cada punto $x \in X$ tiene una vecindad compacta, es decir, existe un conjunto abierto U y un conjunto compacto K, tal que $x \in U \subseteq K$.

Definición 2.2. Sea (X, τ) un espacio topológico, decimos que es *metrizable* si existe una métrica d tal que la topología inducida por d es τ.

Definición 2.3. Sea X un espacio topológico, decimos que es *segundo numerable* si su topología tiene una base numerable.

Definición 2.4. Sea X un espacio topológico, decimos que es un *espacio de Hausdorff* si para cuales quiera x y $y \in X$ existen U y V vecindades alrededor de x y y respectivamente, tal que $U \cap V = \varnothing$.

Un *sistema dinámico topológico* (X, f), es un espacio topológico X con un mapeo continuo $f\colon X \to X$ o un flujo continuo f^t (es decir, una familia de mapeos continuas $f^t\colon X \to X$ indexados por $t \in \mathbb{R}$ tales que $f^t \circ f^s = f^{t+s}$, por lo tanto todas las funciones son invertibles[1]). Supondremos que el espacio topológico X es localmente compacto, metrizable y segundo numerable.

Un conjunto A es f invariante si $f^p(A) \subseteq A$ para todo $p \in \mathbb{Z}$. Si $p \in \mathbb{N}$, entonces es invariante positivo o invariante hacia adelante, y si $p \in \mathbb{Z}^-$, entonces es invariante negativo o invariante hacia atrás.

Dados dos espacios topológicos X y Y, un mapeo continuo $f\colon X \to Y$ es un *homeomorfismo* si es uno a uno (es decir inyectivo) y su función inversa es continua. Dados dos sistemas dinámicos topológicos $f\colon X \to X$ y $g\colon Y \to Y$, una *semiconjugación topológica* de g a f es un mapeo suprayectivo $h\colon Y \to X$ tal que $f \circ h = h \circ g$. Si h cumple ser un homeomorfismo entonces es llamado una *conjugación topológica* y f, g son *topológicamente conjugadas o isomorfas*.

Definición 2.5. Sea $f\colon X \to X$ un sistema dinámico topológico, y sea x un punto en X. Decimos que $y \in X$ es un *punto $\omega-$limite* de x, si existe una sucesión de números naturales $n_k \to \infty$ cuando $k \to \infty$ tal que $f^{n_k}(x) \to y$ y el conjunto $\omega-$límite de x es el conjunto $\omega(x)$

[1] Si no requerimos que las transformaciones sean invertibles, entonces el índice t varía únicamente sobre $\mathbb{R}^+$. En este capítulo, aunque todos los conceptos que veremos se pueden definir tanto como para sistemas discretos como para flujos, nos enfocaremos únicamente en los primeros.

que contiene todo los puntos $\omega-$límite. Más concretamente,

$$\omega(x) = \bigcap_{n\in\mathbb{N}} \overline{\bigcup_{i\geq n} f^i(x)}. \tag{2.1}$$

Si f es invertible, podemos reformular la definición anterior, pero esta vez usando las funciones inversas f^{-i}, y a este nuevo conjunto denotado como $\alpha(x)$ lo llamamos el *conjunto $\alpha-$límite* de x y los puntos en él son los *puntos $\alpha-$límite de x*. Ambos conjuntos son cerrados y $f-$invariantes, y si x pertenece a su propio conjunto $\omega-$límite, entonces decimos que x es *recurrente* y al conjunto de todos los puntos recurrentes lo denotamos por $\mathcal{R}(f)$.

Un punto $x \in X$ es *no errante* si para cualquier vecindad U de x existe un $n \in \mathbb{N}$ tal que $f^n(U) \cap U \neq \varnothing$. Denotamos como $\mathrm{NW}(f)$, que viene del inglés "no wandering", al conjunto de puntos no errantes de f, el cual es cerrado, f-invariante y contiene a los conjuntos $\omega(x)$ y $\alpha(x)$ para todo $x \in \mathrm{NW}(f)$. Recordando el concepto de órbita en los sistemas simbólicos, podemos extenderla a estos sistemas más generales. En este caso la órbita de un punto $x \in X$ es $\mathcal{O}(x) = \bigcup_{n\in\mathbb{Z}}\{f^n(x)\}$ para un mapeo invertible f, y $\mathcal{O}^+(x) = \bigcup_{n\in\mathbb{N}}\{f^n(x)\}$ para un mapeo no invertible.

Definición 2.6. Sea X un espacio topológico compacto y $f\colon X \to X$ un mapeo continuo. Un subconjunto $Y \subseteq X$ cerrado, no vacío y f positivo invariante es un *conjunto minimal* para f si Y no contiene ningún subconjunto propio, cerrado, no vacío y f positivo invariante. Si X es un un conjunto minimal decimos que f es *minimal*.

Proposición 2.1. *Sea $f\colon X \to X$ un sistema dinámico topológico. Si X es compacto, entonces X contiene un conjunto minimal para f.*

Un conjunto compacto X es minimal si y solo si las órbitas positivas de cada punto en X son densas en X. En un espacio topológico compacto, cada punto en un conjunto minimal es recurrente, por lo que la existencia de conjuntos minimales nos da la existencia de puntos recurrentes.

Definición 2.7. Un subconjunto $A \subseteq \mathbb{N}$ es *relativamente denso o sindético*, si existe $k > 0$ tal que $\{n, n+1, n+2, \ldots, n+k\} \cap A \neq \varnothing$ para cualquier n. Decimos que un punto $x \in X$ es *casi periódico* si para cualquier vecindad U de x, el conjunto $\{i \in \mathbb{N} : f^i(x) \in U\}$ es relativamente denso.

Proposición 2.2. *Si X es un espacio compacto de Hausdorff y $f\colon X \to X$ es continua, entonces $\overline{\mathcal{O}^+(x)}$ es minimal para f si y solo si x es casi periódica.*

Definición 2.8. Un sistema dinámico topológico $f\colon X \to X$ es *topológicamente transitivo* si hay un punto $x \in X$ cuya órbita positiva es densa en X.

Proposición 2.3. *Sea $f\colon X \to X$ un mapeo continuo de un espacio X localmente compacto y Hausdorff. Supongamos que para cualesquiera dos conjuntos abiertos no vacíos $U, V \subseteq X$ existe $n \in \mathbb{N}$ tal que $f^n(U) \cap V \neq \varnothing$. Entonces f es topológicamente transitiva.*

Proposición 2.4. *Sea $f\colon X \to X$ un homeomorfismo de un espacio métrico compacto y supongamos que X no tiene puntos aislados. Si existe una órbita (completa) densa $\mathcal{O}(x)$ de algún punto $x \in X$, entonces existe un órbita densa positiva $\mathcal{O}^+(y)$ de algún punto $y \in X$.*

Definición 2.9. Un sistema dinámico topológico $f\colon X \to X$ es topológicamente mezclante si para cualesquiera dos conjuntos abiertos no vacíos $U, V \subset X$ existe una natural $N > 0$ tal que $f^n(U) \cap V \neq \varnothing$ para toda $n \geq N$.

Es claro por la Proposición 2.3 que mezclado topológico implica transitividad topológica, pero el inverso no necesariamente ocurre.

Definición 2.10. Un homeomorfismo $f\colon X \to X$ es *expansivo* si existe $\delta > 0$ tal que para cualesquiera dos puntos distintos $x, y \in X$, existe algún $n \in \mathbb{Z}$ tal que $d(f^n(x), f^n(y)) \geq \delta$. Si el mapeo es no invertible y cumple la propiedad con $n \geq 0$, entonces lo llamas un *mapeo positivamente expansivo.*

Proposición 2.5. *Sea f un homeomorfismo de un espacio métrico, compacto e infinito X. Entonces para cada $\epsilon > 0$, existen dos puntos distinto $x_0, y_0 \in X$ tales que $d(f^n(x_0), f^n(y_0)) \leq \epsilon$ para todo $n \in \mathbb{N}$.*

Corolario 2.1. *Sea f un homeomorfismo expansivo de un espacio métrico, compacto e infinito X. Entonces existen $x_0, y_0 \in X$ tales que $d(f^n(x_0), f^n(y_0)) \to 0$ si $n \to \infty$.*

La entropía topológica es uno de los conceptos más importantes en el estudio de la dinámica topológica. La entropía nos da, asintóticamente, la taza de crecimiento exponencial del número de los segmentos de órbitas esencialmente diferentes de longitud n, es decir, ve que tan rápido crece la cantidad de segmentos de órbitas en el sistema a medida que aumentamos el tamaño de estos, por lo tanto mientras mas pedazos de órbitas existan mas complejo es el sistema.

Sea (X, d) un espacio métrico compacto y $f\colon X \to X$ un mapeo continuo. Empecemos definiendo, para cada $n \in \mathbb{N}$, la función $d_n(x, y) = \text{máx}_{0 \leq k \leq n-1} d(f^k(x), f^k(y))$, la cual mide la máxima distancia entre las primeras n iteraciones de x y y bajo f. Cada d_n es una métrica en X, y son equivalentes pues inducen la misma topología en X.

Fijemos un $\epsilon > 0$. Un subconjunto $A \subseteq X$ es (n, ϵ)-*extendido* si, para cada $x \in X$, existe un $y \in A$ tal que $d_n(x, y) < \epsilon$. Por compacidad, existen conjuntos (n, ϵ)-extendidos

finitos. Definimos entonces a $\operatorname{span}(n, \epsilon, f)$ como la cardinalidad mínima de los conjuntos (n, ϵ)-extendidos.

Decimos que un subconjunto $A \subseteq X$ es (n, ϵ)-*separado* si cualesquiera dos puntos distintos en A están al menos ϵ separados en la métrica d_n. Todos los conjuntos (n, ϵ)-separados son finitos, y por lo tanto, de nuevo por compacidad, la cardinalidad de los conjuntos (n, ϵ)-separados está acotada, por lo que podemos también definir a $\operatorname{sep}(n, \epsilon, f)$ como la cardinalidad máxima de un conjunto (n, ϵ)-separado.

Definimos a $\operatorname{cov}(n, \epsilon, f)$ como la mínima cardinalidad de la cobertura de X por conjuntos de d_n-diámetro menor que ϵ (por compacidad, $\operatorname{cov}(n, \epsilon, f)$ es finito).

Las tres cantidades definidas en los párrafos anteriores están relacionadas con los segmentos de órbitas de tamaño n, pues todas dependen de la medida d_n y son distinguibles en una escala pequeña ϵ. Más concretamente, tenemos el siguiente resultado:

Lema 2.1. $\operatorname{cov}(n, 2\epsilon, f) \leq \operatorname{span}(n, \epsilon, f) \leq \operatorname{sep}(n, \epsilon, f) \leq \operatorname{cov}(n, \epsilon, f)$.

Podemos definir

$$h_\epsilon(f) = \limsup_{n\to\infty} \frac{1}{n} \log(\operatorname{cov}(n, \epsilon, f)). \tag{2.2}$$

Sabemos que $\operatorname{cov}(n, \epsilon, f)$ crece a medida de que ϵ decrece, pues mientras ϵ sea más pequeña, más elementos se necesitan para cubrir al espacio, por lo que $h_\epsilon(f)$ también crece, así que el límite existe.

Definición 2.11. Al límite

$$h_{\text{top}}(f) = h(f) = \lim_{\epsilon\to 0^+} h_\epsilon(f) \tag{2.3}$$

lo llamamos *entropía topológica* de f.

Esta definición también se puede definir usando $\operatorname{span}(n, \epsilon, f)$ y $\operatorname{sep}(n, \epsilon, f)$.

Lema 2.2. *El límite* $\lim_{n\to\infty} \frac{1}{n} \log(\operatorname{cov}(n, \epsilon, f)) = h(f)$ *existe y es finito.*

Proposición 2.6. *La entropía topológica de un mapeo continuo $f\colon X \to X$ no depende de la elección de la métrica particular que genera la topología de X.*

Corolario 2.2. *La entropía topológica es un invariante bajo conjugación topológica.*

Estos últimos tres resultados son muy importantes. El primero nos dice que el límite existe, por lo que siempre podemos obtener la entropía topológica de un mapeo. El segundo nos dice que no depende de la métrica, por lo que el mapeo siempre tendrá la misma entropía y por lo tanto funciona muy bien como invariante topológico.

Proposición 2.7. *Sea $f\colon X \to X$ un mapeo continuo sobre un espacio métrico compacto X.*

1. *$h(f^m) = m \cdot h(f)$ para $m \in \mathbb{N}$.*
2. *Si f es invertible, entonces $h(f^{-1}) = h(f)$. Por lo tanto $h(f^m) = |m| \cdot h(f)$ para todo $m \in \mathbb{Z}$.*
3. *Si A_i, con $i = 1, \ldots, k$, son subconjuntos cerrados f-invariantes de X cuya unión es X, entonces $h(f) = \operatorname{máx}_{1 \leq 1 \leq k} h(f|A_i)$. En particular, si A es un subconjunto cerrado invariante de X, entonces $h(f|A) \leq h(f)$.*

Proposición 2.8. *Sean (X, d^X) y (Y, d^Y) dos espacios métricos compactos, y sean $f\colon X \to X, g\colon Y \to Y$ dos mapeos continuos. Entonces:*

1. *$h(f \times g) = h(f) + h(g)$.*
2. *Si g es un factor de f, entonces $h(f) \geq h(g)$.*

Proposición 2.9. *Sea (X, d) un espacio métrico compacto, y sea $f\colon X \to X$ un homeomorfismo expansivo con constante de expansión δ. Entonces $h(f) = h_\epsilon(f)$ para cualquier $\epsilon < \delta$.*

Si $f\colon X \to X$ es un homeomorfismo en un espacio compacto de Hausdorff, decimos que los puntos $x, y \in X$ son *proximales* si la cerradura $\overline{\mathcal{O}((x,y))}$ de la órbita de (x, y) bajo el mapeo $f \times f$ interseca la diagonal $\Delta = \{(z, z) \in X \times X : z \in X\}$. Si dos puntos no son proximales decimos que son *distales*, y un homeomorfismo $f\colon X \to X$ es *distal* si cualquier par de puntos distintos $x, y \in X$ son distales.

Si (X, d) es un espacio métrico compacto, entonces $(x, y) \in X$ son proximales si existe una sucesión $n_k \in \mathbb{Z}$ tal que $d(f^{n_k}(x), f^{n_k}(y)) \to 0$ si $k \to \infty$, y dos puntos en X son distales si existe $\epsilon > 0$ tal que $d(f^n(x), f^n(y)) > \epsilon$ para todo $n \in \mathbb{Z}$.

Un homeomorfismo $f\colon X \to X$ en un espacio métrico compacto (X, d) es *equicontinuo* si la familia de iteradas de f es una familia equicontinua, es decir, que para cualquier $\epsilon > 0$ existe $\delta > 0$ tal que $d(x, y) < \delta$ implica que $d(f^n(x), f^n(y)) < \epsilon$ para todo $n \in \mathbb{Z}$. Las isometrías preservan distancias y por lo tanto son equicontinuas.

Proposición 2.10. *Un homeomrfismo expansivo de un espacio métrico compacto infinito no es distal.*

Proposición 2.11. *Los homeomorfismos equicontinuos son distales.*

2.2. Shifts como sistemas dinámicos

Definición 2.12. Un espacio métrico (M, ρ) consiste en un conjunto M junto con una métrica o función distancia $\rho : M \times M \to [0, \infty)$ tal que para todos los puntos $x, y, z \in M$, ocurre que

1. $\rho(x, y) = 0$ si y solo si $x = y$,
2. $\rho(x, y) = \rho(y, x)$, y
3. $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$.

Para la dinámica simbólica podemos tomar como espacios métricos a los espacios shift. Teniendo en mente la topología producto, la medida debe mostrar la idea de que los puntos están cerca cuando bloques centrales grandes sean iguales, entonces sea $M = \mathcal{A}^{\mathbb{Z}}$ y sea $\rho(x, y) = 2^{-k}$ donde k es el máximo número tal que $x_{[-k,k]} = y_{[-k,k]}$. Notemos que si $x = y$ entonces se tendría que $\rho(x, y) = 2^{-\infty}$ y entonces definimos la distancia como cero. Evidentemente ρ cumple la segunda condición de la definición de medida pues ya sea para $\rho(x, y)$ o $\rho(y, x)$, el número máximo k del tamaño del bloque central coincidente es el mismo. Ahora sea $\rho(x, y) = 2^{-k}$, lo que quiere decir que $x_{[-k,k]} = y_{[-k,k]}$, y sea $\rho(y, z) = 2^{-n}$ esto es $y_{[-n,n]} = z_{[-n,n]}$. Definamos $m = \text{mín}\{k, n\}$ para entonces tener $x_{[-m,m]} = z_{[-m,m]}$ y por lo tanto obtener $\rho(x, z) \leq 2^{-m} \leq 2^{-k} + 2^{-n} = \rho(x, y) + \rho(y, z)$ y así concluir que ρ cumple la ultima propiedad de métrica.

Definición 2.13. Sea (M, ρ) un espacio métrico y $E \subseteq M$. El *diámetro* de E es

$$\text{diam}(E) = \sup\{\rho(x, y) : x, y \in E\}.$$

Definición 2.14. Sea M un espacio métrico. Una sucesión $\{x_n\}_{n=1}^{\infty}$ en M *converge* a x si $\rho(x_n, x) \to 0$ cuando $n \to \infty$. En este caso escribimos $x_n \to x$ si $n \to \infty$ o $\lim_{n\to\infty} x_n = x$.

Sea X un espacio shift y sea $x \in X$. Una sucesión $x^{(n)} \in X$ de elementos de X es convergente a x, es decir $x^{(n)} \to x$, exactamente cuando para cada $k \geq 0$ existe una n_k tal que $x^{(n)}_{[-k,k]} = x_{[-k,k]}$ para todo $n \geq n_k$.

Lema 2.3. *Sea M un espacio métrico, y supongamos que $x_n \to x$ y $x_n \to y$ cuando $n \to \infty$ entonces $x = y$, es decir el límite es único.*

Dos métricas ρ_1, ρ_2 se dicen *equivalentes* sobre el mismo espacio métrico si preservan la convergencia, es decir $\rho_1(x_n, x) \to 0$ si y solo si $\rho_2(x_n, x) \to 0$.

Definición 2.15. Una función $\phi : M \to N$ de un espacio métrico a otro es continua si para todo $x_n \to x$ en M, entonces $\phi(x_n) \to \phi(x)$ en N, es decir preserva la convergencia. Si ϕ es continua, uno a uno y sobre entonces tiene una inversa continua y llamamos a ϕ un *homeomorfismo*.

Definición 2.16. Un espacio métrico M es *compacto* si para cada sucesión en M tiene una subsucesión convergente.

Definición 2.17. Sea (M, ρ) un espacio métrico. Para $x \in M$ y $r > 0$, la bola abierta de radio r alrededor de x es el conjunto $B_r(x) = \{y \in M : \rho(x, y) < r\}$. Un subconjunto $U \subseteq M$ es abierto si para cada $x \in U$ existe un $r > 0$ tal que $B_r(x) \subseteq U$. Decimos que un subconjunto $V \subseteq M$ es cerrado si su complemento es abierto.

Proposición 2.12. *Sea M un espacio métrico compacto. Un subconjunto $E \subseteq M$ es compacto si y solo si es cerrado.*

Teorema 2.1. *(Teorema de Bolzano-Weierstrass) Un subconjunto de $\mathbb{R}^n$ es compacto si y solo si es cerrado y acotado.*

Teorema 2.2. *Un subconjunto de $\mathcal{A}^{\mathbb{Z}}$ es un espacio shift si y solo si es shift invariante y compacto.*

Teorema 2.3. *(Teorema de Heine-Borel) Un subconjunto E de un espacio métrico (M, ρ) es compacto si y solo si siempre que E esté contenido en la unión de una colección de subconjuntos abiertos de M, entonces E también está contenido en la unión de una subcolección finita de estos conjuntos abiertos.*

Diremos que x es un *punto interior* de E si existe un $r > 0$ para el cual $B_r(x) \subseteq E$. Denotamos por $\text{int}(E)$ al conjunto de todos los puntos interiores de E. Por lo tanto E es abierto si y solo si todos sus puntos son puntos interiores. Definimos un *punto límite* de E como el punto que es el límite de una sucesión de puntos en E. La *cerradura* de E está definida como la unión de E y todos sus puntos límites. Denotamos la cerradura de E como $\bar{E}$. Un conjunto E es *cerrado* si y solo si contiene todos sus puntos límites, es decir si $E = \bar{E}$. Un conjunto E es *denso* en otro conjunto F si $F \subseteq \bar{E}$.

Proposición 2.13. *Sea M un espacio métrico. Supongamos que $E_1, E_2, \ldots, E_n \subseteq M$ son subconjuntos abiertos y densos de M. Entonces $\cap_{k=1}^{n} E_k$ es abierto y denso en M.*

Teorema 2.4. *(Teorema de la categoría de Baire) Sea M un espacio métrico compacto y supongamos que $E_1, E_2, E_3, \ldots,$ son abiertos y densos en M, entonces $\bigcap_{k=1}^{\infty}(E_k)$ es denso en M.*

Definición 2.18. Un *sistema dinámico discreto* (M, ϕ) consiste de un espacio métrico compacto M junto con un mapeo continuo $\phi\colon M \to M$. Si ϕ es un homeomorfismo llamamos a (M, ϕ) un sistema dinámico *invertible*.

Para un sistema dinámico (M, ϕ) la *órbita* del punto $x \in M$ es el conjunto de iteraciones $\{\phi^n(x)\}_{n\in\mathbb{Z}}$ si ϕ es invertible, si no la definimos como $\{\phi^n(x)\}_{n\geq 0}$. Un punto *periódico* es un punto $x \in M$ tal que $\phi^n(x) = x$ para algún $n > 0$. Para un sistema dinámico invertible un punto es periódico si y solo si órbita $\mathcal{O}_\phi(x)$ es finita, por lo tanto, una órbita finita también se llama órbita *periódica*.

Un homomorfismo $\theta\colon (M, \phi) \to (N, \psi)$ de un sistema dinámico a otro es una función continua $\theta\colon M \to N$ que satisface la propiedad conmutativa $\psi \circ \theta = \theta \circ \phi$, es decir, preserva la dinámica, y en este caso decimos que θ *entrecruza* a ϕ y a ψ.

Definición 2.19. Sea (M, ϕ) y (N, ψ) dos sistemas dinámicos, y sea $\theta\colon (M, \phi) \to (N, \psi)$ un homomorfismo. Decimos que θ es *encaje* (o *incrustación*) si es uno a uno. Un mapeo *factor* es un homomorfismo θ que es suprayectivo. Decimos que θ es una *conjugación* si es un homeomorfismo. Si θ es una conjugación topológica escribimos $\theta\colon (M, \phi) \cong (N, \psi)$. Dos sistemas dinámicos son *topológicamente conjugados* si existe una conjugación topológica entre ellos.

Lema 2.4. *Sea M y N dos espacios métricos compactos. Si $\theta\colon M \to N$ es continua y biyectiva, entonces $\theta^{-1}\colon N \to M$ también es continua.*

Teorema 2.5. *(Teorema de Curtis-Lyndon-Hedlund) Supongamos que (X, σ_X) y (Y, σ_Y) son dos sistemas simbólicos, y sea $\theta\colon X \to Y$ una función (no necesariamente continua), entonces θ es un código de bloque deslizante si y solo si es un homomorfismo.*

Un problema muy importante cuando trabajamos con dos sistemas dinámicos es saber si son topológicamente conjugados. Supongamos que podemos dar a cada sistema una propiedad matemática u objeto matemático llamado *invariante* (o invariante de conjugación), de tal forma que cuando dos sistemas dinámicos sean topológicamente conjugados entonces tengan asignado el mismo valor de este invariante. De esta forma si dos sistemas tienen un invariante diferente entonces no son topológicamente conjugados. En general, un invariante no tiene por qué ser *completo*, es decir, si dos sistemas tienen el mismo invariante no necesariamente son conjugados.

Definición 2.20. Sea (M, ϕ) un sistema dinámico para el cual $p_n(\phi) < \infty$ para todo $n \geq 1$. La *función zeta* $\zeta_\phi(t)$ está definida como

$$\zeta_\phi(t) = \exp\left(\sum_{n=1}^{\infty} \frac{p_n(\phi)}{n} t^n\right) \tag{2.4}$$

donde $\exp(x) = e^x$.

La función zeta es una forma de combinar todos los invariantes relacionados con puntos periódicos que vimos anteriormente en un único objeto.

Sabemos que la expansión de Taylor para la función exponencial es $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ y entonces podemos aplicar esta expansión sobre la función zeta para obtener

$$\zeta_\phi(t) = 1 + p_1(\phi)t + \frac{1}{2}[p_2(\phi) + p_1(\phi)^2]t^2 + \frac{1}{6}[2p_3(\phi) + 3p_1(\phi)p_2(\phi) + p_1(\phi)^3]t^3 + \dots. \tag{2.5}$$

Si asumimos que $\zeta_\phi(t)$ tiene un radio de convergencia mayor a cero, se pueden calcular los valores de $p_n(\phi)$ desde la función zeta. Para eso vemos que $\log(\zeta_\phi(t)) = \sum_{n=1}^{\infty} \frac{p_n(\phi)}{n} t^n$ y ocupando la formula de Taylor obtenemos

$$p_n(\phi) = \frac{1}{(n-1)!} \frac{d^n}{dt^n} \log \zeta_\phi(t)|_{t=0}. \tag{2.6}$$

Por ejemplo, supongamos que M tiene solo un punto y ϕ es la identidad. Entonces $p_n(\phi) = 1$ para todo $n \geq 1$ y por lo tanto $\zeta_\phi(t) = \exp(\sum_{n=1}^{\infty} \frac{t^n}{n})$, pero sabemos que $\sum_{n=1}^{\infty} \frac{t^n}{n} = -\log(1-t)$, así que $\zeta_\phi(t) = \exp(-\log(1-t)) = \frac{1}{1-t}$.

Otro ejemplo, sea $M = X_{[2]}$ y $\phi = \sigma_{[2]}$ el shift completo en dos símbolos. Entonces $p_n(\phi) = 2^n$, así que la función zeta es

$$\zeta_\phi(t) = \exp\left(\sum_{n=1}^{\infty} \frac{2^n}{n} t^n\right) = \exp\left(\sum_{n=1}^{\infty} \frac{(2t)^n}{n}\right) = \exp(-\log(1-2t)) = \frac{1}{1-2t}. \tag{2.7}$$

Otro ejemplo, sea $A = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $M = X_A$ y $\phi = \sigma_A$. Es decir, (M, ϕ) es el espacio shift aureo. Entonces $p_n(\phi) = \text{traza}(A^n) = \lambda^n + \mu^n$ donde los eigenvalores son $\lambda = \frac{1+\sqrt{5}}{2}, \mu = \frac{1-\sqrt{5}}{2}$,

de forma que la función zeta es

$$\zeta_\phi(t) = \exp\left(\sum_{n=1}^{\infty} \frac{\lambda^n + \mu^n}{n} t^2\right) = \exp\left(\sum_{n=1}^{\infty} \frac{(\lambda t)^n}{n} + \sum_{n=1}^{\infty} \frac{(\mu t)^n}{n}\right)$$
$$= \exp(-\log(1-\lambda t) - \log(1-\mu t)) = \frac{1}{(1-\lambda t)(1-\mu t)} = \frac{1}{1-t-t^2}. \quad (2.8)$$

Teorema 2.6. *Sea A una matriz entera no negativa de tamaño $r \times r$, $\chi_A(t)$ es el polinomio característico y σ_A es el mapeo shift asociado. Entonces*

$$\zeta_{\sigma_A}(t) = \frac{1}{t^r \chi_A(t^{-1})} = \frac{1}{\det(I - tA)}. \quad (2.9)$$

Por lo tanto la función zeta de un shift de tipo finito no sólo es racional, si no que además es el reciproco de un polinomio.

Dada una matriz cuadrada A, denotamos por $\mathrm{sp}^\times(A)$ al *espectro lejos del cero* de A, es decir, al multiconjunto de eigenvalores de A diferentes de cero, tomando en cuenta su multiplicidad.

Corolario 2.3. *Sea A una matriz entera no negativa. Entonces cada una de los siguientes enunciados determina los otros tres:*

1. $\zeta_{\sigma_A}(t)$.
2. $\mathrm{sp}^\times(A)$.
3. $\{p_n(\sigma_A)\}_{n\geq 1}$.
4. $\{q_n(\sigma_A)\}_{n\geq 1}$.

2.2.1. La función zeta de los espacios shift sóficos

Sea $\mathcal{G} = (G, L)$ una gráfica etiquetada derecha resolvente, donde $L\colon \mathcal{E} \to \mathcal{A}$, r denota el número de vértices en G. Denotamos por B a la matriz de adyacencia "simbólica" de $\mathcal{G}$ cuyas entradas pertenecen formalmente al grupo integral $\mathbb{Z}[\mathcal{A}]$. Sea A la matriz de adyacencia de la gráfica subyacente G.

Si $F = \{f_1, \ldots, f_j\}$ es un conjunto ordenado, podemos representar una permutación π de F por $(f_{i_1}, \ldots, f_{i_j})$ esto es $\pi(f_1) = f_{i_1}, \ldots, \pi(f_j) = f_{i_j}$. Las permutaciones pueden ser pares o impares dependiendo de la paridad del número de intercambios necesarios para generar la permutación. Definimos el signo de una permutación π, denotado por $\operatorname{sgn}(\pi)$, como 1 si es par y -1 si es impar. El signo de una permutación cíclica en un conjunto de tamaño n es $(-1)^{n+1}$. Cada permutación π puede ser expresada como una composición de permutaciones cíclicas de conjuntos ajenos, y el signo de π es el producto de los signos de las permutaciones cíclicas.

Para cada j con $1 \leq j \leq r = |\mathcal{V}|$, podemos construir la gráfica etiquetada $\mathcal{G}_j$ con el alfabeto $\{\pm a : a \in A\}$ de la siguiente manera:

- El conjunto de vértices de $\mathcal{G}_j$ es el conjunto $\mathcal{V}_j$ de todos los subconjuntos de $\mathcal{V}$ que tienen j elementos.
- Como $\mathcal{G}$ es derecha resolvente, para cada $I \in \mathcal{V}$ existe a lo más una arista etiquetada a que comienza en I, en cuyo caso denotamos por $a(I)$ a su vértice terminal. Sean $I_0 = \{I_1, \ldots, I_j\}, J_0 = \{J_1, \ldots, J_j\}$ dos vértices en $\mathcal{V}_j$. En la gráfica $\mathcal{G}_j$ existirá una arista de I_0 a J_0 si todos los $a(I_i)$ están definidos y $(a(I_1), \ldots, a(I_j))$ es una permutación de J_0. En este caso, si la permutación es par, entonces le asignamos la etiqueta a, y si la permutación es impar, entonces le asignamos la etiqueta $-a$.

Denotemos como B_j la matriz de adyacencia simbólica de $\mathcal{G}_j$. Entonces cada entrada de B_j es una combinación de símbolos en $\mathcal{A}$, positivos y negativos. Ahora transformemos todos los símbolos de $\mathcal{A}$ en 1 y sea A_j la matriz que se obtiene al aplicar esta transformación en B_j y la llamaremos la *j-ésima matriz de subconjuntos signados* de $\mathcal{G}$.

Teorema 2.7. *Sea $\mathcal{G}$ una gráfica etiquetada derecha resolvente con r vértices y sea A_j la j-ésima matriz de subconjuntos signados. Entonces*

$$\zeta_{\sigma_{\mathcal{G}}}(t) = \prod_{j=1}^{r} [\det(I - tA_j)]^{-1^j}. \tag{2.10}$$

Por lo tanto, la función zeta de un espacio shift sófico es una función racional.

Demostración. Nuestro cálculo de $\zeta_{\sigma_{\mathcal{G}}}(t)$ está basado en mostrar que

$$p_n(\sigma_{\mathcal{G}}) = \sum_{j=1}^{r} (-1)^{j+1} tr(A_j^n). \tag{2.11}$$

Consideremos la parte simbólica de la suma anterior y tomemos $\sum_{j=1}^{r}(-1)^{j+1}\mathrm{tr}((B_j^n)$. Sea $u = a_1a_2\ldots a_n \in \mathcal{A}^n$. Como $\mathcal{G}$ es derecha resolvente, para cada $I \in \mathcal{V}$ hay a lo más un camino etiquetado u que comienza en I. Si tal camino existe, entonces denotamos por $u(I)$ al estado terminal de dicho camino. Si $u^\infty \notin X_\mathcal{G}$, entonces no existe ningún subconjunto $E \subseteq \mathcal{V}$ en el cual u actúe como una permutación. En efecto, si $u^\infty \in X_\mathcal{G}$ quiere decir que el camino infinito existe en la gráfica que es derecha resolvente. Si este camino inicia en algún vértice $I \in \mathcal{V}$, terminará en algún vértice terminal $u(I)$, de donde iniciará un nuevo camino que también representa a u y que terminará en algún vértice terminal $u(u(I))$, etc. Como hay una cantidad finita de vértices r, eventualmente se forma un ciclo en la secuencia $\{u^n(I) \ : \ n \geq 0\}$, el cual genera la permutación en el conjunto de vértices que lo conforman.

