



Universidad de Murcia

Facultad de Matemáticas

Trabajo Fin de Máster

Curso 2019 / 2020

MEDIDAS INVARIANTES ABSOLUTAMENTE CONTINUAS

Trabajo realizado por:

ALEJO LÓPEZ MOLINA

Dirigido por:

VÍCTOR MANUEL JIMÉNEZ LÓPEZ

Declaración de originalidad

Alejo López Molina, autor del Trabajo de Fin de Máster “Medidas Invariantes Absolutamente Continuas”, bajo la tutela del profesor Víctor Manuel Jiménez López, declara que el trabajo que presenta es original, en el sentido de que ha puesto el mayor empeño en citar debidamente todas las fuentes utilizadas, y que la obra no infringe el copyright de ninguna persona.

En Murcia, a 21 de junio de 2020

A handwritten signature in black ink, appearing to read 'Alejo', written over a horizontal line.

Fdo.: Alejo López Molina

A mis padres por todo su apoyo en la vida, en particular en este máster y más aún en estas circunstancias tan desafortunadas que nos ha tocado vivir. También a mi fiel compañero de aventuras, Coke.

Agradecer a mi tutor, Víctor, su inestimable ayuda y consejos en la realización de este trabajo.

Introducción

Uno de los problemas más difíciles de la teoría de sistemas dinámicos es el de proporcionar información útil acerca del comportamiento asintótico del sistema cuando este es caótico. En el ámbito de la dinámica discreta, que es donde se ubica el presente trabajo, la mejor opción (cuando ello es posible) es encontrar medidas de probabilidad μ no triviales (idealmente, ergódicas) que sean invariantes (es decir, $\mu(\tau^{-1}(A)) = \mu(A)$ para los conjuntos A de la σ -álgebra del espacio de medida en que se trabaja) para la transformación τ que define el sistema, ya que este tipo de medidas dan peso “grande” a los conjuntos con dinámica relevante. Cuando el espacio de fases es el intervalo, que es el caso más estudiado y al que nos limitaremos aquí, es también deseable que tales medidas sean absolutamente continuas con respecto a la medida de Lebesgue (es decir, que los conjuntos de medida de Lebesgue cero también tengan medida μ cero), porque entonces existe una función de densidad con la que podemos computar la medida (con ayuda de herramientas informáticas) en casos concretos.

El objetivo de este trabajo es estudiar las propiedades fundamentales de dichas medidas y probar su existencia y, en los casos que la haya, unicidad para determinados tipos de transformaciones: las expansivas a trozos y las aplicaciones de Markov, en particular las aplicaciones de Markov lineales a trozos.

De esta forma, los resultados principales que presentamos en este trabajo son, por un lado, la existencia de medidas invariantes absolutamente continuas, respecto a la medida de Lebesgue (acims), con densidad de variación acotada para transformaciones expansivas a trozos con buenas propiedades de diferenciabilidad, y por otro la existencia y unicidad de acims equivalentes a la medida de Lebesgue para transformaciones de Markov lineales a trozos con matriz de incidencia primitiva. Más aún, obtendremos en este caso ergodicidad para la medida (esto significa que si $\tau^{-1}(A) = A$, entonces $\mu(A) = 0$ o $\mu(A) = 1$), lo que tendrá importantes consecuencias dinámicas.

En concreto, a través del uso de teoría ergódica, deduciremos que en el sistema dinámico $(I, \mathcal{B}, \mu, \tau)$, donde I es un intervalo compacto y \mathcal{B} la σ -álgebra de Borel, la frecuencia relativa con la que la órbita de un punto $x \in I$, $\{\tau^n(x)\}_{n=0}^{+\infty}$, cae dentro de un conjunto de Borel $A \subset I$ converge a su medida, es decir,

$$\frac{1}{n} \sum_{i=0}^{n-1} \chi_A(\tau^i(x)) \longrightarrow \mu(A), \quad \mu - c.t.p.,$$

y por la equivalencia entre la medida μ y la medida de Lebesgue λ podemos garantizar incluso

$$\frac{1}{n} \sum_{i=0}^{n-1} \chi_A(\tau^i(x)) \longrightarrow \mu(A), \quad \lambda - c.t.p.$$

Se trata de un resultado muy paradójico y muy fuerte: por un lado, el comportamiento de las órbitas es, en apariencia, aleatorio e imprevisible (en particular, son densas en todo I); por otro, “prácticamente” todas se comportan, a largo plazo, del mismo modo. Hay caos y, no obstante, podemos hacer predicciones probabilísticas sobre la dinámica.

Como cabe imaginar, estos resultados no son triviales y para poder llegar hasta ellos es necesario utilizar herramientas potentes pertenecientes a tres ámbitos principales: *teoría de la medida*, *funciones de variación acotada* y, como acabamos de mencionar, *teoría ergódica*. En este proyecto trabajamos algunos de los resultados más importantes de dichos ámbitos que más adelante serán mencionados. Asimismo, también será necesario emplear resultados de la llamada *teoría de Perron-Frobenius*, ya que una de las herramientas técnicas principales del trabajo es el *operador de Perron-Frobenius*, la extensión natural al ámbito de ciertos espacios de funciones del teorema del mismo nombre para matrices no negativas que probamos en el trabajo fin de grado “El teorema de Perron-Frobenius, Google y la Demografía”, [10]. Por lo que, en cierto modo, este proyecto se puede ver como una extensión de dicho trabajo. El texto de referencia que hemos empleado en este proyecto es [1], pero para cada uno de los tres ámbitos mencionados emplearemos textos más específicos.

Para el correcto desarrollo del trabajo y la consecución de los objetivos que nos planteamos lo hemos dividido en cuatro capítulos. A continuación explicamos la distribución de los mismos.

El primer capítulo está dedicado a presentar todos aquellos preliminares necesarios para desarrollar los siguientes capítulos. En él distinguimos tres partes diferentes dedicadas a las tres teorías fundamentales con las que vamos a trabajar: de la medida, de funciones de variación acotadas y ergódica. La primera de ellas, la teoría de la medida, es la rama del Análisis que estudia las medidas y una de las que más importancia recibe en las asignaturas de Análisis Matemático del grado de Matemáticas. En esta parte repasamos algunas de las definiciones básicas de la teoría como la definición de medida, σ -álgebra, función medible, norma \mathcal{L}^1 , función de densidad, etc. Asimismo, supondremos conocidos algunos resultados básicos de la teoría estudiados durante el grado, por ejemplo los teoremas de la convergencia monótona y dominada o el lema de Fatou.

En esta parte el resultado fundamental que demostramos es el teorema de Radon-Nikodym para dos medidas finitas, donde una de ellas es absolutamente continua respecto a la otra. Este resultado nos va a permitir probar la existencia del operador de Perron-Frobenius y, por ende, la expresión de medidas en la forma

$$\mu(A) = \int_A f d\lambda,$$

con $f \in \mathcal{L}^1(I)$. Estas medidas serán de probabilidad cuando f sea una función de densidad en el intervalo, es decir, una función no negativa con norma \mathcal{L}^1 unitaria.

Para la prueba de este teorema vamos emplear medidas reales, conjuntos positivos y negativos y el teorema de descomposición de Hahn.

En la segunda parte estudiamos algunas propiedades importantes de las funciones de variación acotada, aquellas con la propiedad de que el supremo de las sumas

$$\sum_{i=1}^q |f(x_i) - f(x_{i-1})|$$

sobre las particiones $\mathcal{P} = \{x_0, x_1, \dots, x_q\}$ de I (la variación total de f) es finito, e introduciremos el subespacio (en $\mathcal{L}^1(I)$) de las funciones de variación acotada, $BV(I)$, y su norma específica asociada, $\|\cdot\|_{BV}$.

El resultado más relevante de esta parte del trabajo es el teorema de elección o de selección de Helly, el cual nos dice que si tenemos una familia de funciones uniformemente acotada por un valor K y la variación total de cada función está asimismo acotada por K , entonces, para toda sucesión de la familia, se puede extraer una subsucesión de la misma convergente puntualmente a una función de variación acotada. Este resultado nos permitirá probar que un conjunto acotado en la norma $\|\cdot\|_{BV}$ es precompacto en $\mathcal{L}^1(I)$ (de hecho en $BV(I)$) y en particular que, para transformaciones expansivas a trozos, la densidad de la acim encontrada es de variación acotada.

Por otro lado, probamos la densidad de las funciones de paso en $\mathcal{L}^1(I)$. Estas funciones son similares a las funciones simples, pero en este caso son constantes sobre intervalos y no cualquier conjunto medible. Tienen gran importancia en el trabajo, ya que demostraremos que las funciones de densidad de las acims, en el caso de las transformaciones de Markov lineales a trozos, se corresponden, bajo ciertas condiciones, con funciones de paso. Además, como las funciones de paso son de variación acotada, por su densidad en $\mathcal{L}^1(I)$, las funciones de variación acotada también serán densas en $\mathcal{L}^1(I)$ y con ello podremos probar en el capítulo tres que la densidad de cualquier acim, para una transformación expansiva a trozos, es una función de variación acotada.

La tercera parte de este capítulo está dedicada a la teoría ergódica. Esta teoría se dedica principalmente al estudio de la dinámica de una transformación a largo plazo. En particular, en este trabajo nos centraremos en uno de los principales resultados de la teoría, el teorema ergódico de Birkhoff. Este teorema nos dice que si estamos en un sistema dinámico $(\mathbb{X}, \mathcal{A}, \mu, \tau)$, donde τ es una transformación que conserva la medida de probabilidad μ , y $f \in \mathcal{L}^1(\mathbb{X})$, entonces existe una función $f^* \in \mathcal{L}^1(\mathbb{X})$ tal que

$$\frac{1}{n} \sum_{k=0}^{n-1} f(\tau^k(x)) \longrightarrow f^*, \mu - c.t.p.$$

Además, $f^* \circ \tau = f^*$ $\mu - c.t.p.$, $\|f^*\|_1 \leq \|f\|_1$ y $\int_{\mathbb{X}} f^* d\mu = \int_{\mathbb{X}} f d\mu$. En definitiva, el promedio de la órbita de un punto $x \in \mathbb{X}$ bajo f es convergente a una función en $\mathcal{L}^1(\mathbb{X})$.

Sin embargo, este no será el resultado que más nos interese, sino un corolario del mismo que afirma que si τ es ergódica, entonces la frecuencia relativa con la que la órbita de un punto cae

en un conjunto converge a la medida μ de dicho conjunto. De aquí deduciremos la dinámica comentada en los objetivos del trabajo.

Pero para poder probar estos resultados necesitaremos una buena cantidad de resultados previos. Entre ellos podemos destacar que si τ es ergódica y $f \circ \tau = f \mu - c.t.p.$, entonces f es constante $\mu - c.t.p.$, y el teorema ergódico maximal, que garantiza que la integral sobre el conjunto

$$M = \left\{ x : \sup_{n \geq 1} A_n(f)(x) > 0 \right\},$$

con

$$A_n(f) = \frac{1}{n} \sum_{i=0}^{n-1} f \circ \tau^i,$$

es no negativa.

El segundo capítulo de este trabajo está dedicado a la herramienta principal del mismo y que ya ha sido mencionada, el operador de Perron-Frobenius asociado a una transformación τ , P_τ . Hace corresponder a cada $f \in \mathcal{L}^1(I)$ la única función $P_\tau f$ de $\mathcal{L}^1(I)$ que cumple

$$\int_A P_\tau f d\lambda = \int_{\tau^{-1}(A)} f d\lambda$$

para cualquier $A \in \mathcal{B}$. En este capítulo deducimos la existencia de este operador haciendo uso del teorema de Radon-Nikodym tanto en $\mathcal{L}^1(I)$ como en su restricción al espacio de las funciones de densidad $\mathcal{D}(I)$.

Asimismo, recogemos algunas de sus propiedades más importantes, entre las que destacan la positividad del operador, su linealidad, que es una contracción y, por encima de todas, que sus puntos fijos definen medidas τ -invariantes, es decir, las medidas dadas por

$$\mu(A) = \int_A f^* d\lambda,$$

con f^* un punto fijo de P_τ , cumplen $\mu(A) = \mu(\tau^{-1}(A))$ y también es cierto el recíproco. Cuando los puntos fijos son funciones de densidad, que son los que realmente nos interesan, diremos que son densidades de las medidas que definen. Esta propiedad será crucial para probar posteriormente la existencia y unicidad de acims para las transformaciones que nos interesan.

Por último, desarrollamos una representación del operador para transformaciones monótonas a trozos a partir de la partición correspondiente, acompañada de varios ejemplos. Estas transformaciones cumplen que existe una partición \mathcal{P} tal que $|\tau'_i(x)| \geq \alpha > 0$ para todo $i = 1, \dots, q$ y para todo $x \in J_i$, con $\mathcal{P} = \{J_i = (a_{i-1}, a_i), i = 1, \dots, q\}$ y $\tau_i = \tau|_{J_i}$. Esta representación es muy relevante, ya que las transformaciones con las que vamos a trabajar son de dicho tipo, por lo que la usaremos con frecuencia en sucesivos resultados de los últimos capítulos.

El tercer capítulo lo hemos dedicado a las transformaciones expansivas a trozos (esto es, transformaciones monótonas a trozos para las que $\alpha > 1$). En particular, si $\frac{1}{|\tau'(x)|}$ es una función

de variación acotada probaremos que τ admite una acim cuya densidad es de variación acotada. Asimismo, probaremos que todas las posibles acims tienen densidades de variación acotada.

Para poder llegar a esos resultados necesitamos emplear buena parte de la información sobre funciones de variación acotada del primer capítulo, junto con tres lemas en los que se prueban sendas desigualdades que involucran a la variación total de una función y al operador de Perron-Frobenius. Estas desigualdades son una variación de las presentadas por Lasota y Yorke en [8]. Entre ellas podemos destacar que para cualesquiera $f \in BV(I)$ y $n \geq 1$ se cumple

$$\|P_\tau^n f\|_{BV} \leq Cr^n \|f\|_{BV} + R\|f\|_1$$

con $0 < r < 1$, $C > 0$ y $R > 0$ constantes.