Por lo tanto, si $u^\infty \notin X_\mathcal{G}$, entonces u no puede aparecer en $\mathrm{tr}(B_j^n)$ pues no crea una permutación en B_j, así que no aparece en la expresión simbólica anterior, por lo tanto

$$\sum_{j=1}^{r}(-1)^{j+1}\mathrm{tr}((B_j^n) = \sum_{u\in\mathcal{A}^n;u^\infty\in X_\mathcal{G}} c_u \cdot u \tag{2.12}$$

donde c_u son coeficientes en los enteros. Supongamos que $u^\infty \in X_\mathcal{G}$. Ya vimos que debe de haber al menos un subconjunto de $\mathcal{V}$ en el cual u actúe como una permutación. Si dos subconjuntos tienen esta propiedad entonces la tiene la unión, por lo que podemos encontrar el subconjunto mas grande en el que u sea una permutación. Antes de seguir, enunciaremos el siguiente lema.

Lema 2.5. *Sea π una permutación de un conjunto finito F y $\mathcal{C} = \{E \subseteq F : E \neq \varnothing, \pi(E) = E\}$. Entonces*

$$\sum_{E\in\mathcal{C}}(-1)^{|E|+1}\mathrm{sgn}(\pi|_E) = 1. \tag{2.13}$$

Demostración. Sabemos que π descompone al conjunto F en ciclos disjuntos $C_1, C_2, ...C_n$ de forma que $\pi|_{C_k}$ es una permutación cíclica y por lo tanto $\mathrm{sgn}(\pi|_{C_k}) = (-1)^{1+|C_k|}$. Los conjuntos $E \subseteq F$ para los cuales $\pi(E) = E$ son todas las posibles uniones de los ciclos C_k, por lo tanto

$$\sum_{E\in\mathcal{C}}(-1)^{|E|+1}\mathrm{sgn}(\pi|_E) = \sum_{\varnothing\neq P\subseteq\{1,\ldots,n\}}(-1)^{|\cup_{k\in P}C_k|+1}\mathrm{sgn}(\pi|_{\cup_{k\in P}C_k}). \tag{2.14}$$

Ahora como $|\cup_{k\in P} C_k| = \sum_{k\in P}|C_k|$ y $\mathrm{sgn}(\pi|_{\cup_{k\in P}C_k}) = \prod_{k\in P}(-1)^{1+|C_k|} = (-1)^{|P|+\sum_{k\in P}|C_k|}$ pues al tomar el signo de todas las combinaciones, podemos separar en el signo de cada

combinación, entonces obtenemos

$$\begin{aligned}\sum_{\varnothing\neq P\subseteq\{1,\dots,n\}}(-1)^{|\cup_{k\in P}C_k|+1}\mathrm{sgn}(\pi|_{\cup_{k\in P}C_k}) &= \sum_{\varnothing\neq P\subseteq\{1,\dots,n\}}(-1)^{\sum_{k\in P}|C_k|+1}\prod_{k\in P}(-1)^{1+|C_k|}\\ &= \sum_{\varnothing\neq P\subseteq\{1,\dots,n\}}(-1)^{2\sum_{k\in P}|C_k|+1+|P|}\end{aligned}. \tag{2.15}$$

Utilizando métodos de combinatoria podemos representar la anterior suma como

$$\sum_{j=1}^{n}(-1)^{j+1}\binom{n}{j} = 1-(1-1)^n = 1. \tag{2.16}$$

□

Entonces supongamos que $u\in\mathcal{A}^n$ y $u^\infty\in X_{\mathcal{G}}$ y sea F el subconjunto más grande de $\mathcal{V}$ en el cual u actúa como una permutación. Entonces, por la definición de B_j, las aristas tienen un signo y por lo tanto c_u es la suma sobre todos los posibles subconjuntos donde u actúa como permutación tomando en cuenta sus signos, pero sabemos que el signo de una permutación cíclica de tamaño n es $(-1)^{n+1}$ y aplicando el lema obtenemos

$$c_u = \sum_{\varnothing\neq E\subseteq F;u(E)=E}(-1)^{|E|+1}\mathrm{sgn}(u|_E) = 1 \tag{2.17}$$

y por lo tanto

$$\sum_{j=1}^{r}(-1)^{j+1}\mathrm{tr}((B_j^n) = \sum_{u\in A^n;u^\infty\in X_{\mathcal{G}}} u. \tag{2.18}$$

Si convertimos todos los símbolos de $\mathcal{A}$ en 1 obtenemos que B_j^n se convierte en A_j^n y los elementos $u^\infty\in X_{\mathcal{G}}$ son los puntos periódicos de periodo n, así que obtenemos la ecuación (2.11). Luego entonces, sabemos que

$$\exp\left(\sum_{n=1}^{\infty}\frac{\mathrm{tr}(A^n)}{n}t^n\right) = [\det(I-tA)]^{-1} \tag{2.19}$$

y por lo tanto

$$\begin{aligned}\zeta_{\sigma_{\mathcal{G}}}(t) &= \exp\left(\sum_{n=1}^{\infty}\frac{1}{n}\left(\sum_{j=1}^{r}(-1)^{j+1}\mathrm{tr}(A_j^n)\right)t^n\right)\\ &= \prod_{j=1}^{r}\left[\exp\left(\sum_{n=1}^{\infty}\frac{\mathrm{tr}(A_j^n)}{n}t^n\right)\right]^{(-1)^{j+1}} = \prod_{j=1}^{r}[\det(I-tA_j)]^{(-1)^j}.\end{aligned} \tag{2.20}$$

□

2.2.2. Representaciones simbólicas de sistemas dinámicos

Supongamos que queremos estudiar un sistema dinámico (M, ϕ) donde ϕ es invertible y por lo tanto tiene todas sus iteradas tanto positivas como negativas. Para tratar de describir las órbitas $\{\phi^n(y) : n \in \mathbb{Z}\}$ de puntos $y \in M$ usamos una descripción aproximada, dividimos M en un número finito de partes $E_0, E_1, \ldots, E_{r-1}$. Podemos "rastrear" la órbita de un punto y observando únicamente en cuál de estas particiones pertenece $\phi^n(y)$. Esto nos lleva al punto correspondiente $x = \ldots x_{-1}x_0x_1 \ldots \in \{0, 1, \ldots, r-1\}^{\mathbb{Z}} = X_{[r]}$ definido por $\phi^n(y) \in E_{x_n}$ para $n \in \mathbb{Z}$. Por lo que para $y \in M$ obtenemos un punto simbólico x en el r-shift completo con la información del rastreo de la órbita al nivel de la partición dada, aplicando ϕ y todas sus iteraciones, y por lo tanto la imagen de $\phi(y)$ corresponde a $\sigma(x)$. Si el mapeo ϕ no es invertible, entonces solo nos interesan la iteraciones positivas y solo tomamos un lado del r-shift completo $X^+_{[r]}$. Lo anterior determina una asignación $y \mapsto x$, sin embargo la imagen de esta asignación no es necesariamente es un espacio shift, lo cual nos obliga a tener que ser más cuidadosos, pero más adelante, veamos primero unos ejemplos.

Ejemplo 2.1. Sea M un espacio shift sobre $\mathcal{A} = \{0, 1, \ldots, r-1\}$, $\phi = \sigma_M$ el mapeo shift en M y $E_j = \{y \in M : y_0 = j\}$ para algún $0 \leq j \leq r-1$, entonces la representación simbólica de $y \in M$ usando esta subdivisión de M es la sucesión y. Ahora sea $\mathcal{B}_2(M)$ el conjunto de 2-bloques permitidos en M, para $ab \in \mathcal{B}_2(M)$ sea $E_{ab} = \{y \in M : y_0y_1 = ab\}$ entonces la representación simbólica es la representación de bloques de altura dos, $M^{[2]}$ (ver Definición 1.10).

Ejemplo 2.2. Sea $M = \mathbb{T}$ el círculo que se forma al "anudar" el intervalo $[0, 1]$ bajo la identificación $[0, 1]/0 \sim 1 \approx \mathbb{R} \mod 1$, por lo tanto M es el intervalo medio abierto $[0, 1)$. La dinámica la determina la transformación $\phi(y) = 10y \mod 1$. Dividamos M en 10 subintervalos $E_j = [j/10, (j+1)/10]$ para $j = 0, 1, \ldots, 9$. Entonces el alfabeto que podemos definir es $\mathcal{A} = \{0, 1, .., 9\}$. Para $y \in M$ y $n \geq 0$ definimos un símbolo $x_n \in \mathcal{A}$ por $\phi^n(y) \in E_{x_n}$, y el punto resultante $x = .x_0x_1x_2 \ldots$ se encuentra en el espacio shift $X^+_{[10]}$ donde y corresponde al sucesión de dígitos en la expansión decimal $y = .y_0y_1y_2 \ldots = x_0x_1x_2 \ldots$.

Definición 2.21. Una *partición topológica* de un espacio métrico M es un colección finita $\mathcal{P} = \{P_0, P_1, \ldots, P_r - 1\}$ de conjuntos abiertos ajenos cuyas cerraduras $\bar{P}_j$ cumplen que $M = \bar{P}_0 \cup \ldots \cup \bar{P}_{r-1}$.

Supongamos que (M, ϕ) es un sistema dinámico, $\mathcal{P} = \{P_0, P_1, \ldots, P_{r-1}\}$ es una partición topológica de M y sea $\mathcal{A} = \{0, 1, \ldots, r-1\}$. Diremos que una palabra $w = a_1a_2 \ldots a_n$ es *permitida* para $(\mathcal{P}, \phi)$ si se cumple que $\cap_{j=1}^n \phi^{-j}(P_{a_j}) \neq \varnothing$ y definimos como $\mathcal{L}_{\mathcal{P},\phi}$ la colección de palabras permitidas para $(\mathcal{P}, \phi)$. El conjunto de palabras permitidas satisface las propiedades de factor y extensión (ver Proposición 1.1), es decir, es el lenguaje de un espacio shift. Por lo tanto existe un único espacio shift $X_{\mathcal{P},\phi}$ cuyo lenguaje es $\mathcal{L}_{\mathcal{P},\phi}$. Si (M, ϕ)

no es invertible, entonces consideramos únicamente el espacio shift de un lado $X^+_{\mathcal{P},\phi}$.

Consideremos el caso de un sistema dinámico invertible (M, ϕ) y sea $\mathcal{P} = \{P_0, P_1, \ldots, P_{r-1}\}$ un partición topológica de M. Para cada $x \in X_{\mathcal{P},\phi}$ y $n \geq 0$, existe uno conjunto correspondiente no vacío abierto definido por

$$D_n(x) = \bigcap_{k=-n}^{n} \phi^{-k}(P_{x_k}) \subseteq M. \tag{2.21}$$

Las cerraduras de $\overline{D_n(x)}$ son compactas y decrecen con n por lo que $\overline{D_0(x)} \supseteq \overline{D_1} \supseteq \ldots$ así que la intersección no es vacía. Para que los puntos en $X_{\mathcal{P},\phi}$ correspondan a los puntos en M la intersección debe contener solamente un punto.

Definición 2.22. Sea (M, ϕ) un sistema dinámico invertible, una partición topológica $\mathcal{P} = \{P_0, P_1, \ldots, P_{r-1}\}$ de M da una *representación simbólica* de (M, ϕ) si para cada $x \in X_{\mathcal{P},\phi}$ la intersección $\cap_{n=0}^{\infty}\overline{D_n(x)}$ consiste exactamente de un punto. Llamamos a $\mathcal{P}$ una *partición de Markov* para (M, ϕ) si $\mathcal{P}$ da una representación simbólica de (M, ϕ) y además $X_{\mathcal{P},\phi}$ es un espacio shift de tipo finito.

Definición 2.23. Sea (M, ϕ) un sistema dinámico no necesariamente invertible. Una partición topológica $\mathcal{P} = \{P_0, P_1, \ldots, P_{r-1}\}$ de M da una *representación simbólica de un lado* de (M, ϕ) si para cada $x \in X^+_{\mathcal{P},\phi}$ la intersección $\cap_{n=0}^{\infty}\overline{D^+_n(x)}$ consiste en exactamente un punto. Llamamos a $\mathcal{P}$ una *partición de Markov de un lado* para (M, ϕ) si $\mathcal{P}$ da una representación simbólica de (M, ϕ) además $X^+_{\mathcal{P},\phi}$ es un espacio shift de tipo finito.

Supongamos que $\mathcal{P}$ da una representación simbólica de un sistema dinámico invertible (M, ϕ), entonces existe un mapeo bien definido π desde $X = X_{\mathcal{P},\phi}$ a M el cual mapea el punto $x = \ldots x_{-1}x_0x_1 \ldots$ al único punto $\pi(x)$ en la intersección $\cap_{n=0}^{\infty}\overline{D_n(x)}$. Diremos que x es la *representación simbólica* de $\pi(x)$.

Proposición 2.14. *Sea $\mathcal{P}$ una representación simbólica del sistema dinámico invertible (M, ϕ) y sea $\pi : X_{\mathcal{P},\phi} \to M$ el mapeo definido arriba. Entonces π es un mapeo factor de $(X_{\mathcal{P},\phi}, \sigma)$ a (M, ϕ).*

Proposición 2.15. *Supongamos que $\mathcal{P}$ da una representación simbólica del sistema dinámico invertible (M, ϕ). Para cada una de las siguiente propiedades, si $(X_{\mathcal{P},\phi}, \sigma)$ tiene la propiedad, entonces también la tiene (M, ϕ).*

1. *Transitividad topológica.*
2. *Mezcla topológica.*
3. *El conjunto de puntos periódicos es denso.*

2.3. Teoría Ergódica

Sea X un conjunto y tomemos una colección de subconjuntos no vacíos $\mathfrak{A}$. Diremos que $\mathfrak{A}$ es una *σ-álgebra* si es cerrada bajo complementos, uniones contables e intersecciones contables. Una *medida* μ en $\mathfrak{A}$ es una función no negativa $\mu\colon \mathfrak{A} \to \mathbb{R}^+$ que es $\sigma-$aditiva, es decir $\mu(\cup_i A_i) = \sum_i \mu(A_i)$ para cualquier colección contable de conjuntos ajenos $A_i \in \mathfrak{A}$. Un conjunto de medida 0 es llamado un *conjunto nulo*, por lo que si un conjunto tiene un complemento que es conjunto nulo, entonces se dice que tiene medida completa.

Una σ-álgebra es *completa* si contiene cada subconjunto de todos los conjuntos nulos, esto relativo a la medida μ. Para una σ-álgebra $\mathfrak{A}$ podemos obtener la completación $\overline{\mathfrak{A}}$, que es la σ-álgebra mas chica que contiene a $\mathfrak{A}$ y es completa.

Definición 2.24. Un *espacio de medida* $(X, \mathfrak{A}, \mu)$ es una triada donde X es un conjunto, $\mathfrak{A}$ es una σ-álgebra de subconjuntos de X y μ es una medida σ aditiva.

Asumiremos siempre que $\mathfrak{A}$ es completa y que μ es σ finita, es decir que X es la unión contable de subconjuntos de medida finita, los elementos de $\mathfrak{A}$ son llamados *conjuntos medibles*.

Definición 2.25. Si $\mu(X) = 1$, entonces $(X, \mathfrak{A}, \mu)$ es llamado un *espacio de probabilidad* y μ es una *medida de probabilidad*.

Sean $(X, \mathfrak{A}, \mu)$ y $(Y, \mathfrak{B}, \nu)$ dos espacios de medida, un mapeo $T\colon X \to Y$ es llamado *medible* si la preimagen de cualquier conjunto medible es medible. Un mapeo medible T es *no singular* si la preimagen de cualquier conjunto de medida 0 tiene medida 0. Decimos que T *preserva la medida* si $\mu(T^{-1}(B)) = \nu(B)$ para cada $B \in$ B. Le diremos *transformación no singular* a un mapeo no singular de un espacio de medida a sí mismo. Si la transformación T preserva la medida μ, entonces la medida μ es $T-$*invariante*.

Si T es una transformación medible e invertible y su inversa es medible y no singular, entonces sus iteradas T^n forman un grupo de transformaciones medibles a través de la composición como operación. Dos espacios medibles $(X, \mathfrak{A}, \mu)$ y $(Y, \mathfrak{B}, \nu)$ son *isomorfos* si existe un subconjunto X' de medida completa en X, un subconjunto Y' de medida completa en Y y una biyección invertible $T : X' \to Y'$ tal que T, T^{-1} son medibles y preservan la medida con respecto a $(\mathfrak{A}, \mu)$ y $(\mathfrak{B}, \nu)$. Un isomorfismo de un espacio medible a el mismo se denomina *automorfismo*.

Sea X un espacio topológico, la σ-álgebra que contiene todos los subconjuntos abiertos de X es llamada la *σ-álgebra de Borel* de X. Si $\mathfrak{A}$ es la σ-álgebra de Borel, entonces la

medida μ sobre $\mathfrak{A}$ es una medida de Borel, si la medida de cada subconjunto compacto es finita. La medida de Borel es *regular*, ya que la medida de cualquier conjunto es el ínfimo de las medida de los conjuntos abiertos que lo contienen, y el supremo de las medidas de los conjunto compactos contenidos en él.

Un conjunto tiene medida *completa* si su complemento tiene medida 0. Si una propiedad se cumple para un subconjunto de medida completa en X, entonces decimos que es *esencial*, otra forma de denotarlo es que se cumple (mod 0) en X o que se cumple para casi todo-μ en X. Si $(X, \mathfrak{A}, \mu)$ es un espacio de medida, entonces dos funciones medibles son *equivalentes* si coinciden con un conjunto de medida completa.

Teorema 2.8. *(Teorema de recurrencia de Poincaré) Sea T una transformación que preserva la medida de un espacio de probabilidad $(X, \mathfrak{A}, \mu)$. Si A es un conjunto medible entonces para casi todo $x \in A$ existe algún $n \in \mathbb{N}$ tal que $T^n(x) \in A$, de tal manera que para casi todo $x \in A$ hay una infinidad $k \in \mathbb{N}$ para las cuales se cumple $T^k(x) \in A$.*

Si X es un espacio topológico y μ es una medida de Borel en X, entonces el *soporte* de μ, denotado como $\mathrm{sop}(\mu)$, es el complemento de la unión de todos los conjuntos abiertos con medida 0, o la intersección de todos los conjuntos cerrados con medida completa.

Cada sistema dinámico induce una acción sobre las funciones, decimos que T *actúa* sobre una función f denotado por $(T_* f)(x) = f(T(x))$. Dada una función f cualquiera, nos interesa pensar qué tanta diferencia hay entre la función y las acciones que inducen las transformaciones T^m. Estudiar las propiedades ergódicas de un sistema implica estudiar el grado de *independencia estadística*, es decir la medida de los conjuntos f y $T_*^m f$.

Sea T una transformación que preserva la medida en un espacio de medida $(X, \mathfrak{A}, \mu)$. Una función medible $f : X \to \mathbb{R}$ es *esencialmente T-invariante* si $\mu(\{x \in X : f(T^t x) \neq f(x)\}) = 0$ para cada $t \in \mathbb{N}$. Un conjunto A es *esencialmente T-invariante* si la función característica 1_A es esencialmente T-invariante, o equivalentemente si $\mu(T^{-1}(A) \Delta A) = 0$.

Una transformación T que preserva la medida es *ergódica* si cualquier conjunto medible esencialmente T-invariante tiene medida 0 o medida completa.

Proposición 2.16. *Sea $(X, \mathfrak{A}, \mu)$ un espacio de medida y supongamos que $f \colon X \to \mathbb{R}$ es esencialmente invariante para una transformación medible T en X. Entonces hay un función medible estrictamente invariante $\overline{f}$ tal que $f(x) = \overline{f}(x)$ (mod 0).*

Una medida que preserva una transformación T en un espacio de probabilidad $(X, \mathfrak{A}, \mu)$ es llamada *fuertemente mezclada* si para cualesquiera dos conjuntos medibles $A, B \in \mathfrak{A}$,

ocurre que

$$\lim_{t\to\infty} \mu(T^{-t}(A) \cap B) = \mu(A) \cdot \mu(B). \tag{2.22}$$

Decimos que la transformación esta *débilmente mezclada* si únicamente ocurre que

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} |\mu(T^{-i}(A) \cap B) - \mu(A) \cdot \mu(B)| = 0. \tag{2.23}$$

Proposición 2.17. *Fuertemente mezclado implica débilmente mezclado, y débilmente mezclado implica ergodicidad.*

Proposición 2.18. *Sea X un espacio métrico compacto, $T\colon X \to X$ un mapeo continuo y μ una medida de Borel T-invariante en X.*

1. *Si T es ergódica entonces la órbita de casi cada punto en μ es densa en el* $\operatorname{sop}(\mu)$.
2. *Si T es mezclada entonces T esta topológicamente mezclada en el* $\operatorname{sop}(\mu)$.

Teorema 2.9. *(Krylov-Bogolubov) Sea X un espacio métrico compacto y $T\colon X \to X$ un mapeo continuo. Entonces existe una medida μ de probabilidad T-invariante de Borel en X.*

Sea $\mathfrak{M} = \mathfrak{M}(X)$ el conjunto de todas las medidas de probabilidad de Borel en X. Decimos que una sucesión de medidas $\mu_n \in \mathfrak{M}$ converge en la *topología débil** a una medida $\mu \in \mathfrak{M}$ si $\int_X f d\mu_n \to \int_X f d\mu$ para cada $f \in C(X)$. Si tomamos cualquier sucesión en $\mathfrak{M}$ y $F \subseteq C(X)$ es un subconjunto denso y numerable, entonces $\mathfrak{M}$ es compacto en la topología débil*, pues una subsucesión hará converger la integral de cualquier función en F, y por lo tanto la integral de $g \in C(X)$ también lo hará. De igual manera $\mathfrak{M}$ es convexo, es decir $t\mu + (1-t)\nu \in \mathfrak{M}$ para cada $t \in [0,1]$ y $\mu, \nu \in \mathfrak{M}$.

Definimos como *punto extremo* a un punto en un conjunto convexo, el cual no puede ser representado como la combinación convexa no trivial de otros dos puntos. Los puntos extremos de $\mathfrak{M}$ son las medidas de probabilidad soportadas solo en un punto, y las llamamos *medidas de Dirac*.

Sea $\mathfrak{M}_T \subset \mathfrak{M}$ el conjunto de todas las medidas de probabilidad de Borel T-invariantes en X. Entonces $\mathfrak{M}_T$ es cerrado y por lo tanto compacto y convexo en la topología débil*.

Proposición 2.19. *Las medidas ergódicas T-invariantes son precisamente los puntos extremos de $\mathfrak{M}_T$.*

Un espacio medible finito es un *espacio de Lebesgue* si es isomorfo a la unión de un intervalo $[0, a]$ (con la medida de Lebesgue) con un conjunto a lo más numerable de átomos. Recordenos que un *átomo* es un conjunto de un solo elemento con medida no cero. Un espacio de Lebesgue sin átomos es llamado *no atómico* y es isomorfo al intervalo $[0, a]$ con la medida de Lebesgue.

Sea $(X, \mathfrak{A}, \mu)$ un espacio de Lebesgue con $\mu(X) = 1$. Una *partición finita* en X es una colección finita ζ de conjuntos C_i medibles esencialmente ajenos, llamados *elementos o átomos* de ζ, cuya unión cubre X mod 0. Decimos que que una partición ζ' es un *refinamiento* de ζ, si cada elementos de ζ' está contenido mod 0 en un elemento de ζ y se denota como $\zeta \leq \zeta'$. Dos particiones son *equivalentes* si cada una es un refinamiento de la otra.

Definimos el *refinamiento común* $\zeta \vee \zeta'$ de las particiones ζ y ζ', como la partición que resulta de las intersecciones $C_\alpha \cap C'_\beta$ tal que $C_\alpha \in \zeta$ y $C'_\beta \in \zeta'$. Es la partición más pequeña que cumple ser un refinamiento de ζ, ζ'. La *intersección* $\zeta \wedge \zeta'$ es la partición medible más grande para la cual ζ y ζ' son refinamientos. La partición *trivial* ν, es la partición que solo tiene un elemento.

Tomemos $\xi = \{C_i : 1 \leq i \leq m\}$ y $\eta = \{D_j : 1 \leq j \leq n\}$ dos particiones finitas donde podemos asumir que $m = n$ (si es necesario podemos añadir conjuntos vacíos a una de ellas). Entonces definimos la distancia entre dos particiones como

$$d(\xi, \eta) = \min_{\sigma \in S_m} \sum_{i=1}^{m} \mu(C_i \Delta D_{\sigma(i)}). \tag{2.24}$$

Las particiones ζ y ζ' son *independientes* si $\mu(C \cap C') = \mu(C) \cdot \mu(C')$ para todo $C \in \zeta$ y $C' \in \zeta'$, y lo denotamos como $\zeta \perp \zeta'$. Dada una transformación T y una partición $\zeta = \{C_1, \ldots, C_m\}$, definimos $T^{-1}(\zeta) = \{T^{-1}(C_1), \ldots, T^{-1}(C_m)\}$.

2.3.1. Cadenas de Markov y automorfismos de Bernoulli.

Antes de continuar necesitamos estudiar los automorfismos de Bernoulli. Para ésto, sea A una matriz de $m \times m$ irreducible y además estocástica, esto último es, una matriz cuadrada, con entradas reales no negativas y que la suma de cada uno de los renglones sea 1. Supongamos[2] que A tiene un eigenvector izquierdo q no negativo con eigenvalor asociado 1 y la suma de todas sus entradas es 1.

[2]El Teorema de Perron-Frobenius garantiza la existencia de dicho vector: la *distribución estacionaria*.

Definición 2.26. Sea $\mathcal{A} = \{1, 2, \ldots, m\}$ un alfabeto y consideremos el espacio shift completo $\mathcal{A}^{\mathbb{Z}} = \Sigma_m$. Dada una lista $a_1, a_2, \ldots, a_j \in \mathcal{A}$ de símbolos y un conjunto $n_1, \ldots, n_j \in \mathbb{Z}$, definimos el *cilindro* correspondiente como

$$C^{n_1,n_2,\ldots,n_j}_{a_1,a_2,\ldots a_j} = \{x \in \Sigma_m | x_{n_i} = a_i, 1 \leq i \leq j\}. \tag{2.25}$$

Es decir el conjunto de todas las sucesiones en las que, para cada $k = 1, \ldots, j$, el símbolo a_k ocurre en la coordenada n_k. En particular, si $u = a_1 \ldots a_j$ y $n_k = n_1 + (k-1)$, entonces el cilindro $C^{n_1}_u$, denotado también por $C_n(u)$, es el conjunto de todas las sucesiones con la palabra u empezando en la coordenada n_1. Si asumimos que dicha coordenada es cero podemos denotar el cilindro de una palabra u simplemente como $[u]$.

Los conjuntos cilindro generan la σ-álgebra de Borel. Así, para definir una medida basta definirla en los conjuntos cilindro. Definimos una *medida de probabilidad de Borel* $P = P_{A,q}$ en Σ_m de la siguiente manera. Para un cilindro C^n_k de longitud 1, con $k \in \mathcal{A}$ y $n \in \mathbb{Z}$, definimos la función P como $P(C^n_j) = q_j$, y en general definimos

$$P(C^n_u) = q_{a_1} \prod_{i=1}^{j-1} A_{a_i a_{i+1}}. \tag{2.26}$$

Entonces podemos interpretar a q como la distribución de probabilidad inicial en el conjunto $\{1, \ldots, m\}$ y A es la matriz de transición de probabilidades. El número $P(C^n_k)$ es la probabilidad de observar el símbolo k en el n-ésimo lugar, y $A_{k,k'}$ es la probabilidad de pasar de k a k'. El hecho de que $qA = q$ significa que la distribución de probabilidad q es invariante bajo la acción de A, esto es

$$q_j = P(C^{n+1}_k) = \sum_{i=1}^{m} P(C^n_i) A_{ik}, \tag{2.27}$$

razón por la cual se le conoce como *distribución estacionaria.*

Al par (A, q) le llamaremos *cadena de Markov* en el conjunto $\{1, \ldots, m\}$.

Los cilindros son conjuntos que cumplen con ser conjuntos abiertos y cerrados al mismo tiempo en los espacios shift. Pues dado un elemento x en el espacio shift, entonces podemos tomar el cilindro $C_{-n}(x_{[-n,n]})$, el cual es el conjunto que contiene a todas las sucesiones con la palabra central de tamaño $2n+1$ igual a la del elmento x. Por nuestra definición de distancia en los espacios shift, este conjunto es precisamente la bola de radio 2^{-n+1} centrada en x. Por lo tanto para un cilindro cualquiera $C_n(u)$ siempre existe una bola, de radio $2^{-n-|u|+1}$, que lo va a contener.

Por otro lado, el complemento de un cilindro C_u^n, son todas las sucesiones que *no* tienen la palabra u en esa posición, y como u es de tamaño finito y el alfabeto es finito, solo existen una cantidad finita de palabras distintas a u que pueden tomar su lugar, por lo tanto su complemento es la unión finita de dichos cilindros cilindros, es decir,

$$\Sigma \setminus C_u^n = \bigcup_{w \in \mathcal{A}^j \setminus \{u\}} C_w^n$$

y como la unión de abiertos es abierta, entonces el complemento es abierto, y por lo tanto C_u^n es cerrado.

Ls topología que inducen los conjuntos cilindro es, por definición, la topología producto en $\mathcal{A}$, en el entendido de que $\mathcal{A}$ es un espacio topológico discreto. Por lo tanto los conjuntos cilindro son una base de nuestra topología, por lo tanto la σ-álgebra $\mathfrak{C}$ inducida por los conjuntos cilindro es precisamente la σ-álgebra de Borel. La medida P en $\mathfrak{C}$ es una extensión única de una medida σ-aditiva y la llamamos *medida de Markov correspondiente a A y q*, y el espacio de medida $(\Sigma_m, \mathfrak{C}, P)$ es un espacio de probabilidad no átomico de Lebesgue.

Si la transición de probabilidad no depende del estado inicial tenemos un caso particular, en este caso cada renglón de A es el eigenvector q, entonces la medida shift invariante P es llamada la *medida de Bernoulli* y el espacio shift es llamado *automorfismo de Bernoulli*.

Consideremos un automorfismo de Bernoulli en Σ_m con probabilidades $q_i > 0, q_1 + \ldots + q_m = 1$. Sea ξ la partición de Σ_m en m conjuntos $C_i = \{x \in \Sigma_m : x_0 = i\}$ y $\mu(C_i) = q_i$. Definamos $\eta_n = \vee_{k=0}^{n-1} \sigma^{-k}(\xi)$ y sea $\eta_n(x)$ el elemento de η_n que contiene a x. Para $x \in \Sigma_m$ sea $f_i^n(x)$ la frecuencia relativa del símbolo i en la palabra $x_1 \ldots x_n$. Dado que σ es ergódica con respecto a μ por el teorema ergódico de Birkhoff, para cada $\epsilon > 0$ existen una $N \in \mathbb{N}$ y un subconjunto $A_\epsilon \subseteq \Sigma_m$ con $\mu(A_\epsilon) > 1 - \epsilon$ tal que $|f_i^n(x) - q_i| < \epsilon$ para cada $x \in A_\epsilon$ y $n \geq N$. Por lo tanto si $x \in A_\epsilon$, entonces

$$\mu(\eta_n(x)) = \prod_{i=1}^{m} q_i^{(q_i+\epsilon_i)n} = 2^{n \sum_{i=1}^{m} (q_i+\epsilon_i) \log q_i} \tag{2.28}$$

donde $|\epsilon_i| < \epsilon$ y el log denota el logaritmo en base 2 con $0 \log 0 = 0$. Sacando el logaritmo, dividiendo entre n y tomando el límite con $n \to \infty$ obtenemos

$$\lim_{n \to \infty} \frac{1}{n} \log \mu(\eta_n(x)) = \sum_{i=1}^{m} q_i \log q_i. \tag{2.29}$$

Lo que todo esto quiere decir es que podemos encontrar conjuntos A_ϵ de medida casi 1, tal que para n suficientemente grandes, las frecuencia relativas de los elementos x son básicamente

el valor de la medida de probabilidad q_i. Así que la medida de los conjuntos η_n que contienen estas palabras, es decir que tienen una frecuencia de símbolos correcta, están relacionadas con $h = -\sum_{i=1}^{m} q_i \log q_i$.

2.3.2. Particiones y entropía métrica

Sea $(X, T, \mathfrak{A}, \mu)$ una transformación medible de un espacio de Borel. Para ser más concretos: X es un espacio métrico compacto, $T\colon X \to X$ es continua e invertible, $\mathfrak{A}$ es la σ-álgebra de Borel, $\mu\colon \mathfrak{A} \to [0,1]$ es una medida de probabilidad invariante bajo T.

Dada una partición $\zeta = \{C_1, \ldots, C_n\}$ de X, definimos la entropía de ζ por

$$H(\zeta) = -\sum_{i=1}^{n} \mu(C_i) \log \mu(C_i). \tag{2.30}$$

Para $x \in X$ sea $m(x, \zeta)$ la medida de los elementos de ζ que contienen a x. Entonces claramente la siguiente expresión no da una interpretación de la definición de entropía como la esperanza de una cierta variable aleatoria, la *función de información*:

$$H(\zeta) = -\int_X \log m(x, \zeta) d\mu. \tag{2.31}$$

Proposición 2.20. *Sea ζ y η dos particiones finitas. Entonces*

1. *$H(\zeta) \geq 0$ y $H(\zeta) = 0$ si y solo si $\zeta = \nu$.*
2. *Si $\zeta \leq \eta$, entonces $H(\zeta) \leq H(\eta)$ y la igualdad se cumple si y solo si $\zeta = \eta$.*
3. *Si ζ tiene n elementos, entonces $H(\zeta) \leq \log n$ y la igualdad se cumple si y solo si cada elemento de ζ tiene medida $1/n$.*
4. *$H(\zeta \vee \eta) \leq H(\zeta) + H(\eta)$ con la igualdad si y solo si $\zeta \perp \eta$.*

Como ya lo mencionamos, una interpretación de la entropía de una partición es el promedio de la información de los elementos de la partición. Por ejemplo, supongamos que X representa el conjunto de todos los posibles resultados de un experimento y μ es la probabilidad de distribución de esos resultados. Para extraer información de esos experimentos

creamos un sistema de medida que parte X en un número finito de subconjuntos observables o eventos, $C_1, C_2, \ldots, C_n$. Para encontrar una función $I\colon \mathfrak{A} \to \mathbb{R}^+$ que nos de la información de un evento tenemos que debe cumplir las siguiente propiedades:

- La información es una función no negativa.
- La información es una función decreciente de la probabilidad de un evento, pues menor sea la probabilidad de un evento, más grande es la información que se obtiene del evento. De igual manera la información que se obtiene de un evento trivial X es 0.
- Para dos eventos independientes C y D, la información es aditiva, es decir, $I(C \cap D) = I(C) + I(D)$.

Se puede demostrar que la "única" función que nos permite estas propiedades es $I(C) = -\log \mu(C)$. Por lo tanto la entropía de la partición es el promedio de la información de los elementos en la partición.

Para dos subconjuntos medibles $C, D \subseteq X$ con $\mu(D) > 0$, definimos la medida condicional $\mu(C|D) = \mu(C \cap D)/\mu(D)$. Sean $\xi = \{C_i : i \in I\}$ y $\eta = \{D_j : j \in J\}$ dos particiones. Podemos definir la *entropía condicional* de ξ con respecto a η, como la entropía promedio de la partición inducida por ξ en un elemento de η, definida por la fórmula

$$H(\xi|\eta) = -\sum_{j \in J} \mu(D_j) \sum_{i \in I} \mu(C_i|D_j) \log \mu(C_i|D_j). \tag{2.32}$$

Proposición 2.21. *Sean ξ, η y ζ particiones finitas. Entonces*

1. $H(\xi|\eta) \geq 0$ *y la igualdad se cumple si y solo si* $\xi \leq \eta$.
2. $H(\xi|\nu) = H(\xi)$.
3. *Si* $\eta \leq \zeta$, *entonces* $H(\xi|\eta) \geq H(\xi|\zeta)$.
4. *Si* $\eta \leq \zeta$, *entonces* $H(\xi \vee \eta|\zeta) \geq H(\xi|\zeta)$.
5. *Si* $\xi \leq \eta$, *entonces* $H(\xi|\zeta) \leq H(\eta|\zeta)$ *y la igualdad se cumple si y solo si* $\xi \vee \zeta = \eta \vee \zeta$.
6. $H(\xi \vee \eta|\zeta) = H(\xi|\zeta) + H(\eta|\xi \vee \zeta)$ *y* $H(\xi \vee \eta) = H(\xi) + H(\eta|\xi)$.
7. $H(\xi|\eta \vee \zeta) \leq H(\xi|\zeta)$.

8. *$H(\xi|\eta) \leq H(\xi)$ y la igualdad se cumple si y solo si $\xi \perp \eta$.*

Definición 2.27. Dadas dos particiones finitas ξ y η, definimos $\rho(\xi,\eta) = H(\xi|\eta) + H(\eta|\xi)$ como la *métrica de Rokhlin*, la cual define una métrica en el espacio de clases de equivalencia de particiones.

Proposición 2.22. *Para cada $\epsilon > 0$ y $m \in \mathbb{N}$, existe $\delta > 0$ tal que si ξ y η son particiones finitas con a lo más m elementos y $d(\xi,\eta) < \delta$, entonces $\rho(\xi,\eta) < \epsilon$.*

Supongamos que $\zeta = \{C_\alpha : \alpha \in I\}$ es una partición en X con entropía finita. Para $k, n \in \mathbb{N}$, definimos $T^{-k}(\zeta) = \{T^{-k}(C_\alpha) : \alpha \in I\}$ y $\zeta^n = \zeta \vee T^{-1}(\zeta) \vee \ldots \vee T^{-n+1}(\zeta)$.

Si tomamos

$$\begin{aligned} H(\zeta^{m+n}) &= H(\zeta \vee T^{-1}(\zeta) \vee \ldots \vee T^{-n+1}(\zeta) \vee \ldots \vee T^{-(n+m)+1}(\zeta) \\ &\leq H(\zeta \vee T^{-1}(\zeta) \vee \ldots \vee T^{-n+1}(\zeta)) + H(T^{-n}\zeta \vee T^{-n-1}(\zeta) \vee \ldots \vee T^{-n-m+1}(\zeta)) \end{aligned} \tag{2.33}$$

y aplicando que $H(T^{-k}(\zeta)) = H(\zeta)$, podemos entonces concluir que $H(\zeta^{m+n}) \leq H(\zeta^m) + H(\zeta^n)$. Por lo tanto el límite

$$h(T,\zeta) = \lim_{n\to\infty} \frac{1}{n} H(\zeta^n) \tag{2.34}$$

existe y se define como la *entropía de medida*, o *entropía de métrica* de T relativa a ζ.

Proposición 2.23. *$h(T,\zeta) = \lim_{n\to\infty} H(\zeta|T^{-1}(\zeta^n))$. Es decir que $h(T,\zeta)$ es la información promedio añadida por el estado presente condicionada a que todos los estados pasados sean conocidos.*

Proposición 2.24. *Sea ξ y η particiones finitas. Entonces*

1. *$h(T,T^{-1}(\xi)) = h(T,\xi)$, y si T es invertible, entonces $h(T,T(\xi)) = h(T,\xi)$.*
2. *$h(T,\xi) = h\big(T, \bigvee_{i=0}^{n} T^{-i}(\xi)\big)$ para $n \in \mathbb{N}$, y si T es invertible, entonces $h(T,\xi) = h(T, \bigvee_{i=-n}^{n} T^{-i}(\xi))$.*
3. *$h(T,\xi) \leq h(T,\eta) + H(\xi|\eta)$, y si $\xi \leq \eta$, entonces $h(T,\xi) \leq h(T,\eta)$.*
4. *$|h(T,\xi) - h(T,\eta)| \leq \rho(\xi,\eta) = H(\xi|\eta) + H(\eta|\xi)$.*
5. *$h(T,\xi \vee \eta) \leq h(T,\xi) + h(T,\eta)$.*

Definimos la *entropía métrica de T* como el supremo de todas las entropías métricas $h(T,\zeta)$ sobre todas las particiones finitas y medibles ζ de X. Si dos transformaciones que preservan la medida son isomorfas, entonces las entropías de métricas son iguales.

Si una sucesión (ζ_n) de particiones finitas es tal que $\zeta_{n+1} \geq \zeta_n$ para todo $n \in \mathbb{N}$, entonces decimos que es una sucesión *refinadora*. La sucesión es *generadora* si para cada partición finita ξ y cada $\delta > 0$, existe $n_0 \in \mathbb{N}$ tal que, para cada $n \geq n_0$, existe una partición ξ_n con $\xi_n \leq \bigvee_{i=1}^{n} \zeta_i$ y $d(\xi_n, \xi) < \delta$.

Definición 2.28. Si (ζ_n) es una sucesión de particiones finitas que es refinadora y generadora, entonces $h(T) = h_\mu(T) = \lim_{n\to\infty} h(T,\zeta_n)$.

Un *generador de un lado* para una transformación T que preserva la medida (no necesariamente invertible), es una partición finita ξ tal que la sucesión $\xi^n = \bigvee_{k=0}^{n} T^{-k}(\xi)$ es generadora. Si la transformación es invertible, entonces el *generador de dos lados* es la partición finita ζ tal que la sucesión $\bigvee_{k=-n}^{n} T^k(\zeta)$ es generadora.

Teorema 2.10. *(Kolmogorov-Sinai) Sea ξ un generador para T. Entonces $h(T) = h(T,\xi)$.*

Proposición 2.25. *Sean T y S dos transformaciones que preservan la medida de dos espacios de medida $(X,\mathfrak{A},\mu)$ y $(Y,\mathfrak{B},\nu)$ respectivamente. Entonces*

1. *$h(T^k) = k \cdot h(T)$ para cada $k \in \mathbb{N}$, y si T es invertible, entonces $h(T^{-1}) = h(T)$ y $h(T^k) = |k| \cdot h(T)$ para cada $k \in \mathbb{Z}$.*
2. *Si T es un factor de S, entonces $h_\mu(T) \leq h_\nu(S)$.*
3. *$h_{\mu\times\nu}(T \times S) = h_\mu(T) + h_\nu(S)$.*

Sea $f\colon X \to X$ un homeomorfismo de un espacio métrico compacto X y se $\mathcal{M}$ el espacio de medidas de probabilidad de Borel en X.

Lema 2.6. *Sean $\mu,\nu \in \mathcal{M}$ y $t \in (0,1)$. Entonces para cada partición medible de ζ en X ocurre que*

$$tH_\mu(\zeta) + (1-t)H_\nu(\xi) \leq H_{t\mu+(1-t)\nu}(\zeta). \tag{2.35}$$

Para una partición $\xi = \{A_1,\ldots,A_k\}$ definimos la *frontera* de ξ como el conjunto $\delta\xi = \bigcup_{i=1}^{k} \partial A_i$ donde $\partial A = \overline{A} \cap \overline{X \setminus A}$.

Lema 2.7. *Sea $\mu \in \mathcal{M}$. Entonces*

1. *Para cualquier* $x \in X$ *y* $\delta > 0$*, existe* $\delta' \in (0, \delta)$ *tal que* $\mu(\partial B(x, \delta')) = 0$.
2. *Para cualquier* $\delta > 0$*, existe una partición medible finita* $\xi = \{C_1, \ldots, C_k\}$ *con* $\operatorname{diam}(C_i) < \delta$ *para todo* i *y* $\mu(\partial \xi) = 0$.
3. *Si* $\{\mu_n\} \subset \mathcal{M}$ *es una sucesión de medidas de probabilidad que convergen a* μ *en la topología débil* y* A *es un conjunto medible con* $\mu(\partial A) = 0$*, entonces* $\mu(A) = \lim_{n\to\infty} \mu_n(A)$.

Teorema 2.11. *(Principio Variacional) sea* f *un homeomorfismo de un espacio métrico compacto* (X, d)*. Entonces* $h_{\text{top}}(f) = \sup\{h_\mu(f) | \mu \in \mathcal{M}_f\}$.

El Principio Variacional es uno de los teoremas más importantes en teoría ergódica ya que nos da una relación entre las entropías topológica y métricas, a saber, la entropía topológica es el supremo de todas las entropías métricas. Como vimos, la entropía topológica es una propiedad importante para entender qué tan complejo es nuestro sistema, pues depende de las órbitas: entre mayor entropía, mayor será el número de fragmentos de órbitas. La entropía topológica también es un invariante de conjugación, así que nos es muy útil para encontrar conjugaciones topológicas en sistemas dinámicos.

Si tenemos diferentes medidas de Borel, sabemos, por el Principio Variacional, que ninguna entropía métrica será estrictamente mayor que la entropía topológica, pero puede pasar que exista μ tal que $h_\mu = h_{\text{top}}$. Si esto pasara, sería muy importante, pues la medida μ sería una buena medida para trabajar: la información sería igual a la información que nos dan sus órbitas (topológicas).

Es por todo esto que un problema fundamental en teoría ergódica es determinar cuándo existe una medida de entropía máxima (es decir que su entropía métrica coincide con la entropía topológica), y más aún, si este es el caso, determinar cuándo es única. Si la medida cuya entropía métrica es igual a la topológica es única, entonces se vuelve la mejor medida para trabajar, por lo anterior.

Este capítulo nos dio la información que necesitábamos para entender los diferentes tipos de entropía y qué relación tienen entre ellos, así como un principio de la teoría ergódica y cómo la medida juega un papel importante en el estudio de los sistemas dinámicos. En general todo esto nos lleva a que los sistemas pueden ser caóticos, que en términos generales significa entropía estrictamente positiva (mayor que cero), aunque no existe un acuerdo general para esto último (por ejemplo, los sistemas de entropía cero, también conocidos como de "baja complejidad", pueden exhibir comportamientos que se consideran caóticos,

El trabajo que haremos será intentar responder la pregunta de cuándo existe una única medida de entropía máxima. Para esto estudiaremos la propiedad de especificación, la cual se

ha mostrado que puede ayudar a encontrar sistemas donde ocurra ésto. En especial veremos que le ocurre a los sistema simbólicos bajo algunas otras propiedades mas débiles que la especificación.

Capítulo 3

Especificación

Buscar una medida cuya entropía de medida teórica sea igual que la medida topológica y que sea única es un problema bastante complicado. Por suerte existe la especificación. La especificación es una propiedad de los sistemas dinámicos que se asemeja mucho a la mezcla topológica, pero que bajo ciertas condiciones está relacionada con nuestro problema.

A lo largo de este capítulo usaremos espacios shift de un solo lado, es decir que las sucesiones de símbolos estarán indexadas por $\mathbb{N}$ en lugar de $\mathbb{Z}$. Esto solo para hacerlo más sencillo, aunque en general todo se va cumplir para los espacios shift de dos lados (sucesiones bi-infinitas).

Antes de hablar de especificación daremos unas cuantas definiciones que nos ayudarán a llegar a este concepto.

3.1. Propiedades básicas

Definición 3.1. Sea X un espacio shift. Sean $a, b \in \mathbb{N}_0$ tales que $a \leq b$. Definimos el *segmento de órbita* de $x \in X$ sobre $[a, b]$ como la sucesión $T^{[a,b]}(x) = (T^a(x), T^{a+1}(x), \ldots, T^b(x))$. También definimos $T^{[a,b)}(x) = T^{[a,b-1]}(x)$.

Definición 3.2. Una *especificación* es una familia de segmentos de órbita $\xi = \{T^{[a_j,b_j]}(x^j)\}_{j=1}^n$, donde $n \in \mathbb{N}$ y $b_j < a_{j+1}$ para todo $1 \leq j < n$. Al número de segmentos de órbita en la especificación lo llamamos *rango*.

Notemos que, en la especificación, el rango tiene que ser finito. De igual manera en cada segmento de órbita el elemento $x^j \in X$ puede ser distinto, esto nos será de mucha utilidad una vez que estemos trabajando con la propiedad de especificación en los espacios shift.

Definición 3.3. La *distancia de Bowen* entre $x, y \in X$ a lo largo de un conjunto finito $\Lambda \subset \mathbb{N}$, es $\rho_\Lambda^T(x,y) = \text{máx}\{\rho(T^j(x), T^j(y)) : j \in \Lambda\}$. Si $\Lambda = \{0, 1, \ldots, n\}$ entonces se denota $\rho_\Lambda^T(x,y) = \rho_n^T(x,y)$.

La distancia de Bowen nos da la distancia más grande que hay entre todas las iteraciones de dos elementos en un conjunto específico de iteraciones. Se puede definir una *bola de Bowen* de radio ϵ centrada en $x \in X$ a lo largo de Λ como el conjunto $B_\Lambda(x,\epsilon) = \{y \in X : \rho_\Lambda^T(x,y) < \epsilon\}$, es decir, el conjunto de elementos y tal que, para $i \in \Lambda$, ninguna iteración $T^i(y)$ se aleja mucho (más que ϵ) de $T^i(x)$.

El *espacio de límite inverso* de un sistema dinámico sobreyectivo (X,T) es el espacio $X_T = \{(x_1, x_2, x_3, \ldots) \in X^\infty : T(x_{i+1}) = x_i$ para todo $i \in \mathbb{N}\}$. Podemos interpretar a los elementos en X_T como la órbita $\mathcal{O}^{[-\infty,n-1]}$ para algún x_n pues cada elemento lo podemos ver e.g. como

$$(x_1, x_2, x_3, x_4, x_5, \ldots) = (T(x_2), T(x_3), T(x_4), T(x_5), T(x_6), \ldots) \tag{3.1}$$
$$= (T^3(x_4), T^2(x_4), T(x_4), x_4, T^{-1}(x_4), \ldots). \tag{3.2}$$

Podemos equipar a X_T con la topología inducida como subespacio de X^∞, este último equipado con la topología producto. El mapeo T es llamado el mapeo *unión* y definimos al mapeo $\sigma_T \colon X_T \to X_T$ tal que $\sigma_T(x_0, x_1, x_2, \ldots) = (T(x_0), T(x_1), T(x_2), \ldots) = (T(x_0), x_0, x_1, \ldots)$ como el *homomorfismo shift* y el sistema dinámico invertible (X_T, σ_T) es la *extensión natural* de (X,T). Estos dos sistemas dinámicos están íntimamente relacionados, si uno de ellos es transitivo o está mezclado, el otro necesariamente lo es, de igual manera comparten la misma entropía topológica.

Un sistema dinámico invertible (X,T) es *positivamente expansivo* si existe una constante $c > 0$ tal que, si $x, y \in X$ satisfacen que $d(T^n(x), T^n(y)) < c$ para toda $n \in \mathbb{N}_0$ entonces $x = y$. Es decir que en un espacio positivamente expansivo, por muy cerca que estén dos puntos, eventualmente se separan bajo iteraciones de T. Veamos que los espacios shift de un lado son positivamente expansivos.

Sea (X, σ) un espacio shift de un lado y $x, y \in X$. Procedemos por contradicción. Supongamos que $x \neq y$ y $d(\sigma^n(x), \sigma^n(y)) < c$ para algún $c > 0$ y $n \in \mathbb{N}$. Sabemos que en los espacios shift el número más grande que alcanza la distancia entre dos puntos es 1, y esto es si el primer símbolo en cada uno de los elementos es diferente, así que pediremos que $c < 1$ pues si no cualquier par de elementos lo cumpliría sin problema. Supongamos que la distancia entre los dos elementos es $d(x, y) = 2^{-k}$ donde k es tal que $x_{[0.k]} = y_{[0.k]}$ pero $x_{k+1} \neq y_{k+1}$, esto por nuestra definición de distancia en los espacios shift. Ahora $d(\sigma(x), \sigma(y)) = 2^{-(k-1)}$, pues al usar la transformación σ se mueve un espacio a la izquierda la sucesión. Así podemos hacerlo sucesivamente hasta que $d(\sigma^k(x), \sigma^k(y)) = 2^{-1}$ y entonces tenemos que $d(\sigma^{k+1}(x), \sigma^{k+1}(y)) = 2^0 = 1 > c$, llegando a la contradicción y por lo tanto un espacio shift es positivamente expansivo.

A continuación definiremos la especificación tal como Kwetniak y Oprocha lo definieron en [11].

Definición 3.4. Sea $\nu \colon \mathbb{N} \to \mathbb{N}$ cualquier función. Una especificación $\xi = \{T^{[a_j, b_j]}(x^j)\}_{j=1}^n$ es es *ν-espaciada* si $a_i - b_{i-1} \geq \nu(b_i - a_i + 1)$ para $2 \leq i \leq n$. Dada una constante $N \in \mathbb{N}$ decimos que una especificación es N-espaciada si la función $\nu(n) = N$ para toda $n \in \mathbb{N}$.

Definición 3.5. Decimos que la especificación $\xi = \{T^{[a_j, b_j]}(x^j)\}_{j=1}^n$ es *ϵ-rastreable* por $y \in X$ si $\rho(T^k(y), T^k(x^j)) \leq \epsilon$ para $a_j \leq k \leq b_j$ y $1 \leq j \leq n$.

Definición 3.6. (X, T) tiene la *propiedad de especificación* si para cualquier $\epsilon > 0$, existe una constante $N = N(\epsilon) \in \mathbb{N}$, tal que cualquier especificación N-espaciada $\xi = \{T^{[a_j, b_j]}(x^j)\}_{j=1}^n$ es ϵ-rastreable para algún $y \in X$. Si adicionalmente y puede ser elegida de tal manera que $T^{b_n - a_0 + N}(y) = y$, entonces tiene la *propiedad de especificación periódica*.

Una vez que tenemos definida la propiedad de especificación, queremos observar cómo se comportaría un espacio shift con esta propiedad. Supongamos que tenemos un espacio shift X y el sistema dinámico (X, σ) tiene la propiedad de especificación, por lo tanto sabemos que para cualquier ϵ podemos obtener el número $N = N(\epsilon)$ y cualquier especificación N-espaciada $\xi = \{T^{[a_j, b_j]}(x^j)\}_{j=1}^n$ es ϵ-rastreable para algún $y \in X$.

Definamos $x^j_{[a_j, b_j]} = u_j$ como la palabra que se encuentra entre los lugares a_j y b_j en el elemento x^j. Evidentemente esta palabra se encuentra en el lenguaje del espacio shift, esto es $u_j \in \mathcal{L}(X)$, y podemos definirlo para cada $1 \leq j \leq n$. Como la especificación es ϵ-rastreable por y, entonces tenemos que $\rho(\sigma^k(y), \sigma^k(x^j)) \leq \epsilon$, para $a_j \leq k \leq b_j$. Dado que estamos tomando la distancia en un espacio shift, los bloques $[2a_j - b_j - m, b_j + m]$ de y y x^j deben coincidir, donde $m > 0$ es constante, pues cuando se aplique la transformación σ^{a_j} este bloque pasará a la posición $[a_j - b_j - m, b_j + m - a_j]$ el cual es el bloque central de ambos y por lo tanto la distancia es a lo más $2^{-(b_j - a_j + m)}$. Cuando apliquemos las subsecuentes

iteraciones, seguirán teniendo bloques centrales iguales cada vez más pequeños, hasta que en σ^{b_j} el bloque pasará a ser $[2a_j - 2b_j - m, m]$ y como sabemos que $b_j > a_j$, entonces este intervalo contiene al bloque central $[-m, m]$ y por lo tanto la distancia será a lo más 2^{-m}, así que haciendo que m sea tal que $2^{-m} \leq \epsilon$, obtenemos $m \geq -\log_2(\epsilon)$, y esto se cumple para cada $j \leq n$.

La palabra u_j está contenida en $[2a_j - b_j - m, b_j + m]$ y en especial podemos decir que $u_j = x^j_{[a_j,b_j]} = y_{[a_j,b_j]}$ para cada j, y sabemos que $y_{[b_j,a_{j+1}]} = w_j \in \mathcal{L}_N(X)$ pues, por como se construyó la especificación, sabemos que la distancia entre b_j y a_{j+1} es N. Por lo tanto la palabra $u_1w_1u_2w_2 \ldots w_{n-1}u_n = y_{[a_1,b_n]} \in \mathcal{L}(X)$ pues es una palabra que se encuentra en y, el cual es un punto del espacio shift. Como esto se hizo para cualquier especificación, entonces las palabras u_j's pueden ser cualesquiera, pues cualquier palabra que queramos se encuentra en algún punto x^j en el bloque $[a_j, b_j]$, por lo tanto la propiedad de especificación implica que para cualesquiera n palabras $u_1, u_2, \ldots, u_n \in \mathcal{L}(X)$, existen $w_1, w_2, \ldots, w_{n-1} \in \mathcal{L}(X)$ tales que $u_1w_1u_2w_2 \ldots w_{n-1}u_n \in \mathcal{L}(X)$.

Nos falta probar la inversa, lo cual es muy sencillo pues se hace la misma construcción pero en reversa. Supongamos que para un espacio shift X y un sistema dinámico (X, σ) tenemos que para cualesquiera n palabras $u_1, u_2, \ldots, u_n \in \mathcal{L}(X)$, existen $w_1, w_2, \ldots, w_{n-1} \in \mathcal{L}(X)$ tales que $u_1w_1u_2w_2 \ldots w_{n-1}u_n = p \in \mathcal{L}(X)$. Entonces como $p \in \mathcal{L}(X)$, podemos construir una sucesión y donde la palabra p se encuentre en el bloque que queramos de y, digamos que el bloque es $[a_1, b_n]$ y está conformado de tal forma que $u_1 = y_{[a_1,b_1]}, w_1 = y_{[b_1,a_2]}$ y así sucesivamente para todo $1 \leq j \leq n$. Además, $a_1 \in \mathbb{N}$ puede ser el valor que queramos y todos los demás dependen de él. De igual manera construimos x^j tal que $x^j_{[a_j-b_j-m,b_j+m]} = y_{[a_j-b_j-m,b_j+m]}$, como los espacios shift son positivamente expansivos podemos encontrar x^j tal que $x^j \neq y$. Al tomar $\rho(\sigma^k(y), \sigma^k(x^i))$ para todo $k \in [a_j, b_j]$ tenemos que, como vimos más arriba, la distancia es 2^{-m} y eligiendo el epsilon adecuado para cada $1 \leq j \leq n$ es fácil observar que y $\epsilon-$rastrea a los segmentos de x^j, los cuales se pueden representar como la especificación N-espaciada $\xi = \{T^{[a_j,b_j]}(x^j)\}_{j=1}^n$. Dado que esto es para cada palabra, podemos crear cualquier especificación N-espaciada y ϵ el cual solo depende del valor que tomemos para m.

Teorema 3.1. *Un espacio shift tiene la propiedad de especificación si existe un entero $N \geq 0$ tal que para cualesquiera n palabras $u_1, u_2, \ldots, u_n \in \mathcal{L}(X)$, existen $n-1$ palabras $w_1, w_2, \ldots, w_{n-1} \in \mathcal{L}_N(X)$ tales que $u_1w_1u_2w_2 \ldots w_{n-1}u_n \in \mathcal{L}(X)$.*

El siguiente teorema nos da ciertas propiedades que tienen los sistemas dinámicos con la propiedad de especificación y son igualmente válidos con la propiedad de especificación periódica.

Teorema 3.2. *1. Si (X,T) tiene la propiedad de especificación, entonces (X,T^k) tiene la propiedad de especificación para cada $k \geq 1$.*

2. Si $(X,T),(Y,S)$ tienen la propiedad de especificación, entonces el sistema producto $(X \times Y, T \times S)$ también tiene la propiedad de especificación.

3. Cada factor de un sistema con la propiedad de especificación tiene la propiedad de especificación.

4. Cada mapeo sobreyectivo con la propiedad de especificación es topológicamente mezclado.

Demostración. 1) Sea $\xi = \{T^{k[a_j,b_j]}(x_j)\}_{j=1}^n$ una especificación N-espaciada para la transformación T^k. Entonces sabemos que el segmento de órbita $T^{k[a_j,b_j]}(x_j)$ está contenido en el segmento de órbita $T^{[ka_j,kb_j]}(x_j)$ y comparten el primer y último elemento. Como la especificación es N-espaciada, entonces sabemos que $a_{j+1} - b_j = N$ por lo que $ka_{j+1} - kb_j = kN$ y por lo tanto podemos construir para T la especificación $\zeta = \{T^{[ka_j,kb_j]}(x^j)\}_{j=1}^n$ la cual es kN-espaciada y donde para cada j el segmento $T^{[ka_j,kb_j]}(x_j)$ contiene al segmento $T^{k[a_j,b_j]}(x_j)$ y no interseca a ningún otro segmento de la especificación ξ. Ahora como ζ es kN-espaciada y (X,T) tiene la propiedad de especificación, sabemos que podemos tomar ϵ tal que existe una $y \in X$ que la ϵ-rastrea, y por lo tanto también rastrea a la especificación ξ. Como esto es para cada especificación N-espaciada, concluimos que (X,T^k) tiene la propiedad de especificación.

2) Sean dos sistemas dinámicos $(X,T),(Y,S)$ con la propiedad de especificación. Entonces tomemos el sistema dinámico $(X \times Y, T \times S)$ y la especificación N-espaciada $\xi = \{(T \times S)^{[a_j,b_j]}(x^j,y^j)\}_{j=1}^n$. Como (X,T) tiene la propiedad de especificación, podemos tomar la especificación $\zeta = \{T^{[a_j,b_j]}(x^j)\}_{j=1}^n$ donde los intervalos $[a_j,b_j]$ son los mismos que en ξ, por lo tanto ζ es N-espaciada y existe un $u \in X$ y $\epsilon_X > 0$ tal que ζ es ϵ_X-rastreable por u. Hacemos lo mismo para (Y,S), definimos la especificación $\nu = \{S^{[a_j,b_j]}(y^j)\}_{j=1}^n$ con los intervalos $[a_j,b_j]$ iguales que en ξ, por lo tanto ν es N-espaciada y existen $w \in Y$ y $\epsilon_Y > 0$ tales que ν es ϵ_Y-rastreable por w.

Si tomamos el elemento $(u,w) \in X \times Y$ y $\rho((T \times S)^k(x^j,y^j),(T \times S)^k(u,w))$ para $a_j \leq k \leq b_j$ y $1 \leq j \leq n$, entonces al tomar la métrica natural del producto cruz tenemos que

$$\rho((T \times S)^k(x^j,y^j),(T \times S)^k(u,w)) = \text{máx}\{\rho_X(T^k(x^j),T^k(u)),\rho_Y(S^k(x^j),S^k(w))\} \tag{3.3}$$

$$\leq \text{máx}\{\epsilon_X,\epsilon_Y\}. \tag{3.4}$$

Por lo tanto podemos ver que definiendo $\epsilon = \text{máx}\{\epsilon_X, \epsilon_Y\}$ podemos encontrar ξ la cual es N-espaciada y la cual es ϵ-rastreable por (u, w).

3) Sea (Y, S) un sistema dinámico tal que (X, T) es un factor de él, esto quiere decir que existe una semiconjugación continua $\pi\colon Y \to X$ tal que $T^k \circ \pi = \pi \circ S^k$ para toda $k \in \mathbb{N}$. (Y, S) tiene la propiedad de especificación, es decir que para todo ϵ existe $N(\epsilon)$ tal que, para toda especificación $\xi = \{S^{[a_j,b_j]}(y^j)\}_{j=1}^n$ N-espaciada existe $u \in Y$ tal que, $\rho_Y(S^k(y^i), S^k(u)) \leq \epsilon$. Ahora para (X, T) definamos una especificación $\nu = \{T^{[a_j,b_j]}(x^j)\}_{j=1}^n$ N-espaciada, entonces aplicando que la semiconjugación es sobre podemos escribir $\nu = \{T^{[a_j,b_j]}(\pi(y_j))\}_{j=1}^n$ y definimos $w = \pi(u)$. Luego entonces

$$\rho_X(T^k(x_j), T^k(w)) = \rho_X(T^k \circ \pi(y_j), T^k \circ \pi(u)) = \rho_X(\pi \circ S^k(y_j), \pi \circ S^k(u)) \leq \delta(\epsilon)$$

pues sabemos que la semiconjugación π es continua y $\rho_Y(S^k(y_j), S^k(u)) \leq \epsilon$, por lo tanto la especificación ν es δ-rastreable por $w \in X$ y esto para $\delta > 0$. Concluimos entonces que (X, T) tiene la propiedad de especificación.

4) Sabemos que un sistema dinámico es topológicamente mezclado si para $U, V \subset X$ abiertos existe $N \in \mathbb{N}$ tal que, para todo $n \geq N$ se cumple que $U \cap T^{-n}(V) \neq \varnothing$.