El último capítulo está dedicado en exclusiva a las transformaciones de Markov, en particular a las transformaciones de Markov lineales a trozos. Una transformación de Markov es una transformación monótona a trozos que se caracteriza por llevar cada intervalo de su partición asociada a una unión de intervalos de la partición. A dichas transformaciones se asocia de manera natural una matriz A_τ , su matriz de incidencia, cuyas entradas a_{ij} valen 1 o 0 según que $J_i \subset \tau(J_j)$ o $J_i \cap \tau(J_j) = \emptyset$, y es en este punto donde adquiere gran importancia la teoría de Perron-Frobenius desarrollada en el trabajo [10]. En concreto, nos va interesar el caso en que la matriz de incidencia es primitiva, es decir, existe una potencia de la misma con todas sus entradas positivas. Entonces podremos aplicar sobre ellas el teorema de Perron-Frobenius para matrices primitivas, que garantiza la existencia de un vector propio positivo asociado al autovalor de módulo máximo y además que el espacio propio correspondiente tiene multiplicidades algebraica y geométrica 1.

Más aún, probamos que para una variación de la matriz de incidencia dada por $M_\tau = (m_{ij})$ con $m_{ij} = \frac{a_{ij}}{|\tau_j|}$, el valor propio de módulo máximo es 1 y, cuando A_τ es primitiva, las multiplicidades algebraica y geométrica de este valen 1, existiendo un vector propio positivo. Además, para esta matriz se cumple que si f es una función de paso (constante a trozos en los intervalos de la partición asociada) $P_\tau f = M_\tau \pi^f$, donde identificamos f con el vector columna π^f cuyos coeficientes son los valores de f sobre los intervalos de la partición asociada. Por lo tanto, todo punto fijo de P_τ constante a trozos será un vector propio de valor propio 1 asociado a la matriz M_τ .

Asimismo, probamos que si $\inf |\tau'| > 1$, entonces todo punto fijo de P_τ será constante a trozos y por tanto solución de $M_\tau \pi^f = \pi^f$. Este resultado es de vital importancia, ya que es una de las claves de la unicidad de la medida invariante que buscamos.

A su vez, probamos varios resultados relacionados con la irreducibilidad (una versión débil de la primitividad) de las matrices de incidencia de las transformaciones de Markov lineales a trozos, que combinados con la condición de que una transformación sea completa (para cualquier intervalo A existe k , dependiente del conjunto, tal que $\overline{\tau^k(A)} = I$), nos permite probar la existencia de m_0 y $\kappa > 1$ tales que para todo $m \geq m_0$

$$\inf_x |(\tau^m(x))'| \geq \kappa.$$

Así, tomando la función $\phi = \tau^m$ para un $m \geq m_0$, si la matriz de incidencia de τ es primitiva la de ϕ también lo será, y podremos garantizar la existencia y unicidad de acims para ϕ , ya que todos los puntos fijos de P_ϕ serán funciones de paso y la primitividad garantiza la unicidad.

Finalmente, como las medidas τ -invariantes también lo son para ϕ , necesariamente esa medida será la única acim para τ . Además, probamos que dicha medida es equivalente a la de Lebesgue y τ es ergódica respecto a ella. Por tanto podremos aplicar el corolario del teorema ergódico de Birkhoff deduciendo la dinámica que buscábamos y con ello alcanzando los objetivos del trabajo. Además, acompañamos algunos de estos resultados con ejemplos ilustrativos.

Concluimos esta introducción con unas breves notas históricas y comentando en que ha consistido nuestra aportación al trabajo.

El teorema 3.4 se demostró más o menos a la vez, y de manera independiente, en [8] y [7], y fue refinado en [9], donde se dilucidó la naturaleza y el total de posibles acims para la transformación. A partir de estos resultados, en [3] se mostró que si además la transformación era de Markov y la matriz de incidencia era irreducible, entonces existía una única acim, ergódica y equivalente a la de Lebesgue. Nótese que la expansividad (en realidad, en el artículo se usa la condición, ligeramente más débil, de que alguna iterada de la transformación es expansiva) y el teorema del valor medio implican que la matriz de incidencia no puede ser una matriz de permutación (es decir, con exactamente un 1 en cada fila y columna). En [3] se probó asimismo que si además la matriz era primitiva y la transformación lineal a trozos, la densidad de la acim era una función de paso en la partición asociada a la transformación. Los resultados de [3] se refinaron en [2] (usando herramientas de [5]) de dos modos. Por un lado, se probó que todas las densidades de las acims de transformaciones de Markov lineales y expansivas son funciones de paso en la partición asociada. Asimismo, se mostró que en el segundo resultado de [3] la hipótesis de expansividad es redundante, es decir: toda transformación de Markov lineal a trozos, cuya matriz de incidencia es primitiva, tiene exactamente una acim, ergódica y equivalente respecto a la medida de Lebesgue, y su densidad es una función de paso en la partición asociada. Este es nuestro teorema 4.21, el resultado principal de este trabajo fin de máster.

Un asunto muy notable a destacar, que no se menciona en ninguna de las fuentes bibliográficas que hemos consultado, incluyendo los anteriores artículos, es que el teorema 4.21 puede demostrarse con una hipótesis más débil sobre la matriz de incidencia: basta con que sea irreducible y no sea de permutación. En realidad, nuestra demostración también es válida en este caso, con dos salvedades. Necesitamos la primitividad para, primero, garantizar que la iterada $\phi = \tau^m$ expansiva a trozos tiene una única acim, pero este es justo el primer resultado de [3], que no es difícil de probar con las técnicas desarrolladas en este trabajo (a pesar de que no puede asegurarse que la matriz de incidencia de A_ϕ sea irreducible). Por otro lado necesitamos que esta acim (que automáticamente será la única de τ) venga dada por una función de paso f con vector π^f positivo, pero esto lo tenemos asegurado porque el teorema de Perron-Frobenius (nuestro teorema 4.12) también es válido para matrices irreducibles. Como nuestra intención era que este trabajo sirviese como prolongación de nuestro trabajo fin de grado, pero por lo demás

fuese autocontenido, y allí solo probamos el teorema para matrices primitivas, hemos preferido enunciar el teorema 4.21 para matrices primitivas, tal y como se hace en [2], [3] o [1].

Otra cuestión que deseamos enfatizar es que, estrictamente hablando, la parte del teorema 3.4 que realmente necesitábamos para nuestros propósitos era la de que las densidades de las acims son de variación acotada, porque entonces, la existencia (y la unicidad) están garantizadas, vía el teorema 4.13, por la teoría de Perron-Frobenius. Sorprendentemente, nuestro principal texto de referencia, [1], omite (aunque la usa) justo la prueba de dicha parte, y el artículo [8] utiliza para su demostración el teorema de Kakutani-Yosida, un resultado bastante profundo de teoría de operadores. Por tanto nos hemos visto forzados a elaborar una demostración alternativa usando solo las herramientas de los capítulos 1 y 2. Esta es (junto a la generalización del teorema 4.21 mencionada en el párrafo anterior) la principal aportación del presente trabajo.

Índice general

Introducción	I
1. Preliminares	3
1.1. Teoría de la medida	4
1.1.1. El teorema de Radon-Nikodym	8
1.2. Funciones de variación acotada en una dimensión	12
1.2.1. El teorema de elección de Helly	17
1.3. Teoría ergódica	24
1.3.1. El teorema Ergódico de Birkhoff	25
2. El operador de Perron-Frobenius	35
2.1. Propiedades del operador de Perron-Frobenius	36
2.2. Representación del operador de Perron-Frobenius	39
3. Transformaciones expansivas a trozos	43
4. Transformaciones de Markov	51
4.1. Transformaciones de Markov lineales a trozos	52
4.2. Funciones propias asociadas a las matrices inducidas por transformaciones de Markov lineales a trozos	54
4.3. Existencia y unicidad de medidas invariantes absolutamente continuas	61
Bibliografía	71
Índice alfabético	72

Preliminares

Este primer capítulo va a ir dedicado a presentar todos aquellos resultados preliminares necesarios para probar los principales teoremas que se van a tratar en este trabajo. En él podemos distinguir tres partes bien diferenciadas. La primera de ellas está destinada a elementos de teoría de la medida, dónde se recuerdan definiciones vistas durante el grado y este máster. En ella se prueba el teorema de Radon-Nikodym para medidas reales finitas, el cual es necesario para fundamentar la existencia del principal instrumento de este trabajo, el operador de Perron-Frobenius.

La segunda está dirigida a presentar las funciones de variación acotada y una serie de resultados sobre ellas, necesarios para probar la existencia de medidas invariantes absolutamente continuas con funciones de densidad con buenas propiedades. Destaca el llamado principio de selección de Helly, el cual, bajo determinadas condiciones que más adelante veremos, garantiza la existencia de una sucesión de funciones de variación acotada convergente a otra de variación acotada.

La tercera y última parte va dirigida a introducir elementos de la conocida como teoría ergódica, una disciplina nueva, en el sentido de que no ha sido estudiada ni durante el grado ni durante el máster. El principal resultado de esta parte será el teorema ergódico de Birkhoff, que garantiza la convergencia de los promedios ergódicos de una función f en un sistema dinámico. Pero lo que de verdad nos resultará útil será un corolario que garantiza que cuando nos encontramos en un espacio de medida finita, en particular un espacio de probabilidad, y tenemos una transformación τ ergódica en ese espacio, entonces, para cualquier subconjunto E del espacio \mathbb{X} , la frecuencia relativa con la que la órbita de un punto $x \in \mathbb{X}$, $\{\tau^n(x)\}_{n=0}^{+\infty}$, cae en E , es decir $\frac{1}{n} \sum_{i=0}^{n-1} \chi_E(\tau^i(x))$, viene dada por $\mu(E)$, siendo μ la medida invariante asociada a τ .

En lo que sigue vamos a dar por conocidos algunos resultados de teoría de la medida, por ejemplo los teoremas de convergencia monótona y dominada o el lema de Fatou, que son empleados en varias pruebas de este trabajo, por entender que quedan suficientemente cubiertos en los estudios de grado y máster de esta facultad.

1.1. Teoría de la medida

Comenzamos con las primeras definiciones de teoría de la medida. Se pueden encontrar en [1, cap 2].

De ahora en adelante denotaremos por \mathbb{X} a un conjunto.

Definición 1.1. Una familia \mathcal{A} de subconjuntos de \mathbb{X} se llama σ -álgebra si y solo si:

- 1) $\mathbb{X} \in \mathcal{A}$;
- 2) para todo $A \in \mathcal{A}$ se cumple $\mathbb{X} \setminus A \in \mathcal{A}$;
- 3) si $A_n \in \mathcal{A}$, para $n = 1, 2, \dots$, entonces $\bigcup_{n=1}^{+\infty} A_n \in \mathcal{A}$.

A los elementos de \mathcal{A} se les suele llamar *conjuntos medibles* y a $(\mathbb{X}, \mathcal{A})$ se le llama un *espacio medible*.

Definición 1.2. Sea $(\mathbb{X}, \mathcal{A})$ un espacio medible. Una transformación $\tau : \mathbb{X} \rightarrow \mathbb{X}$ se dice que es *medible* si $\tau^{-1}(A) \in \mathcal{A}$, es decir, $A \in \mathcal{A}$ implica $\tau^{-1}(A) \in \mathcal{A}$, donde $\tau^{-1}(A) \equiv \{x \in \mathbb{X} : \tau(x) \in A\}$.

Para cualquier familia de subconjuntos de \mathbb{X} , \mathcal{J} , existe la menor σ -álgebra, \mathcal{B} , que contiene a \mathcal{J} . Decimos entonces que \mathcal{J} genera \mathcal{B} y escribimos $\mathcal{B} = \sigma(\mathcal{J})$.

Veamos ahora la definición de medida en su sentido habitual.

Definición 1.3. Sea $(\mathbb{X}, \mathcal{A})$ un espacio medible. Una función $\mu : \mathcal{A} \rightarrow \mathbb{R}^+ \cup \{+\infty\}$ se llama *medida* en $(\mathbb{X}, \mathcal{A})$ si y solo si:

- 1) $\mu(\emptyset) = 0$;
- 2) para toda familia numerable $\{A_n\}_{n=1}^{+\infty}$ de conjuntos medibles disjuntos,

$$\mu \left(\bigcup_{n=1}^{+\infty} A_n \right) = \sum_{n=1}^{+\infty} \mu(A_n).$$

Aquí \mathbb{R}^+ representa a los números reales no negativos.

A la terna $(\mathbb{X}, \mathcal{A}, \mu)$ se le conoce por *espacio de medida* y en función de las propiedades de la medida μ se tienen las siguientes precisiones: si $\mu(\mathbb{X}) < +\infty$ la medida se llama *finita*; si $\mu(\mathbb{X}) = 1$, \mathbb{X} se llama *espacio de probabilidad* y μ recibe el nombre de *medida de probabilidad o medida normalizada*.

Definición 1.4. Sea \mathbb{X} un espacio topológico y \mathcal{O} la familia de conjuntos abiertos de \mathbb{X} . Entonces la σ -álgebra $\mathcal{B} = \sigma(\mathcal{O})$ se llama *la σ -álgebra de Borel de \mathbb{X}* y sus elementos los *conjuntos de Borel de \mathbb{X}* .

Si \mathbb{X} es un espacio topológico (en particular $\mathbb{X} = \mathbb{R}$, $[a, b]$ o $[-\infty, +\infty]$, con sus topologías habituales), el espacio medible en el que trabajaremos será siempre $(\mathbb{X}, \mathcal{B})$.

Definición 1.5. Sea $(\mathbb{X}, \mathcal{A}, \mu)$ un espacio de medida. Una función $f : \mathbb{X} \rightarrow [-\infty, +\infty]$ se dice *medible* si para todo $\alpha \in \mathbb{R}$, $f^{-1}(\alpha, +\infty] \in \mathcal{A}$, o equivalentemente si $f^{-1}(A) \in \mathcal{A}$ para todo conjunto de Borel $A \subset [-\infty, +\infty]$.