Tomemos $U, V \subset X$ abiertos y dos elementos $x \in U$ y $y \in V$, elijamos $\epsilon > 0$ de forma que el diámetro de U y el diámetro de V sean mayores que ϵ. Como el sistema tiene la propiedad de especificación, existe $N = N(\epsilon)$ tal que cualquier especificación N espaciada es rastreable.

Sea $n \geq N$ y $\zeta = \{T^0(x), T^n(y')\}$ una especificación donde $y' \in T^{-n}(\{y\})$, lo cual es válido porque la transformación es sobreyectiva. Nuestra especificación solo tiene dos elementos, y la distancia entre los dos intervalos es mayor que N, por lo tanto es rastreable y existe $w \in X$ tal que $d(w, x) \leq \epsilon$ y $d(T^n(w), T^n(y')) = d(T^n(w), y) \leq \epsilon$. Así que $w \in U$ y $w \in T^{-n}(V)$, por lo tanto (X, T) es topológicamente mezclado. □

Lema 3.1. *Si (X, T) tiene la propiedad de especificación y su extensión natural es expansiva, entonces (X, T) tiene la propiedad de especificación periódica.*

3.2. Especificación y ergodicidad

En teoría ergódica se pueden definir los k-sistemas, estos son transformaciones invertibles que preservan la medida en un espacio de probabilidad, donde su entropía de medida teórica

es completamente positiva. Un sistema tiene entropía completamente positiva si la entropía de Kolmogorov-Sinai de cada partición no trivial es positiva y la entropía de cada factor que preserva la medida es positiva. Esto quiere decir que no importa cómo dividamos al sistema, ni cómo lo expandamos, la entropía siempre es positiva. Al momento de querer pasar esta idea de tener entropía siempre positiva sin importar las divisiones a la dinámica topológica, debemos hacer algunos cambios pues las entropías no están definidas de la misma manera, por lo que usaremos las siguientes definiciones.

Sea (X, T) un sistema dinámico y $\mathscr{A}$ una cubierta de abiertos para X. Si tomamos el conjunto $E \subset \mathscr{A} \times \ldots \times \mathscr{A}$ (n veces) y para cada $z \in X$ existe una n-tupla de conjuntos abiertos $(A_0, \ldots, A_{n-1}) \in E$ tal que $T^k(z) \in A_k$ para todo $0 \leq k < n$, entonces decimos que E es una *n-cubierta*. Para una n-cubierta definimos su cardinalidad como la cantidad diferente de n-tuplas que puede tener. Básicamente decimos que podemos encontrar una lista de conjuntos abiertos que corresponden a una cubierta de X tal que los pedazos de órbita de tamaño n vivan en ellos. Esta definición nos será de mucha ayuda para demostrar algunas de las siguientes proposiciones.

Denotemos como $M_n(T, \mathscr{A})$ a la cardinalidad mínima que puede tener una n-cubierta, para la cubierta abierta $\mathscr{A}$. Bowen [2] define la siguiente cantidad

$$h(T, \mathscr{A}) = \lim_{n \to \infty} \frac{1}{n} \log M_n(T, \mathscr{A}) \tag{3.5}$$

y finalmente define la entropía topológica como

$$h_{top}(T) = \sup_{\mathscr{A}} h(T, \mathscr{A}). \tag{3.6}$$

Esta forma de definir la entropía es equivalente a la Definición 2.11, pues ambas dependen de coberturas abiertas y de las órbitas de los puntos, solo cambia un poco la forma de escribirlo.

Definición 3.7. Un sistema dinámico (X, T) tiene *entropía uniformemente positiva* si cualquier cubierta por dos conjuntos abiertos no densos tiene entropía positiva.

Definición 3.8. Un sistema dinámico (X, T) tiene *entropía completamente positiva* si cualquier factor no trivial tiene entropía positiva.

Como podemos ver, estas dos definiciones son relativamente parecidas a la definición para los k-sistemas, solo que en este caso en vez de tener cualquier partición no trivial tenemos una especie de partición en solo dos conjuntos abiertos.

Proposición 3.1. *Sea un sistema dinámico (X, T) con entropía uniformemente positiva. Entonces el sistema es transitivo y tiene entropía completamente positiva.*

Demostración. Supongamos que T no es transitiva, entonces existe un abierto $C \neq \varnothing$ con $T^{-1}(C) \subset C$ y $\overline{C} \neq X$. Sea $B \neq \varnothing$ un abierto tal que $\overline{B} \subset C$ y definimos $D = X \setminus \overline{B}$ el cual es abierto, así que $\{C, D\}$ es una cubierta abierta de X.

Sea $E_n = \{(C_1, \ldots, C_i, D_1, \ldots, D_j) : i + j = n, i, j \geq 0\}$ y queremos ver que E_n es una n-cubierta de $(T, \{C, D\})$. Como la preimagen de C está contenida en C, entonces ningún elemento fuera de C puede ir a C. D puede intersecar a C pero no está completamente contenido en C, así que si $x \in X$, entonces puede ser que $T^k(x) \in D$ para todo $0 \leq k < n$, o existe una $k = k(x)$ suficientemente grande tal que, $0 \leq k < n$ y $T^k(x) \notin D$. Como $\{C, D\}$ es una cubierta abierta, entonces $T^{k(x)}(x) \in C$, por lo tanto $T^m(x) \in C$ para todo $m \leq k(x)$ pues $T^{-1}(C) \subset C$ y $T^p(x) \in D$ para $p > k(X)$, lo cual demuestra que E_n es una n-cubierta.

La cardinalidad de E_n es $n + 1$, entonces $M_n(T, \{C, D\}) \leq n + 1$ y por lo tanto la entropía $h(T, \{C, D\}) = \lim_{n\to\infty} \frac{1}{n} \log M_n(T, \{C, D\}) = 0$ lo cual es una contradicción pues la entropía es positiva. Por lo tanto T es transitiva.

Tomemos ahora un factor no trivial y veamos si tiene entropía positiva. Sea (G, Y) un sistema dinámico tal que $G\colon Y \to Y$ y $H\colon X \to Y$ son mapeos continuos, H es sobre y $G \circ H = H \circ T$. Supongamos que $\{C, D\}$ es una cubierta abierta para Y con $\overline{C} \neq \overline{Y} \neq \overline{D}$. Como H es continua, $\{H^{-1}(C), H^{-1}(D)\}$ es una cubierta abierta para X y entonces tiene entropía topológica positiva. H y H^{-1} inducen una biyección entre las n-coberturas de $(T, \{H^{-1}(C), H^{-1}(D)\})$ y $(G, \{C, D\}) = H(T\{H^{-1}(C), h^{-1}(D)\})$ por lo que la entropía topológica es positiva.

□

Teorema 3.3. *Si (X, T) es un sistema dinámico con la propiedad de especificación, entonces T tiene entropía uniformemente positiva.*

Demostración. Sea $\{C, D\}$ una cubierta abierta de X, $\epsilon > 0$ y $p, q \in X$ de forma que $B_\epsilon(p) \subset C \setminus D$ y $B_\epsilon(q) \subset D \setminus C$. Sea $N = N(\epsilon)$ el entero que forma una especificación N espaciada y definimos $M = N + 1$. Ahora definamos el conjunto $A = \{x = (a_0, \ldots, a_{n-1}) \in \prod_{k=0}^{n-1} X$ tal que $T^{kM}(a_k) = p$ o $T^{kM}(a_k) = q$ para todo $a_k\}$.

Para cada $x \in A$, podemos construir la especificación $\xi = \{T^{kM}(a_k)\}_{k=0}^{n}$ donde cada a_k es uno de los elementos que forman a x. Esta especificación es N-espaciada pues

$$(k + 1)M - kM = (k + 1)(N + 1) - k(N + 1) = N + 1.$$

Entonces existe un $y \in X$ tal que $\rho(T^{kM}(y), T^{kM}(a_k)) < \epsilon$.

Tomemos $x_p = (a_0, \ldots, a_{n-1}), x_q = (b_0, \ldots, b_{n-1}) \in A$ tal que $x_p \neq x_q$. Esto quiere decir que existe k tal que $a_k \neq b_k$ y sin pérdida de generalidad podemos tomar a $T^{kM}(a_k) = p$ y $T^{kM}(b_k) = q$. Usando la especificación formada por x_p obtenemos que existe y_p tal que $\rho(T^{kM}(y_p), T^{kM}(a_k)) < \epsilon$ y por lo tanto $\rho(T^{kM}(y_p), p) < \epsilon$, así que $T^{kM}(y_p) \in B_\epsilon(p) \subset C \setminus D$. Podemos repetir este proceso para x_q y obtenemos un y_q tal que $T^{kM}(y_q) \in B_\epsilon(q) \subset D \setminus C$.

Sea E_n una $nM-$cubierta de $(T, \{C, D\})$ para $x \in X$. Elijamos un elemento $F_n(X) \in E_n$ donde $F_n(x) = (F_n^0(x), \ldots, F_n^{nM-1}(x))$ y $T^j(x) \in F_n^j(x)$ para $0 \leq j < nM$. Como sabemos que $T^{kM}(y_p) \in B_\epsilon(p)$ y $\{C, D\}$ es una cubierta, entonces el abierto que cubre a $F_n^{kM}(y_p)$ es C y por lo tanto $F_n^{kM}(y_p) = C$, y análogamente $F_n^{kM}(y_q) = D$ por lo que $F_n^{kM}(y_p) \neq F_n^{kM}(y_q)$. Obtuvimos dos abiertos posibles para cada uno de los F_n y card$E_n \geq 2^n$, por lo tanto $M_{nM}(T, \{C, D\}) \geq 2^n$ así que $h(T, \{C, D\}) \geq \lim \frac{1}{nM} \log 2^n = \frac{1}{M} \log 2 > 0$. Por lo tanto T tiene entropía positiva para cualquier cubierta abierta $\{C, D\}$, y por lo tanto tiene entropía uniformemente positiva.

□

Definición 3.9. Dado un sistema dinámico (X, T) decimos que es *débilmente mezclado* si el sistema producto $(X \times X, T \times T)$ es topológicamente transitivo.

Lema 3.2. *Un sistema dinámico (X, T) no es débilmente mezclado si y solo si existen dos conjuntos abiertos no vacíos U, V tal que para toda $n \in \mathbb{N}$ no se puede tener simultáneamente $U \cap T^n(U) \neq \varnothing$ y $U \cap T^n(V) \neq \varnothing$ (ver [1]).*

Proposición 3.2. *Un sistema dinámico con entropía uniformemente positiva es débilmente mezclado.*

Demostración. Supongamos que el sistema (X, T) no es débilmente mezclado. Entonces podemos tomar dos conjuntos abiertos, no vacíos U, V como en el lema anterior. Definamos la cubierta abierta U', V' para X tal que $V^c \subset U'$ y $U^c \subset V'$. Si $n \in \mathbb{N}$, $i \in \{0, 1, \ldots, n\}$, entonces, por el lema anterior, tenemos dos posibilidades: Por un lado, puede suceder que $U \cap T^i U \neq \varnothing$, por lo que $U \cap T^i V = \varnothing$, en cuyo caso $U \subset (T^i V)^c \subset T^i U'$. La otra posibilidad es que $U \cap T^i U = \varnothing$, en cuyo caso $U \subset T^i V'$. El conjunto U está contenido en la intersección de estos conjuntos para cada i, es decir que $U \subset W_0 \cap TW_1 \cap \ldots \cap T^n W_n = W$ donde cada W_i puede ser U' o V'.

Tomemos $x \in X$ de forma que $x \in U$, por lo tanto $x \in W$. Podemos iterar a x tal que $T(x) \in T(W)$ y así sucesivamente hasta que $T^n(x) \in T^n(W)$, por lo que si tomamos la n-cubierta E_n de $(T, \{U', V'\})$ entonces el elemento $F_n(x)$ para toda $x \in U$ es $F_n(x) = (W, T(W), \ldots, T^n(W))$. Ahora si $y \notin U$ entonces $y \in V'$ por lo que puede ser que toda la órbita de tamaño n de y se encuentre en V' o que exista un $m < n$

tal que $T^n(y) \in U$, por lo que en este caso el elemento de la n-cubierta es $F_n(y) = (V', T(V'), \ldots, T^{m-1}(V'), W, T(w), \ldots, T^{n-m}(W))$. La cardinalidad de la n-cubierta E_n es $n+1$ y $M_n(T, \{U', V'\}) \leq n+1$ y por lo tanto $h(T, \{U', V'\}) = 0$ lo cual contradice que nuestro sistema dinámico tiene entropía uniformemente positiva, así que el sistema es débilmente mezclado.

□

El hecho de que un sistema dinámico con la propiedad de especificación tenga entropía uniformemente positiva nos dice que la especificación implica importantes propiedades topológicas como lo es ser débilmente mezclado o tener entropía completamente positiva. Partiendo de estas definiciones, introducimos la definición de los sistemas k-topológicos, de forma que se tenga cierta relación con la definición de los k-sistemas. [20]

Definición 3.10. Sea (X, T) un sistema dinámico, decimos que es un *k-sistema topológico* si para cada cubierta finita abierta de conjuntos no densos de X tiene entropía topológica positiva.

Dependiendo del tamaño de la cubierta finita decimos que el k-sistema topológico es un sistema con entropía uniforme positiva de orden n, de esta forma un sistema con entropía uniformemente positiva es un sistema con entropía uniforme positiva de orden 2. Es sencillo ver que si tenemos un sistema dinámico sobreyectivo con la propiedad de especificación entonces es un k-sistema topológico, y cada k-sistema topológico es mezclado.

Ya que estudiamos varias relaciones de la especificación con la entropía de sus coberturas, estudiaremos la ergodicidad de los sistemas con especificación y su relación con la entropía de medida teórica

Recordemos que una medida μ es ergódica si se cumple que para cada conjunto A tal que $T^{-1}(A) = A$ entonces $\mu(A) = 0$ o 1.

Teorema 3.4. *Sea (X, T) un sistema dinámico invertible y expansivo con la propiedad de especificación. Entonces $h(T) = \lim_{n\to\infty} \frac{1}{n} \log(\mathrm{Per}_n(T))$.*

Demostración. Sea $\mathscr{A}$ una cubierta finita con $\mathrm{diam}(A) < \delta$ para todo $A \in \mathscr{A}$ y δ es la constante de expansividad. Definimos el *número de Lebesgue* $\beta > 0$ como el número tal que cualquier bola cerrada de radio β está completamente contenido en algún elemento de la cubierta. Sea Q un conjunto (n, β)-separado maximal, esto quiere decir que, para todo $x, y \in Q$ y $0 \leq t \leq n$ se tiene que $d(T^t(x), T^t(y)) > \beta$. Para $z \in Q$ definamos $B(z) = (A_0(z), \ldots, A_{n-1}(z))$ donde $A_k(z) \in \mathscr{A}$ y $\mathrm{cl}(B_\beta(T^k(z))) \subset A_k(z)$.

Veamos que $E_n = \{B(z)|z \in Q\}$ es una n-cubierta para $(T, \mathscr{A})$, para cada $x \in X$. Supongamos que $x \in Q$. Entonces por la definición de $B(x) \in E_n$ tenemos que cumple las características que pedimos para una n-cubierta. Ahora sea $x \in X \setminus Q$. Entonces existe $z_x \in Q$ tal que $d(T^k(x), T^k(z_x)) \leq \beta$ para todo $0 \leq k < n$, pues si no existiera este z_x entonces $Q \cup \{x\}$ sería un conjunto (n, β)-separado más grande que Q lo cual no es posible pues Q es maximal. Dado que $T^k(x) \in B_\beta(T^k(z_x)) \subset A_k(z_x)$, entonces $B(z_x) \in E_n$ es el elemento que cumple la propiedad de una n-cubierta para $x \notin Q$, por lo tanto E_n es una n-cubierta. Definamos como $N(n, \beta)$ la cardinalidad de un conjunto (n, β)-separado maximal. Como este último es precisamente Q y la cardinalidad de E_n es el número de elementos en Q, la cardinalidad de E_n es entonces $N(n, \beta)$ y por lo tanto $M_n(T, \mathscr{A}) \leq N(n, \beta)$.

Sea E una n-cubierta para $(T, \mathscr{A})$ y R un conjunto (n, δ)-separado. Para $x \in R$ definamos a la función $g(x) = (A_0(x), \ldots, A_{n-1}(x)) \in E$, es decir que $T^k(x) \in A_k(x)$ para todo $0 \leq k < n$. Si $g(x) = g(y)$, entonces $A_k(x) = A_k(y)$ y por lo tanto $d(T^k(X), T^k(y)) \leq \text{diam} A_k(X) < \delta$ para $0 \leq k < n$. Entonces, como R es un conjunto (n, δ)-separado, solo nos queda que $x = y$. Entonces la función $g\colon R \to E$ es una función inyectiva y card$E \geq$card$R$ y $M_n(T, \mathscr{A}) \geq N(n, \delta)$.

Sea $\text{Per}_n(U) = \{x \in U : T^n(x) = x\}$ el conjunto de puntos periódicos de T en el conjunto U y también definimos $N_n(U) =$card$(\text{Per}_n(U))$. Usaremos el siguiente lema debido a Bowen [2]

Lema 3.3. *Para cualquier* $\alpha > 0$ *se cumple que* $\liminf_{d\to\infty} \frac{N_d(t)}{N(d,\alpha)} > 0$.

Aplicando el lema a nuestras cantidades inferimos que existe un $S > 0$ y n_0 tal que $N_n(T) \geq SN(n, \beta)$ para $n \geq n_0$. Como $M_n(T, \mathscr{A}) \leq N(n, \beta)$, deducimos que $SM_n(T, \mathscr{A}) \leq SN(n, \beta)$ y $N_n(T) \geq SM_n(T, \mathscr{A})$.

Ahora tomemos dos elementos $p \neq q$ que pertenezcan a $\text{Per}_n(T)$. Entonces como el sistema es expansivo y son elementos diferentes, existe $t \in \mathbb{N}$ tal que $d(T^t(p), T^t(q)) > \epsilon$, pero al ser puntos periódicos de periodo n, tenemos que $t < n$ y por lo tanto pertenecen al conjunto (n, δ)-separado, así que $N(n, \delta) \geq N_n(T)$ y por lo tanto $N_n(T) \leq M_n(T, \mathscr{A})$ para todo $n \geq n_0$.

Dado que $(1/n) \log M_n(T, \mathscr{A})$ aproxima el valor del límite $h(T, \mathscr{A})$ entonces también lo hace $(1/n) \log N_n(T)$ pues $SM_n(T, \mathscr{A}) \leq N_n(T) \leq M_n(T, \mathscr{A})$ y el límite para $\frac{1}{n} \log(SM_n(T, \mathscr{A}))$ es igual que el límite para $\frac{1}{n} \log(M_n(T, \mathscr{A}))$. Por lo tanto obtenemos

$$h(T) = h(T, \mathscr{A}) = \lim_{n\to\infty} \frac{1}{n} \log N_n(T). \tag{3.7}$$

□

Un sistema dinámico que tiene una única medida de entropía máxima se dice que es *intrínsecamente ergódico*. Nuestra meta es demostrar que un sistema dinámico expansivo con la propiedad de especificación es intrínsecamente ergódico. Para esto, primero debemos demostrar que al menos existe una medida con entropía máxima: veremos que cualquier sistema dinámico expansivo lo cumple, es decir, tiene al menos una medida con entropía máxima. Para llegar a esta propiedad debemos estudiar antes los generadores. [6]

Definición 3.11. Una cubierta finita $\mathscr{A} = \{A_i \text{ tal que } 1 \leq i \leq s\}$ de X es un *generador de* T si para cada sucesión $\{k_i\}_{i\in\mathbb{Z}}$ con $1 \leq k_i \leq s$ la intersección $\bigcap_{i\in\mathbb{Z}} T^i\overline{A_{k_i}}$ contiene a lo más un punto. Definiremos a s como la *potencia* del generador de $\mathscr{A}$.

Teorema 3.5. *Sea (X, T) un sistema dinámico expansivo, y sea $\epsilon > 0$ la constante de expansión para T. Entonces cada cubierta abierta $\mathscr{A} = \{A_i$ tal que $1 \leq i \leq s\}$ de X tal que* $\operatorname{diam} A_i < \epsilon$ *es un generador de T. De igual manera si $\mathscr{A} = \{A_i$ tal que $1 \leq i \leq s\}$ es un generador de T, entonces T es expansivo y cada número de Lebesgue es una constante expansiva para T.*

Demostración. Si (X, T) es un sistema dinámico expansivo, entonces existe $\epsilon > 0$ tal que para cualquier par de puntos distintos $x, y \in X$ existe una $n \in \mathbb{Z}$ tal que $d(T^n(x), T^n(y)) \geq \epsilon$. Entonces no existen dos puntos $x, y \in \bigcap_{i\in\mathbb{Z}} T^i\overline{A_{k_i}}$ pues x, y pertenecerían a A_{k_1} y por lo tanto la distancia entre ellos sería menor o igual a ϵ. Pero de igual manera x, y pertenecen a todo los conjuntos $T^n(A_{i_n})$ y por ser expansivo en algún momento la distancia entre uno de los conjuntos tiene que ser mayor a ϵ por lo tanto no puede contener los dos puntos de la intersección y entonces es un generador.

Ahora para el regreso sea $\mathscr{A} = \{A_i \text{ tal que } 1 \leq i \leq s\}$ un generador de T y $\epsilon > 0$ el número de Lebesgue para $\mathscr{A}$, es decir, el diámetro mas pequeño que debe de tener un conjunto para estar en uno de los conjuntos A_is. Supongamos que existen $x \neq y \in X$ tales que $d(T^n(x), T^n(y)) < \epsilon$ para toda $n \in \mathbb{Z}$. Entonces para cada $n \in \mathbb{Z}$ existe $A_{i_n} \in \mathscr{A}$ tal que $B_\epsilon(T^{-n}(x)) \subset A_{i_n}$ por ser número de Lebesgue, y esto también se cumple para y. Por lo tanto $x, y \in \bigcap_{i\in\mathbb{Z}} T^i\overline{A_{k_i}}$ lo cual es una contradicción pues $\mathscr{A}$ es generador.

□

Definición 3.12. Sea $X_0 \subset X$. Una partición α de X a lo más contable es llamada un *generador topológico de* X_0 si las siguientes condiciones se satisfacen:

1. $\bigcup_{A\in\alpha} \mathrm{int}A$ es denso en X e $\mathrm{int}A \neq \varnothing$ para cada $A \in \alpha$.
2. Para cada $x \in X_0$ y para cada sucesión $A_{i_k} \in \alpha$ con $(k \in \mathbb{Z})$ tales que

$$x \in \bigcap_{n\in\mathbb{Z}} \overline{\bigcap_{k=-n}^{n} T^k \mathrm{int} A_{i_k}}$$

se sigue que $\{x\} = \bigcap_{n\in\mathbb{Z}} \overline{\cap_{k=-n}^{n} T^k \mathrm{int} A_{i_k}}$.

Los generadores topológicos lo que hacen es separar los puntos de un conjunto T-invariante. Son objetos de estudios muy grandes en la teoría de los sistemas dinámicos, aunque eso queda fuera del alcance del trabajo. Uno de los resultados más importantes es que cuando el sistema tiene un generador topológico finito, entonces es factor de un sistema simbólico. La razón de que estudiemos los generadores de T es para mostrar que podemos encontrar generadores topólogicos finitos, mas aún acabamos de mostrar que los generadores de T están relacionados con los sistemas expansivos y vimos que ser expansivo depende de la cubierta y no de la medida.

Proposición 3.3. *Sean T una transformación expansiva y $\epsilon > 0$ la constante de expansión para T. Para cada partición $\alpha = (A_1, \ldots, A_s)$ de X, si $X = \overline{\cup_{i=1}^{s} \mathrm{int} A_i}$ y $\mathrm{diam} A_i < \epsilon$, entonces α es un generador topológico de X*

Demostración. Sabemos que existe una $\delta > 0$ tal que el $\max_i \mathrm{diam}\ A_i < \epsilon - \delta$. Definamos la cubierta abierta $\mathscr{A} = \{B_\delta(\mathrm{int}A_i) | 1 \leq i \leq s\}$ de X. Como $\delta < \epsilon$, podemos aplicar el teorema anterior y $\mathscr{A}$ es un generador de T, entonces la intersección $\bigcap_{m\in\mathbb{Z}} T^m \overline{B_\delta(\mathrm{int}A_{i_m})}$ contiene a lo más un elemento, y como $\bigcap_{n\in\mathbb{Z}} \overline{\bigcap_{k=-n}^{n} T^{-k} \mathrm{int} A_{i_k}} \subset \bigcap_{m\in\mathbb{Z}} T^m \overline{B_\delta(\mathrm{int}A_{i_m})}$ obtenemos que es un generador topológico, ya que tiene a lo más un elemento.

□

Definición 3.13. La σ-álgebra más chica de subconjuntos de X que contiene todos los subconjuntos abiertos es denotado por $\mathfrak{B}(X)$ y es llamada la *σ-álgebra de Borel*. Sus elementos son llamados los *conjuntos de Borel*. Los conjuntos de Borel también pueden ser definidos como los conjuntos que están formados por uniones contables, intersecciones contables y complementos de abiertos. Una *partición de Borel* es una partición hecha por conjuntos de Borel.

Definición 3.14. Sea $m \in \mathfrak{M}_T$ y A un conjunto medible, donde $\mathfrak{M}_T$ en el conjunto de las medidas T-invariantes . Decimos que es un *conjunto m-continuo* si la medida de la frontera de A es cero.

Lema 3.4. *Sea X un espacio métrico compacto y $\mathscr{U} = \{U_i$ tal que $1 \leq i \leq s\}$ una cubierta abierta de X. Entonces existe una partición de Borel $\alpha = \{A_i$ tal que $1 \leq i \leq s\}$ tal que $A_i \subset U_i$. También es posible encontrar una partición $\alpha = \{A_i$ tal que $1 \leq i \leq s\}$ tal que para cada $1 \leq i \leq s$ se tiene que $A_i \subset U_i$ y $A_i \subset \overline{intA_i}$. Más aún, si $m \in \mathfrak{M}_T(X)$ es una medida dada, podemos elegir la partición de Borel α tal que α es una partición m-continua.*

Demostración. Definamos la cubierta abierta $\mathscr{U}' = \{U_i'$ tal que $1 \leq i \leq s\}$ y $\delta > 0$ tal que $\overline{B_\delta(U_i')} \subset U_i$ para cada $1 \leq i \leq s$. Esta cubierta está bien definida porque para cualquier cubierta podemos encontrar bolas que están completamente contenidas en cada conjunto de la cubierta.

Dada $m \in \mathfrak{M}_T(X)$ podemos dar un $0 < \delta_0 < \delta$ tal que $m(\overline{B_{\delta_0}(U_i')}) = m(B_{\delta_0}(U_i'))$ para toda $1 \leq i \leq s$, ya que los conjuntos son abiertos y la medida es regular. Sea $U = \{B_{\delta_0}(U_i')$ tal que $1 \leq i \leq s\}$ y definamos de manera inductiva $A_1 = \overline{B_{\delta_0}(U_1')}$ y $A_i = \overline{B_{\delta_0}(U_i')} - \bigcup_{k=1}^{i-1} A_k$ con $2 \leq i \leq s$. La cerradura de A_i está contenida en U_i pues está formada por un conjunto más chico que $\overline{B_{\delta_0}(U_i')}$ el cual está contenido en U_i.

A_i es medible por Borel, ya que está formado por conjuntos medibles de Borel. En A_1 su medida es $m(A_1) = m(\overline{B_{\delta_0}(U_1')}) = m(B_{\delta_0}(U_1'))$ por lo cual la medida de su frontera es cero y por lo tanto es un conjunto m-continuo. Por inducción podemos ver que A_i es un conjunto m-continuo, pues en cada A_i le estamos restando a un conjunto m-continuo, la unión de conjuntos m-continuos, por lo que al tomar la cerradura tenemos puntos de frontera que por ser m-continuos su medida es cero. Esta es la partición de Borel, pues está formada por conjuntos medibles por Borel y m-continuos. □

Definición 3.15. Una función f es *semi-continua por arriba en el punto x* si para todo $\epsilon > 0$ existe $\delta > 0$ tal que $f(y) < f(x) + \epsilon$ para todo $y \in B_\delta(x)$, y decimos que la función es *semi-continua por arriba* si es semi-continua por arriba en todo punto x.

Proposición 3.4. *Si T es expansiva entonces la función $h_m(T)\colon \mathfrak{M}_T(X) \to \mathbb{R}$ es semi-continua por arriba con respecto a la topología débil en $\mathfrak{M}_T(X)$.*

Demostración. Tomemos la sucesión de medidas $\{m_n\}_{n\in\mathbb{N}} \in \mathfrak{M}_T(X)$ y sea $m \in \mathfrak{M}_T(X)$ un límite débil de la sucesión. Usando el lema anterior sabemos que existe una partición de Borel $\alpha = (A_1, \ldots, A_s)$, la cual es un generador para cada medida m y m_n tal que, cada

A_i es un conjunto m-continuo, entonces el $\lim_{n\to\infty} m_n(A) = m(A)$ para cada $A \in (\alpha)_0^{k-1} = \bigvee_{0\leq i\leq k-1} T^{-i}\alpha$, es decir un refinamiento de los conjuntos abiertos y sus órbitas.

Sea $\epsilon > 0$ entonces $h_m(T) = h_m(T,\alpha) \geq \frac{1}{k}H_m((\alpha)_0^{k-1}) - \epsilon$ ya que la igualdad la alcanzaría si tomáramos el límite. Dado que $(\alpha)_0^{k-1}$ es finito existe una $N \in \mathbb{N}$ tal que para todo $n \geq N$ $\frac{1}{k}|H_{m_n}((\alpha)_0^{k-1}) - H_m((\alpha)_0^{k-1})| < \epsilon$.

Por lo tanto $h_m(T) \geq \frac{1}{k}H_m((\alpha)_0^{k-1}) - \epsilon \geq \frac{1}{k}H_{m_n}((\alpha)_0^{k-1}) - 2\epsilon \geq h_{m_n}(T,\alpha) - 2\epsilon = h_{m_n}(T) - 2\epsilon$ Lo cual implica semicontinuidad por arriba pues $h_m(T) + \epsilon > h_{m_n}(T)$.

□

Lo importante de que una función sea semi-continua por arriba es que la función tiene una cota superior y alcanza su supremo, por lo tanto existe una medida de entropía maximal y esto ocurre para cualquier sistema dinámico expansivo, solo nos falta ver que esta medida sea única, y para esto vamos a necesitar la especificación.

Definición 3.16. Dado un espacio medible X, un subconjunto A y un elemento $x \in X$, definimos la *medida de Dirac* como

$$\delta_x(A) = \left\{ \begin{array}{ll} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{array} \right\}. \tag{3.8}$$

Definición 3.17. Denotemos por μ_n la medida $\frac{1}{\mathrm{Per}_n(T)}\sum_{x\in P^n(T)} \delta_x(A)$, y por μ_B el punto límite de la sucesión μ_n, donde δ_x es la medida de Dirac.

Las medidas μ_n son T-invariantes ya que si tomamos el conjunto B y obtenemos su medida $\mu_n(B)$, entonces por cada punto de periodo menor o igual a n, su preimagen $T^{-1}(B)$ también tiene un punto de periodo menor o igual a n. La medida μ_B, la cual llamaremos *medida de Bowen*, también es T-invariante pues es límite de medidas T-invariantes.

Lema 3.5. *Para $\epsilon > 0$ suficientemente pequeña y para un sistema (X,T) expansivo y con la propiedad de especificación, existe un $A_\epsilon > 0$ tal que $\mu_B(B_\epsilon^n(x)) \geq A_\epsilon e^{-nh}$ para todo $x \in X$ y $n \geq 1$.*

Demostración. Definimos h como la entropía topológica y el conjunto $B_\epsilon^n(x) = \{y \in X$ tal que $d(T^j(x), T^j(y)) \leq \epsilon$ para $0 \leq j < n\}$. Sea $E \subset X$ un conjunto $(m, 3\epsilon)-$separado, y definamos $r = n + m + 2M$ con $M(\epsilon)$ siendo el tamaño de la especificación de acuerdo a la definición. Dado que el sistema es expansivo y tiene especificación, sabemos que el sistema

tiene especificación periódica, entonces para cualquier $z \in E$ podemos tomar la especificación $\{(T^{[0,n]}(x)), (T^{[n+M,n+M+m]}(z))$ donde existe $x_z \in P^r(T)$ tal que $d(T^j(x_z), T^j(x)) \leq \epsilon$ para $0 \leq j < n$ y $d(T^{n+M+j}(x_z), T^{n+M+j}(z)) \leq \epsilon$ para $0 \leq j < m$.