Veamos ahora las definiciones de norma, espacio $\mathcal{L}^1(\mathbb{X}, \mathcal{A}, \mu)$ y norma \mathcal{L}^1 . De nuevo, pueden encontrarse en [1, cap 2].

Definición 1.6. Sea V un espacio vectorial sobre un cuerpo \mathbb{K} . Una *norma* es una aplicación $\|\cdot\| : V \rightarrow [0, +\infty)$ que cumple:

- 1) $\|x\| \geq 0$ y $\|x\| = 0$ si y solo si $x = 0$;
- 2) $\|x + y\| \leq \|x\| + \|y\|$ para todo $x, y \in V$;
- 3) $\|\lambda x\| = |\lambda| \|x\|$ para todo $x \in V$ y $\lambda \in \mathbb{K}$.

Definición 1.7. Sea $(\mathbb{X}, \mathcal{A}, \mu)$ un espacio de medida. La familia de funciones medibles $f : \mathbb{X} \rightarrow [-\infty, +\infty]$ cumpliendo

$$\int_{\mathbb{X}} |f(x)| d\mu < +\infty$$

se llama *espacio* $\mathcal{L}^1(\mathbb{X}, \mathcal{A}, \mu)$ o, más abreviadamente si no hay ambigüedad sobre \mathcal{A} y μ , $\mathcal{L}^1(\mathbb{X})$. En particular, en el caso de que $\mathbb{X} = \mathbb{R}$ o $[a, b]$, $\mathcal{L}^1(\mathbb{X})$ siempre denotará a $\mathcal{L}^1(\mathbb{X}, \mathcal{B}, \lambda)$, con λ la medida de Lebesgue. Si $f \in \mathcal{L}^1(\mathbb{X})$ se dice que es *integrable*.

Definición 1.8. Dada $f \in \mathcal{L}^1(\mathbb{X}, \mathcal{A}, \mu)$, llamamos *norma* \mathcal{L}^1 de f a

$$\|f\|_1 = \int_{\mathbb{X}} |f(x)| d\mu.$$

Como bien es sabido con esta norma el espacio $\mathcal{L}^1(\mathbb{X}, \mathcal{A}, \mu)$ es un espacio de Banach. Recordamos que para que $\|\cdot\|_1$ sea una auténtica norma consideramos iguales a las funciones que lo son en casi todo punto.

Necesitaremos usar funciones de densidad para el correcto desarrollo de muchos de los resultados del trabajo, en particular en nuestro resultado principal, por lo que es importante introducirlas.

Definición 1.9. Sea $(\mathbb{X}, \mathcal{A}, \mu)$ un espacio de medida. El espacio

$$\mathcal{D}(\mathbb{X}) = \mathcal{D}(\mathbb{X}, \mathcal{A}, \mu) = \{f \in \mathcal{L}^1(\mathbb{X}, \mathcal{A}, \mu) : f \geq 0 \text{ y } \|f\|_1 = 1\}$$

recibe el nombre de *espacio de las funciones de densidad* de $(\mathbb{X}, \mathcal{A}, \mu)$. Una función $f \in \mathcal{D}(\mathbb{X})$ se llama *función de densidad*.

De cara al teorema de Radon-Nikodym necesitamos introducir las medidas reales, que son similares a las usuales salvo que pueden tomar valores en la recta real. Véase [6, cap 11].

Definición 1.10. Sea $(\mathbb{X}, \mathcal{A})$ un espacio medible. Una *medida real* en $(\mathbb{X}, \mathcal{A})$ es una función $\mu : \mathcal{A} \rightarrow \mathbb{R}$ tal que

- 1) $\mu(\emptyset) = 0$;
- 2) para toda familia numerable $\{A_n\}_{n=1}^{+\infty}$ de conjuntos medibles disjuntos,

$$\mu \left(\bigcup_{n=1}^{+\infty} A_n \right) = \sum_{n=1}^{+\infty} \mu(A_n).$$

Llamamos ahora a la terna $(\mathbb{X}, \mathcal{A}, \mu)$ un *espacio de medida real*.

La siguiente definición de medida absolutamente continua, y con ella la de medidas equivalentes, tendrá una gran importancia en este trabajo, ya que justamente nuestro objetivo será garantizar la existencia y unicidad de medidas que sean absolutamente continuas (en nuestro caso respecto a la medida de Lebesgue) y, en particular, equivalentes a ella. La primera se puede encontrar en [1, cap 2].

Definición 1.11. Dadas ν y μ dos medidas reales en el mismo espacio medible $(\mathbb{X}, \mathcal{A})$, se dice que ν es *absolutamente continua* respecto a μ si para todo $A \in \mathcal{A}$, tal que $\mu(A) = 0$, se tiene que $\nu(A) = 0$. Lo denotaremos como $\nu \ll \mu$.

Observación 1. Si $(\mathbb{X}, \mathcal{A}, \mu)$ es un espacio de medida, entonces se puede asociar de manera natural a cada función $f \in \mathcal{L}^1(\mathbb{X})$ una medida real mediante

$$\mu_f(A) = \int_A f d\mu.$$

Claramente, esta medida μ_f es absolutamente continua respecto a μ .

Definición 1.12. Dadas dos medidas reales ν y μ en $(\mathbb{X}, \mathcal{A})$, se dice que μ es *equivalente* a ν cuando

$$\mu \ll \nu \text{ y } \nu \ll \mu.$$

Las siguientes definiciones se pueden encontrar en [1, cap 2]. Llamaremos a $\mathcal{M}(\mathbb{X})$ al espacio de medidas reales en $(\mathbb{X}, \mathcal{A})$. Dada $\tau : \mathbb{X} \rightarrow \mathbb{X}$ una transformación medible, τ induce una transformación en $\mathcal{M}(\mathbb{X})$ dada por $(\tau_*\mu)(A) = \mu(\tau^{-1}(A))$. Como τ es medible entonces $\tau_*\mu \in \mathcal{M}(\mathbb{X})$. Por lo tanto, τ_* está bien definido.

Definición 1.13. Sea $(\mathbb{X}, \mathcal{A}, \mu)$ un espacio de medida real. Diremos que la transformación medible $\tau : \mathbb{X} \rightarrow \mathbb{X}$ es *no singular* si y solo si $\tau_*\mu \ll \mu$, es decir, para cualquier $A \in \mathcal{A}$, tal que $\mu(A) = 0$, se tiene que $\tau_*\mu(A) = \mu(\tau^{-1}(A)) = 0$.

El siguiente resultado relaciona la continuidad absoluta de medidas con la de las medidas inducidas por esas transformaciones.

Proposición 1.14. Sean $(\mathbb{X}, \mathcal{A}, \mu)$ un espacio de medida real y $\tau : \mathbb{X} \rightarrow \mathbb{X}$ una transformación no singular. Si $\nu \ll \mu$ entonces $\tau_*\nu \ll \tau_*\mu \ll \mu$.

Demostración. Como $\nu \ll \mu$, dado $A \in \mathcal{A}$, $\mu(A) = 0$ implica $\nu(A) = 0$. Por otro lado, como τ es no singular, por definición, $\mu(A) = 0$ implica $\mu(\tau^{-1}(A)) = 0$ y por lo anterior se tiene $\nu(\tau^{-1}(A)) = 0$. Por tanto, $\tau_*\nu \ll \tau_*\mu$, y al ser τ no singular, $\tau_*\mu \ll \mu$ obteniéndose el resultado buscado. ■

Concluimos con un resultado muy útil de teoría de la medida, el teorema del cambio de variable. La siguiente prueba es una variación de la que puede encontrarse en [12, cap 1]. Antes, vendrá bien recordar la siguiente definición, que se puede encontrar en [4, cap 3].

Definición 1.15. Sea $(\mathbb{X}, \mathcal{A}, \mu)$ un espacio de medida. Una *función simple* es una función de la forma

$$\sum_{i=1}^n c_i \chi_{A_i},$$

donde los valores c_i son constantes y los A_i conjuntos medibles.

Teorema 1.16 (Cambio de variable). Sean $(\mathbb{X}, \mathcal{A}, \mu)$ un espacio de medida y $\tau : \mathbb{X} \rightarrow \mathbb{X}$ una transformación medible. Si f es una función medible, entonces

$$\int_{\mathbb{X}} f d(\tau_*\mu) = \int_{\mathbb{X}} f \circ \tau d\mu,$$

donde la igualdad se entiende en el sentido de que la integral de la izquierda está bien definida si y solo si lo está la de la derecha, y en tal caso ambas coinciden.

Demostración. Vamos a proceder siguiendo una de las estrategias más utilizadas en teoría de la medida, estudiar casos desde funciones características hasta funciones medibles pasando por funciones simples.

Primer caso, cuando $f = \chi_A$ es la función característica de un conjunto medible $A \in \mathcal{A}$. Por un lado, tenemos

$$\int_{\mathbb{X}} \chi_A d(\tau_*\mu) = \tau_*\mu(A) = \mu(\tau^{-1}(A))$$

y por otro

$$\int_{\mathbb{X}} \chi_A \circ \tau d\mu = \int_{\mathbb{X}} \chi_{\tau^{-1}(A)} d\mu = \mu(\tau^{-1}(A)),$$

por lo tanto el resultado es cierto para funciones características.

Al ser cierto para funciones características, necesariamente, es cierto para funciones simples, ya que, como es bien sabido, estas son combinaciones lineales de funciones características de

conjuntos medibles. Por lo tanto, si $f = \sum_{i=1}^n \alpha_i \chi_{A_i}$ resulta

$$\begin{aligned} \int_{\mathbb{X}} f d(\tau_*\mu) &= \int_{\mathbb{X}} \sum_{i=1}^n \alpha_i \chi_{A_i} d(\tau_*\mu) = \sum_{i=1}^n \alpha_i \int_{\mathbb{X}} \chi_{A_i} d(\tau_*\mu) \\ &= \sum_{i=1}^n \alpha_i \int_{\mathbb{X}} \chi_{A_i} \circ \tau d\mu = \int_{\mathbb{X}} \left(\sum_{i=1}^n \alpha_i \chi_{A_i} \right) \circ \tau d\mu = \int_{\mathbb{X}} f \circ \tau d\mu. \end{aligned}$$

Por otro lado, como toda función medible no negativa f , es límite puntual de una sucesión monótona creciente de funciones simples no negativas, el teorema de la convergencia monótona garantiza el resultado, ya que si $f = \lim_n s_n$, con $(s_n)_n$ una sucesión monótona creciente de funciones simples no negativas, se obtiene

$$\begin{aligned} \int_{\mathbb{X}} f d(\tau_*\mu) &= \int_{\mathbb{X}} \lim_{n \rightarrow +\infty} s_n d(\tau_*\mu) = \lim_{n \rightarrow +\infty} \int_{\mathbb{X}} s_n d(\tau_*\mu) \\ &= \lim_{n \rightarrow +\infty} \int_{\mathbb{X}} s_n \circ \tau d\mu = \int_{\mathbb{X}} \left(\lim_{n \rightarrow +\infty} s_n \right) \circ \tau d\mu = \int_{\mathbb{X}} f \circ \tau d\mu. \end{aligned}$$

Finalmente, como cualquier función medible f la podemos expresar como diferencia de sus partes positiva y negativa, $f^+ = \max\{0, f(x)\}$ y $f^- = \max\{0, -f(x)\}$, con $f = f^+ - f^-$, resulta

$$\begin{aligned} \int_{\mathbb{X}} f d(\tau_*\mu) &= \int_{\mathbb{X}} (f^+ - f^-) d(\tau_*\mu) = \int_{\mathbb{X}} f^+ d(\tau_*\mu) - \int_{\mathbb{X}} f^- d(\tau_*\mu) \\ &= \int_{\mathbb{X}} f^+ \circ \tau d\mu - \int_{\mathbb{X}} f^- \circ \tau d\mu = \int_{\mathbb{X}} (f^+ - f^-) \circ \tau d\mu = \int_{\mathbb{X}} f \circ \tau d\mu. \end{aligned}$$

■

1.1.1. El teorema de Radon-Nikodym

En esta sección se presenta la prueba del teorema de Radon-Nikodym. Este resultado es clave para introducir la herramienta principal de este trabajo, el operador de Perron-Frobenius, que ya hemos mencionado al inicio del capítulo.

Antes de comenzar su prueba es necesario incluir una serie de resultados previos y definiciones que serán empleados en ella. Destacar que el teorema de Radon-Nikodym que aquí va a ser demostrado tiene como hipótesis que las medidas sean finitas, ya que en el trabajo vamos a trabajar sobre intervalos que tienen medida finita, pero existen versiones más generales. La mayor parte de los resultados mostrados a continuación pueden encontrarse en [6, cap 11].

Una novedad de las medidas reales respecto a las usuales son los llamados conjuntos positivos y negativos que presentamos a continuación.

Definición 1.17. Sea $(\mathbb{X}, \mathcal{A}, \mu)$ un espacio de medida real. Se dice que un subconjunto A de \mathbb{X} es un *conjunto positivo* para μ si $A \in \mathcal{A}$ y cada subconjunto medible de A tiene medida mayor o igual que 0. Análogamente, un subconjunto A de \mathbb{X} se dice que es un *conjunto negativo* para μ si $A \in \mathcal{A}$ y cada subconjunto medible de A tiene medida menor o igual que 0.

Veamos un lema necesario para la prueba del teorema de descomposición de Hahn, que emplearemos en la demostración del teorema de Radon-Nikodym. La prueba de la primera parte se puede encontrar en [6, cap 11] y la segunda es de elaboración propia.

Lema 1.18. Sean $(\mathbb{X}, \mathcal{A}, \mu)$ un espacio de medida real y A un conjunto medible tal que $\mu(A) < 0$. Entonces A contiene un conjunto negativo B tal que $\mu(B) \leq \mu(A)$. De igual manera, si A es un conjunto medible tal que $\mu(A) > 0$, entonces A contiene un conjunto positivo E tal que $\mu(E) \geq \mu(A)$.