Tomemos dos elementos diferentes $z, w \in E$. Entonces $x_z \neq x_w$. Estos puntos que se construyen por la especificación pertenecen a $B^n_\epsilon(x)$ y también a $P^r(T)$ y como para cada punto diferente en E obtenemos un punto diferente que lo rastree por la especificación, entonces $\text{card}(B^n_\epsilon(x) \cap P^r(T)) \geq \text{card}E = s_m(3\epsilon)$.

Sabemos que $\mu_r(B^n_\epsilon(x)) = \frac{1}{\text{Per}_r(T)} \sum_{x \in P^r(T)} \delta_x(B^n_\epsilon)$ de tal forma que los puntos que tienen medida de Dirac 1 son aquellos que viven en nuestra bola y tienen periodo r, es decir, $\mu_r(B^n_\epsilon(x)) = \frac{1}{\text{Per}_r(T)} \text{ card}(B^n_\epsilon(x) \cap P^r(T)) \geq \frac{s_m(3\epsilon)}{\text{Per}_r(T)}$. Los siguientes lemas debidos a Denker nos ayudarán a acotar la medida del conjunto.

Lema 3.6. *Para ϵ suficientemente pequeño, tenemos que*

$$D_\epsilon^{-1} e^{hn} \leq s_n(\epsilon) \leq E_\epsilon^{-1} e^{hn} \tag{3.9}$$

para toda $n \geq 1$.

Lema 3.7. *Existen constantes $0 < D < E$ tales que para n suficientemente grande se cumple que*

$$De^{hn} \leq \text{Per}_n(T) \leq Ee^{hn}. \tag{3.10}$$

Entonces combinando estos lemas obtenemos la desigualdad $\frac{s_m(3\epsilon)}{\text{Per}_r(T)} \geq \frac{e^{-hr}}{E} \frac{e^{hm}}{D_{3\epsilon}} = A_\epsilon e^{-hn}$ con $A_\epsilon = (ED_{3\epsilon} e^{2Mh})^{-1}$.

Veamos ahora que $B^n_\epsilon(x)$ es cerrado. Supongamos que no lo es, por lo tanto existe una sucesión convergente $\{y_m\}$ tal que su límite y no está en $B^n_\epsilon(x)$, es decir que existe $0 \leq k < n$ tal que $d(T^k(x), T^k(y)) > \epsilon$. Como la sucesión converge y la transformación es continua, $d(T^k(y), T^k(y_m)) \leq \epsilon_0$ para m suficientemente grande, tomando ϵ_0 suficientemente chico y dado que la sucesión está en el conjunto y $d(T^k(x), T^k(y_m) \leq \epsilon$, obtenemos la contradicción.

Si dejamos que $r \to \infty$ y como el conjunto $B^n_\epsilon(x)$ es cerrado, entonces como la sucesión de las medidas tiende a la medida de Bowen obtenemos que $\mu_B(B^n_\epsilon(x)) \geq \limsup_{r \in S} \mu_r(B^n_\epsilon(x)) \geq A_\epsilon e^{-hn}$.

□

Lema 3.8. *Para una $\epsilon > 0$ suficientemente pequeña existe un $\delta = \delta(\epsilon)$ y una constante B_ϵ tal que $\mu_B(B^n_\delta(y)) \leq B_\epsilon e^{-nh}$ para todo $y \in X$ y $n \geq 1$.*

Demostración. Sea ϵ dado y $M = M(\epsilon)$ la constate que se obtiene por la especificación. Elijamos $\delta = \delta(\epsilon) < \frac{\epsilon}{6}$ de forma que si $d(x,y) < 6\delta$ entonces $d(T^j(X), T^j(y)) \leq \epsilon$ para $-M \leq j \leq M$. Definamos $r = n + m + 2M$. Dado que el sistema es expansivo, cualesquiera dos puntos x, z en $P^r(T) \cap B^n_{3\delta}(y)$ son (r,ϵ)-separados. Dado que $d(T^j(x), T^j(z)) \leq \epsilon$ para $-M \leq j \leq M + n$, por la forma en la que lo hemos definido, $T^{n+M}(x)$ y $T^{n+M}(z)$ son (m,ϵ)-separados, y por lo tanto $\mathrm{card}(B^n_{3\delta}(y) \cap P^r(T)) \leq s_m(\epsilon)$. Entonces

$$\mu_r(B^n_{3\delta}(y)) = \frac{1}{\mathrm{Per}_r(T)}\mathrm{card}(B^n_{3\delta}(y) \cap P^r(T)) \leq \frac{1}{\mathrm{Per}_r(T)} s_m(\epsilon).$$

Ocupando los Lemas 3.7 y 3.7 obtenemos

$$\frac{1}{\mathrm{Per}_r(T)} s_m(\epsilon) \leq e^{hm}[De^{hr}E_\epsilon]^{-1} = B_\epsilon e^{-hn}$$

con $B_\epsilon = (De^{2hM}E_\epsilon)^{-1}$.

Ahora sea V un conjunto abierto $\{x \in X$ tal que $d(T^j(x), T^j(y)) < 2\delta$ para $0 \leq j < n\}$. Entonces $B^n_\delta(y) \subset V \subset B^n_{3\delta}(y)$, así que $\mu_r(V) \leq B_\epsilon e^{-hn}$, por lo que si hacemos $r \to \infty$ obtenemos

$$\mu_B(B^n_\delta(y)) \leq \mu_B(V) \leq \lim\inf \mu_r(V) \leq B_\epsilon e^{-hn}.$$

□

Definición 3.18. Una medida $m \in \mathfrak{M}_T(X)$ se llama *T-homogénea* si y solo si para cada $\epsilon > 0$ existe $\delta > 0$ y $c > 0$ tal que la siguiente propiedad se cumple: para cada $n \in \mathbb{N}$ y $x, y \in X$,

$$c \cdot m\left(\bigcap_{k=0}^{n-1} T^{-k}\overline{B_\epsilon(T^k(x))}\right) \geq m\left(\bigcap_{k=0}^{n-1} T^{-k}\overline{B_\delta(T^k(y))}\right). \tag{3.11}$$

Lema 3.9. *Si $m \in \mathfrak{M}_T(X)$ es T-homogénea, entonces $m \in \mathfrak{M}_{\text{máx}}T$, es decir, es una medida maximal (de entropía máxima).*

Demostración. Empecemos definiendo el conjunto $B^n_\epsilon(x) = \bigcap_{k=0}^{n-1} T^{-k}\overline{B_\epsilon(T^k(x))}$ y elijamos $\delta > 0$ y $c > 0$ tal que $c \cdot m(B^n_{\epsilon/2}(x)) \geq m(B^n_\delta(y))$.

Sea E un conjunto (n,ϵ)-separado maximal. Para dos elementos $x, y \in E$ obtenemos que $B^n_{\epsilon/2}(x) \cap B^n_{\epsilon/2}(y) = \varnothing$, pues dependen de la vecindades de las órbitas y sabemos que están separadas por al menos ϵ. Tomando $s_n(\epsilon, T)$ como la cardinalidad de un conjunto

maximal (n, ϵ)-separado y como sin pérdida de generalidad la medida de todo el espacio es 1, obtenemos

$$1 \geq \sum_{x \in E} m(B^n_{\epsilon/2}(x)) \geq s_n(\epsilon, T) \frac{\sup_{z \in X} m(B^n_\delta(z))}{c} \tag{3.12}$$

y por lo tanto $s_n(\epsilon, t) \leq \frac{c}{\sup_{z \in X} m(B^n_\delta(z))}$.

Sea $\alpha = \{A_1, \ldots A_N\}$ una partición m medible tal que el diámetro de cada uno de los conjuntos A_i sea menor a δ. Para cada A que pertenece a $(\alpha)^{n-1}_0$ existe una $z \in X$ tal que $A \subset B^n_\delta(z)$, pues A es parte de un refinamiento de conjuntos con diámetro menor a δ y sus órbitas, así que el diámetro tiene que ser menor a δ y $B^n_\delta(z)$ es el seguimiento de las preimágenes de la órbita de una vecindad de tamaño δ. Entonces $m(A) \leq \sup_{z \in X} m(B^n_\delta(z))$. Luego, por propiedades de los logaritmos obtenemos

$$-\sum_{A \in (\alpha)^{n-1}_0} m(a) \log m(A) \geq \log(\sup_{z \in X} m(B^n_\delta(z)))^{-1} \geq \log \frac{s_n(\epsilon, T)}{c}. \tag{3.13}$$

Dividiendo entre n y tomando el límite con n tendiendo al infinito se obtiene $h_m(T) \geq \lim\sup_{n \to \infty} \frac{1}{n} \log \frac{s_n(\epsilon, T)}{c}$ y si $\epsilon \to 0$, entonces $h_m(T) \geq h_{top}(T)$ y por el principio variacional (el supremo de las entropías métricas es la entropía topológica), tiene que ser igual.

□

Teorema 3.6. *La medida de Bowen tiene entropía maximal.*

Demostración. Para probar que la medida de Bowen tiene entropía maximal vamos a ver que es una medida T-homogénea. Los lemas anteriores nos han dado dos desigualdades, $\mu_B(B^n_\epsilon(x)) \geq A_\epsilon e^{-nh}$ y $B_\epsilon e^{-nh} \geq \mu_B(B^n_\delta(y))$, donde $A_\epsilon = (ED_{3\epsilon} e^{2Mh})^{-1}$ y $B_\epsilon = (DE_\epsilon e^{2hM})^{-1}$.

De acuerdo con el Lema 3.6 el valor de E_ϵ depende de forma inversamente proporcional a los valores de $s_n(\epsilon)$, entonces mientras más grande sea la cardinalidad del conjunto (n, ϵ)-separado, más pequeño será E_ϵ. Como en un caso tenemos ϵ y en el otro tenemos 3ϵ, entonces el conjunto $(n, 3\epsilon)$-separado es más grande y por lo tanto su cardinalidad es mayor.

Aplicando todo esto, obtenemos que $\mu_B(B^n_\epsilon(x)) \geq A_\epsilon e^{-nh} \geq B_\epsilon e^{-nh} \geq \mu_B(B^n_\delta(y))$ y como esto se aplica para todas las bolas de Bowen, entonces se cumple que la medida con respecto a las bolas de radio ϵ multiplicadas por una constante es mayor que la medida de las bolas de un radio δ y por lo tanto es T-homogénea y es una media de entropía maximal. □

Estamos muy cerca de lograr demostrar el resultado más importante de especificación. Para llegar, ya tenemos que la medida de Bowen es una medida de entropía maximal, nos falta ver que la media sea ergódica y que sea una medida única.

Lema 3.10. *Sea (X, T) un sistema dinámico expansivo con la propiedad de especificación. Entonces μ_B es ergódica.*

Demostración. Sean $A, B \subset X$ conjuntos compactos y sean U, V dos 2δ-vecindades μ_B-continuas de A, B para algún δ. Recordemos que un conjunto es μ-continuo si la medida de la frontera es cero. Sea $U' = B_\delta(A)$ y $V' = B_\delta(B)$. Usaremos la siguiente proposición, la cual es un corolario de otra proposición más fuerte [6].

Proposición 3.5. *Si T es una transformación expansiva con constante de expansión $\epsilon > 0$, entonces para cada $0 < \delta < \epsilon$ existe una $n(\delta) \in \mathbb{N}$ tal que $d(x, y) \geq \delta$ implica la existencia de $n \in \{-n(\delta), \ldots, n(\delta)\}$ que satisface $d(T^n(x), T^n(y)) \geq \epsilon$.*

Si δ^* es la constante de expansión y elegimos $\epsilon < \delta^*/2$, entonces existe un $N = N(\delta)$ tal que $d(T^j(x), T^j(y)) < \epsilon$ para $-N \leq j \leq N$ implica que $d(x, y) < \delta$. Asumamos $n > N$ y sean s, t enteros positivos. También definamos $M = M(\epsilon)$ dado por la propiedad de especificación como el tamaño del espacio entre las órbitas. Sea $E_s \subset X$ un conjunto $(s, 2\epsilon)$-separado y $E_t \subset X$ un conjunto $(t, 2\epsilon)$-separado.

Sean $I_i = [a_i, b_i) \cap \mathbb{Z}$ para $i = 1, 2, 3, 4$ intervalos enteros con $a_1 = -n, b_1 = n, a_2 = b_1 + M, b_2 = a_2 + s, a_3 = b_2 + M = n + s + 2M, b_3 = a_3 + 2n, a_4 = b3 + M, b_4 = a_4 + t$. Entonces el tamaño entre cada uno de los intervalos es M y definimos a $m = b_4 - a_1 + M = t + s + 4n + 4M$. Ahora asumamos $\bar{z} = (z_1, z_2, z_3, z_4)$ tal que $z_i \in P^{2n}(T) \cap T^{-n}U', z_2 \in E_s, z_3 \in P^{2n}(T) \cap T^{-n}V', z_4 \in E_t$.

Como la distancia entre cada uno de los intervalos es M, podemos tomar la especificación $\zeta = \{T^{[a_i,b_i]}(w_i)\}_{i=1}^4$ tal que existe un x que rastrea la especificación, es decir $d(T^k(x), T^k(w_i)) \leq \epsilon$ para $a_i \leq k \leq b_i$ y $i \in \{1, 2, 3, 4\}$. Luego, como el sistema es expansivo, la especificación puede ser elegida de forma periódica y $T^m(x) = x$. Esto también lo podemos escribir como $d(T^{a_i+j}(x), T^{a_i+j}(w_i)) < \epsilon$ para $0 \leq j \leq b_i - a_i$ con $i \in \{1, 2, 3, 4\}$. Si definimos w_i tal que $T^{a_i}(w_i) = z_i$, obtenemos que $d(T^{a_i+j}(x), T^j(z_i)) < \epsilon$ para $0 \leq j \leq b_i - a_i$ y por lo tanto $x = x(\overline{z})$ es decir que puede ser elegida en función de $\overline{z}$.

Por especificación tenemos que $d(T^{a_1+n}(x), T^n(z_1)) = d(x, T^n(z_1)) < \epsilon$, como $n > N$ y $T^n(z_1) \in U'$ pues $z_1 \in T^{-n}U'$, por lo tanto $x \in U$. De la misma manera podemos ver que, como $z_3 \in T^{-n}V'$, entonces $T^{s+2M+2n}(x) \in V$. Si definimos $r = s + 2(n + M)$, entonces $x \in P^m(T) \cap U \cap T^{-r}V$.

Sea $\bar{z} \neq \bar{z'}$. Demostremos que $x(\bar{z}) \neq x(\bar{z'})$. Supongamos que son iguales. Si suponemos que $z_1 \neq z_1'$, por especificación, ambas órbitas estarían muy cerca pues $d(T^j(z_1), T^j(z_1')) < 2\epsilon < \delta^*$ para todo $0 < j < 2n$, pero sabemos que los dos elementos están en $P^{2n}(T)$ y por lo tanto tiene periodo $2n$ así que estamos tomando sus órbitas completas, por lo tanto $d(T^j(z_1), T^j(z_1')) < 2\epsilon < \delta^*$ para todo $j \in \mathbb{Z}$, pero esto contradice que el sistema sea expansivo. Si suponemos $z_2 \neq z_2'$, entonces, de nuevo por especificación, tenemos que $d(T^j(z_2), T^j(z_2')) < 2\epsilon$ para $0 \leq j < s$, lo cual es una contradicción pues los $z_2 \in E_s$. Esto se puede extender de la misma manera para z_3, z_4 y por lo tanto sabemos que la cantidad de los diferentes $\bar{z}$ depende de la cantidad de cada z_i, por lo tanto serán al menos $\mathrm{Per}_{2n}(T)\mu_{2n}(U')s_s(2\epsilon)\mathrm{Per}_{2n}(T)\mu_{2n}(V')s_t(2\epsilon)$.

Sabemos que $\mu_m(U \cap T^{-r}V) = \frac{1}{\mathrm{Per}_m(T)}\mathrm{card}(P^m(T) \cap U \cap T^{-r}V)$ pues la medida μ_m depende de los puntos m-periódicos que estén en el conjunto. Como $x \in P^m(T) \cap U \cap T^{-r}V$ y encontramos la cota para la cantidad posible de diferentes x,

$$\frac{1}{\mathrm{Per}_m(T)}\mathrm{card}(P^m(T) \cap U \cap T^{-r}V) \geq \frac{1}{\mathrm{Per}_m(T)}(\mathrm{Per}_{2n}(T))^2 s_s(2\epsilon)s_t(2\epsilon)\mu_{2n}(U')\mu_{2n}(V'). \quad (3.14)$$

Ahora aplicando los Lemas 3.6 y 3.7 para $s_s(2\epsilon)$ y $s_t(2\epsilon)$ obtenemos

$$\frac{1}{\mathrm{Per}_m(T)}(\mathrm{Per}_{2n}(T))^2 s_s(2\epsilon)s_t(2\epsilon)\mu_{2n}(U')\mu_{2n}(V') \geq \frac{e^{-hm}}{E}\left(\frac{De^{h2n}}{D_{2\epsilon}}\right)^2 e^{hs}e^{ht}\mu_{2n}(U')\mu_{2n}(V'). \quad (3.15)$$

Por lo tanto

$$\mu_m(U \cap T^{-r}V) \geq R\mu_{2n}(U')\mu_{2n}(V') \quad (3.16)$$

con $R = \frac{D^2}{E(D_{2\epsilon}^2)}e^{h(4n+s+t-m)} = \frac{D^2}{E(D_{2\epsilon}^2)}e^{-4Mh} > 0$ pues $m = t + s + 4n + 4M$.

Dado que t fue elegida de forma independiente de todas las demás constantes y conjuntos, podemos tomar $t \to \infty$ y dejar que m corra a lo largo de toda una sucesión de enteros S con la constante n y s fijas y por lo tanto r fija. Entonces obtenemos que $\mu_B(U \cap T^{-r}V) = \lim_{m \in S} \mu_m(U \cap T^{-r}V) \geq R\mu_{2n}(U')\mu_{2n}(V')$, pues R no depende ni de t ni de m.

Ahora $n > N$ es independiente de δ y ϵ, por lo tanto si hacemos que $n \to \infty$ tal que ahora $2n$ corra a lo largo de la sucesión de enteros S, y dado que $r = s + 2(M + n)$ entonces también tiende al infinito por lo tanto $\liminf_r \mu_B(U \cap T^{-r}V) \geq \liminf_n R\mu_{2n}(U')\mu_{2n}(V') \geq R\mu_B(U')\mu_B(V') \geq R\mu_B(A)\mu_B(B)$. Esto debido a que si tendemos r a infinito es igual a tender a n a infinito y aplicamos la definición de medida de Borel, además por definición $A \subset U'$ y $B \subset V'$.

Dado que $\mu_B(A \cap T^{-r}B) \geq \mu_B(U \cap T^{-r}V) - \mu_B(U \setminus A) - \mu_B(V \setminus B)$ pues $A \subset U$ y $B \subset V$ se sigue que $\liminf_r \mu_B(A \cap T^{-r}B) \geq R\mu_B(A)\mu_B(B) - \mu_B(U \setminus A) - \mu_B(V \setminus B)$.

Como ϵ fue elegido independiente de δ, M depende de ϵ, y R depende de M entonces tampoco dependen de δ, así que podemos tomar $\delta \to 0$ y los conjuntos U y V son prácticamente A y B por lo tanto $\mu_B(U \setminus A) \to 0$ y $\mu_B(V \setminus B) \to 0$. Finalmente $\liminf_r \mu_B(A \cap T^{-r}B) \geq R\mu_B(A)\mu_B(B)$.

Ahora sean $P, Q \subset X$ tales que tienen medidas de Borel positivas. Por regularidad existen conjuntos compactos $A \subset P$ y $B \subset Q$ tales que sus medidas de Borel son igualmente positivas, por lo tanto $\liminf_r \mu_B(P \cap T^{-r}Q) > 0$ lo cual dice que la órbita de un elemento pasa por todos los conjuntos de medida positiva, por lo tanto la medida es ergódica.

□

A continuación daremos unos lemas que nos ayudarán para demostrar nuestro teorema importante

Lema 3.11. *Sean X un espacio métrico compacto y $U = \{U_i$ tal que $1 \leq i \leq s\}$ una cubierta abierta de X. Entonces existe una partición de Borel $\alpha = (A_i$ tal que $1 \leq i \leq s)$ de forma que $A_i \subset U_i$.*

Lema 3.12. *Si $a_1, \ldots, a_m$ son reales mayores que cero tales que $0 < a_1 + \ldots + a_m = s \leq 1$ entonces*

$$-\sum a_i \log a_i \leq s \log m - s \log s \leq s \log m + \frac{1}{2}. \tag{3.17}$$

Lema 3.13. *Sea ϵ una constante de expansión, $n \in \mathbb{Z}$ y α una partición de X tal que diam $T^j A < \epsilon$ para $A \in \alpha$ y $0 \leq j < n$ entonces*

$$\frac{1}{n} H_\mu(\alpha) \geq h_\mu(T). \tag{3.18}$$

Teorema 3.7. *Si (X, T) satisface expansividad y la propiedad de especificación entonces (X, T) es intrínsecamente ergódico.*

Demostración. Recordemos que un sistema es intrínsecamente ergódico si tiene una única medida de entropía maximal ergódica. Sabemos que μ_B es maximal y ergódica entonces asumamos que existe $\nu \in \mathfrak{M}_T(X)$ tal que $h_\nu(T) = h$, así que queremos mostrar que $\nu = \mu_B$.

Dado que la entropía es un mapeo afín, podemos suponer que ν es ergódica, como suponemos que ν es diferente a μ_B, entonces existe un conjunto T-invariante $C \subset X$ tal que $\mu_B(C) = 0$ y $\nu(C) = 1$.

Sea $E \subset X$ un conjunto $(n, 2\epsilon)$-separado maximal y ϵ suficientemente pequeño. Como E es maximal entonces es $(n, 2\epsilon)$-extendido por lo tanto $X \subset \bigcup_{x \in E} B^n_{2\epsilon}(x)$ y si $x \neq y$ pertenecen a E entonces $B^n_\epsilon(x) \cap B^n_\epsilon(y) = \emptyset$.

Por el Lema 3.11, para cada x existe un conjunto de Borel A_x con $B^n_\epsilon(x) \subset A_x \subset B^n_{2\epsilon}(x)$ tal que $\alpha_n = (A_x \text{ con } x \in E)$ es una partición de Borel. Ahora por expansibilidad tenemos que mientras $n \to \infty$, diam $T^{[\frac{n}{2}]} B^n_{2\epsilon}(x) \to 0$ pues podemos tomar ϵ lo suficientemente chico tal que 2ϵ sea menor que la constante de expansión. Por lo anterior diam$T^{[\frac{n}{2}]}\alpha_n \to 0$.

Como los diámetros de las bolas se van haciendo cada vez más chicos, están contenidas unas en otras y por lo tanto la partición de Borel es un generador y existen conjuntos C_n los cuales son uniones de átomos de α_n tales que $\nu(T^{[\frac{n}{2}]} C_n \Delta C) \to 0$ y $\mu_B(T^{[\frac{n}{2}]} C_n \Delta C) \to 0$ cuando $n \to \infty$. Dado que C es T-invariante, tomar la inversa nos deja en el mismo conjunto, y por lo tanto al tomar $T^{-[\frac{n}{2}]}$ obtenemos $\nu(C_n \Delta C) \to 0$ y $\mu_B(C_n \Delta C) \to 0$.

Por el Lema 3.13, $h_\nu(T) \leq \frac{1}{n} H_\nu(\alpha_n)$ y como ν es una medida de entropía maximal, por el principio variacional, su entropía tiene que ser igual a la entropía topológica. Entonces

$$nh \leq -\sum_{A_x \in \alpha_n} \nu(A_x) \log \nu(A_x) = -\Bigg(\sum_{A_x \subset C_n} \nu(A_x) \log \nu(A_x) + \sum_{A_x \cap C_n = \varnothing} \nu(A_x) \log \nu(A_x) \Bigg). \tag{3.19}$$

Por estar nuestras medidas normalizadas $\nu(C_n)$ es a lo más 1, lo mismo ocurre para C_n^c. Si tomamos todos los conjuntos $A_x \subset C_n$, como generan a C_n, la suma de sus medidas son la medida de C_n, por lo tanto podemos aplicar el Lema 3.12 y

$$nh \leq \nu(C_n) \log \text{card}\{A_x | A_x \subset C_n\} + \nu(C_n^c) \log \text{card}\{A_x | A_x \cap C_n = \varnothing\} + 1. \tag{3.20}$$

Como la unión de C_n y C_n^c forma todo el espacio, la suma de sus medidas debe de ser 1, y por lo tanto podemos reescribir la ecuación como

$$-1 \leq \nu(C_n)[\log \text{card}\{A_x | A_x \subset C_n\} - nh] + \nu(C_n^c)[\log \text{card}\{A_x | A_x \cap C_n = \emptyset\} - nh]. \tag{3.21}$$

Por el Lema 3.5 obtenemos $e^{-nh} \leq A_\epsilon^{-1} \mu_B(B^n_\epsilon(X)) \leq A_\epsilon^{-1} \mu_B(A_x)$ y entonces

$$-1 \leq \nu(C_n) \log(A_\epsilon^{-1} \mu_B(C_n)) + \nu(C_n^c) \log(A_\epsilon^{-1} \mu_B(C_n^c)). \tag{3.22}$$

Si hacemos tender n al infinito, como $\mu_B(C) = 0$, entonces $\mu_B(C_n) \to 0$ y como $\nu(C) = 1$, $\nu(C_n^c) \to 0$ lo cual hace que el lado derecho de la ecuación tienda a $-\infty$ por lo que tenemos una contradicción y por lo tanto es intrínsecamente ergódica.

$\square$

Los sistemas dinámicos intrínsecamente ergódicos tienen la ventaja de que al existir una medida única de entropía maximal, ésta se vuelve la medida más natural para usar.

En los sistemas simbólicos solo necesitamos la propiedad de especificación para saber que es intrínsecamente ergódico pues todos los sistemas simbólicos son expansivos. Mostrar esto es muy sencillo, sea X un shift finito en un alfabeto $\mathscr{A}$ y la distancia entre dos puntos está definida como $d(x, y) = 2^{-k}$ con $k = \sup\{j \geq 0 | x_j = y_j\}$. En este caso la distancia más grande que puede haber entre dos elementos es cuando $k = 0$ y $d(x, y) = 2^0 = 1$ entonces es expansivo si al tomar $0 < c < 1$ y x, y satisfacen que $d(\sigma^n x, \sigma^n y) < c$ para todo $n \in \mathbb{N}$ entonces $x = y$.

Supongamos que no lo cumple y que $x \neq y$ entonces existe algún $i \in \mathbb{N}$ tal que $x_i \neq x_j$ y para todo $j < i$ se tiene que $x_j = y_j$. Por la definición de la función shift tenemos que $\sigma^{i-1}(x)_1 = x_i$ y de la misma forma $\sigma^{i-1}(y)_1 = y_i$ por lo que $d(\sigma^{i-1}(x), \sigma^{i-1}(y)) = 1 > c$. Esto es una contradicción y nos dice que un sistema simbólico es expansivo.

Con esto acabamos con la propiedad de especificación donde finalmente pudimos ver que está fuertemente relacionada con la ergodicidad de un sistema dinámico. En un principio mostrar que un sistema dinámico tiene la propiedad de especificación puede ser bastante complicado, pero una vez que la logramos mostrar obtenemos mucha información del sistema y su entropía.

En los sistemas simbólicos es más sencillo estudiar la especificación pues se puede ver como unión de palabras y son expansivos lo cual da mucha mas información del sistema. Recordemos que en un shift X la propiedad de especificación nos dice que dadas dos palabras u, v que pertenezcan al lenguaje del shift existe una palabra w de tamaño igual a g tal que uwv está en el lenguaje. Estudiar este comportamiento puede resultar sencillo dependiendo de cómo esté construido el shift, por ejemplo si decidimos tomar el shift áureo, para cualesquiera dos palabras que tomemos siempre las podremos pegar con una palabra de tamaño M formada por M ceros y por lo tanto cumple con la propiedad de especificación.

3.3. Casi especificación y especificación no uniforme

Es natural pensar en diferentes formas de modificar la propiedad de especificación con la intención de hacerla mas débil y de esa forma más sistemas la pueda tener. Un primer intento dado por Marcus [13] modifica el tamaño del espacio entre los pedazos de órbitas. En la propiedad de especificación tenemos que la distancia entre cada una de las órbitas esta dada por una constante $M(\epsilon)$, la forma propuesta para modificar esto, es que la distancia entre las órbitas sea variable y dependa del tamaño de la órbita.

Definición 3.19. Un sistema dinámico (X, T) tiene la *propiedad de especificación débil* si para cada $\epsilon > 0$ existe una función no decreciente $M_\epsilon \colon \mathbb{N} \to \mathbb{N}$ con $M_\epsilon(n)/n \to 0$ cuando $n \to \infty$, tal que cualquier especificación M_ϵ-espaciada esta ϵ-rastreada por algún punto en X. Diremos que M_ϵ es una función ϵ-espaciada para (X, T).

Esta definición ha tenido varios nombres, incluso hoy en día se le llama de muchas maneras así que siempre hay que tener especial cuidado en la nomenclatura que usamos. En muchos lugares usan el término "casi especificación débil" o simplemente "casi especificación", pero Pavlov [14] usa el término "especificación no uniforme con función espacio M" para esta propiedad y "casi especificación" lo utiliza para una propiedad diferente mientras que Kwetniak [10] utiliza "especificación débil". En este trabajo nos quedaremos con los términos de Pavlov pero es importante saber que en la literatura lo podemos encontrar de muchas maneras diferentes.

Dada nuestra definición de propiedad de especificación no uniforme y la definición de especificación ν-espaciada podemos ver que la función M depende del tamaño de la órbita anterior al espacio. Esto es solo una convención ya que en un principio no hay nada que nos impidiera decidir que la función M dependiera del segmento de órbita posterior al espacio y no al anterior.

Si suponemos que depende del segmento posterior entonces para la especificación M espaciada la condición que tenemos es $a_i - b_{i-1} \geq M(b_i - a_1)$, para $2 \leq i \leq n$, es decir el tamaño del espacio entre el segmento i y el segmento $i+1$ debe ser mayor o igual que la función aplicada en el tamaño del segmentos i, en este caso diremos que el sistema tiene especificación no uniforme hacia adelante. Pero podemos considerar que la función dependa del segmento anterior y obtener la siguiente propiedad $a_i - b_{i-1} \geq M(b_{i-1} - a_{i-1})$ y diremos que tiene la propiedad de especificación no uniforme hacia atrás.

Esta dualidad de entender la especificación no uniforme es interesante ya que no son equivalentes, es decir, un sistema que tiene especificación no uniforme hacia atrás no necesa-

riamente la tiene hacia adelante y viceversa. En este texto utilizaremos siempre la definición hacia atrás, así que cuando digamos que un sistema tiene especificación no uniforme naturalmente estaremos diciendo que es hacia atrás. A continuación mostraremos ciertas propiedades que cumplen los sistemas con especificación no uniforme, sin perdida de generalidad podemos cambiar nuestra definición de especificación no uniforme y tomarla hacia adelante y los seguirán cumpliendo, por lo tanto al hablar de especificación no uniforme tendríamos que pensar en ambas posibles definiciones.

Proposición 3.6. *Un subshift tiene especificación no uniforme con función espacio $f(n)$ si*

1. *$f(n)$ es positiva y no decreciente.*
2. $\frac{f(n)}{n} \to 0$
3. *Para cuales quiera palabras $w^1, w^2, \ldots, w^k \in \mathscr{L}(X)$ y para cualesquiera $n_1, n_2, \ldots, n_{k-1}$ enteros tal que $n_i \geq f(|w^i|)$ para todo i, entonces existen palabras $v^1 \in \mathscr{L}_{n_1}(X), \ldots, v^{k-1} \in \mathscr{L}_{n_{k-1}}(X)$ tal que la palabra $w^1v^1w^2v^2 \ldots w^{k-1}v^{k-1}w^k \in \mathscr{L}(X)$.*

Demostración. Ya tenemos la función $f(n)$ la cual es no decreciente, positiva y $\frac{f(n)}{n} \to 0$ cuando $n \to \infty$. Si tiene la propiedad de especificación no uniforme sabemos que para cualquier especificación que este f-espaciada existirá un elemento x que la rastrea. Entonces demos la especificación $\xi = \{T^{[a_j,b_j]}(x_j)\}_{j=1}^{n}$ donde la distancia entre cada pedazo de la órbita esta dada por $a_i - b_{i-1} \geq f(b_{i-1} - a_{i-1})$.