Demostración. Veamos la primera parte.

Definimos $\delta_1 = \sup\{\mu(E) : E \in \mathcal{A}, E \subset A\}$ y tomamos un conjunto $A_1 \in \mathcal{A}$ cualquiera tal que $A_1 \subset A$ y $\mu(A_1) \geq \min\{\delta_1/2, 1\}$, donde el “1” en el mínimo anterior se introduce para controlar la posibilidad de que $\delta_1 = +\infty$.

En caso de que ese conjunto no existiera, se tendría que $\mu(E) < \min\{\delta_1/2, 1\}$, para cada subconjunto medible E de A , pero por la definición de supremo, δ_1 no puede ser mayor que 0, por tanto $\delta_1 = 0$ (el vacío es un subconjunto medible de A y $\delta_1 \geq \mu(\emptyset) = 0$), y el conjunto A sería un subconjunto negativo de sí mismo con $\mu(A) \leq \mu(A)$.

Para llevar a cabo la prueba, procederemos por inducción formando las sucesiones $(\delta_n)_n$ y $(A_n)_n$, donde para $n \geq 2$ definimos

$$\delta_n = \sup \left\{ \mu(E) : E \in \mathcal{A}, E \subset A \setminus \left(\bigcup_{i=1}^{n-1} A_i \right) \right\}$$

y seleccionamos $A_n \in \mathcal{A}$, $A_n \subset A \setminus \left(\bigcup_{i=1}^{n-1} A_i \right)$ tal que $\mu(A_n) \geq \min\{\delta_n/2, 1\}$. Como ocurría anteriormente, si ese conjunto A_n no existiera, necesariamente $\delta_n = 0$, al ser el supremo, y el conjunto $A \setminus \left(\bigcup_{i=1}^{n-1} A_i \right)$ sería un subconjunto negativo de A con $\mu(A) = \sum_{i=1}^{n-1} \mu(A_i) + \mu(A \setminus \left(\bigcup_{i=1}^{n-1} A_i \right)) \geq \mu(A \setminus \left(\bigcup_{i=1}^{n-1} A_i \right))$, ya que $\mu(A_i) \geq 0$ para cada $1 \leq i \leq n-1$.

De esta forma, en el caso de que si existieran los conjuntos A_n , podemos definir los conjuntos $A_\infty = \bigcup_{n=1}^{+\infty} A_n$ y $B = A \setminus A_\infty$ y vamos a ver que B es un subconjunto negativo de A . Por definición, los A_n son disjuntos dos a dos y como $\mu(A_n) \geq \min\{\delta_n/2, 1\}$, se tiene que $\mu(A_n) \geq 0$, por tanto, $\mu(A_\infty) \geq 0$ y $\mu(B) \leq \mu(B) + \mu(A_\infty) = \mu(A)$.

Para terminar, nos falta ver que B es negativo. Como la serie $\mu(A_\infty) = \sum_n \mu(A_n)$ es convergente, su término general tiende a cero, es decir,

$$\lim_{n \rightarrow +\infty} \mu(A_n) = 0.$$

Por lo tanto,

$$\lim_{n \rightarrow +\infty} \delta_n = 0$$

al ser el supremo. Como cualquier subconjunto medible E de B cumple $\mu(E) \leq \delta_n$, para cada n , se tiene $\mu(E) \leq 0$ y, en definitiva, B sería negativo.

Veamos la segunda parte.

Sea A un conjunto medible tal que $0 < \mu(A) < +\infty$. Si consideramos la medida real $\tilde{\mu} = -\mu$, tendremos $\tilde{\mu}(A) < 0$. Aplicando la primera parte del lema con esta medida resulta que existe un conjunto negativo E tal que $\tilde{\mu}(E) \leq \tilde{\mu}(A)$. Por lo tanto, $-\mu(E) \leq -\mu(A)$ y por consiguiente $\mu(E) \geq \mu(A)$. Además, como E es negativo para $\tilde{\mu} = -\mu$, necesariamente, E es positivo para μ , con lo que queda probado el resultado. ■

A continuación se presenta el teorema de descomposición de Hahn. Se puede encontrar en [6, cap 11]

Teorema 1.19 (Teorema de descomposición de Hahn). *Sea $(\mathbb{X}, \mathcal{A}, \mu)$ un espacio de medida real. Entonces existen dos conjuntos disjuntos P y N de \mathbb{X} , tales que P es un conjunto positivo para μ y N es un conjunto negativo para μ con $\mathbb{X} = P \cup N$.*

Demostración. En primer lugar, definimos el número

$$L = \inf\{\mu(A) : A \text{ es un conjunto negativo para } \mu\},$$

es decir, el ínfimo de las medidas de los conjuntos negativos para μ . Este conjunto es no vacío, ya que por lo menos contiene a $\mu(\emptyset) = 0$. Consideremos una sucesión $(A_n)_n$ de conjuntos negativos tal que $L = \lim_{n \rightarrow +\infty} \mu(A_n)$ y sea $N = \bigcup_n A_n$. Cada subconjunto medible de N puede escribirse como unión disjunta de conjuntos medibles, estando cada uno de ellos contenido en alguno de los A_n , por lo tanto N es un conjunto negativo para μ . Por la definición de L , se tiene $L \leq \mu(N) = \mu(A_n) + \mu(N \setminus A_n) \leq \mu(A_n)$ para cada n , al ser N negativo. Por tanto $L = \mu(N)$, por la definición de ínfimo.

Sea $P = \mathbb{X} \setminus N$ y vamos a ver que es positivo para μ . Si no lo fuera, existiría un conjunto $A \subset P$ tal que $\mu(A) < 0$ y, por el lema anterior 1.18, A contendría un conjunto negativo B tal que $\mu(B) < 0$. Entonces $N \cup B$ sería un conjunto negativo con $\mu(N \cup B) = \mu(N) + \mu(B) < \mu(N) = L$, que contradice la definición de L como ínfimo. Por lo tanto P es positivo. ■

A partir del teorema de descomposición podemos definir lo que entendemos por descomposición de Hahn.

Definición 1.20. Una *descomposición de Hahn* de una medida real μ es una partición medible (P, N) de \mathbb{X} tal que P es positivo para μ y N es negativo para μ .

Llegamos así al resultado que queríamos probar, el teorema de Radon-Nikodym, primero para medidas finitas y, a partir de ese caso, para medidas finitas con la segunda de ellas real. La prueba puede encontrarse en [6, cap 11].

Teorema 1.21 (Teorema de Radon-Nikodym). *Sean $(\mathbb{X}, \mathcal{A})$ un espacio medible, μ una medida finita y ν una medida real en $(\mathbb{X}, \mathcal{A})$. Si $\nu \ll \mu$ entonces existe una única $f \in \mathcal{L}^1(\mathbb{X}, \mathcal{A}, \mu)$ tal que para todo $A \in \mathcal{A}$,*

$$\nu(A) = \int_A f d\mu.$$

Demostración. De momento suponemos que ν es finita (no negativa).

Definimos el conjunto

$$\mathcal{F} := \left\{ g : \mathbb{X} \longrightarrow [0, +\infty] : \int_A g d\mu \leq \nu(A), \forall A \in \mathcal{A} \right\}.$$

Vamos a ver que \mathcal{F} contiene una función f tal que $\int_{\mathbb{X}} f d\mu = \sup\{\int_{\mathbb{X}} g d\mu : g \in \mathcal{F}\}$ y que esa función f es en efecto la que buscamos.

En primer lugar, \mathcal{F} es un conjunto no vacío ya que la función nula, denotémosla por 0 , pertenece a \mathcal{F} , al ser $\int_A 0 d\mu = 0 \leq \nu(A)$, para todo $A \in \mathcal{A}$.

Por otro lado, si $g_1, g_2 \in \mathcal{F}$ entonces $\sup\{g_1, g_2\} \in \mathcal{F}$. Esto es debido a que dado $A \in \mathcal{A}$, si consideramos los conjuntos $A_1 = \{g_1 > g_2\} \cap A$ y $A_2 = \{g_1 \leq g_2\} \cap A$ se cumple

$$\int_A \sup\{g_1, g_2\} d\mu = \int_{A_1} g_1 d\mu + \int_{A_2} g_2 d\mu \leq \nu(A_1) + \nu(A_2) = \nu(A).$$

Sea $(g_n)_n$ una sucesión en \mathcal{F} tal que

$$\lim_{n \rightarrow +\infty} \int_{\mathbb{X}} g_n d\mu = \sup \left\{ \int_{\mathbb{X}} g d\mu : g \in \mathcal{F} \right\}.$$

Notemos que la podemos suponer creciente, sustituyendo g_n si es necesario por el supremo de $\{g_1, \dots, g_n\}$, que como acabamos de ver está en \mathcal{F} . Sea $f = \lim_n g_n$. El teorema de la convergencia monótona garantiza que

$$\int_A f d\mu = \int_A \lim_{n \rightarrow +\infty} g_n d\mu = \lim_{n \rightarrow +\infty} \int_A g_n d\mu \leq \nu(A),$$

para cada $A \in \mathcal{A}$. Por tanto $f \in \mathcal{F}$ y $\int_{\mathbb{X}} f d\mu = \sup\{\int_{\mathbb{X}} g d\mu : g \in \mathcal{F}\}$.

Pasemos a ver que $\nu(A) = \int_A f d\mu$ para cada $A \in \mathcal{A}$. Como $f \in \mathcal{F}$, la expresión $\nu_0(A) := \nu(A) - \int_A f d\mu$ define una medida finita (positiva) en \mathcal{A} y tenemos que probar que $\nu_0 = 0$. Si no lo fuera, como μ es finita, existe un $\varepsilon > 0$ tal que

$$\nu_0(\mathbb{X}) > \varepsilon \mu(\mathbb{X}).$$

Sea (P, N) una descomposición de Hahn para la medida real $\nu_0 - \varepsilon \mu$. Para cada $A \in \mathcal{A}$ se tiene que $\nu_0(A \cap P) \geq \varepsilon \mu(A \cap P)$, por tanto

$$\begin{aligned} \nu(A) &= \int_A f d\mu + \nu_0(A) \geq \int_A f d\mu + \nu_0(A \cap P) \\ &\geq \int_A f d\mu + \varepsilon \mu(A \cap P) = \int_A (f + \varepsilon \chi_P) d\mu. \end{aligned}$$

Además, $\mu(P) > 0$, ya que si $\mu(P) = 0$ entonces $\nu_0(P) = 0$, al ser $\nu \ll \mu$ y $0 \leq \nu_0(P) \leq \nu(P)$, y $\nu_0(\mathbb{X}) - \varepsilon \mu(\mathbb{X}) = \nu_0(N) - \varepsilon \mu(N) \leq 0$, en contra de la elección realizada de ε . Por lo tanto,

$f + \varepsilon \chi_P \in \mathcal{F}$ y $\int_{\mathbb{X}} (f + \varepsilon \chi_P) d\mu > \int_{\mathbb{X}} f d\mu$, esto contradice la maximalidad de f en \mathcal{F} y por tanto $\nu_0 = 0$. Así, $\nu(A) = \int_A f d\mu$ para cada $A \in \mathcal{A}$.

Notemos que, como ν es finita, f es μ -integrable y finita $\mu - c.t.p.$ y puede modificarse de forma que solo tome valores reales.

Para terminar, vamos a estudiar el caso en que μ es finita y ν una medida real. Entonces podemos descomponer ν en sus partes positivas y negativas, $\nu = \nu^+ - \nu^-$, donde ν^+ y ν^- son finitas. Entonces, por el caso anterior se pueden encontrar funciones medibles reales integrables, g_1 y g_2 , tales que $\nu^+(A) = \int_A g_1 d\mu$ y $\nu^-(A) = \int_A g_2 d\mu$ para cada $A \in \mathcal{A}$. De esta forma, la función $g = g_1 - g_2$ cumple que $\nu(A) = \int_A g d\mu$.

Por último veamos la unicidad. Si f_1 y $f_2 \in \mathcal{L}^1(\mathbb{X})$ cumplen $\nu(A) = \int_A f_1 d\mu = \int_A f_2 d\mu$ para todo $A \in \mathcal{A}$, $f = f_1 - f_2$ cumple

$$\int_A f d\mu = 0 \quad \forall A \in \mathcal{A}.$$

Pero entonces $f = 0$ $\mu - c.t.p.$, por lo tanto $f_1 = f_2$ $\mu - c.t.p.$ ■

Observación 2. La función f recibe el nombre de *derivada de Radon-Nikodym* de ν (respecto a μ) y la denotamos por $\frac{d\nu}{d\mu}$. Si ν es una medida finita entonces $\frac{d\nu}{d\mu} \geq 0$, y si ν es de probabilidad entonces $\frac{d\nu}{d\mu}$ es una función de densidad.

Sean $\nu \ll \mu$ con μ finita y ν real. Por la proposición 1.14, $\tau_*\nu \ll \mu$. Por lo tanto, la derivada de Radon-Nikodym de ν se transforma en la derivada de Radon-Nikodym de $\tau_*\nu$. Esta transformación, denotada por P_τ (es decir $P_\tau \left(\frac{d\nu}{d\mu} \right) = \frac{d\tau_*\nu}{d\mu}$), se llama *operador de Perron-Frobenius* asociado a τ y es la base sobre la que va a girar este trabajo. Notemos que si $\frac{d\nu}{d\mu}$ es una densidad, es decir, ν es de probabilidad, entonces $\frac{d\tau_*\nu}{d\mu}$ también lo será.

1.2. Funciones de variación acotada en una dimensión

En esta sección se van a establecer resultados relacionados con las funciones de variación acotada. En particular, serán necesarios para probar la existencia de una medida invariante absolutamente continua, asociada a una transformación expansiva a trozos, cuya densidad es de variación acotada. El principal resultado que vamos a ver en esta sección es el principio de elección de Helly.

Comenzamos con las definiciones de partición, función de variación acotada en una dimensión y su variación total, que se pueden encontrar en [1, cap 2].

Sea $[a, b] \subset \mathbb{R}$ un intervalo compacto. Para toda secuencia de puntos $a = x_0 < x_1 < \dots < x_{q-1} < x_q = b$, $q \geq 1$, definimos la partición $\mathcal{P} = \{J_i = (x_{i-1}, x_i) : i = 1, \dots, q\}$ de $[a, b]$. Los puntos $\{x_0, x_1, \dots, x_q\}$ son llamados *puntos finales* de la partición \mathcal{P} . La partición también puede representarse por los puntos de la secuencia, $\mathcal{P} = \{x_0, x_1, \dots, x_q\}$.

Definición 1.22. Sea $f : [a, b] \rightarrow \mathbb{R}$. Si existe un número positivo M tal que

$$\sum_{i=1}^q |f(x_i) - f(x_{i-1})| \leq M$$

para toda partición $\mathcal{P} = \{x_0, x_1, \dots, x_q\}$ de $[a, b]$, entonces f se dice que es de *variación acotada* en $[a, b]$.

Definición 1.23. Sea $f : [a, b] \rightarrow \mathbb{R}$. El valor

$$V_{[a,b]}f = \sup_{\mathcal{P}} \left\{ \sum_{i=1}^q |f(x_i) - f(x_{i-1})| \right\} \leq +\infty$$

se llama *variación total o variación* de f en $[a, b]$. Así pues, las funciones de variación acotada serán aquellas cuya variación total es finita.

Definimos ahora el espacio de funciones de variación acotada siguiendo la definición de [1, cap 2]. Notemos que “*c.t.p.*” se refiere a la medida de Lebesgue λ .

Definición 1.24. El *espacio de funciones de variación acotada en $[a, b]$* , $BV([a, b])$, viene dado por

$$BV([a, b]) = \{f \in \mathcal{L}^1([a, b]) : \inf_{f_1=f \text{ c.t.p.}} V_{[a,b]}f_1 < +\infty\}.$$

Sobre este espacio podemos definir la siguiente norma:

$$\|f\|_{BV} = \|f\|_1 + \inf_{f_1=f \text{ c.t.p.}} V_{[a,b]}f_1.$$

Más adelante veremos que $\|\cdot\|_{BV}$ es, efectivamente, una norma.

En la norma $\|\cdot\|_{BV}$ es necesaria la introducción de la norma $\|\cdot\|_1$, ya que si no no sería una norma. Esto es debido a que puede haber funciones que no sean completamente nulas y sin embargo su variación sí lo es, incumpléndose las propiedades de una norma, en concreto las funciones constantes.

Comenzamos con unas primeras propiedades que permiten identificar tipos de funciones de variación acotada, en particular, las funciones monótonas crecientes, así como la suma y la diferencia de funciones de variación acotada. Se pueden encontrar en [11, cap 8, sec 4]. En la primera se muestra aquí una prueba más completa y en la segunda la diferencia es de elaboración propia.

Proposición 1.25. *Una función monótona creciente en $[a, b]$ es de variación acotada.*

Demostración. Sea f una función monótona creciente en $[a, b]$ y consideramos una partición cualquiera $a = x_0 < x_1 < \dots < x_{q-1} < x_q = b$. Al ser creciente, para cada par $x_i > x_{i-1}$ resulta $f(x_i) \geq f(x_{i-1})$ y por tanto $f(x_i) - f(x_{i-1})$ es no negativo. Calculamos la variación de f en esa

partición:

$$\sum_{i=1}^q |f(x_i) - f(x_{i-1})| = \sum_{i=1}^q f(x_i) - f(x_{i-1}) = f(b) - f(a) = M.$$

Como $M \geq 0$, es finito y la partición escogida es arbitraria, se deduce el resultado buscado. ■

Proposición 1.26. *La suma y diferencia de dos funciones de variación acotada son funciones de variación acotada. Más aún, $V_{[a,b]}(f + g), V_{[a,b]}(f - g) \leq V_{[a,b]}f + V_{[a,b]}g$.*

Demostración. Sean f y g dos funciones de variación acotada en $[a, b]$ y consideramos la función suma $s = f + g$. Escogemos una partición arbitraria $a = x_0 < x_1 < \dots < x_{q-1} < x_q = b$, entonces para cada par x_i, x_{i-1} resulta

$$|s(x_i) - s(x_{i-1})| = |f(x_i) - f(x_{i-1}) + g(x_i) - g(x_{i-1})| \leq |f(x_i) - f(x_{i-1})| + |g(x_i) - g(x_{i-1})|.$$

Tomando sumatorios se obtiene

$$\sum_{i=1}^q |s(x_i) - s(x_{i-1})| \leq \sum_{i=1}^q |f(x_i) - f(x_{i-1})| + \sum_{i=1}^q |g(x_i) - g(x_{i-1})| \leq M_1 + M_2 = M,$$

donde M_1 y M_2 son las respectivas variaciones de f y g . Como la partición escogida es arbitraria se deduce que $V_{[a,b]}s \leq V_{[a,b]}f + V_{[a,b]}g$ y la suma de funciones de variación acotada es de variación acotada.

De forma análoga, si consideramos la función diferencia $h = f - g$ y escogemos la misma partición arbitraria $a = x_0 < x_1 < \dots < x_{q-1} < x_q = b$, entonces para cada par x_i, x_{i-1} resulta

$$|h(x_i) - h(x_{i-1})| = |f(x_i) - f(x_{i-1}) - (g(x_i) - g(x_{i-1}))| \leq |f(x_i) - f(x_{i-1})| + |g(x_i) - g(x_{i-1})|.$$

Tomando sumatorios se obtiene

$$\sum_{i=1}^q |h(x_i) - h(x_{i-1})| \leq \sum_{i=1}^q |f(x_i) - f(x_{i-1})| + \sum_{i=1}^q |g(x_i) - g(x_{i-1})| \leq M_1 + M_2 = M,$$

donde M_1 y M_2 son las respectivas variaciones de f y g . Como la partición escogida es arbitraria se deduce que $V_{[a,b]}h \leq V_{[a,b]}f + V_{[a,b]}g$ y la diferencia de funciones de variación acotada es de variación acotada. ■

Asimismo, se prueba fácilmente que un múltiplo de una función de variación acotada también es de variación acotada, esto es $BV([a, b])$ es un subespacio vectorial de $\mathcal{L}^1([a, b])$. Veamos que $\|\cdot\|_{BV}$ es una norma en dicho espacio. La prueba es de elaboración propia.

Proposición 1.27. $\|\cdot\|_{BV}$ es una norma en $BV([a, b])$.

Demostración. Veamos que $\|\cdot\|_{BV}$ cumple las tres condiciones que definen a una norma.

Sea $f \in BV([a, b])$. Entonces $\|f\|_{BV} \geq 0$, al ser $\|f\|_1 \geq 0$ y $V_{[a,b]}f \geq 0$ para cualquier función f .

Si $\|f\|_{BV} = 0$ entonces $\|f\|_1 = 0$, por tanto $f = 0$. Si $f = 0$ entonces $\|f\|_1 = 0$ y $\inf_{f_1=f \text{ c.t.p.}} V_{[a,b]}f_1 = 0$, por lo tanto $\|f\|_{BV} = 0$.

Sean $\alpha \in \mathbb{R}$ y $f \in BV([a, b])$. Entonces se tiene

$$\|\alpha f\|_{BV} = \|\alpha f\|_1 + \inf_{\alpha f_1 = \alpha f \text{ c.t.p.}} V_{[a,b]}\alpha f_1 = |\alpha| \|f\|_1 + |\alpha| \inf_{f_1=f \text{ c.t.p.}} V_{[a,b]}f_1 = |\alpha| \|f\|_{BV}.$$

Por último, comprobamos la desigualdad triangular. Sean f y $g \in BV([a, b])$. Entonces por la proposición anterior, 1.26, se tiene, fijada $\bar{f} = f$ c.t.p.,

$$\begin{aligned} \|f + g\|_{BV} &= \|f + g\|_1 + \inf_{h=f+g \text{ c.t.p.}} V_{[a,b]}h \\ &\leq \|f\|_1 + \|g\|_1 + \inf_{\bar{g}=g \text{ c.t.p.}} V_{[a,b]}(\bar{f} + \bar{g}) \\ &\leq \|f\|_1 + \|g\|_1 + \inf_{\bar{g}=g \text{ c.t.p.}} V_{[a,b]}\bar{f} + V_{[a,b]}\bar{g} \\ &= \|f\|_1 + \|g\|_1 + V_{[a,b]}\bar{f} + \inf_{\bar{g}=g \text{ c.t.p.}} V_{[a,b]}\bar{g} \\ &= \|f\|_1 + V_{[a,b]}\bar{f} + \|g\|_{BV}. \end{aligned}$$

Por tanto, tomando el ínfimo,

$$\|f + g\|_{BV} \leq \|f\|_1 + \inf_{\bar{f}=f \text{ c.t.p.}} V_{[a,b]}\bar{f} + \|g\|_{BV} = \|f\|_{BV} + \|g\|_{BV}.$$

■

El siguiente resultado puede encontrarse en [1, cap 2].

Proposición 1.28. Sea $f \in BV([a, b])$. Entonces $|f(x)| \leq V_{[a,b]}f + \frac{\|f\|_1}{b-a}$ para todo $x \in [a, b]$.

Demostración. Para comenzar la prueba vamos a garantizar la existencia de un punto $y \in [a, b]$ tal que $|f(y)| \leq \frac{\|f\|_1}{b-a}$. Si esto no ocurriera entonces para todo $x \in [a, b]$ se tendría lo contrario,

$$|f(x)| > \frac{\|f\|_1}{b-a}.$$

De esta forma

$$\|f\|_1 = \int_a^b |f(x)| d\lambda > \int_a^b \frac{\|f\|_1}{b-a} d\lambda = \|f\|_1,$$

lo que es una contradicción.

Por lo tanto, existe $y \in [a, b]$ tal que $|f(y)| \leq \frac{\|f\|_1}{b-a}$. Además, para todo $x \in [a, b]$ se cumple

$$|f(x)| \leq |f(x) - f(y)| + |f(y)|,$$

usando lo anterior

$$|f(x)| \leq |f(x) - f(y)| + \frac{\|f\|_1}{b-a},$$

y por la definición de $V_{[a,b]}f$

$$|f(x)| \leq V_{[a,b]}f + \frac{\|f\|_1}{b-a}.$$

■

El siguiente teorema es de gran utilidad a la hora de trabajar con funciones de variación acotada, ya que garantiza que la variación en un intervalo es igual a la suma de las variaciones en subintervalos. La prueba es una variación de la propuesta en [11, cap 8].

Teorema 1.29. Sean $f : [a, b] \rightarrow \mathbb{R}$ una función de variación acotada y $c \in (a, b)$. Entonces f es de variación acotada en $[a, c]$ y $[c, b]$ cumpliendo

$$V_{[a,b]}f = V_{[a,c]}f + V_{[c,b]}f.$$

Demostración. Vamos a comenzar tomando dos particiones de los intervalos $[a, c]$ y $[c, b]$. Sean

$$y_0 = a < y_1 < \dots < y_m = c, \quad z_0 = c < z_1 < \dots < z_q = b,$$

y consideramos los sumatorios

$$V_1 = \sum_{k=1}^m |f(y_k) - f(y_{k-1})|, \quad V_2 = \sum_{k=1}^q |f(z_k) - f(z_{k-1})|.$$

En particular, los puntos $\{y_k\}_{k=1}^m \cup \{z_k\}_{k=1}^q$ forman una partición del intervalo completo $[a, b]$. Si consideramos V como la suma de los sumatorios anteriores

$$V = V_1 + V_2,$$

claramente por la definición de $V_{[a,b]}f$

$$V = V_1 + V_2 \leq V_{[a,b]}f$$

y tomando supremos

$$V_{[a,c]}f + V_{[c,b]}f \leq V_{[a,b]}f. \tag{1.1}$$

Veamos la otra desigualdad. Consideramos una partición del intervalo $[a, b]$

$$x_0 = a < x_1 < \dots < x_q = b,$$

de manera que el punto c está incluida en ella, es decir $c = x_m$ para algún $0 < m < q$. Ahora podemos considerar la suma, V , asociada a dicha partición,

$$V = \sum_{k=1}^m |f(x_k) - f(x_{k-1})| + \sum_{k=m+1}^q |f(x_k) - f(x_{k-1})| = V_1 + V_2,$$

donde V_1 y V_2 se corresponden con las sumas de los intervalos $[a, c]$ y $[c, b]$ respectivamente, por la elección de c . Entonces, por la definición de variación total en un intervalo, claramente

$$V \leq V_{[a,c]}f + V_{[c,b]}f. \quad (1.2)$$

La partición del intervalo $[a, b]$ la hemos tomado incluyendo el punto c , lo que implica que para cualquier otra partición que no contenga a c la suma también cumple 1.2 (porque si se añade un punto a la partición la suma nunca decrece). De aquí se deduce

$$V_{[a,b]}f \leq V_{[a,c]}f + V_{[c,b]}f. \quad (1.3)$$

Finalmente, combinando 1.1 y 1.3 se obtiene la igualdad. ■

Observación 3. Del teorema anterior se puede deducir fácilmente, por inducción, que si tenemos una partición del intervalo $[a, b]$, $\mathcal{P} = \{x_0, \dots, x_q\}$, considerando los intervalos cerrados $I_i = [x_{i-1}, x_i]$ se tiene que para toda $f : [a, b] \rightarrow \mathbb{R}$ de variación acotada

$$V_{[a,b]}f = \sum_{i=1}^q V_{I_i}f.$$

Este resultado será aplicado constantemente en la prueba de diferentes resultados relacionados con funciones de variación acotada.

1.2.1. El teorema de elección de Helly

Para probar un resultado posterior, que garantiza que un conjunto acotado de funciones de variación acotada es precompacto para la norma \mathcal{L}^1 , es necesario introducir el teorema de elección o selección de Helly. Para ello se requieren una serie de lemas previos, por lo que hemos considerado adecuado dedicar un apartado para él. Los siguientes resultados pueden encontrarse en [11, cap 8, sec 4].