Como estamos en un subshift entonces cada elemento es una sucesión infinita de elementos en el abecedario de X. Definamos como $x_{i[a_i,b_i]} = w^i$ como la palabra que forman los elementos de a_i a b_i en el elemento x. Como esta palabra se encuentra en un elemento del subshift por lo tanto es una palabra válida.

Después de esto podemos aplicar el mismo proceso que usamos cuando estudiamos la especificación de un sistema simbólico para decir que $w^i = x_{i[a_i,b_i]} = y_{[a_i,b_i]}$ para toda $i < n$ con y siendo el elemento que rastrea la especificación. Sabemos que la distancia entre cada uno de los pedazos de órbitas está definida por la órbita anterior, entonces en el elemento y tenemos la palabra $y_{[a_{i+1}-b_i]} = v^i$ de tamaño $f(b_i - a_i)$ es decir depende del tamaño de la palabra w^i. Como esto se cumple para toda la especificación entonces tenemos las palabras $w^1, \ldots, w^n$ que son parte del elemento y y entre cada una de ellas existen las palabras v^i que dependen del tamaño de la palabra w^{i-1}.

□

Con esto en mente veamos un ejemplo sobre cómo las dos posibles definiciones de especificación no uniforme no se implican la una a la otra necesariamente. Consideremos el abecedario $\{0,1\}$ y el conjunto de palabras $\mathcal{F} = \{10^b1^a : a,b \in \mathbb{N}, b < \log_2(a)\}$ y $\mathcal{G} = \{1^a0^b1 : a,b \in \mathbb{N}, b < \log_2(a)\}$. Sean los shifts de tipo finito $X = X_{\mathcal{F}}$ y $Y = X_{\mathcal{G}}$. Sean u, w dos palabras en el diccionario de X entonces la palabra $u0^{\lceil \log_2 |w| \rceil}w \in \mathcal{B}(X)$ pues si suponemos que w esta formada por solo la letra 1 y la ultima letra en v es 1, entonces para que sea una palabra válida el número de ceros que hay en medio debe de ser mayor a $\log_2(|w|)$. Por la misma razón si $u, w \in \mathcal{B}(Y)$ entonces $u0^{\lceil \log_2 |u| \rceil}w \in B(Y)$.

Si tomamos el sistema simbólico (X, σ) podemos ver que tiene especificación no uniforme hacia adelante pues acabamos de ver que para cualquiera dos palabras siempre existe una palabra formada de puros 0 tal que puede pegar ambas palabras y depende del tamaño de la segunda palabra. En este caso la función que nos da la especificación no uniforme es $\lceil \log_2(x) \rceil$. El mismo proceso sirve para ver que (Y, σ) tiene especificación no uniforme hacia atrás con la misma función.

Ambos espacios shift contienen los puntos $x_1 = 1^\infty$ y $x_2 = 01^\infty$ por lo que las palabras 1^n y 01^n también pertenecen a los shifts. Si queremos que la palabra $1w01^n$ este en el diccionario de X entonces w debe de estar formada por ceros tal que $|w| + 1 > \log_2(n)$ por lo que $|w| = \lfloor \log_2 n \rfloor = p$ lo cumple por lo tanto $w = 0^p$.

Sabemos que X tiene especificación no uniforme hacia adelante, ahora supongamos que también la tiene hacia atrás, entonces existe la función f que nos da la especificación no uniforme. Definamos como $k = f(1)$ entonces para cada $n \in \mathbb{N}$ existe una palabra w de longitud k tal que $1w01^n$ es admisible en X. Pero el tamaño de w no depende de n, entonces al tomar n suficientemente grande, tal que, $\log_2(n) \geq k + 1$ generará una contradicción con nuestra definición del shift, por lo tanto no tiene especificación no uniforme hacia atrás. El ejemplo para mostrar que Y no tiene especificación no uniforme hacia adelante se hace de la misma forma con las palabras 1^n y i^n0. Por lo tanto las dos posibles definiciones de especificación no uniforme no son equivalentes.

La especificación no uniforme cumple con ciertas propiedades que la especificación también cumple.

Teorema 3.8. *1. Si (X, T) tiene la propiedad de especificación no uniforme entonces (X, T^k) tiene la propiedad de especificación no uniforme para todo $k \geq 1$.*

2. Si $(X, T), (Y, S)$ tienen la propiedad de especificación no uniforme entonces $(X \times Y, T \times S)$ tiene la propiedad de especificación no uniforme.

3. Cada factor de un sistema con la propiedad de especificación no uniforme tiene la

propiedad de especificación no uniforme.

4. *Cada mapeo $T\colon X \to X$ sobre con la propiedad de especificación no uniforme es topológicamente mezclado.*

Demostración. 1) Sea $\zeta = \{T^{[a_i,b_i]}(x^i)\}_{i=1}^n$ una especificación tal que $a_{i+1} - b_i \geq f(b_i - a_i)$ con f la función que nos permite rastrear esta especificación con un elemento x ya que el sistema tiene especificación no uniforme. Queremos demostrar que para cualquier k la especificación $\xi = \{T^{k[a_i,b_i]}(x^i)\}_{i=1}^n$ es también rastreable con un elemento $x^k \in X$ cuando la distancia entre cada pedazo de órbita está acotado por abajo por una función g.

La especificación $\xi \subset \{T^{[ka_i,kb_i]}(x^i)\}_{i=1}^n = \zeta_k$, pues $T^{k[a_i,b_i]}(x^i) \subset T^{[ka_i,kb_i]}(x^i)$ pero su primera y última iteración es la misma. Para la especificación ζ_k queremos ver la distancia entre sus órbitas, entonces $ka_{i+1} - kb_i = k(a_{i+1} - b_i) \geq kf(b_i - a_i) \geq f(k(b_i - a_i)) = f(kb_i - ka_i)$, pues sabemos que la función debe de crecer más lento que la identidad, por lo tanto $ka_{i+1} - kb_i \geq f(kb_i - ka_i)$. Entonces la especificación ζ_k cumple en estar espaciada por f y por lo tanto es rastreable por un elemento x^k y como $\xi \subset \zeta_k$ también es rastreable por x^k con una función espacio $kf(x)$.

2) Sea $\zeta_1 = \{T^{[a_i,b_i]}(x^i)\}_{i=1}^n$ una especificación en el sistema (X,T) tal que la distancia $a_{i+1} - b_i \geq f(b_i - a_i)$ donde f es la función definida por la especificación no uniforme, por lo tanto existe $w \in X$ tal que rastrea la especificación. De igual manera sea $\zeta_2 = \{S^{[a_i,b_i]}(y^i)\}_{i=1}^n$ una especificación en el sistema (Y,S) tal que la distancia $a_{i+1} - b_i \geq g(b_i - a_1)$ donde g es la función definida por la especificación no uniforme, así que existe $u \in Y$ la cual rastrea toda la especificación.

Construyamos una especificación para el sistema $(X \times Y, T \times S)$, sea $\xi = \{(T \times S)^{[a_i,b_i]}(x^i, y^i)\}_{i=1}^n$ donde la distancia $a_{i+1} - b_i \geq \text{máx}\{f(b_i - a_i), g(b_i - a_i)\}$, veamos que el elemento (w,u) rastrea la especificación. Tomemos la distancia usual para un producto cruz $d((T \times S)^k(x^i, y^i), (T \times S)^k(w,u))$ para $a_i \leq k \leq b_i$ con $1 \leq i \leq n$, entonces $d((T \times S)^k(x^i, y^i), (T \times S)^k(w,u)) = \text{máx}\{d_X(T^k(x^i), T^k(w)), d_Y(S^k(y^i), S^k(u))\} < \epsilon$. Por lo tanto tiene la propiedad de especificación no uniforme con la función $\text{máx}\{f(x), g(x)\}$.

3) Sea (Y,S) un sistema dinámico tal que (X,T) es un factor de él, es decir, existe la semiconjugación continua $\pi\colon Y \to X$ tal que $T^k \circ \pi = \pi \circ S^k$ para toda $k \in \mathbb{N}$. Sabemos que (Y,S) tiene la propiedad de especificación no uniforme entonces para toda especificación $\zeta = \{S^{[a_i,b_i]}(y^i)\}_{j=1}^n$ tal que $a_{i+1} - b_i \geq f(b_i - a_i)$ existe $u \in Y$ tal que $d_Y(S^k(y^i), S^k(u)) \leq \epsilon$.

Para el sistema (X,T), definamos la especificación $\xi = \{T^{[a_i,b_i]}(x^i)\}_{j=1}^n$ donde $x^i = \pi(y^i)$

y $w = \pi(u)$. Entonces $d_X(T^k(x^i), T^k(w)) = d_X(T^k \circ \pi(y^i), T^k \circ \pi(u)) = d_X(\pi \circ S^k(y^i), \pi \circ S^k(u)) \leq \delta(\epsilon)$ pues la función π es una semiconjugación y es continua y $d_Y(S^k(y^i), S^k(u)) \leq \epsilon$. Por lo tanto la especificación ξ es rastreable por $w \in X$. Por lo tanto tiene la propiedad de especificación no uniforme.

4) Un sistema dinámico es topológicamente mezclado si para $U, V \subset X$ entonces existe $N \in \mathbb{N}$ tal que para todo $n \geq N$ $U \cap T^{-n}(V) \neq \varnothing$. Entonces tomemos $U, V \subset X$ abiertos y $x \in U$, $y \in V$ y sea $\epsilon > 0$ tal que tanto diam$U$ como diam$V$ son ambos mayores que ϵ. Como el sistema tiene especificación no uniforme, podemos encontrar la función f_ϵ que define el tamaño de los agujeros.

Sea $n \geq f_\epsilon(1)$ y sea $\zeta = \{T^0(x), T^n(y')\}$ una especificación donde $y' \in T^{-n}(\{y\})$. Como la distancia entre las dos órbitas es mayor o igual que $f(1)$ entonces es rastreable y existe $w \in X$ tal que $d(w, x) \leq \epsilon$ y $d(T^n(w), T^n(y')) = d(T^n(w), y) \leq \epsilon$. Por lo tanto $z \in U$ y $z \in T^{-n}(V)$ por lo tanto el sistema es topológicamente mezclado.

□

Lema 3.14. *Si (X, T) tiene la propiedad de especificación no uniforme y su extensión natural es expansiva entonces (X, T) tiene la propiedad de especificación no uniforme periódica.*

Como podemos ver en muchos sentidos, la especificación no uniforme se comporta muy parecido a la especificación pues comparten muchas propiedades. Más adelante veremos que es cuando hablamos de ergodicidad que empiezan a comportarse de manera diferente.

Sabemos que si un sistema dinámico tiene especificación eso quiere decir que tenemos un conjunto de pedazos de órbitas para distintos elementos del sistema que son rastreables por un elemento x, esto es, x está cerca de la órbita de un elemento x_1 a lo largo de un pedazo, después necesita un tiempo N para llegar a la órbita de x_2 y estar cerca de él por todo un pedazo. Por lo tanto no importa qué pedazos del sistema tomemos, siempre podremos visitarlos todos.

Con la especificación no uniforme se cumple lo mismo, a excepción de que el tiempo que nos toma llegar de un pedazo de órbita al otro depende de cuánto tiempo estuvimos cerca de la órbita. Como la función $f(x)$ es una función creciente, mientras más tiempo nuestro elemento x esté cerca de x_1, más tiempo tardará en poder estar cerca de x_2. Seguimos visitando todo el sistema, solo que necesitamos más tiempo para hacerlo, esto será lo que nos complique demostrar que el sistema sea intrínsecamente ergódico.

Otra forma de buscar propiedades parecidas a la especificación se debe a Thompson [18]

al estudiar los $\beta-$shifts, se dio cuenta de que no cumplen con la propiedad de especificación, pero sí con lo que él definió como casi especificación.

Sea $X = [0,1]$, para cualquier $\beta > 1$ tomemos la β-transformación $f_\beta \colon X \to X$ tal que $f_\beta(x) = \beta x \mod 1$. Para $\beta \notin \mathbb{N}$ sea $b = \lfloor \beta \rfloor$, y consideramos la partición del intervalo $[0,1]$ en $J_i = [\frac{i}{\beta}, \frac{i+1}{\beta})$ con $0 \leq i \leq b-1$ y $J_b = [\frac{b}{\beta}, 1)$.

Para $x \in [0,1)$ definamos $w(x) = (w_j(x))_{j=1}^{\infty}$ como la sucesión donde $w_j(x) = i$ cuando $f^{j-1}(x) \in J_i$. $w(x)$ es conocida como la β-expansión de x y tenemos que $x = \sum_{j=1}^{\infty} w_j(x)\beta^{-j}$.

Entonces el β-shift $(\Sigma_\beta, \sigma_\beta)$ es la cerradura de todas estas sucesiones en el shift completo del abecedario $\{0, \ldots, b\}$.

Sabemos que la β-expansión de 1 está definida como $1 = \sum_{i=1}^{\infty} c_i \beta^{-i}$, cuyas constantes están dadas por el algoritmo $r_0 = 1, c_{i+1} = \lceil \beta r_i \rceil, r_{i+1} = \beta r_i - c_{i+1}$ con $i \in \mathbb{Z}_+$. Con esto podemos asegurar que $r_i > 0$ para todo $i \in \mathbb{Z}_+$ por lo tanto la sucesión $c = \{c_i\}$ no puede acabar solo en ceros.

Dadas dos sucesiones $(a_i)_{i\geq 1}$ y $(b_i)_{i\geq 1}$, el orden lexicográfico está definido como $(a_i) < (b_i)$ si y solo si para el menor índice i tal que $a_i \neq b_i$ ocurre que $a_i < b_i$. La sucesión (c_i) tiene que ser el supremo lexicográfico de todas las sucesiones que pertenecen al β-shift. Si $w \in \Sigma_\beta$, entonces evidentemente $w \leq c$ pues la suma donde los c_i son coeficientes dan 1 y donde w_i son coeficientes da un número menor que 1. Lo mismo ocurre para toda $\sigma^k(w) \leq c$ pues cada iteración de sigma estamos viendo en qué intervalo cae la función $\beta^k x \mod 1$, entonces estamos tomando la sucesión para otro número menor que 1.

Utilizando la idea de que c es el supremo lexicográfico de todas las sucesiones, podemos representarlo como un gráfica. Sea $v_i, v_2, \ldots$ un conjunto contable de vértices, dibujaremos una arista dirigida entre v_i y v_{i+1} para toda i y la etiquetaremos con el valor de c_i. Este camino representa la sucesión c. Si $c_i \geq 1$, entonces dibujaremos $c_i - 1$ aristas de v_i a v_1 y las etiquetaremos con todos los valores de 0 a $c_i - 1$. De esta forma estamos tomando los elementos que son menores a c y siempre regresamos al principio, lo cual nos confirma que al tomar las iteraciones seguimos siendo más chicos que c y por lo tanto $w \in \Sigma_\beta$ si y solo si w etiqueta un camino infinito en $\mathcal{G}_\beta$ que empieza en v_1. Esta será la presentación gráfica del β-shift y lo denotamos como $\mathcal{G}^\beta$.

Sea $z_n = \min\{i \geq 0 : c_{n+i} \neq 0\}$. Entonces $z_n + 1$ es el mínimo número de aristas que se requieren para viajar de v_n a v_1, esto debido a que si $c_{n+i} = 0$, entonces no existen aristas que regresen a v_1. Tomemos dos palabras u, w que pertenezcan al β-shift. Entonces son pedazos

de caminos en $\mathcal{G}^\beta$, por lo tanto la palabra u es un camino que acaba en un vértice v_n y la palabra w empieza en el vértice v_1. Si queremos unir ambas palabras, necesitamos encontrar un camino que una v_n con v_1, este camino esta dado por el camino que va de v_n a v_{n+i}, con $c_{n+i} \neq 0$ pues de ahí podemos regresar a v_1 por como construimos la gráfica. Entonces el tamaño del camino es z_n.

Para que un sistema tenga especificación necesitamos que exista M tal que para cualesquiera dos palabras, exista una palabra de al menos tamaño M que las una. Pero no hay razón para poder decir que el conjunto formado por todos los z_n esté acotado y por lo tanto no existe la especificación. Para que un β-shift tenga la propiedad de especificación tiene que ocurrir que la sucesión c, que es la expansión en base de β de 1, sea eventualmente periódica, de esta forma la gráfica es una gráfica finita y $\{z_n\}$ es acotado.

Ya que los β-shifts no tienen la propiedad de especificación, Pfister y Sullivan [16] propusieron una nueva propiedad parecida a la especificación llamada propiedad de g-casi producto.

Definición 3.20. Sea $g\colon \mathbb{N} \to \mathbb{N}$ un mapeo no decreciente y no acotado con la propiedad $g(n) < n$ y $\lim_{n\to\infty} \frac{g(n)}{n} = 0$. La función g es llamada la función error. Sea $x \in X$ y $\epsilon > 0$. El *g-error de $B_n(x,\epsilon)$* es el conjunto cerrado

$$B_n(g;x,\epsilon) = \{y \in X : \exists \Lambda \subset \Lambda_n, |\Lambda_n \setminus \Lambda| \leq g(n) \text{ y } \operatorname{máx}\{d(T^j(x), T^j(y)) : j \in \Lambda\} \leq \epsilon\}$$

donde $\Lambda_n = [-n, n]$.

Recordemos que $B_n(x,\epsilon)$ es una bola de Bowen y está definida como el conjunto de todos los puntos $y \in X$ tales que $\operatorname{máx}\{d(T^j(x), T^j(y)) \leq \epsilon : j \in \Lambda_n\}$. Entonces al tomar el g-error de $B_n(g,x,\epsilon)$ estamos tomando un superconjunto de la bola de Bowen pues elegimos un subconjunto Λ del intervalo Λ_n donde la condición de Bowen no necesariamente se satisface.

Definición 3.21. El sistema dinámico (X,T) tiene *la propiedad de g-casi producto con función error g* si existe una función no creciente $m\colon \mathbb{R} \to \mathbb{N}$ tal que para cualquier $k \in \mathbb{N}$, $x^1, \ldots, x^k \in X$, cualesquiera $\epsilon_1, \ldots, \epsilon_k > 0$ y $n_1 \geq m(\epsilon_1), \ldots, n_k \geq m(\epsilon_k)$, se tiene que

$$\bigcap_{j=1}^{k} T^{-M_{j-1}} B_{n_j}(g; x^j, \epsilon_j) \neq \varnothing, \tag{3.23}$$

donde $M_0 = 0, M_i = n_1 + \ldots + n_i, i = 1, \ldots, k-1$.

Thompson y Kwetniak [10] renombraron esta propiedad como casi especificación con función error g y cambiaron un poco las definiciones para hacerla más general.

Definición 3.22. Sea $g : \mathbb{N}_0 \times (0, \epsilon_0) \to \mathbb{N}$ una función, donde $\epsilon_0 > 0$. Decimos que g es una *función error* si para todo $\epsilon < \epsilon_0$ y para todo $n \in \mathbb{N}_0$ tenemos que $g(n, \epsilon) \leq g(n+1, \epsilon)$ y $\lim_{n\to\infty} \frac{g(n,\epsilon)}{n} = 0$.

Dada una función error g, definimos la función $k_g : (0, \infty) \to \mathbb{N}$ declarando que $k_g(\epsilon)$ es la $n \in \mathbb{N}$ más chica tal que $g(m, \epsilon) < m\epsilon$ para todo $m \geq n$.

Definición 3.23. Dada una función error g, $0 < \epsilon < \epsilon_0$ y $n \geq k_g(\epsilon)$, definimos el conjunto

$$I(g; n, \epsilon) = \{\Lambda \subset \{0, 1, \ldots, n-1\} : |\Lambda| \geq n - g(n, \epsilon)\}. \tag{3.24}$$

Decimos que un punto $y \in X$ $(g; \epsilon, n)$-rastrea el segmento de órbita $T^{[a,b]}(x)$ si para algún $\Lambda \in I(g; \epsilon, n)$ tenemos que $\rho^T_\Lambda(T^a(x), T^a(y)) \leq \epsilon$. Por $B_n(g; x, \epsilon)$ denotamos el conjunto de todos los puntos que $(g; \epsilon, n)$-rastrea un segmento de órbita $T^{[0,n)}(x)$.

Definición 3.24. Un sistema dinámico (X, T) tiene la propiedad de casi especificación si existe una función error g tal que para cualquier $m \geq 1$, cualquier $\epsilon_1, \ldots, \epsilon_m > 0$ y para cualquier especificación $\{T^{[a_j,b_j]}(x_j)\}_{j=1}^m$ con $b_j - a_j + 1 \geq k_g(\epsilon)$ para cada $j = 1, \ldots, m$ podemos encontrar un punto $z \in X$ tal que $(g; b_j - a_j + 1, \epsilon_j)$-rastrea el segmento de órbita $T^{[a_j,b_j]}(x_j)$ para cada $j = 1, \ldots, m$

Estas dos definiciones son muy parecidas y representan las misma idea, dada una especificación cualquiera, existen palabras que van a rastrear la especificación salvo un error, es decir que cuando recorramos todos los pedazos de órbitas, existirá y que rastreara el elemento excepto en algunos puntos, esos son nuestro error. La cantidad de puntos en la órbita que no están propiamente rastreados está controlada por la función error g. Notemos que en nuestra definición no nos importa el tamaño entre el espacio de la especificación como sí nos importaba en la propiedad de especificación o especificación no uniforme, lo que nos importa es el tamaño de los pedazos de órbita porque de este tamaño depende la cantidad de puntos donde hay un error en el rastreo. A partir de este momento cuando nos referiremos a esta propiedad, diremos que tiene casi especificación.

Cuando hablemos de sistemas simbólicos es mucho mas fácil usar la definición que da Pavlov [14], basado en las anteriores definiciones.

Definición 3.25. Un subshift X tiene la propiedad de casi especificación con función error $g(n)$ si:

1. $g(n)$ es una función positiva no decreciente.
2. $\frac{g(n)}{n} \to 0$.

3. Para cuales quiera palabras $w^1, w^2, \ldots, w^k \in \mathcal{L}(X)$, existen palabras $v^1, v^2, \ldots, v^k \in \mathcal{L}(X)$ tal que $|w^i| = |v^i|$ para toda i. w^i y v^i deben de diferir en a lo más $g(|w^i|)$ letras para cada i y la concatenación $v^1 v^2 \ldots v^k \in \mathcal{L}(X)$.

La forma para ver que la definición de Pavlov es la misma que Kwetniak y Thompson es relativamente sencilla. La función error se comporta de la misma manera con la única diferencia de que para Pavlov igual que con Thompson no depende de ϵ. Por lo tanto podemos tomar cualquier especificación sin importarnos de su tamaño, en este caso tomemos la especificación $\zeta = \{\sigma^{[a_j,b_j]}(x_j)\}_{j=1}^m$ donde $b_j = a_{j+1}$. Esta especificación se convierte, como ya hemos visto, en un conjunto de palabras $w^1, w^2, \ldots, w^m \in \mathcal{L}(X)$. Entonces si el sistema tiene casi especificación quiere decir que para cada intervalo $[a_j, b_j]$ existe un subconjunto $\Lambda_j \subset \{a_j, a_{j+1}, \ldots, b_j\}$ de tamaño $g(b_j - a_j)$ tal que existe un elemento y que rastrea toda la especificación excepto en el conjunto Λ_j.

Al convertir esto al sistema simbólico obtenemos que para cada palabra w^i existe el subconjunto de elementos cuya cardinalidad depende del tamaño de la palabra, es decir $g(|w^i|)$ tal que rastreará la especificación excepto en esos puntos. Como sabemos que rastrear la especificación implica que la unión de las palabras con sus espacios forman una palabra válida en el lenguaje de X. En este caso los espacios no existen y solo estamos hablando de la unión de las palabras excepto en algunos puntos, por lo tanto existe el elemento $y \in X$ tal que el bloque $[a_1, b_m]$ es un bloque válido y los sub bloques de este bloque $v^i = [a_i, b_i]$ son iguales que la palabra w^i excepto por los puntos dados por el subconjunto Λ. El regreso es igual y por lo tanto podemos usar la definición de Pavlov para hablar de casi especificación.

Ya sabemos que los β-shifts no tienen en general la propiedad de especificación, pero podemos mostrar que todos tienen la propiedad de casi especificación con función error $g(n) = 1$. Por la construcción de la gráfica $\mathcal{G}^\beta$ sabemos que cualquier elemento $x \in X_\beta$ es un camino enumerado que empieza en v_1 por lo tanto podemos decir que para cualquier palabra en el lenguaje del β-shift, va a ser un pedazo de camino que empiece en v_1. Sean dos palabras $w^1, w^2 \in \mathcal{L}(X_\beta)$. Entonces sabemos que w^2 empieza en v_1 y supongamos que w^1 acaba en un v_n. Si acabara en v_1, entonces la unión de $w^1 w^2$ es una palabra válida por que representa un camino en la gráfica. Ahora supongamos que $v_n \neq v_1$. Entonces $w^1 = u^k c^{m-k}$ donde $m - k = n$, esto es, la concatenación de dos caminos, uno es u^k y empieza y termina en v_1 y termina en v_1, y el otro camino que recorre los c_i hasta c_n. Sea c_j el último elemento en la palabra c^{m-k} el cual es diferente de cero y lo cambiamos por el elemento $c_j - 1$, por como está construida la gráfica, este camino existe y nos regresa a v_1 y como el resto son ceros, nos está regresando a v_1, llamamos a esta nueva palabra p^1 y por lo tanto $p^1 w^2$ se pueden concatenar.

Aplicando este argumento de manera inductiva para $w^1, w^2, w^3, \ldots, w^n$ podemos ver que cumple la propiedad de casi especificación donde a las palabras w^i a lo más hay que cambiarles un elemento para que se puedan pegar, por lo tanto su función error es la función constante $g(n) = 1$.

La propiedad de casi especificación es una propiedad mucho más débil que la especificación o la especificación no uniforme, por lo que no podemos obtener mucho más de ella, como sí lo lográbamos en las otras propiedades.

3.4. Ergodicidad de casi especificación y especificación no uniforme

Cuando analizamos la propiedad de especificación pudimos ver que obteníamos información muy importante sobre la entropía del sistema y sobre su ergodicidad, nos gustaría saber si la especificación no uniforme o la casi especificación nos dan tanta información. Esta pregunta es mucho más complicada de lo que parece y aún es un problema abierto en el campo, cuándo una función error o una función gap implican ergodicidad intrínseca. Ahora revisaremos el artículo de Pavlov [14] donde de muestran algunos ejemplos de shifts que cumplen la propiedad de especificación no uniforme y la propiedad de casi especificación y que no son intrínsecamente ergódicos. De la misma forma, Pavlov muestra que existen maneras de acotar las funciones error y gap para acercarnos mucho a ergodicidad intrínseca, aunque aún no completamente.

Teorema 3.9. *Para cualquier función $f(n)$ positiva y no decreciente con* $\liminf_{n\to\infty} \frac{f(n)}{\log n} > 0$*, existe un subshift con especificación no uniforme con función gap $f(n)$ con exactamente dos medidas ergódicas de entropía maximal cuyo soportes son ajenos.*

Demostración. Definamos X con el alfabeto $\mathcal{A} = \{-N, \ldots -1, 0, 1, \ldots N\}$ con $N > e^6 \approx$ 403,4 y vamos a definir una lista de palabras prohibidas $\mathcal{F}$ que consiste en todos los pares adyacente ij cuyo producto es negativo, es decir que uno sea negativo y el otro positivo y todas las palabras $v_1 v_2 \ldots v_j 0^m v_{j+1}$ donde $v_i \neq 0$ y $m \leq \log j$.

Esto quiere decir que X consiste de puntos de la siguiente forma $\ldots w^{-1} 0^{m_{-1}} w^0 0^{m_0} w^1 0^{m_1} \ldots$ tal que cada palabra w^i debe de estar formada por letras diferentes de cero y del mismo signo para que así no tengan problemas con la palabras prohibidas, y para cada $i, m_i > \log |w^i|$ es decir la cantidad de ceros en la palabra 0^{m_i} debe de estar dada por la palabra anterior. Otros

puntos que viven en el shift son puntos degenerados, el punto formado por una sucesión de puros ceros, los puntos formados por una sucesión bi-infinita de letras positivas, los puntos formados por letras negativas. De igual manera si w es una palabra que no contiene ninguna palabra prohibida entonces $\ldots 000w000\ldots$ es un elemento en el shift y $w \in \mathcal{L}(X)$.

Veamos que X tiene especificación no uniforme con función gap $1+\lfloor \log n \rfloor$. Consideremos cualesquiera palabras $w^1, \ldots, w^k \in \mathcal{L}(X)$ y $m_1, \ldots, m_{k-1}$ positivos con $m_i \geq 1 + \lfloor \log |w^i| \rfloor$ para cada i. Si queremos ver que tiene especificación no uniforme debemos de dar palabras v_i de tamaño m_i tal que $w = w^1 v^1 w^2 v^2 \ldots . w^{n-1} v^{n-1} w^n \in \mathcal{L}(X)$, propongamos que $v^i = 0^{m_i}$. Como m_i es positivo para toda i entonces las palabras agregadas de puros ceros no pueden formar un par adyacente de letras de signos distintos. Entonces nos falta ver que w no contiene ninguna palabra de la forma $u_1 u_2 \ldots u_j 0^m u_{j+1}$ con $m \leq \log j$.

Supongamos que existe una palabra u prohibida de esta forma, u no puede estar en ninguna de las palabras w^i pues estas son palabras válidas en el lenguaje. Por como construimos la palabra w, el bloque de ceros 0^m en u debe de ser uno de los bloques de cero 0^{m_i} y las letras $u_1 u_2 \ldots u_j$ deben de pertenecer a w^i pero esto significa que $m \geq 1+\lfloor \log |w^i| \rfloor > \log |w^i|$.Como asumimos que $m \leq \log j$ pues es una palabra no válida entonces $j > |w^i|$ pero en la palabra w sabemos que w^i está precedida por una cadena de ceros para $i > 1$ y por nada por $i = 1$ lo cual es una contradicción, por lo tanto w no tiene palabras prohibidas y $w \in \mathcal{L}(X)$. Por lo tanto X tiene especificación no uniforme con función gap $1 + \lfloor \log n \rfloor$.

Ahora mostraremos que para cualquier $C > 0$ existe un shift potencia X^k con la propiedad de especificación no uniforme con función gap $\text{máx}(1, \lceil C \log n \rceil)$. Cualquier función f positiva tal que $\liminf_{n\to\infty} \frac{f(n)}{\log n} > 0$ está acotada por abajo por $\text{máx}(1, \lceil C \log n \rceil)$ para algún $C > 0$ esto debido a que $f(n)$ debe de crecer mas rápido que $\log n$. Queremos mostrar que esto implique que para cualquier de estos f el shift X^k tenga especificación no uniforme con función gap f.

Sea $C > 0$ y definamos $k > 4C^{-1}$, sea X^k el shift potencia de altura k de X, veamos que X^k tiene especificación no uniforme con función gap $\lceil (1 + \lfloor \log(kn) \rfloor)/k \rceil$. Consideremos las palabras $w_k^1, \ldots, w_k^n$ en el shift X^k, esto quiere decir que cada letra es un bloque de tamaño k en X tal que la palabra en X que genera la palabra w_k^i es una palabra válida en X de tamaño $|w_k^i|k$. Definamos $w = w_k^1 0_k^{m_1} w_k^2 0_k^{m_2} \ldots 0_k^{m_{n-1}} w_k^n$ donde $0_k^{m_i}$ es una palabra formada por bloques de tamaño k con puros ceros. Queremos ver que esta es una palabra válida, supongamos que tiene una palabra que no es válida, esto quiere decir que al regresar a su forma original forma una palabra no válida en el shift X. Sea la palabra que se forma en X como $u = u_1 u_2 \ldots u_j 0^m u_{j+1}$ tal que $m \leq \log j$, las palabras $0_k^{m_i}$ forman una palabra en X formada por una cadena de ceros de tamaño $m_i \geq \lceil (1 + \lfloor \log(k|w_k^i|) \rfloor)/k \rceil k$. La cadena

de ceros de la palabra u debe de contener a una de estas cadenas de ceros por lo tanto $m \geq \lceil(1 + \lfloor\log(k|w_k^i|)\rfloor)/k\rceil k > \log(k|w_k^i|)$ entonces como $m \leq \log j$ implica que $j > |w_k^i|k$. Pero esto genera una contradicción directamente por la especificación no uniforme en X por lo tanto la palabra no existe y X^k tiene la propiedad de especificación no uniforme.