Definición 1.30. Diremos que una familia de funciones \mathcal{H} definidas en $[a, b]$ está *uniformemente acotada* si existe K tal que

$$|f(x)| \leq K$$

para toda $f \in \mathcal{H}$ y $x \in [a, b]$.

Lema 1.31. *Sea \mathcal{H} una familia uniformemente acotada. Entonces para todo subconjunto numerable E de $[a, b]$ y cada sucesión F en \mathcal{H} , existe una subsucesión de F que converge en todo punto del conjunto E .*

Demostración. Sea $E = \{x_k\}_{k=1}^{+\infty}$ un conjunto numerable. Consideremos el conjunto formado por los valores que toman las funciones de la sucesión F en el punto x_1 , es decir

$$\{f(x_1) : f \in F\}.$$

Como toda función de la familia está acotada por el mismo K , el conjunto anterior también lo está y por el teorema de Bolzano-Weierstrass podemos extraer una subsucesión convergente, digamos

$$f_1^{(1)}(x_1), f_2^{(1)}(x_1), f_3^{(1)}(x_1), \dots \text{ tal que } \lim_{n \rightarrow +\infty} f_n^{(1)}(x_1) = A_1.$$

Consideramos ahora la sucesión

$$f_1^{(1)}(x_2), f_2^{(1)}(x_2), f_3^{(1)}(x_2), \dots$$

formada por los valores que toman las funciones del conjunto $\{f_n^{(1)}(x)\}_n$ en x_2 . Esta sucesión, de nuevo, es acotada y el teorema de Bolzano-Weierstrass garantiza que podemos obtener una subsucesión convergente de ella, sea

$$f_1^{(2)}(x_2), f_2^{(2)}(x_2), f_3^{(2)}(x_2) \text{ tal que } \lim_{n \rightarrow \infty} f_n^{(2)}(x_2) = A_2,$$

donde es importante remarcar que el orden relativo entre dos funciones $f_n^{(2)}$ y $f_m^{(2)}$ en la sucesión anterior es el mismo que en la sucesión $(f_r^{(1)}(x_1))_r$. Repitiendo este proceso de obtención de subsucesiones convergentes de manera indefinida, podemos construir un conjunto numerable de sucesiones convergentes como sigue:

$$f_1^{(1)}(x_1), f_2^{(1)}(x_1), f_3^{(1)}(x_1), \dots \text{ tal que } \lim_{n \rightarrow +\infty} f_n^{(1)}(x_1) = A_1,$$

$$f_1^{(2)}(x_2), f_2^{(2)}(x_2), f_3^{(2)}(x_2), \dots \text{ tal que } \lim_{n \rightarrow +\infty} f_n^{(2)}(x_2) = A_2,$$

⋮

$$f_1^{(k)}(x_k), f_2^{(k)}(x_k), f_3^{(k)}(x_k), \dots \text{ tal que } \lim_{n \rightarrow +\infty} f_n^{(k)}(x_k) = A_k,$$

⋮

de manera que cada sucesión es una subsucesión de la anterior evaluada en el punto x_k correspondiente y el orden de los elementos no se ve alterado. Si consideramos ahora la sucesión formada por las funciones en la diagonal, es decir

$$\left(f_n^{(n)} \right)_n,$$

esta sucesión converge en cada punto del conjunto E . De hecho, para todo k fijo, la sucesión

$$\left(f_n^{(n)}(x_k) \right)_{n \geq k}$$

es una subsucesión de $\left(f_n^{(k)}(x_k) \right)_n$ y converge a A_k . ■

Lema 1.32. *Sea \mathcal{H} una familia uniformemente acotada de funciones crecientes en $[a, b]$. Entonces, para cada sucesión F en \mathcal{H} existe una subsucesión de F que converge a una función creciente g en todo punto de $[a, b]$.*

Demostración. Si aplicamos el lema anterior 1.31 a la sucesión F , tomando como conjunto numerable E el conjunto de todos los puntos racionales en $[a, b]$ junto con el punto a , en el caso que a sea irracional, podemos encontrar una subsucesión de F , digamos F_0 con

$$F_0 = \left(f^{(n)} \right)_n,$$

tal que su límite

$$\lim_{n \rightarrow +\infty} f^{(n)}(x_k)$$

existe y es finito en cada punto $x_k \in E$.

Vamos a definir una función ϕ de la siguiente manera. En primer lugar, definimos

$$\phi(x_k) = \lim_{n \rightarrow +\infty} f^{(n)}(x_k)$$

para todo $x_k \in E$, quedando la función ϕ ya definida para todo punto de E . Definida de esa forma, la función ϕ es creciente en E , ya que si $x_i, x_j \in E$ con $x_i < x_j$ se tiene

$$\phi(x_i) \leq \phi(x_j),$$

al ser las funciones $f^{(n)}$ crecientes.

Para $x \in [a, b] \setminus E$ definimos $\phi(x)$ como

$$\phi(x) = \sup_{x_k < x} \{\phi(x_k)\}$$

donde $x_k \in E$. De esta forma, $\phi(x)$ es claramente una función creciente en el intervalo $[a, b]$ y el conjunto de los puntos donde ϕ es una función discontinua, digamos B , es a lo sumo numerable.

El siguiente paso va a ser ver que

$$\lim_{n \rightarrow +\infty} f^{(n)}(x_0) = \phi(x_0)$$

en cada punto x_0 donde ϕ es continua. Para ello, tomamos $\varepsilon > 0$ arbitrario y x_i, x_j puntos de E tales que

$$x_j < x_0 < x_i, \quad \phi(x_i) - \phi(x_j) < \frac{\varepsilon}{2}.$$

Al ser x_i, x_j puntos de E , por su definición, podemos encontrar un natural n_0 tal que para todo $n \geq n_0$ se cumple

$$|f^{(n)}(x_j) - \phi(x_j)| < \frac{\varepsilon}{2} \quad \text{y} \quad |f^{(n)}(x_i) - \phi(x_i)| < \frac{\varepsilon}{2}.$$

De esta forma, se puede deducir que

$$|f^{(n)}(x_i) - \phi(x_0)| \leq |f^{(n)}(x_i) - \phi(x_i)| + |\phi(x_i) - \phi(x_0)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

y

$$|f^{(n)}(x_j) - \phi(x_0)| \leq |f^{(n)}(x_j) - \phi(x_j)| + |\phi(x_j) - \phi(x_0)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

obteniéndose la cadena de desigualdades

$$\phi(x_0) - \varepsilon < f^{(n)}(x_j) \leq f^{(n)}(x_i) < \phi(x_0) + \varepsilon,$$

para $n \geq n_0$. Como

$$f^{(n)}(x_j) \leq f^{(n)}(x_0) \leq f^{(n)}(x_i),$$

al ser funciones crecientes, resulta

$$\phi(x_0) - \varepsilon < f^{(n)}(x_0) < \phi(x_0) + \varepsilon$$

para $n \geq n_0$. Con lo que queda probado que

$$\lim_{n \rightarrow +\infty} f^{(n)}(x) = \phi(x),$$

en todos los puntos donde hay continuidad. Lo anterior solo puede fallar en el conjunto B , donde no hay continuidad. Para terminar la prueba, vamos a aplicar el lema 1.31 a la sucesión F_0 tomando como conjunto E los puntos de B donde no se cumple el límite anterior, que a lo sumo es numerable. Esto nos lleva a una subsucesión de F_0 , digamos $F_1 = (f_n)_n$, que converge en todos los puntos de $[a, b]$, ya que es una subsucesión de una sucesión convergente en los puntos que no pertenecen al nuevo E y el lema 1.31 garantiza la convergencia en E . Estableciendo la función

$$g(x) = \lim_{n \rightarrow +\infty} f_n(x),$$

al ser las f_n crecientes, g será creciente y el resultado queda probado. ■

Con estos dos lemas preliminares ya estamos en disposición de probar el teorema de Helly.

Teorema 1.33 (Teorema de Helly). *Sea \mathcal{H} una familia uniformemente acotada de funciones en $[a, b]$, con constante de acotación K , y tal que $V_{[a,b]}f \leq K$ para toda $f \in \mathcal{H}$. Entonces, para cada sucesión F en \mathcal{H} existe una subsucesión de F que converge en todo punto de $[a, b]$ a una función g de variación acotada.*

Demostración. Para cada función f de la sucesión F definimos las funciones

$$\pi_f(x) = V_{[a,x]}f \quad \text{y} \quad \nu_f(x) = \pi_f(x) - f(x).$$

Estas funciones son crecientes, veámoslo. En primer lugar, $\pi_f(x)$ es creciente, ya que, por el teorema 1.29, si $x < y$ entonces

$$\pi_f(x) = V_{[a,x]}f \leq V_{[a,x]}f + V_{[x,y]}f = V_{[a,y]}f = \pi_f(y).$$

Por otro lado, $\nu(x)$ también lo es, ya que si $x < y$ se tiene

$$\nu_f(y) = \pi_f(y) - f(y) = \pi_f(x) + V_{[x,y]}f - f(y)$$

y

$$\nu_f(y) - \nu_f(x) = V_{[x,y]}f - (f(y) - f(x)).$$

Ahora bien, por la definición de variación total

$$f(y) - f(x) \leq V_{[x,y]},$$

luego

$$\nu_f(y) \geq \nu_f(x).$$

Observemos que las familias $\mathcal{U} = \{\pi_f : f \in \mathcal{H}\}$ y $\mathcal{V} = \{\nu_f : f \in \mathcal{H}\}$ están uniformemente acotadas por K y $2K$ respectivamente, por lo que podemos aplicar el lema 1.32 a ambas familias. Aplicándose a $A = \{\pi_f : f \in F\}$ podemos encontrar una subsucesión convergente, $(\pi_k)_k$, tal que

$$\lim_{n \rightarrow +\infty} \pi_k(x) = \alpha(x),$$

con α creciente. Por otro lado, para cada función $\pi_k(x)$ corresponde una función $\nu_k(x)$ y de la misma manera una función $f_k(x)$ de F . Aplicando el lema 1.32 a la sucesión $(\nu_k)_k$, podemos encontrar una subsucesión convergente, $(\nu_{k_i})_i$ de $(\nu_k)_k$, tal que

$$\lim_{i \rightarrow +\infty} \nu_{k_i}(x) = \beta(x).$$

Entonces la sucesión de funciones

$$f_{k_i}(x) = \pi_{k_i}(x) - \nu_{k_i}(x),$$

converge a la función

$$g(x) = \alpha(x) - \beta(x).$$

Como g es diferencia de dos funciones monótonas crecientes, por tanto diferencia de dos funciones de variación acotada, por 1.25, la proposición 1.26 garantiza que g es de variación acotada. ■

A continuación, se presentan las funciones constantes a trozos o funciones de paso, que tendrán especial importancia en la consecución del objetivo del trabajo, ya que, bajo las condiciones adecuadas, la función de densidad de una medida absolutamente continua será una función de paso. Además, a partir de ellas podemos probar la densidad de las funciones de variación acotada en $\mathcal{L}^1([a, b])$. Se puede encontrar en [4, cap 2].

Definición 1.34. Una *función constante a trozos* (o de paso) es una función real de la forma

$$\sum_{i=1}^n c_i \chi_{A_i},$$

donde los valores c_i son constantes y los A_i son intervalos de longitud finita.

El siguiente resultado establece la densidad de las funciones de paso en $\mathcal{L}^1(\mathbb{R})$. Puede encontrarse en [4, cap 5, prop 5.1.2].

Proposición 1.35. *El espacio de las funciones de paso es denso en $\mathcal{L}^1(\mathbb{R})$.*

Demostración. Sea $f \in \mathcal{L}^1(\mathbb{R})$. Vamos a ver que la podemos aproximar por una función de paso tanto como se quiera. Consideramos las partes positiva y negativa de f , f^+ y f^- , con $f = f^+ - f^-$, para las que existen sucesiones monótonas crecientes de funciones simples no negativas, $(f_n^+)_n$ y $(f_n^-)_n$, tales que $(f_n^+)_n \uparrow f^+$ y $(f_n^-)_n \uparrow f^-$. Por el teorema de la convergencia monótona, se cumple

$$\int_{\mathbb{R}} f_n^+ - \int_{\mathbb{R}} f^+ \longrightarrow 0 \text{ cuando } n \longrightarrow +\infty$$

y

$$\int_{\mathbb{R}} f_n^- - \int_{\mathbb{R}} f^- \longrightarrow 0 \text{ cuando } n \longrightarrow +\infty,$$

por tanto $\|f_n^+ - f^+\|_1 \longrightarrow 0$ y $\|f_n^- - f^-\|_1 \longrightarrow 0$ cuando n tiende a infinito. Así, dado $\varepsilon > 0$ se puede encontrar un cierto n_ε tal que para todo $n \geq n_\varepsilon$ se cumple $\|f_n^+ - f^+\|_1 < \frac{\varepsilon}{4}$ y $\|f_n^- - f^-\|_1 < \frac{\varepsilon}{4}$. Sea $f_n = f_n^+ - f_n^-$, que es una función simple al serlo tanto f_n^+ como f_n^- . Por las desigualdades anteriores resulta

$$\|f - f_n\|_1 = \|f^+ - f_n^+ + f_n^- - f^-\|_1 \leq \|f^+ - f_n^+\|_1 + \|f^- - f_n^-\|_1 < \frac{\varepsilon}{2}.$$

Si fijamos un valor de n , como cada f_n es simple, será de la forma

$$f_n = \sum_{j=1}^r c_j \chi_{E_j},$$

con E_j conjuntos de Borel de medida finita. Sea $\kappa = r \cdot \max\{|c_j| : j = 1, 2, \dots, r\}$ y definimos $\varepsilon' = \frac{1}{4} \frac{\varepsilon}{\kappa}$. Consideremos un conjunto E de los E_j . Al ser de Borel existe una sucesión de intervalos $(J_l)_l$ tal que $E \subset \bigcup_{l=1}^{+\infty} J_l$ y $\sum_{l=1}^{+\infty} \lambda(J_l) < \lambda(E) + \varepsilon'$. Como $\lambda(E)$ es finita entonces $\sum_{l=1}^{+\infty} \lambda(J_l)$ es convergente y existe un N tal que

$$\sum_{l=1}^N \lambda(J_l) > \sum_{l=1}^{+\infty} \lambda(J_l) - \varepsilon'.$$

Definimos

$$A = \bigcup_{l=1}^N J_l \quad \text{y} \quad B = \bigcup_{l=1}^{+\infty} J_l$$

y consideramos la diferencia en norma de las funciones características de A y E . Entonces

$$\begin{aligned} \|\chi_E - \chi_A\|_1 &= \lambda((E \setminus A) \cup (A \setminus E)) \leq \lambda(E \setminus A) + \lambda(A \setminus E) \\ &\leq \lambda(B \setminus A) + \lambda(B \setminus E) < 2\varepsilon' = \frac{1}{2} \frac{\varepsilon}{\kappa}, \end{aligned}$$

por las desigualdades anteriores.

Repetiendo el mismo proceso para cada E_j , podemos encontrar conjuntos A_j que sean unión finita de intervalos y la función $g = \sum_{j=1}^r c_j \chi_{A_j}$ es una función de paso. Haciendo la diferencia entre f y g en norma, resulta que

$$\begin{aligned} \|f - g\|_1 &\leq \|f - f_n\|_1 + \|f_n - g\|_1 < \frac{\varepsilon}{2} + \left\| \sum_{j=1}^r c_j \chi_{E_j} - \sum_{j=1}^r c_j \chi_{A_j} \right\|_1 \\ &< \frac{\varepsilon}{2} + \sum_{j=1}^r |c_j| \|\chi_{E_j} - \chi_{A_j}\|_1 < \frac{\varepsilon}{2} + \kappa \cdot \frac{\varepsilon}{2\kappa} = \varepsilon. \end{aligned}$$

Como $\varepsilon > 0$ es arbitrario, tomando el límite cuando tiende a 0 se obtiene el resultado buscado. ■

Del resultado anterior podemos deducir la densidad, no solo en $\mathcal{L}^1(\mathbb{R})$, sino también en $\mathcal{L}^1([a, b])$ y, a partir de ella, la de las funciones de variación acotada en $\mathcal{L}^1([a, b])$. El siguiente resultado puede encontrarse en [1, cap 2].

Proposición 1.36. *El conjunto $BV([a, b])$ es denso en $\mathcal{L}^1([a, b])$.*

Demostración. Como las funciones constantes a trozos o de paso son densas en $\mathcal{L}^1([a, b])$, al estar dichas funciones contenidas en el espacio de las de variación acotada $BV([a, b])$ (las funciones características de intervalos obviamente lo están, úsese también la proposición 1.26), se deduce que $BV([a, b])$ es denso en $\mathcal{L}^1([a, b])$. ■

Veamos la definición de conjunto precompacto en $\mathcal{L}^1([a, b])$.

Definición 1.37. Un conjunto \mathcal{U} se dice *precompacto* en $\mathcal{L}^1([a, b])$ si su clausura es compacta en $\mathcal{L}^1([a, b])$.

La siguiente proposición garantiza que si tenemos un conjunto acotado en el espacio de funciones de variación acotada, entonces es precompacto. La prueba presentada aquí es una variación de la que se puede encontrar en [1, cap 2, prop 2.3.4]. En nuestro caso, se emplea el teorema de la convergencia dominada para probar la convergencia en $\mathcal{L}^1([a, b])$, mientras que en la referencia anterior se emplea otro teorema auxiliar.

Proposición 1.38. *Si \mathcal{U} es un conjunto acotado en $BV([a, b])$, entonces es precompacto en $\mathcal{L}^1([a, b])$.*

Demostración. Basta probar que toda sucesión $(f_n)_n$ en \mathcal{U} admite una subsucesión convergente en $\mathcal{L}^1([a, b])$. Por la acotación, existe $K_1 > 0$ tal que

$$\|f_n\|_{BV} \leq K_1 \quad \text{para todo } n.$$

Por la definición de la norma $\|\cdot\|_{BV}$, al ser la suma de la norma $\|\cdot\|_1$ y el ínfimo de las variaciones totales de las funciones iguales en casi todo punto a la función, la proposición 1.28 garantiza (reemplazando si es necesario f_n por otra función igual a ella en casi todo punto) que la familia $\{f_n\}_n$ está acotada uniformemente, es decir existe una constante $K_2 > 0$ tal que

$$|f_n(x)| \leq K_2 \quad \text{para todo } n \text{ y todo } x \in [a, b].$$

Sea $K = \max\{K_1, K_2\}$. Entonces

$$|f_n(x)| \leq K \quad \text{y} \quad V_{[a,b]} f_n \leq K \quad \text{para todo } n \text{ y todo } x \in [a, b],$$

por lo que estamos en las condiciones para aplicar el teorema de Helly 1.33, el cual garantiza la existencia de una subsucesión $(f_{n_k})_k$ tal que

$$f_{n_k} \longrightarrow f^*$$

en todo punto, con f^* de variación acotada y $|f^*(x)| \leq K$ para todo $x \in [a, b]$, por lo que $f^* \in \mathcal{L}^1([a, b])$.

Como $(f_n)_n$ es uniformemente acotada y f_{n_k} tiende a f^* en todo punto, podemos aplicar el teorema de la convergencia dominada a la sucesión $(|f_{n_k} - f^*|)_k$ y deducir que $f_{n_k} \longrightarrow f^*$ en $\mathcal{L}^1([a, b])$. ■

La siguiente proposición establece que si tenemos una sucesión de funciones acotada en $BV([a, b])$ y convergente en $\mathcal{L}^1([a, b])$ a una función f , entonces $f \in BV([a, b])$. Se puede encontrar en [1, cap 2, prop 2.3.5].

Proposición 1.39. *Si existe $K > 0$ tal que $\|f_n\|_{BV} \leq K$, para cada n y $f_n \longrightarrow f$ en $\mathcal{L}^1([a, b])$, entonces $f \in BV([a, b])$.*

Demostración. Reemplazando cada f_n si es necesario por una función igual en casi todo punto, podemos suponer que $V_{[a,b]} f_n \leq K$ para cada n y argumentar como en la proposición anterior para encontrar $(f_{n_k})_k$ y $f^* \in BV([a, b])$ tales que $f_{n_k} \longrightarrow f^*$ en $\mathcal{L}^1([a, b])$, por lo que $f = f^*$ estará en $BV([a, b])$. ■

1.3. Teoría ergódica

En esta sección se realiza una introducción a algunos de los resultados más importantes y conocidos de la llamada teoría ergódica. Estos serán necesarios para obtener la descripción de la dinámica que andamos buscando: que la frecuencia relativa con la que una órbita cae dentro de un conjunto en un espacio de probabilidad, converja a la medida de dicho conjunto en casi todo punto, pero no solo en casi todo punto respecto a la medida, sino en casi todo punto respecto a la medida de Lebesgue. Los resultados más importantes son el teorema ergódico de Birkhoff y un corolario del mismo relativo a la convergencia en media a la que acabamos de referirnos.

La siguiente noción es la gran protagonista de la sección. La idea básica tras la misma es que si un conjunto A es irrelevante para la dinámica de una transformación τ , en el sentido de que las trayectorias $\{\tau^n(x)\}_n$ no visitan dicho conjunto (lo que ocurrirá, por ejemplo, si $\tau^{-1}(A) = \emptyset$), entonces su medida $\mu(A)$ sea nula. Se puede encontrar en [12, cap 1].

Definición 1.40. Sea $(\mathbb{X}, \mathcal{A}, \mu)$ un espacio de medida. Una transformación medible $\tau : \mathbb{X} \longrightarrow \mathbb{X}$ se dice que *conserva la medida* μ o que μ es τ -invariante si $\mu(\tau^{-1}(A)) = \mu(A)$ para todo $A \in \mathcal{A}$.

Observación 4. En las condiciones del teorema 1.16, si μ es τ -invariante entonces

$$\int_{\mathbb{X}} f d\mu = \int_{\mathbb{X}} f \circ \tau d\mu,$$

al ser $\mu(\tau^{-1}(A)) = \mu(A)$ para todo $A \in \mathcal{A}$.

A partir de aquí nos concentramos en transformaciones que conservan la medida de un espacio de probabilidad. Las siguientes nociones pueden encontrarse en [1, cap 3].

Definición 1.41. Si $(\mathbb{X}, \mathcal{A}, \mu)$ es un espacio de probabilidad y $\tau : \mathbb{X} \rightarrow \mathbb{X}$ conserva la medida μ , llamamos a la cuaterna $(\mathbb{X}, \mathcal{A}, \mu, \tau)$ *sistema dinámico*.

Definición 1.42. Sea $(\mathbb{X}, \mathcal{A}, \mu, \tau)$ un sistema dinámico. Un conjunto $A \in \mathcal{A}$ se dice τ -invariante si $\tau^{-1}(A) = A$. Se dice *casi τ -invariante* si $\mu(\tau^{-1}(A) \Delta A) = 0$, donde $B \Delta C = (B \setminus C) \cup (C \setminus B)$. Una función medible $f : \mathbb{X} \rightarrow \mathbb{R}$ se dice τ -invariante si $f \circ \tau = f$ y se dice *casi τ -invariante* si $f \circ \tau = f$ μ -c.t.p.

Definición 1.43. Sea $(\mathbb{X}, \mathcal{A}, \mu, \tau)$ un sistema dinámico. Se dice que τ es *ergódica* (o que μ es *ergódica* para τ) si para todo $A \in \mathcal{A}$ τ -invariante se cumple $\mu(A) = 0$ o $\mu(A) = 1$.

1.3.1. El teorema Ergódico de Birkhoff

En este apartado se va a probar el teorema ergódico de Birkhoff junto con todos aquellos resultados necesarios para su prueba. El resultado más importante es en realidad un corolario de este teorema, que garantiza que si la transformación τ es ergódica, entonces para cualquier subconjunto E del espacio \mathbb{X} , la frecuencia relativa con la que la órbita de un punto $x \in \mathbb{X}$, $\{\tau^n(x)\}_{n=0}^{+\infty}$, cae en E , es decir $\frac{1}{n} \sum_{i=0}^{n-1} \chi_E(\tau^i(x))$, viene dada por $\mu(E)$ μ -c.t.p.

El primer teorema que necesitamos probar requiere de una serie de lemas previos. Aquí se presenta una versión reducida del mismo. Puede encontrarse una prueba en [1, cap 3, teo 3.2.2].

Teorema 1.44. Sea $(\mathbb{X}, \mathcal{A}, \mu, \tau)$ un sistema dinámico. Si τ es ergódica y A es casi τ -invariante, entonces $\mu(A) = 0$ o 1.

Lema 1.45. Si $\tau^{-1}(A) \subset A$, entonces existe un conjunto $A_1 \subset A$ tal que $\mu(A \setminus A_1) = 0$ y $\tau^{-1}(A_1) = A_1$.

Demostración. La inclusión $\tau^{-1}(A) \subset A$ implica $\dots \subset \tau^{-2}(A) \subset \tau^{-1}(A) \subset A$, es decir, obtenemos una familia monótona decreciente de conjuntos. Sea $A_1 = \bigcap_{k=0}^{+\infty} \tau^{-k}(A)$. Claramente $A_1 \subset A$ y $\mu(A_1) = \mu(\bigcap_{k=0}^{+\infty} \tau^{-k}(A)) = \lim_{k \rightarrow +\infty} \mu(\tau^{-k}(A))$ y al ser μ τ -invariante se tiene $\mu(\tau^{-k}(A)) = \mu(A)$. Por tanto, $\mu(A_1) = \mu(A)$ y $\mu(A \setminus A_1) = 0$. Además $\tau^{-1}(A_1) = \tau^{-1}(\bigcap_{k=0}^{+\infty} \tau^{-k}(A)) = \bigcap_{k=0}^{+\infty} \tau^{-k-1}(A) = \bigcap_{k=1}^{+\infty} \tau^{-k}(A) = A_1$. ■

Lema 1.46. Si A es casi τ -invariante, entonces existe un conjunto A_1 tal que $\mu(A \Delta A_1) = 0$ y $\tau^{-1}(A_1) = A_1$.

Demostación. Si $\mu(\tau^{-1}(A) \triangle A) = \mu((\tau^{-1}(A) \setminus A) \cup (A \setminus \tau^{-1}(A))) = 0$, en particular $\mu(\tau^{-1}(A) \setminus A) = 0$. Por otro lado, sea $A_2 = \bigcup_{k=0}^{+\infty} \tau^{-k}(A)$, que claramente cumple $A \subset A_2$ y podemos considerar

$$A_2 \setminus A = \left(\bigcup_{k=0}^{+\infty} \tau^{-k}(A) \right) \setminus A = \bigcup_{k=0}^{+\infty} (\tau^{-k-1}(A) \setminus \tau^{-k}(A)).$$

Al ser μ τ -invariante y $\mu(\tau^{-1}(A) \setminus A) = 0$, se cumple

$$\begin{aligned} \mu(A_2 \setminus A) &= \mu \left(\bigcup_{k=0}^{+\infty} (\tau^{-k-1}(A) \setminus \tau^{-k}(A)) \right) \leq \sum_{k=0}^{+\infty} \mu(\tau^{-k-1}(A) \setminus \tau^{-k}(A)) \\ &= \sum_{k=0}^{+\infty} \mu(\tau^{-k}(\tau^{-1}(A) \setminus A)) = \sum_{k=0}^{+\infty} \mu(\tau^{-1}(A) \setminus A) = 0. \end{aligned}$$

Además, $\tau^{-1}(A_2) = \bigcup_{k=1}^{+\infty} \tau^{-k}(A) \subset A_2$, por lo que el lema 1.45 garantiza la existencia de un conjunto $A_1 \subset A_2$ cumpliendo $\mu(A_2 \setminus A_1) = 0$ y $\tau^{-1}(A_1) = A_1$. Como tenemos que $\mu(A_1 \setminus A) \leq \mu(A_2 \setminus A) = 0$ y $\mu(A \setminus A_1) \leq \mu(A_2 \setminus A_1) = 0$, se tiene $\mu(A \triangle A_1) = 0$. ■

Ya podemos ver la prueba del teorema.