Ahora probemos que

$$\lceil(1 + \lfloor(kn)\rfloor)/k\rceil \leq \text{máx}(1, \lceil C \log n\rceil). \tag{3.25}$$

Ambos lados de la ecuación son positivos enteros y en ambos casos su mínimo valor posible es 1, por lo tanto con 1 es válido. Supongamos que el lado izquierdo es al menos 2, esto quiere decir que, como k es un entero positivo, $(1+\lfloor(kn)\rfloor)/k > 1$ entonces $1+\lfloor(kn)\rfloor > k$ y $\log(kn) \geq k$ y $n \geq \frac{e^k}{k} > k$. Como estamos suponiendo que la ecuación es al menos 2 entonces $\lfloor\log(nk)\rfloor \geq 1$.

$$\frac{1 + \lfloor\log(kn)\rfloor}{k} \leq \frac{2}{k}\log(kn) < \frac{2}{k}\log(n^2) \leq \frac{4}{k}\log n < C \log n. \tag{3.26}$$

Por lo tanto tenemos la desigualdad deseada, como el shift X^k tiene especificación no uniforme con función gap $\lceil(1 + \lfloor(kn)\rfloor)/k\rceil$ quiere decir que existen palabras v^i tal que $|v^i| \geq \lceil(1+\lfloor(kn)\rfloor)/k\rceil$ cumplen las propiedades que queremos. En especial como $\lceil(1+\lfloor(kn)\rfloor)/k\rceil \leq \text{máx}(1, \lceil C \log n\rceil)$ quiere decir que lo cumple para $|v^i| \geq \text{máx}(1, \lceil C \log n\rceil)$ y por lo tanto se puede decir que tiene especificación no uniforme con función gap $\text{máx}(1, \lceil C \log n\rceil$. Más aún, como habíamos visto, existe $C > 0$ tal que $\text{máx}(1, \lceil C \log n\rceil \leq f(n)$ para cualquier función positiva $f(n)$ con $\liminf_{n\to\infty} \frac{f(n)}{\log n} > 0$ por lo tanto existe $k > 4C^{-1}$ tal que X^k tiene especificación no uniforme con función gap $f(n)$.

Ahora comprobaremos que $h(X) = \log N$. Como X contiene al shift completo de N elementos y sabemos que la entropía de un shift completo es $\log N$ entonces $h(X) \geq \log N$. Por lo tanto solo necesitamos mostrar que $h(X) \leq \log N$, así que buscaremos acotar por arriba la cardinalidad de $\mathcal{L}_n(X)$.

Sabemos que para cada $w \in \mathcal{L}_n(X)$ podemos descomponerlo como $w^1 0^{m_1} w^2 \ldots w^k 0^{m_k}$ donde cada w^i es una palabra de símbolos diferentes de cero con el mismo signo, además el tamaño de las corridas de ceros debe de estar dado tal que $m_i > \log |w^i|$. Definamos $|w^i| = p_i$ y $n_i = p_i + m_i$ entonces $\sum : i = 1^k n_i = n$. Si fijamos un k y n_i entonces la palabra w está determinada por la elección del $|w^i|$ los signos de sus letras y sus letras.

Para cada $1 < i < k$ sabemos que $m_i > \log p_i$ entonces $p_i + \log p_i < n_i$. Como $x + \log x$ es una función creciente cuyo rango son los reales entonces existe x tal que $x + \log x = n_i$, donde $x < n_i$ y $x = n_i - \log x > n_i + \log n_i$. Pero como $x + \log(n_i - \log n_i) < n_i$ entonces $x < n_i - \log(n_i - \log n_i)$, dado que $p_i + \log p_i < n_i$ entonces $p_i < x < n_i - \log(n_i - \log n_i)$.

Entonces para cada uno de los n_i sabemos que el número de elecciones posibles para la palabra $w^i 0^{m_i}$ depende primero del tamaño de w_i el cual sabemos que es p_i y como esta acotado por arriba tenemos las palabras posibles desde 1 hasta $\lfloor n_i - \log(n_i - \log n_i)\rfloor$. Estas palabras posibles son cualesquiera en el alfabeto $\{1, 2, \ldots N\}$ por lo tanto para un p_i la cantidad de palabras posibles es N^{p_i} y esto es por dos considerando que podemos tomar el alfabeto negativo, así que el número de palabras está acotado por arriba por

$$2\left(\sum_{p_i=1}^{\lfloor n_i - \log(n_i - \log n_i)\rfloor} N^{p_i}\right) < 4N^{n_i - \log(n_i - \log n_i)}. \tag{3.27}$$

La desigualdad se debe a que $\sum_{i=1}^{n} x^i < 2x^n$. Sea $M_j = 4N^{j-\log(j-\log j)}$ para $j \geq 2$. Para $i = 1$ y $i = k$ es más fácil acotar el número de posibilidades, pues no tenemos necesidad de ser tan puntuales en los tamaños ya que nada nos impide empezar la palabra w directamente con la cadena de ceros, es decir $w^1 0^{m_1}$ puede ser simplemente unos cuantos ceros o al contrario puede ser una cadena de palabras de tamaño n_i que se junte con w^2 y va a ser completamente válida. De igual manera la palabra $w^k 0^{m_k}$ puede acabar con menos ceros de los necesarios y aun así es válida. Esto nos permite relajar más la cantidad posible de palabras y n_i es nuestro máximo posible de longitud

$$2\left(\sum_{p_i=0}^{n_i} N^{p_i}\right) < 4N^{n_i}. \tag{3.28}$$

Entonces si buscamos encontrar todas las palabras posibles para $|\mathcal{L}(X)|$ debemos de pasar por todos los posibles k desde 1 hasta n. Si $k = 1$ entonces $w = w^1 0^{m_1}$ por lo que la cantidad de palabras posibles esta dado por $2\sum_{i=0}^{n} N^i < 4N^n$. Si $k = 2$ entonces $w = w^1 0^{m_1} w^2 0^{m_2}$ por lo visto anteriormente sabemos que los dos están acotados por $4N^{n_i}$ y debemos considerar todas las combinaciones posibles para n_1 y n_2 las cuales son $n-1$ por lo tanto la cantidad de palabras está acotada por $4N^{n_1}4N^{n_2}(n-1) = 16N^n(n-1)$. Finalmente si tomamos $k \geq 3$ debemos de tomar todas las combinaciones posibles de n_i y usar las cotas que vimos por lo tanto obtenemos la siguiente desigualdad

$$|\mathcal{L}(X)| < 4N^n + 16(n-1)N^n + \sum_{k=3}^{n} \sum_{\substack{n_1,\ldots,n_k \\ \sum n_i = n}} 4N^{n_1}\left(\prod_{i=2}^{k-1} M_{n_i}\right)4N^{n_k}. \tag{3.29}$$

Como $n_1 + n_2 + \ldots n_k = n$, en el tercer término de la desigualdad para cada M_{n_i} podemos factorizar N^{n_i} tal que $4N^{n_1}(\prod_{i=2}^{k-1} M_{n_i})4N^{n_k} = 16N^{n_1}N^{n_2}\ldots N^{n_k}(\prod_{i=2}^{k-1} M_{n_i}N^{-n_i})$ y

$\sum_{i=1}^{k} N^{n_k} = N^n$ por lo tanto reescribimos la desigualdad como

$$|\mathcal{L}(X)| < 4N^n + 16(n-1)N^n + 16N^n \sum_{k=3}^{n} \sum_{\substack{n_1,\dots,n_k \\ \sum n_i = n}} \left(\prod_{i=2}^{k-1} M_{n_i} N^{-n_i} \right). \tag{3.30}$$

Ahora buscaremos acotar aún más el tercer término. Al momento de factorizar los N^{n_i} tomamos N^{n_1} y N^{n_k} para formar N^n, por lo tanto en el segundo sumando ya no tiene caso tomar desde n_1 hasta n_k. Ahora la suma se toma desde n_2 hasta n_{k-1} y la suma de estos ya no debe dar n si no cualquier número menor que n. Como la suma es menor a n esto quiere decir que $n - \sum_{j=2}^{k-1} n_j = n_1 + n_k$ y debemos considerar en nuestra suma todas las diferentes opciones que tenemos para $n_1 + n_k$ las cuales son $n - \sum_{j=2}^{k-1} n_j - 1$ pues ni n_1 ni n_k pueden ser cero, por lo tanto obtenemos que

$$16N^n \sum_{k=3}^{n} \sum_{\substack{n_1,\dots,n_k \\ \sum n_i = n}} \left(\prod_{i=2}^{k-1} M_{n_i} N^{-n_i} \right) = 16N^n \sum_{k=3}^{n} \sum_{\substack{n_2,\dots,n_{k-1} \\ \sum n_1 < n}} \left(n - \sum_{j=2}^{k-1} n_j - 1 \right) \left(\prod_{i=2}^{k-1} M_{n_i} N^{-n_i} \right) \tag{3.31}$$

y como $(n - \sum_{j=2}^{k-1} n_j - 1) < n$ y recordando nuestra definición de M_j tenemos la siguiente desigualdad

$$16N^n \sum_{k=3}^{n} \sum_{\substack{n_2,\dots,n_{k-1} \\ \sum n_1 < n}} \left(n - \sum_{j=2}^{k-1} n_j - 1 \right) \left(\prod_{i=2}^{k-1} M_{n_i} N^{-n_i} \right) < 16nN^n \sum_{k=3}^{n} \sum_{\substack{n_2,\dots,n_{k-1} \\ \sum n_i < n}} \left(\prod_{i=2}^{k-1} \frac{4}{N^{\log(n_i - \log n_i)}} \right). \tag{3.32}$$

Si tomamos $(\sum_{t=2}^{\infty} \frac{4}{N^{\log(t-\log t)}})^{k-2}$ obtenemos la suma de todas las combinaciones posibles de $\prod_{i=2}^{k-1} \frac{4}{N^{\log(n_i - \log n_i)}}$ por lo tanto

$$16nN^n \sum_{k=3}^{n} \sum_{\substack{n_2,\dots,n_{k-1} \\ \sum n_i < n}} \left(\prod_{i=2}^{k-1} \frac{4}{N^{\log(n_i - \log n_i)}} \right) \leq 16nN^n \sum_{k=3}^{n} \left(\sum_{t=2}^{\infty} \frac{4}{N^{\log(t - \log t)}} \right)^{k-2}. \tag{3.33}$$

Entonces podemos acotar aún más dado que $\sum_{t=2}^{\infty} \frac{4}{N^{\log(t-\log t)}} = \sum_{t=2}^{\infty} \frac{4}{(t-\log t)^{\log N}}$ y sabiendo que los elementos de la suma tienden a cero,

$$\sum_{t=2}^{\infty} \frac{4}{(t - \log t)^{\log N}} < \frac{4}{(2 - \log 2)^{\log N}} + \frac{4}{(3 - \log 3)^{\log N}} + \int_3^{\infty} \frac{1}{(x - \log x)^{\log N}} dx. \tag{3.34}$$

Como $2 - \log 2 > 1{,}3$, $3 - \log 3 > 1{,}8$ y $x - \log x > 0{,}5x$ podemos acotar aún más la ecuación integrando la ecuación que es más fácil

$$\frac{4}{1{,}3^{\log N}} + \frac{4}{1{,}8^{\log N}} + \int_3^{\infty} \frac{1}{(0{,}5x)^{\log N}} = \frac{4}{1{,}3^{\log N}} + \frac{4}{1{,}8^{\log N}} + \frac{3}{((\log N) - 1)1{,}5^{\log N}} \tag{3.35}$$

Ahora dado que $N > e^6$, entonces $\log N > 6$. Si damos que $N = e^6$ entonces la ecuación nos da 0,9989845177 < 1 y como la función es decreciente, sabemos que siempre es menor que 1 y la llamaremos α. Aplicando todas estas ecuaciones obtenemos la siguiente desigualdad y viendo que la suma es una suma geométrica pues $\alpha < 1$, obtenemos

$$|\mathcal{L}_n(X)| < 4N^n + 16(n-1)N^n + 16nN^n \sum_{k=3}^{n} \alpha^{k-2} < \frac{16n}{1-\alpha} N^n. \tag{3.36}$$

Si tomamos el logaritmo de esta cantidad obtenemos $\log(\frac{16n}{1-\alpha}N^n) = \log(\frac{16n}{1-\alpha}) + \log(N^n)$, después al dividir entre un n y tomamos el límite al infinito nos queda $\log(N) + \lim_{n\to\infty} \frac{1}{n}\log(\frac{16n}{1-\alpha}) = \log(N)$ y del otro lado obtenemos la entropía por lo tanto $h(X) \leq \log N$ y por lo tanto $h(X) = \log N$.

Sea μ una medida ergódica de entropía maximal de X. Mostremos que $\mu([a0]) = 0$ para cualquier $\alpha \in \{1, \ldots, N\}$. Recordemos que para una palabra w el cilindro $[w]$ es el conjunto de $x \in X$ tal que $x_1x_2x_3 \ldots x_n = w$. Asumamos que $\mu([a0]) = \beta > 0$ para alguna a.

Antes daremos un Lema que se da en [14] y es demostrado de manera general en [15]

Lema 3.15. *Para cualquier subshift X, cualquier medida ergódica μ en X, cualquier palabra $w \in \mathcal{L}_n(X)$, cualquier $n \in \mathbb{N}$ y $\epsilon > 0$, definimos el conjunto $C_{n,\epsilon,w}(X)$ como el conjunto de todos los $w \in \mathcal{L}_n(X)$ que tienen entre $n(\mu([w]) - \epsilon)$ y $n(\mu([w]) + \epsilon)$ ocurrencias de w. Entonces*

$$\liminf_{n\to\infty} \frac{\log |C_{n,\epsilon,w}(X)|}{n} \geq h(\mu). \tag{3.37}$$

Definamos para toda n el conjunto $C_{n,0,5\beta,a0}(X)$ de palabras de n letras con alrededor de 0,5$n\beta$ y 1,5$n\beta$ ocurrencias de $a0$. Por el lema anterior $\liminf_{n\to\infty} \frac{\log |C_{n,0,5\beta,a0}(X)}{n} \geq h(\mu) = h(X) = \log N$. Ahora sabemos que para una palabra de tamaño n se construye a partir de k palabras $w^i 0^{m_i}$ por lo tanto para encontrar $a0$ debe de estar a lo más k veces, y por los mismos argumentos que desarrollamos, tenemos que

$$|C_{n,0,5\beta,a0}(X)| < 16nN^n \sum_{k=\lceil 0,5n\beta \rceil}^{n} \alpha^{k-2} \leq \frac{16n}{1-\alpha} N^n \alpha^{0,5n\beta}. \tag{3.38}$$

De nuevo si sacamos el logaritmo, dividimos entre n y sacamos el límite superior cuando $n \to \infty$ obtenemos $\limsup_{n\to\infty} \frac{\log |C_{n,0,5\beta,a0}|}{n} \leq \log N + 0{,}5\beta \log \alpha < \log N$ lo cual es una contradicción, por lo tanto $\mu([a0]) = 0$, y de forma similar se obtiene que $\mu([b0]) = 0$ para $b \in \{-N, \ldots, -1\}$. Por lo tanto μ debe de estar completamente en $0^{\mathbb{Z}}, \{1, 2, \ldots, N\}^{\mathbb{N}}$ o $\{-1, -2, \ldots, -N\}^{\mathbb{N}}$. Pero si el primer caso se cumpliera implicaría que $h(X) = h(\mu) = 0$

pues es una medida ergódica de entropía maximal y la entropía topológica de $0^{\mathbb{Z}}$ es 0, por lo tanto $\mu([0]) = 0$.

Las únicas medidas con entropía igual a $\log N$ son las medidas de Bernoulli sobre $\{1, \ldots, N\}$ y $\{-N, \ldots, -1\}$, por lo tanto X tiene dos medidas ergódicas de entropía maximal con soportes ajenos.

Consideremos cualquier medida ergódica de entropía maximal ν para X^k. La medida ν induce una nueva medida ν^* en X, definida de la siguiente manera $\nu^*([w^1 \ldots w^n]) = \nu([w^1 \ldots w^n])$ donde cada palabra w^i es una palabra de tamaño k tal que la palabra $w^1 \ldots w^n$ es una palabra válida en X. Por ser una medida ergódica ν es σ-invariante en X^k es decir que dado un conjunto B tal que $\sigma^{-1}B = B$ entonces $\nu(B) = 0$ o $\nu(B) = 1$. Como en este caso cada letra es una palabra de tamaño k entonces ν^* debe de ser invariante bajo la transformación σ^k. Si definimos la medida $\mu = \frac{1}{k}\sum_{i=0}^{k-1} \sigma^i \nu^*$ entonces para ver que es σ invariante debemos mostrar que $\mu(\sigma^{-1}(A)) = \mu(A)$, entonces usando que $\nu^*(\sigma^{-k}(A)) = \nu^*(A)$ tenemos que $\nu^*(\sigma^{-1}(A)) = \nu^*(\sigma^{k-1}(A))$. Por lo tanto la suma pasa a ser $\mu(A) = \frac{1}{k}\sum_{i=-1}^{k-2} \sigma^i \nu^*(A) = \mu(\sigma^{-1}(A))$, por lo tanto es σ invariante, y es una medida de entropía maximal, pues ν lo es en X^k y la entropía es una transformación afín entonces $h_\mu = \frac{1}{k}\sum_{1=0}^{k-1} h_{\sigma^i \nu^*}$. Como (X^k, σ) es topológicamente isomorfo a (X, σ^k) entonces su entropía topológica es la misma y entonces ν^* es una medida de entropía maximal para (X, σ^k), por lo tanto μ es una medida de entropía maximal para X.

Entonces $\mu([0]) = 0$ lo cual implica que $\nu([u]) = 0$ para cualquier palabra de tamaño k que contenga a 0. Dado que ν es ergódica entonces ν está soportado enteramente por el shift completo en $\{1, \ldots, N\}^k$ o en los negativos por lo tanto tenemos dos medidas de entropía maximal ergódicas con soportes ajenos.

□

Con esto encontramos muchos sistemas que tiene la propiedad de especificación no uniforme y que no son intrínsecamente ergódicos. Es importante notar que estos sistemas deben de tener una función gap que crezca mas rápido que logaritmo y que, a la vez, tengan un abecedario muy grande. Nos interesaría ver que pasa si alguna de estas cosas no se cumple y si podemos hablar mas de ergodicidad del sistema, es así como Pavlov da el siguiente teorema.

Teorema 3.10. *Si un subshift tiene especificación no uniforme con función gap $f(n)$ donde $\liminf_{n\to\infty} \frac{f(n)}{\log n} = 0$ entonces no puede tener dos medidas de entropía maximal con soportes ajenos.*

Demostración. Supongamos por contradicción que X es un subshift con especificación no uniforme con función gap $f(n)$ tal que $\liminf_{n\to\infty} \frac{f(n)}{\log n} = 0$ el cual tiene dos medidas de entropía maximal ν y μ con soportes ajenos.

Para cada n definamos el conjunto $\mathcal{M}_n(X) = \{w \in \mathcal{L}_n(X) | \mu([x]) > 0\}$ y el conjunto $\mathcal{N}_n(X) = \{w \in \mathcal{L}_n(X) | \nu([w]) > 0\}$ es decir los conjuntos de las palabras tal que la medida de su cilindro no es cero, si $n = 0$ los conjunto solo están dados por la palabra vacía. Para cualquier $k < n$, todas las sub palabras de tamaño k de una palabra en $\mathcal{M}_n(X)$ deben de pertenecer a $\mathcal{M}_n(X)$ y lo mismo para $\mathcal{N}_n(X)$ y $\mathcal{N}_k(X)$, pues si v es la sub palabra de tamaño k de la palabra w entonces $[w] \subset [v]$.

Ahora demostraremos que $|\mathcal{M}_n(X)| \geq e^{nh(\mu)}$. Para cualquier vector de probabilidad $(x_1, \ldots, x_n)$ tenemos que $-\sum_{i=1}^{n} x_i \log x_i \leq \log n$ ya que la función $x \log x$ es estrictamente convexa. Sin pérdida de generalidad podemos hacer que la medida de todo el sistema sea 1 y los elementos de $\mathcal{M}_n(X)$ son tal que sus cilindros forman todo X y por lo tanto la suma de todas las medidas da 1, aplicando esto y dividendo entre n obtenemos que

$$h(\mu) \leq \frac{-1}{n} \sum_{w \in A^n} \mu([w]) \log \mu([w]) \leq \frac{\log |\mathcal{M}_n(X)}{|} n \tag{3.39}$$

y por lo tanto $|\mathcal{M}_n(X)| \geq e^{nh(\mu)} = e^{nh(X)}$ y lo mismo se cumple para $|\mathcal{N}_n(X)| \geq e^{nh(\nu)} = e^{nh(X)}$ pues son maximales para toda n pero como tienen soportes ajenos existe N tal que $\mathcal{M}_n(X) \cap \mathcal{N}_n(X) = \varnothing$ para todo $n \geq N$, pues entre más grandes sean las palabras entonces los cilindros se vuelven mas pequeños y dejan de vivir en ambos soportes.

Elijamos n tal que $n > f(n)$, entonces para cada $i \in [0, n - f(n)]$ tal que sea múltiplo de $f(n) + N$. Escojamos las palabras $w \in \mathcal{M}_i(X)$ y $v \in \mathcal{N}_{n-f(n)-i}(X)$, por definición $i \leq n$ y dado que $f(n)$ es no decreciente se cumple que $n - i - f(n) \leq n$. Sabemos que X tiene especificación no uniforme por lo tanto existe u con longitud $f(n)$ tal que wuv es una palabra válida de tamaño $i + f(n) + n - f(n) - i = n$, ya que para que se cumpla la especificación no uniforme, el tamaño de u debe de ser de tamaño al menos $f(i)$ y $f(n) > f(i)$. Ahora queremos ver que el mapeo de (i, w, v) a wuv, donde u es la opción minimal en el orden lexicográfico que hace este mapeo una función, es uno a uno. Supongamos por contradicción que no lo es, entonces para $(i, w, v) \neq (i', w', v')$ se cumple que $wuv = w'u'v'$. Si $i = i'$ entonces $w \neq w'$ o $v \neq v'$, y por lo tanto no es posible que $wuv = w'u'v'$, por lo tanto tenemos que tomar $i \neq i'$ pero ambas son múltiplos de $f(n) + N$ así que $|i - i'| \geq f(n) + N$ pero wuv y $w'u'v'$ son de tamaño n y el tamaño de u y u' es $f(n)$, así que suponiendo que $i > i'$ entonces al menos w y v' deben de compartir una palabra de tamaño N y si $i < i'$ entonces w' y v lo comparten. Pero recordemos que $w \in \mathcal{M}_i(X)$ y $v \in \mathcal{N}_{n-f(n)-i}(X)$ y por lo tanto implicaría que viven en $\mathcal{M}_n(X) \cap \mathcal{N}_N(X)$ lo cual es una contradicción pues habíamos visto que la intersección es vacía, así pues es uno a uno.

Como es uno a uno la siguiente suma que representa toda la cantidad posible de w y v es la cantidad posible de palabras de tamaño n

$$\sum_{i\in[0,n-f(n)],(f(n)+N)|i} |\mathcal{M}_i(X)||\mathcal{N}_{n-f(n)-i}(X)| \tag{3.40}$$

Entonces usando que $|\mathcal{M}_n(X)| \geq e^{nh(X)}$ y $|\mathcal{N}_n(X)| \geq e^{nh(X)}$ tenemos que es mayor o igual a

$$\sum_{i\in[0,n-f(n)],(f(n)+N)|i} e^{ih(X)}e^{(n-f(n)-i)h(X)} = \sum_{i\in[0,n-f(n)],(f(n)+N)|i} e^{(n-f(n))h(X)} \geq \frac{n-f(n)}{f(n)+N}e^{(n-f(n))h(X)} \tag{3.41}$$

Por lo tanto

$$|\mathcal{L}_n(X)| \geq \frac{n-f(n)}{f(n)+N}e^{(n-f(n))h(X)}. \tag{3.42}$$

Notemos que para cualquier t y cualquier $w^1, \ldots, w^t \in \mathcal{L}_n(X)$ podemos usar la especificación no uniforme para crear la palabra $w^1v^1 \ldots w^tv^t$ de tamaño $t(n+f(n))$ pues todas las v son de tamaño $f(n)$. Esto implica que $|\mathcal{L}_{t(n+f(n))}(X)| \geq |\mathcal{L}_n(X)|^t$.

Entonces como $|\mathcal{L}_{t(n+f(n))}(X)| \leq |\mathcal{L}_t(X)|^{n+f(n)}$ entonces sacando el logaritmo, dividiendo entre t y tomando el límite de t cuando va a infinito obtenemos $\lim_{t\to\infty} \frac{1}{t}(n+f(n))\log|\mathcal{L}_t(X)| \geq \lim_{t\to\infty}\log|\mathcal{L}_n(X)|$ por lo tanto

$$(n+f(n))h(X) \geq \log|\mathcal{L}_n(X)| \tag{3.43}$$

Aplicando las Ecuaciones 3.42 y 3.43 tenemos que para n suficientemente grande

$$(n+f(n))h(X) \geq \log|\mathcal{L}_n(X)| \geq \log(n-f(n)) - \log(f(n)+N) + (n-f(n))h(X), \tag{3.44}$$

lo cual se puede reescribir como

$$2f(n)h(X) + \log(f(n)+N) \geq \log(n-f(n)). \tag{3.45}$$

Ahora si tomamos una sucesión n_k tal que $\frac{f(n_k)}{\log n_k} \to 0$ y $k \to \infty$ entonces si tomamos el lado izquierdo y lo dividimos entre $\log(n_k)$ tenemos que $2f(n_k)h(X)/\log(n_k) \to 0$ y $\log(f(n_k)+N)/\log(n_k) \to 0$ pues $f(n)+N$ crece mas lento que la identidad, pro lo tanto todo el lado derecho es dominado por $\log(n_k)$ asintoticamente, mientras que el lado derecho $\log(n_k - f(n_k)) \to \log(n_k)$ por la izquierda pues $f(n_k)$ crece más lento que la identidad, lo cual es una contradicción pues una función dominada por $\log(n_k)$ tiene que ser menor a $\log(n_k)$. Por lo tanto no puede tener las dos medidas de entropía maximal.

□

Es importante notar que esto no quiere decir que estemos comprobando que el subshift sea intrínsecamente ergódico, pues solo estamos demostrando que no tiene dos medidas ergódicas de entropía maximal con soportes ajenos, pero aún podría tener dos medidas ergódicas de entropía maximal cuyos soportes sean ajenos.

Estas dos demostraciones nos muestran cómo la especificación no uniforme tiene cierta transición de fase de acuerdo a qué tan rápido crece la función gap, de tal forma que permite medidas ergódicas con soportes ajenos, pero aún queda abierto si esta transición de fase también nos permite encontrar subshifts intrínsecamente ergódicos cuando la función gap crezca lentamente.

Ahora nos interesa saber si este tipo de teoremas e ideas pueden ser extendidas a la propiedad de casi especificación. Pavlov encontró que sí y demostró dos teoremas que contienen la misma idea pero antes de demostrarlos necesitamos algunas definiciones extra y unos lemas.

Definición 3.26. Para cualquier alfabeto $\mathcal{A}$ y $n \in \mathbb{N}$ la distancia de Hamming d en $\mathcal{A}^n$ esta dado por $d(v,w) = |\{i : v_i \neq w_i\}|$ es decir el número de lugares donde son diferentes.

Lema 3.16. *Para cada alfabeto $\mathcal{A}$ y un enetro positivo n, existe un conjunto $T_{\mathcal{A},n} \subset \mathcal{A}^n$ tal que $|T_{\mathcal{A},n}| \leq \frac{1}{2^{\lfloor \log_2 n \rfloor}} |\mathcal{A}|^n$ y $T_{\mathcal{A},n}$ es 1-extendido con respecto a la distancia de Hamming d, es decir, para cualquier $w \in \mathcal{A}^n$ existe $t \in T_{\mathcal{A},n}$ tal que $d(t,w) \leq 1$.*

Demostración. Sea $\mathcal{A}$ un alfabeto y n un entero positivo. Podemos asumir sin pérdida de generalidad que $\mathcal{A} = \{0, 1, \ldots, |\mathcal{A}| - 1\}$, definamos $m = \lfloor \log_2 n \rfloor$ tal que $2^m \leq n < 2^{m+1}$. Entonces para cualquier $v = v_0 v_1 \ldots v_{m-1} \in \{0,1\}^m$ definimos $T_{\mathcal{A},n,v}$ como el conjunto de todas las $w = w_0 w_1 \ldots w_{n-1} \in \mathcal{A}^n$ tal que para cada $j \in [0, m)$ la suma de w_i sobre todos los $i \in [0, 2^m - 1]$ cuya expansión binaria tiene un 0 en el lugar 2^j es igual a $v_j \mod 2$. Esta definición es bastante complicada de entender por lo que es más fácil entenderla con un ejemplo. Sea $\mathcal{A} = \{0,1,2\}$ y $n = 10$ entonces $m = \lfloor \log_2(10) \rfloor = 3$ y sea $v = 010$. Ahora $i \in [0,7]$ y sus expansiones binarias son $0 = 000, 1 = 100, 2 = 010, 3 = 110, 4 = 001, 5 = 101, 6 = 011, 7 = 111$ y $j \in [0,3)$. Entonces $T_{\mathcal{A},n,v}$ es el conjunto de todas las $w \in \mathcal{A}^n$ tal que $w_0 + w_2 + w_4 + w_6 = 0 \mod 2$ pues son los números que tienen cero en la coordenada 2^0 y la coordenada 0 de v es 0 pues $j = 0$. Repitiendo lo mismo para $j = 1$ tenemos $w_0 + w_1 + w_4 + w_5 = 1 \mod 2$ y para $j = 2$ tenemos $w_0 + w_1 + w_2 + w_3 = 0$ mód 2, entonces una posible solución podría ser $w = 0121200111$ pues $0 + 2 + 2 + 0 = 4 = 0$ mód $2, 0 + 1 + 2 + 0 = 3 = 1 \mod 2$ y $0 + 1 + 2 + 1 = 4 = 0 \mod 2$.

Mostremos que cualquier conjunto $T_{\mathcal{A},n,v}$ es 1-extendido. Sea $w \in \mathcal{A}^n$. Para algún valor $j \in [0, m)$, la suma de los w_i sobre todas las is cuya expansión binaria tenga un 0 en 2^j es

igual a $v_j \mod 2$ y para algunas no lo es. Definamos $J \subset [0, m)$ como el conjunto de j para las cuales la suma sí sea igual a $v_j \mod 2$. Ahora elijamos $i \in [0, 2^m)$ tal que la expansión binaria de i tenga el valor 0 precisamente en los lugares pertenecientes a los índices 2^j para $j \notin J$, esto es $i = \sum_{j \in J} 2^j$ y $0 \leq i \leq 2^m - 1$. Ahora definamos w' como la palabra obtenida al cambiar w_i a cualquier letra en $\mathcal{A}$ con la paridad opuesta, entonces si buscamos hacer todas estas sumas para w' sabemos que las sumas en J no las estamos modificando por que en la letra que cambiamos su índice es tal que no tiene un 0 en la posición 2^j para $j \in J$, en cambio modifica a todos los otros $j \notin J$ pero como cambia su paridad, cambia la paridad de toda la suma, y como estamos tomando que no es igual a $v_j \mod 2$, entonces al cambiar la paridad por el modulo 2, obtenemos la igualdad, y por lo tanto $w' \in T_{\mathcal{A},n,v}$.

Si regresamos al ejemplo anterior, tomemos la palabra $w = 0211221100$, la cual no pertenece a $T_{\mathcal{A},n,v}$ pues $0 + 1 + 2 + 1 = 4 = 0 \mod 2$, $0 + 2 + 2 + 2 = 6 \neq 1 \mod 2$ y $0 + 2 + 1 + 1 = 4 = 0 \mod 2$ por lo tanto $J = \{0, 2\}$ así que $i = 2^0 + 2^2 = 5$ por lo tanto cambiamos la letra que está en la posición 5 la cual es 2 por una de paridad diferente, en este caso 1, por lo tanto obtenemos la palabra $w' = 0211211100$ y en este caso nos da $0 + 2 + 2 + 1 = 5 = 1 \mod 2$ y por lo tanto está en $T_{\mathcal{A},n,v}$.

Por lo tanto para cualquier palabra $w \in \mathcal{A}$ existe $w' \in T_{\mathcal{A},n,v}$ tal que solo difieren por una letra, es decir su distancia de Hamming es 1 y por lo tanto $T_{\mathcal{A},n,v}$ es 1-extendido. Dado que $T_{\mathcal{A},n,v}$ es una partición de $\mathcal{A}^n$ con 2^m opciones para v, pues por construcción v es una palabra en $\{0, 1\}^m$ con $m = \lfloor \log_2 n \rfloor$, existe un $T_{\mathcal{A},n,v}$ cuya cardinalidad debe de ser menor o igual a $\frac{|\mathcal{A}|^n}{2^m}$ pues es dividir toda la cardinalidad del sistema $|\mathcal{A}|^n$ entre los 2^m elementos, por lo que o todos son iguales y su cardinalidad es $\frac{|\mathcal{A}|^n}{2^m}$ o algunos son más grandes y otros más pequeños, a estos son los que definiremos como $T_{\mathcal{A},n}$.

□

Lema 3.17. *Para cada alfabeto $\mathcal{A}$ y un positivo entero n existe un conjunto $U_{\mathcal{A},n} \subset \mathcal{A}^n$ tal que $|U_{\mathcal{A},n}| \leq \frac{16}{n^2}|\mathcal{A}|^n$ y $U_{\mathcal{A},n}$ es 2-extendido con respecto a la distancia de Hamming d.*

Demostración. Para cualquier n definamos el conjunto $T_{\mathcal{A},\lfloor 0{,}5n \rfloor}$ y $T_{\mathcal{A},\lceil 0{,}5n \rceil}$ como en el lema anterior, entonces definamos $U_{\mathcal{A},n} = \{uv : u \in T_{\mathcal{A},\lfloor 0{,}5n \rfloor}, v \in T_{\mathcal{A},\lceil 0{,}5n \rceil}\}$, entonces $U_{\mathcal{A},n}$ es 2-extendido, pues para cualquier $w \in \mathcal{A}^n$ solo se necesita un cambio en la palabra $w_1 w_2 \ldots w_{\lfloor 0{,}5n \rfloor}$ que tiene tamaño $\lfloor 0{,}5n \rfloor$ a una palabra en $T_{\mathcal{A},\lfloor 0{,}5n \rfloor}$ pues es 1-extendido, lo mismo para la palabra $w_{\lfloor 0{,}5n \rfloor + 1} \ldots w_n$ que tiene tamaño $\lceil 0{,}5n \rceil$ solo se necesita un cambio para llevar a una palabra en $T_{\mathcal{A},\lceil 0{,}5n \rceil}$. Por lo tanto solo se necesitan a lo más dos cambios en la palabra w para tener una palabra en $U_{\mathcal{A},n}$ y por lo tanto es 2-extendido.

Ahora usando que $2^{\lfloor \log_2(\lfloor 0{,}5n \rfloor)\rfloor} \geq 0{,}25n$ entonces por el lema anterior

$$|U_{\mathcal{A},n}| = |T_{\mathcal{A},\lfloor 0{,}5n\rfloor}||T_{\mathcal{A},\lceil 0{,}5n\rceil}| \leq \frac{1}{0{,}25n}|\mathcal{A}|^{\lfloor 0{,}5n\rfloor}\frac{1}{0{,}25n}|\mathcal{A}|^{\lceil 0{,}5n\rceil} = \frac{16}{n^2}|\mathcal{A}|^n.$$

□

Lema 3.18. *Para cada alfabeto $\mathcal{A}$ y entero positivo $n > 1$ existe un conjunto $V_{\mathcal{A},n} \subset \mathcal{A}^n$ tal que $|V_{\mathcal{A},n}| = \frac{1}{|\mathcal{A}|^2}|\mathcal{A}|^n$ y $V_{\mathcal{A},n}$ es 2-extendido con respecto a la distancia de Hamming d.*

Demostración. Definiremos el conjunto $V_{\mathcal{A},n} = \{w = w_1w_2\ldots w_n \in \mathcal{A}^n : w_1 = w_2 = 1\}$. Entonces evidentemente es 2-extendido pues si tomamos un elemento $u \in \mathcal{A}^n$ tal que $u = u_1u_2\ldots u_n$ entonces solo hay que cambiar el valor de u_1 y u_2 a 1, y el elemento vivirá en $V_{\mathcal{A},n}$. El tamaño de $V_{\mathcal{A},n}$ es $|\mathcal{A}|^{n-2} = \frac{|\mathcal{A}|^n}{|\mathcal{A}|^2}$.

□

Con estos lemas y los conjuntos que definen se pueden demostrar los siguientes teoremas que tiene que ver con la ergodicidad de un sistema con casi especificación.

Teorema 3.11. *Existe un subshift con casi especificación y función error $g(n) = 4$ con exactamente dos medidas ergódicas de entropía maximal cuyos soportes son ajenos.*

Demostración. Definiremos X como el alfabeto $\mathcal{A} = \{-N, \ldots, -1, 1, \ldots, N\}$, para un parámetro $l \geq 2$ definiremos

$$P_n = \left\{ \begin{array}{ll} \{1\}, & \text{si } n = 1 \\ V_{\{1,\ldots,N\},n}, & \text{si } 1 < n \leq l \\ U_{\{1,\ldots,N\},n}, & \text{si } n > l \end{array} \right\} \tag{3.46}$$

y

$$N_n = \left\{ \begin{array}{ll} \{-1\}, & \text{si } n = 1 \\ V_{\{-N,\ldots,-1\},n}, & \text{si } 1 < n \leq l \\ U_{\{-N,\ldots,-1\},n}, & \text{si } n > l \end{array} \right\} \tag{3.47}$$

donde $U_{\mathcal{A},n}$ y $V_{\mathcal{A},n}$ son como los definimos en los lemas anteriores. Ahora por simetría tenemos que $|N_n| = |P_n|$ para toda n, entonces les daremos el valor M_n. Como $U_{\mathcal{A},n}$ y $V_{\mathcal{A},n}$ son 2-extendidos entonces los conjuntos P_n y N_n son 2-extendidos en sus respectivos abecedarios.

Definimos X por el conjunto de palabras prohibidas $\mathcal{F}$ el cual contiene todas las palabras nPn' donde $n, n' < 0$ y P es una palabra de letras positivas tal que $P \notin \cap P_n$, de igual manera

contiene a las palabras pNp' tal que $p.p' > 0$, N es una palabra de letras negativas tal que $N \notin \cap N_n$.

Entonces los puntos $x \in X$ pueden estar formados por concatenaciones de palabras de signos constantes tal que pertenezcan a P_n o N_n, se pueden tener punto sin restricciones donde todo el punto esta formado por letras del mismo signo y ya no restringimos a estar en un conjunto. Para cualquier palabra permitida en $\mathcal{L}(X)$, es posible construir un punto que la contenga, simplemente extendiendo la palabra por la izquierda con letras del mismo símbolo que la primera letra, y por la derecha con letras del mismo símbolo que la última.

Consideremos las palabras $w^1, w^2, \ldots, w^k \in \mathcal{L}(X)$ entonces la concatenación de estas palabras w no necesariamente va a estar en $\mathcal{L}(X)$, pues puede pasar que W^i sea una palabra de letras positivas tal que no viva en $\cap P_n$, pero $w^{i-1}yw^{i+1}$ sean palabras válidas de signo negativo, por lo tanto $w^{i-1}w^iw^{i+1}$ no pertenece a $\mathcal{L}(X)$. Sea $w = w^1w^2 \ldots w^k$, si suponemos que no está en el diccionario de X, entonces tiene al menos una de las palabras nPn' o pNp', donde P, N no viven en los conjuntos $\cap P_n, \cap N_n$ respectivamente, pero los conjuntos son 2-extendidos, por lo que solo necesitamos cambiar a lo más dos letras de la palabra P o N para que se vuelva una palabra válida. Esto lo podemos aplicar a cualquier palabra de signo constante que se encuentre en w, en algunas ya será válida la palabra pues vive en algún P_n o N_n y en algunas otras necesitara uno o los dos cambios.

Si alguna de estas palabras son prefijos o sufijos de alguna de los w^i, éstas evidentemente ya no necesitarán cambio pues $w^i \in \mathcal{L}(X)$, ahora es posible que w^i sea de la forma $w^i = rst$ con s una palabra negativa que viva en algún N_n pero a r y t solo tenemos que pedirles que sean de signo positivo en orden de que w^i sea una palabra válida, entonces para que la concatenación con otras posibles palabras sea válida, debemos de aplicar a lo más dos cambios en r y a lo más dos cambios en t. Por lo tanto para palabra w^i que forma w, debemos de cambiar a lo más 4 letras para que toda la concatenación sea válida, por lo tanto tiene casi especificación con función error $g(x) = 4$.

Ahora de forma parecida a la demostración del teorema 3.9 vamos a probar que X tiene dos medidas ergódicas de entropía maximal con soportes ajenos, y la forma de probarlo es muy similar, empecemos probando que $h(X) = \log N$. Como X contiene al shift entero del abecedario $\{1, 2, \ldots, N\}$ y la entropía de este shift completo es $\log N$ entonces $h(X) \geq \log N$. De nuevo vamos a buscar que $h(X) \leq \log N$ y para eso acotaremos $\mathcal{L}(X)$ por arriba. Una palabra w es una unión de palabras de signo constante, entonces definamos a w por el número $k \geq 1$ de palabras de signo constante que la forman y sus longitudes $n_1, n_2, \ldots, n_k$. Si recordamos en el teorema 3.9, las palabras también estaban caracterizadas por un número k de subpalabras de signo constante, por lo tanto podemos recrear la demostración.

Sea

$$|\mathcal{L}_n(X)| = 2N^n + 2(n-1)N^n + \sum_{k=3}^{n} \sum_{\substack{n_1,\dots,n_k \\ \sum n_i = n}} 2N^{n_1} \left(\prod_{j=2}^{k-1} M_{n_j} \right) N^{n_k} \tag{3.48}$$

donde $2N^n$ es el caso $k = 1$ donde es una sola palabra de signo constante, el segundo término es para $k = 2$ donde $(n-1)$ nos da los dos posibles tamaños de las palabras. Y los otros casos cuando $k \geq 3$ y todos las combinaciones de tamaño de palabras y M_{n_j} representa el número de elecciones de subpalabras que puede haber entre la palabra prefijo N^{n_1} y la palabra sufijo N^{n_k}. Haciendo las mismas cotas que en el teorema 3.9 obtenemos que

$$\sum_{k=3}^{n} \sum_{\substack{n_1,\dots,n_k \\ \sum n_i = n}} 2N^{n_1} \left(\prod_{j=2}^{k-1} M_{n_j} \right) N^{n_k} = 2N^n \sum_{k=3}^{n} \sum_{\substack{n_1,\dots,n_k \\ \sum n_1 = n}} \prod_{j=2}^{k-1} (M_{n_j} N^{-n_j}) \tag{3.49}$$

Esto debido a que la suma sobre todas las n_i debe de ser n, ahora aplicando que en el segundo solo podemos usar los valores desde n_2 hasta n_{k-1} obtenemos

$$\begin{aligned} 2N^n \sum_{k=3}^{n} \sum_{\substack{n_1,\dots,n_k \\ \sum n_1 = n}} \prod_{j=2}^{k-1} (M_{n_j} N^{-n_j}) &= 2N^n \sum_{k=3}^{n} \sum_{\substack{n_2,\dots,n_{k-1} \\ \sum n_1 < n}} \left(n - \sum_{j=2}^{k-1} n_j - 1 \right) \prod_{j=2}^{k-1} (M_{n_j} N^{-n_j}) \\ &\leq 2nN^n \sum_{k=3}^{n} \sum_{\substack{n_2,\dots,n_{k-1} \\ \sum n_1 < n}} \prod_{j=2}^{k-1} (M_{n_j} N^{-n_j}) \end{aligned} \tag{3.50}$$

y aplicando el mismo argumento que en 3.9

$$2nN^n \sum_{k=3}^{n} \sum_{\substack{n_2,\dots,n_{k-1} \\ \sum n_1 < n}} \prod_{j=2}^{k-1} (M_{n_j} N^{-n_j}) \leq 2nN^n \sum_{k=3}^{n} (\sum_{t=1}^{\infty} M_t N^{-t})^{k-2} \tag{3.51}$$

Ahora usando los lemas anteriores podemos encontrar una forma de acotar $\sum_{t=1}^{\infty} M_t N^{-t}$, pues M_t es la cantidad de palabras de signo constante que podemos elegir, entonces $M_t = P_t$ o $M_t = N_t$, como en la construcción de las suma existe un 2 el cual representa esta dualidad nos basta trabajar con una sola entonces

$$\sum_{t=1}^{\infty} M_t N^{-t} = \frac{1}{N} + \sum_{t=2}^{l} |V_{\{1,\dots,N\},t}| N^{-t} + \sum_{t=l+1}^{\infty} |U_{\{1,\dots,N\},t}| N^{-t} \tag{3.52}$$

Ahora usando que $|U_{\{1,\dots,N\},t}| \leq \frac{16}{t^2}N^t$ y $|V_{\{1,\dots,N\},t}| = \frac{N^2}{N^t}$ obtenemos

$$\frac{1}{N} + \sum_{t=2}^{l} |V_{\{1,\dots,N\},t}|N^{-t} + \sum_{t=l+1}^{\infty} |U_{\{1,\dots,N\},t}|N^{-t} \leq N^{-1} + \sum_{t=2}^{l} N^{-2} + \sum_{t=l+1}^{\infty} 16t^{-2} \tag{3.53}$$

$$\leq N^{-1} + (l+1)N^{-2} + 16l^{-1} \tag{3.54}$$

Nos interesa elegir N y l de tal manera que esta suma final sea menor que 1, de esta manera si renombramos $\sum_{t=1}^{\infty} M_t N^{-t} = \alpha < 1$ y de esta forma podemos volver a aplicar lo que usamos en el teorema 3.9 tenemos

$$|\mathcal{L}(X)| \leq 2N^n + 2(n-1)N^n + 2nN^n \sum_{k=3}^{\infty} \alpha^{k-2} = \frac{2n}{1-\alpha} N^n \tag{3.55}$$

Y tomando el logaritmo de ambos lados, dividiendo entre n y tomando el límite obtenemos $h(X) \leq \log N$ y por lo tanto $h(X) = \log N$.

Ahora tomemos una medida ergódica de entropía maximal μ en X, esta vez queremos mostrar que la medida del cilindro $[ij]$ es cero para cada par i, j de signo contrario, entonces por contradicción asumamos que $\mu([ij]) = \beta > 0$. Para cada n, tomemos el conjunto $C_{n,0,5\beta,ij}(X)$, al aplicar el lema 3.15 obtenemos $\liminf_{n\to\infty} \frac{\log |C_{n,0,5\beta,ij}|}{n} \geq h(\mu) = h(X) = \log N$. Recordando que el conjunto $C_{n,0,5\beta,ij}(X)$ esta formado por las palabras de tamaño n que tienen entre $0{,}5n\beta$ y $1{,}5n\beta$ ocurrencias de ij, y sabemos que al ser palabras de tamaño n están formadas por k palabras de signo constante, y ocupando los mismo argumentos de las desigualdades anteriores y sabiendo que al menos $k = 0{,}5n\beta$ pues tiene al menos esa cantidad de ij obtenemos

$$|C_{n,0,5\beta,ij}(X)| < 2nN^n \sum_{k=\lfloor 0,5n\beta \rfloor}^{\infty} \alpha^{k-2} \leq \frac{2n}{1-\alpha} N^n \alpha^{0,5n\beta} \tag{3.56}$$

Entonces si tomamos el logaritmo del lado derecho tenemos $\log(\frac{2n}{1-\alpha} N^n \alpha^{0,5n\beta}) = \log(2n) - \log(1-\alpha) + \log(N^n) + 0{,}5n\beta \log(\alpha)$ y dividiendo entre n y sacando el límite, y haciendo lo mismo del otro lado obtenemos

$$\liminf_{n\to\infty} \frac{\log |C_{n,0-5\beta,ij}(X)|}{n} \leq \log N + 0{,}5\beta \log(\alpha) < \log N \tag{3.57}$$

Esto es una contradicción pues ya habíamos dicho que era mayor o igual a $\log N$, por lo tanto $\mu([ij]) = 0$ para todo i, j de signos contrarios. Entonces μ debe de estar completamente soportada por $\{1, \ldots, N\}^{\mathbb{Z}}$ o por $\{-N, \ldots, -1\}^{\mathbb{Z}}$ y las únicas medidas que cumplen esto cuya entropía es $\log N$ son las medidas de Bernoullli, por lo tanto tenemos una para el shift formado por la parte positiva y otra para el formado por la parte negativa, lo cual las hace que estas dos medidas que acabamos de definir tengan soportes ajenos.

□

Con este tenemos demostrado que se pueden encontrar shifts con casi especificación tal que su función error sea muy simple y acotada tal que no cumpla ser intrínsecamente ergódico. A diferencia del caso con especificación no uniforme no necesitamos que el alfabeto sea muy grande, solo hay que saber cómo acotar l tal que a partir de cierto tamaño de palabra cambiemos entre los conjuntos $U_{\mathcal{A},n}$ y $V_{\mathcal{A},n}$.

Nos interesa ahora saber si hacer mas chica la función error sí nos lleva a poder mostrar que el shift es intrínsecamente ergódico, esto es aún un problema abierto, pero Pavlov encontró una cota que a pesar de no dar ergodicidad intrínseca nos acerca mucho, pero antes de estudiar esa cota necesitamos el siguiente lema.

Lema 3.19. *Para cualquier alfabeto $\mathcal{A}$, entero positivo n y conjunto $W \subset \mathcal{A}^n$, existe un conjunto $S \subset W$ tal que $|S| \geq \frac{|W|}{4n|\mathcal{A}|^2}$ y S es 3-separado con respecto a la distancia de Hamming, es decir, para todo $w, w' \in S$ $d(w, w') \geq 3$.*

Demostración. Elijamos cualquier $\mathcal{A}$ y n y sin pérdida de generalidad $\mathcal{A} = \{0, 1, \ldots, |\mathcal{A}|-1\}$ entonces para cualquier $i \in [0, 2|\mathcal{A}|]$ y $j \in [0, 2|\mathcal{A}|n]$ definimos $S_{n,i,j} = \{w = w_1w_2w_3 \ldots w_n \in W : \sum_{k=1}^{n} w_k = i \mod 2|\mathcal{A}|, \sum_{k=1}^{n} kw_k = j \mod 2|\mathcal{A}|n\}$. Queremos mostrar que $S_{n,i,j}$ es 3-separado con respecto a la distancia de Hamming, es decir que para 2 palabras en el conjunto, al menos tiene 3 letras diferentes. Si tomamos una palabra en $S_{n,i,j}$ y le cambiamos una sola letra, entonces ya no vivirá en el conjunto pues cambiará la suma $\sum_{k=1}^{n} w_k$ con una distancia de a lo más $|\mathcal{A}|-1$ por lo que no podría seguir siendo igual a $i \mod 2|\mathcal{A}|$. Ahora supongamos por contradicción que existe $w \neq w' \in S_{n,i,j}$ que difieren de a lo más 2 letras, entonces $\sum_{k=1}^{n} w_k$ y $\sum_{k=1}^{n} w'_k$ son iguales a $i \mod 2|\mathcal{A}|$, pero a lo más pueden diferir por $2(|\mathcal{A}|-1) < 2|\mathcal{A}|$ por lo tanto tienen que ser iguales. De la misma manera $\sum_{k=1}^{n} kw_k$ y $\sum_{k=1}^{k} k'_w$ deben de ser iguales a $j \mod 2|\mathcal{A}|n$ pero difieren de a lo mas $n(|\mathcal{A}|-1) + (n-1)(|\mathcal{A}|-1) < 2|\mathcal{A}|n$ por lo tanto también tiene que ser iguales. Pero w y w' solo tienen dos letras diferentes supongamos que en a, b entonces $w_a + w_b = w'_a + w'_b$ y $aw_a + bw_b = aw'_a + bw'_b$ y la única forma de que se cumpla este sistema es que $w_a = w'_a$ y $w_b = w'_b$ lo cual es una contradicción, por lo tanto $S_{n,i,j}$ es 3-separado.

Dado que $S_{n,i,j}$ nos genera una partición del conjunto W, pues un elemento que vive en $S_{n,i,j}$ no puede vivir en $S_{n,i',j'}$. Como depende de los valores que tomemos para i, j hay $4n|\mathcal{A}|^2$ pues $i \in [0, 2|\mathcal{A}|]$ y $j \in [0, 2|\mathcal{A}|n]$, entonces uno de los $S_{n,i,j}$ tendrá al menos cardinalidad de $\frac{|W|}{4n|\mathcal{A}|^2}$, y definimos a ese conjunto como S.

□

Con este lema, estamos listos para demostrar el último teorema el cual nos da una cota para acercarnos a un sistema con casi especificación intrínsecamente ergódico, este teorema es debido a Pavlov.

Teorema 3.12. *Si un subshift tiene casi especificación con función error $g(n) = 1$ entonces no puede tener dos medidas de entropía maximal con soportes ajenos.*

Demostración. Supongamos por contradicción que X es un subshift con casi especificación y función error $g(n) = 1$ y dos medidas de entropía maximal μ, ν con soportes ajenos. De nuevo como en el teorema 3.10 para cada n vamos a definir los conjuntos $\mathcal{M}_n(X) = \{w \in \mathcal{L}_n(X) : \mu([w]) > 0\}$ y $\mathcal{N}_n(X) = \{w \in \mathcal{L}_n(X) : \nu([w]) > 0\}$ y como ya vimos en 3.10 tenemos que $|\mathcal{M}_n(X)| \geq e^{nh(\mu)} = e^{nh(X)}$ y $|\mathcal{N}_n(X)| \geq e^{nh(X)}$ para todo n y por lo tanto existe N tal que $\mathcal{M}_n(X) \cap \mathcal{N}_n(X) = \varnothing$ para todo $n \geq N$.

Para cada n podemos usar el lema anterior y encontrar subconjuntos $\mathcal{M}'_n \subset \mathcal{M}_n(X)$ y $\mathcal{N}'_n \subset \mathcal{N}_n(X)$ tal que sean 3-separados por la distancia de Hamming y que su norma esté acotada por abajo por $\frac{e^{nh(X)}}{4n|\mathcal{A}|^2}$. Definamos como $S_n = \text{mín}(|\mathcal{M}'_n|, |\mathcal{N}'_n|)$. Elijamos t tal que $\sum_{j=1}^{t} \frac{1}{12jN|\mathcal{A}|^2} > 1$ y denotamos esta suma como α.

Ahora tomaremos cualquier $n > 3tN$ y construiremos palabras en $\mathcal{L}_n(X)$ de la siguiente manera, sea $k = \lfloor \frac{n}{3tN} \rfloor$ y definamos cualquier $n_i \in [1, t]$ para cualquier i entre 1 y k. Ahora dado que $n > 3tN$ y por lo tanto $n/3tN > 1$ y $\lfloor n/3tN \rfloor \leq 1$. Ahora elijamos cualesquiera palabras $w_1 \in \mathcal{M}'_{3Nn_1}$, $w_2 \in \mathcal{N}'_{3Nn_2}$ y así seguimos alternando entre los conjuntos hasta que w_k pertenece a $\mathcal{M}'_{3Nn_k}$ si k es impar, y si k es par pertenece a $\mathcal{N}'_{3Nn_k}$, y el término w_{k+1} lo elegimos en $\mathcal{N}'_{n-\Sigma_{i=1}^{k} 3Nn_i}$ o en $\mathcal{M}'_{n-\Sigma_{i=1}^{k} 3Nn_i}$ de acuerdo a su paridad. Como el sistema tiene casi especificación con función error $g(n) = 1$ podemos construir la palabra $f(w_1 \dots w_{k+1}) = v_1 v_2 \dots v_{k+1} \in \mathcal{L}_n(x)$ pues la suma de la norma de todas es $\sum_{i=1}^{k} 3Nn_i + n - \sum_{i=1}^{k} 3Nn_i = n$, y cada v_i difiere de w_i por a lo más una letra.

Al igual que en el teorema 3.10 vamos a probar que esta función f es inyectiva, es decir que $f(w_1, \dots w_{k+1}) \neq f(w'_1, \dots, w'_{k+1})$ a menos que $w_i = w'_i$ para toda i. Por contradicción

asumamos que no es inyectiva, es decir existen la tuplas de tamaño $k+1$ diferentes, (w_i) y (w'_i) tal que $f(w_1, \dots w_{k+1}) = f(w'_1, \dots, w'_{k+1})$ y para esto hay dos posibles casos.

Primero puede suceder que todas las palabras tengan el mismos tamaño, es decir que $|w_i| = |w'_i|$ para toda i, como la norma es $3Nn_i$ esto quiere decir que $n_i = n'_i$ para todo i y para que sea diferente la tupla entonces al menos existe j tal que $w_j \neq w'_j$, pero como $f(w_1, \dots, w_{k+1}) = f(w'_1, \dots w'_{k+1})$ entonces tanto w_j y w'_j se deben de convertir en v_j al cambiar una sola de sus letras, entonces a lo más difieren entre sí de dos letras, lo cual es una contradicción pues $\mathcal{M}'_{n_j}$ y $\mathcal{N}'_{n_j}$ son 3-separados.

El segundo caso sucede cuando $n_j \neq n'_j$ para algún j, es decir cuando las palabras no son del mismo tamaño. Entonces elijamos el j más chico tal que $n_j \neq n'_j$ y sin pérdida de generalidad asumamos que $n_j < n'_j$. Dado que todos los w_i y w'_i tienen longitudes múltiplos de $3N$, entonces las sub palabras u y u' de $w_1 w_2 \dots w_{k+1}$ y $w'_1 w'_2 \dots w'_{k+1}$ de longitud $3N$ que comienzan en el índice $\sum_{i=1}^{j} 3Nn_i + 1$ son sub palabras de w_{j+1} y w'_j respectivamente. pues el tamaño de w'_j es mas grande que w_j por al menos $3N$..

Entonces ya sea que $u \in \mathcal{M}'_{n_{j+1}}$ y $u' \in \mathcal{N}'_{n_j}$ o $u \in \mathcal{N}'_{n_{j+1}}$ u $u' \in \mathcal{M}'_{n_j}$. Pero sabemos que $f(w_1, \dots, w_{k+1}) = f(w'_1, \dots, w'_{k+1})$ por lo que u y u' se convierten en la misma palabra u'' después de hacerles a lo más un cambio a cada una de ellas. Dado que $|u''| = 3N$ entonces u'' debe de contener al menos una palabra de tamaño N la cual no cambia en u y u' y por lo tanto es una palabra que viven en ambas, lo cual es una contradicción pues ninguna palabra de tamaño N puede vivir en $\cap \mathcal{M}_i(X)$ y en $\cap \mathcal{N}_i(X)$ pues habíamos visto que su intersección es vacía. Por lo tanto la función es inyectiva.

Como la función es inyectiva podemos generar una partición y ver que genera

$$|\mathcal{L}_n(X)| \geq \sum_{\substack{n_1, \dots n_k \\ 1 \leq n_i \leq t}} \left(\prod_{i=1}^{k} S_{3Nn_i} \right) S_{n - \sum_{i=1}^{k} 3Nn_i}. \tag{3.58}$$

Utilizando el lema 3.19 y la cota que genera tenemos

$$\begin{aligned} \sum_{\substack{n_1, \dots n_k \\ 1 \leq n_i \leq t}} \left(\prod_{i=1}^{k} S_{3Nn_i} \right) S_{n - \sum_{i=1}^{k} 3Nn_i} &\geq \sum_{\substack{n_1, \dots, n_k \\ 1 \leq n_i \leq t}} \left(\prod_{i=1}^{k} \frac{e^{3Nn_i h(X)}}{12N|\mathcal{A}|^2 n_i} \right) \frac{e^{(n - \sum 3Nn_i)h(X)}}{4|\mathcal{A}|^2 (n - \sum 3Nn_i)} \\ &\geq \frac{e^{nh(X)}}{4n|\mathcal{A}|^2} \sum_{\substack{n_1, \dots, n_k \\ 1 \leq n_i \leq t}} \prod_{i=1}^{k} \frac{1}{12N|\mathcal{A}|^2 n_i} \end{aligned} \tag{3.59}$$

Ahora como estamos tomando todas las posibles sumas de k elementos sobre todos los elementos $i \in [1, t]$ y recordado el valor de α y k tenemos

$$\frac{e^{nh(X)}}{4n|\mathcal{A}|^2} \sum_{\substack{n_1,\ldots,n_k \\ 1 \leq n_i \leq t}} \prod_{i=1}^{k} \frac{1}{12N|\mathcal{A}|^2 n_i} = \frac{e^{nh(X)}}{4n|\mathcal{A}|^2} \left(\sum_{j=1}^{t} \frac{1}{12N|\mathcal{A}|^2 j} \right)^k \geq \frac{e^{nh(X)}}{4n|\mathcal{A}|^2} \alpha^{n/6tN} \tag{3.60}$$

Por todo esto $\mathcal{L}_n(X) \geq \frac{e^{nh(X)}}{4n|\mathcal{A}|^2} \alpha^{n/6tN}$ para todo $n > 3tN$, pero si tomamos el logaritmo dividimos entre n y sacamos el límite cuando n tiende a infinito la ecuación nos queda $h(X) \geq h(X) + \log(\alpha)/6tN$ lo cual es una contradicción pues $\alpha > 1$. Por lo tanto es falso suponer la existencia de dos medidas de entropía maximal con soportes ajenos.

□

De nuevo esto no nos está diciendo que el sistema sea intrínsecamente ergódico, pero nos acerca mucho a pensar que esta función error $g(n) = 1$ podría dar la propiedad. De hecho Pavlov y Climenhaga demostraron en [4] que la función error $g(n) = 1$ sí implica que el sistema es intrínsecamente ergódico, para esto usaron una nueva propiedad que llamaron casi especificación de un lado, la cual es un punto medio entre casi especificación y especificación no uniforme, pues solo cambias letras de la primera palabra. En este caso si la función error está acotada entonces el sistema es intrínsecamente ergódico. Como podemos ver que un sistema tenga casi especificación de un lado nos da más información que las otras propiedades, pues si podemos encontrar un grupo de funciones error que nos generen una medida de entropía maximal única. En el mismo artículo encontraron que un sistema con casi especificación y función error constante igual a 1 implica que el sistema tiene casi especificación del lado izquierdo con una función error acotada y por lo tanto es intrínsecamente ergódica. Este estudio queda fuera de nuestro alcance, pero muestra que aún hay mucho más rango de investigación en cómo la especificación y sus formas más débiles están relacionadas con los sistemas intrínsecamente ergódicos.

Bibliografía

[1] Blanchard, F. Fully Positive Topological Entropy and Topological Mixing, Symbolic Dynamics and its Applications **135** (1991).

[2] Bowen, R. Periodic Points and Measures for Axiom A Diffeomorphism. Transactions of the American Mathematical Society **154** (1971).

[3] Brin, M. y Stuck, G. Introduction to Dynamical Systems. Cambridge University Press 2004.

[4] Climenhaga, V. y Pavlov, R. One sided Almost specification and intrinsic ergodicity. Ergodic Theory and Dynamical Systems **39**. (2019).

[5] Climenhaga, V. y Thompson, D. Intrinsic Ergodicity Beyond Specification: β-shifts, S-gap Shifts and their Factors. Israel Journal of Mathematics **192** (2012).

[6] Denker, M. Grillenberg, C y Sigmund, K. Ergodic Theory in Compact Spaces. Springer-Verlag Berlin 1976.

[7] Devaney, R. An Introduction to Chaotic Dynamical Systems. Addison-Wesley Publishing Company 1989.

[8] Downarowicz, T. Entropy in Dynamical Systems. Cambridge University Press 2011.

[9] Johnson, K. Beta-shift Dynamical Systems and their Associated Languages. Chapell Hill, 1999.

[10] Kwietniak, D. Lacka, M. y Oprocha, P. A Panorama of Specification-like properties and their Consequences. Contemporary Mathematics **669** (2016).

[11] Kwietniak, D. y Oprocha, P. A note on the average shadowing property for expansive maps. Topology and Applications **159** (2012).

[12] Lind, D. y Marcus, B. An Introduction to Symbolic Dynamics and Coding. Cambridge University Press 1995.

[13] Marcus, B. A Note on Periodic Points for Ergodic Toral Automorphisms. Monaths. Math **89** (1980).

[14] Pavlov, R. On Intrinsic Ergodicity and Weakenings of the Specification Property. Advances in Mathematics **295** (2016).

[15] Pavlov, R. Perturbations of Multidimensional Shifts of Finite Type, Ergodic Theory and Dynamical Systems **31** (2011).

[16] Pfister, C y Sullivan, W. On the Topological Entropy of Saturated Sets. Ergodyc Theory and Dynamic Sets **0** (2005).

[17] Pfiester, C y Sullivan, W. Large Deviations Estimates for Dynamical Systems without the Specification Property. Applications to the β-shifts. Nonlinearity **18** (2005).

[18] Thompson, D. Irregular Sets, the β-Transformation and the Almost Specification Property, Trans. Amer. Math. Soc. **384** (2012).

[19] Weiss, B. Intrisically Ergodic Systems, Hebrew University, 1970.

[20] Wen Huang y Xiangdong Ye. A Local Variational Relation and Aplications, Israel J. Math, **151** (2006).