Demostación. (Prueba del teorema 1.44). En las condiciones del teorema, si $\mu(\tau^{-1}(A) \triangle A) = 0$, por el lema 1.46, existe un conjunto A_1 tal que $\mu(A \triangle A_1) = 0$ (por lo tanto, $\mu(A) = \mu(A_1)$) y $\tau^{-1}(A_1) = A_1$. Además, por la ergodicidad de τ , $\tau^{-1}(A_1) = A_1$ implica $\mu(A_1) = 0$ o $\mu(A_1) = 1$ y por consiguiente $\mu(A) = 0$ o $\mu(A) = 1$. ■

El siguiente teorema es una variación del que podemos encontrar en [14, cap 1, teo 1.6].

Teorema 1.47. *Sea $(\mathbb{X}, \mathcal{A}, \mu, \tau)$ un sistema dinámico. Si τ es ergódica y $f : \mathbb{X} \rightarrow \mathbb{R}$ es medible y casi τ -invariante, entonces f es constante $\mu - c.t.p.$*

Demostación. Comenzamos definiendo para $k \in \mathbb{Z}$ y $n > 0$ el conjunto

$$X(k, n) = \{x : k/2^n \leq f(x) < (k+1)/2^n\} = f^{-1}([k/2^n, (k+1)/2^n))$$

y por tanto

$$\tau^{-1}(X(k, n)) = \{x : k/2^n \leq f(\tau(x)) < (k+1)/2^n\} = \tau^{-1}f^{-1}([k/2^n, (k+1)/2^n)).$$

Por otro lado,

$$\tau^{-1}(X(k, n)) \triangle X(k, n) \subset \{x : (f \circ \tau)(x) \neq f(x)\},$$

de donde podemos deducir que $\mu(\tau^{-1}(X(k, n)) \triangle X(k, n)) = 0$ al ser $f \circ \tau = f$ $\mu - c.t.p.$ Por el teorema 1.44, $\mu(X(k, n)) = 0$ o 1.

Ahora, para cada n fijo, la unión $\bigcup_{k \in \mathbb{Z}} X(k, n) = \mathbb{X}$ y por definición cada elemento es disjuncto con cualquier otro. Así, como $\mu(X(k, n)) = 0$ o 1, existe un único k_n tal que $\mu(X(k_n, n)) = 1$, teniendo el resto medida nula. Sea $Y = \bigcap_{n=1}^{+\infty} X(k_n, n)$. Como todos tienen medida 1, necesari-

riamente $\mu(Y) = 1$. Y como Y es la intersección infinita de esos conjuntos y cada $X(k_n, n) \subset X(k_{n+1}, n+1)$, f tiene que ser constante en Y . Como $\mu(Y) = 1$, f es constante $\mu - c.t.p.$ ■

Ya estamos en disposición de presentar el teorema ergódico de Birkhoff, que garantiza la convergencia de los promedios ergódicos de una función f en un sistema dinámico.

Teorema 1.48 (Teorema ergódico de Birkhoff). Sean $(\mathbb{X}, \mathcal{A}, \mu, \tau)$ un sistema dinámico y $f : \mathbb{X} \rightarrow \mathbb{R}$ integrable. Entonces existe una función $f^* \in \mathcal{L}^1(\mathbb{X})$ tal que

$$\frac{1}{n} \sum_{k=0}^{n-1} f(\tau^k(x)) \rightarrow f^*, \quad \mu - c.t.p.$$

Además, $f^* \circ \tau = f^*$ $\mu - c.t.p.$, $\|f^*\|_1 \leq \|f\|_1$ y $\int_{\mathbb{X}} f^* d\mu = \int_{\mathbb{X}} f d\mu$.

Antes de poder abordar la demostración de este teorema es necesario introducir una serie de conceptos y resultados que serán empleados en su prueba. En primer lugar introducimos el concepto de promedio ergódico, que va a ser utilizado con asiduidad en los siguientes resultados.

Definición 1.49. Sean $(\mathbb{X}, \mathcal{A}, \mu, \tau)$ un sistema dinámico y $f : \mathbb{X} \rightarrow \mathbb{R}$ integrable. Se llaman *promedios ergódicos* de f a las sumas

$$A_n(f) = \frac{1}{n} \sum_{i=0}^{n-1} f \circ \tau^i, \quad n = 1, 2, \dots$$

Además, necesitamos el siguiente resultado que puede encontrarse en [14, cap 1].

Lema 1.50. Sean $(\mathbb{X}, \mathcal{A}, \mu, \tau)$ un sistema dinámico, $U : \mathcal{L}^1(\mathbb{X}) \rightarrow \mathcal{L}^1(\mathbb{X})$ un operador lineal positivo (es decir, $f \geq 0$ implica $Uf \geq 0$) con $\|U\| \leq 1$, N un entero positivo y $f : \mathbb{X} \rightarrow \mathbb{R}$ integrable. Si definimos $f_0 = 0$, $f_n = f + Uf + U^2f + \dots + U^{n-1}f$ para $n \geq 1$ y $F_N = \max_{0 \leq n \leq N} f_n$, entonces

$$\int_{\{x: F_N(x) > 0\}} f d\mu \geq 0.$$

Demostración. Claramente F_N es integrable, al ser el máximo de funciones integrables. Por otro lado, al ser el máximo, $F_N \geq f_n$ para $0 \leq n \leq N$, por lo tanto $UF_N \geq Uf_n$ por la positividad del operador. De donde podemos deducir, por la definición de los f_n , que $UF_N + f \geq f_{n+1}$. Por lo tanto, cuando $F_N(x) > 0$,

$$\begin{aligned} UF_N(x) + f(x) &\geq \max_{1 \leq n \leq N} f_n(x) \\ &= \max_{0 \leq n \leq N} f_n(x) \quad (\text{al ser } f_0 = 0) \\ &= F_N(x). \end{aligned}$$

Despejando resulta que $f \geq F_N - UF_N$ cuando $F_N(x) > 0$, es decir, en el conjunto $A = \{x : F_N(x) > 0\}$, por tanto

$$\begin{aligned} \int_A f d\mu &\geq \int_A F_N d\mu - \int_A UF_N d\mu \\ &= \int_{\mathbb{X}} F_N d\mu - \int_A UF_N d\mu \quad (\text{al ser } F_N = 0 \text{ en } \mathbb{X} \setminus A) \\ &\geq \int_{\mathbb{X}} F_N d\mu - \int_X UF_N d\mu \quad (\text{al ser } UF_N \geq 0 \text{ por la positividad del operador}) \\ &= \int_{\mathbb{X}} F_N - UF_N d\mu \geq 0 \quad (\text{por ser } \|U\| \leq 1). \end{aligned}$$

■

El siguiente teorema recibe el nombre de teorema ergódico maximal. Establece que la integral en el conjunto donde el supremo de los promedios ergódicos es positivo es no negativa y lo vamos a probar a partir del anterior. Se puede encontrar enunciado en [12, cap 2, teo 2.1]. La prueba de que M es medible es de elaboración propia.

Teorema 1.51 (Teorema ergódico maximal). Sean $(\mathbb{X}, \mathcal{A}, \mu, \tau)$ un sistema dinámico, $f : \mathbb{X} \rightarrow \mathbb{R}$ integrable y M el conjunto

$$M = \{x : \sup_{n \geq 1} A_n(f)(x) > 0\}.$$

Entonces M es medible y

$$\int_M f d\mu \geq 0.$$

Demostración. El conjunto $M = \{x : \sup_{n \geq 1} A_n(f)(x) > 0\}$, por la definición de $A_n(f)$, puede escribirse como

$$M = \left\{ x : \sup_{n \geq 1} \frac{1}{n} \sum_{i=0}^{n-1} f(\tau^i(x)) > 0 \right\} = \left\{ x : \sup_{n \geq 1} \sum_{i=0}^{n-1} f(\tau^i(x)) > 0 \right\}.$$

Por tanto, puede verse como la unión disjunta de los conjuntos

$$\begin{aligned} M_1 &= \{x : f(x) > 0\} \\ M_2 &= \{x : f(x) \leq 0, f(x) + f(\tau(x)) > 0\} \\ &\vdots \\ M_n &= \{x : f(x) \leq 0, \dots, f(x) + \dots + f(\tau^{n-2}(x)) \leq 0, f(x) + \dots + f(\tau^{n-1}(x)) > 0\} \\ &\vdots \end{aligned}$$

Por ser f y τ medibles, las funciones $f, f \circ \tau, \dots, f \circ \tau^{k-1}$ son medibles y su suma también lo será. Por la definición de los M_k , al ser la suma de dichas funciones medibles, todos los conjuntos M_k son medibles. Al ser $M = \bigcup_{k=1}^{+\infty} M_k$, M es un conjunto medible.

Veamos la no negatividad de la integral. Consideramos el operador $Uf = f \circ \tau$. Entonces es positivo, ya que si $f \geq 0$ necesariamente $Uf = f \circ \tau \geq 0$; lineal, ya que para cualesquiera f, g integrables y α, β escalares se tiene

$$U(\alpha f + \beta g) = (\alpha f + \beta g) \circ \tau = (\alpha f) \circ \tau + (\beta g) \circ \tau = \alpha Uf + \beta Ug;$$

y $\|U\| = 1$, ya que para cualquier función integrable f se tiene

$$\|Uf\|_1 = \int_{\mathbb{X}} |Uf| d\mu = \int_{\mathbb{X}} |f \circ \tau| d\mu = \int_{\mathbb{X}} |f| \circ \tau d\mu = \int_{\mathbb{X}} |f| d\mu = \|f\|_1,$$

conforme al cambio de variable de la observación 4. Ahora podemos aplicar el lema anterior, 1.50, a dicho operador de manera que

$$\int_{\{x: F_N(x) > 0\}} f d\mu \geq 0,$$

con $F_N = \max_{0 \leq n \leq N} f_n$, donde $f_0 = 0$ y $f_n = f + f \circ \tau + \dots + f \circ \tau^{n-1} = \sum_{i=0}^{n-1} f \circ \tau^i$ para $n \geq 1$.

Como $\max_{1 \leq n \leq N} f_n(x) = \max_{0 \leq n \leq N} f_n(x) = F_N(x)$ cuando $F_N(x) > 0$ resulta que

$$\begin{aligned} \bigcup_{N=1}^{+\infty} \{x : F_N(x) > 0\} &= \left\{ x : \lim_{N \rightarrow +\infty} F_N(x) > 0 \right\} \\ &= \left\{ x : \lim_{N \rightarrow +\infty} \max_{1 \leq n \leq N} \sum_{i=0}^{n-1} f(\tau^i(x)) > 0 \right\} \\ &= \left\{ x : \sup_{n \geq 1} \sum_{i=0}^{n-1} f(\tau^i(x)) > 0 \right\} = M. \end{aligned}$$

Por lo tanto,

$$\int_M f d\mu = \int_{\bigcup_{N=1}^{+\infty} \{x: F_N(x) > 0\}} f d\mu = \lim_{N \rightarrow +\infty} \int_{\{x: F_N(x) > 0\}} f d\mu \geq 0,$$

donde en la segunda igualdad hemos utilizado que la unión es creciente. ■

El siguiente lema es un refinamiento del teorema ergódico maximal. Su prueba puede encontrarse en [1, cap 3, lema 3.3.3], donde el apartado b) es de elaboración propia.

Lema 1.52. Sean $(\mathbb{X}, \mathcal{A}, \mu, \tau)$ un sistema dinámico, $f : \mathbb{X} \rightarrow \mathbb{R}$ integrable y $\alpha \in \mathbb{R}$. Si A es un conjunto medible τ -invariante tal que, para cada $x \in A$,

a)

$$\sup_{n \geq 1} A_n(f)(x) > \alpha, \text{ entonces } \int_A f d\mu \geq \alpha \mu(A);$$

b)

$$\inf_{n \geq 1} A_n(f)(x) < \alpha, \text{ entonces } \int_A f d\mu \leq \alpha \mu(A).$$

Demostración. En primer lugar, vamos a probar el apartado a) para el caso $\alpha = 0$. Definimos $g = f \cdot \chi_A$ y el conjunto

$$M = \left\{ x : \sup_{n \geq 1} \frac{1}{n} \sum_{i=0}^{n-1} g(\tau^i(x)) > 0 \right\}.$$

Por el teorema maximal, 1.51, $\int_M g d\mu \geq 0$. Por otro lado, la función característica de A es τ -invariante al serlo el conjunto A , es decir, como $\tau^{-1}(A) = A$, se cumple $\chi_A = \chi_A \circ \tau$. Por tanto, aplicando τ sucesivas veces, $g \circ \tau^i = (f \circ \tau^i) \cdot (\chi_A \circ \tau^i) = (f \circ \tau^i) \cdot \chi_A$, esto es,

$$g(\tau^i(x)) = \begin{cases} f(\tau^i(x)), & x \in A, \\ 0, & x \in \mathbb{X} \setminus A. \end{cases}$$

Por lo tanto,

$$\frac{1}{n} \sum_{i=0}^{n-1} g(\tau^i(x)) = \begin{cases} A_n(f)(x), & x \in A, \\ 0, & x \in \mathbb{X} \setminus A. \end{cases}$$

Por la definición de M y la hipótesis sobre los puntos de A , $M = A$, obteniéndose

$$0 \leq \int_M g d\mu = \int_A g d\mu = \int_A f \cdot \chi_A d\mu = \int_A f d\mu,$$

quedando probado así el apartado a) para $\alpha = 0$.

Para el caso α distinto de cero, podemos transformar las condiciones iniciales en

$$\sup_{n \geq 1} \frac{1}{n} \sum_{i=0}^{n-1} (f(\tau^i(x)) - \alpha) = \sup_{n \geq 1} (A_n(f)(x) - \alpha) > 0$$

y aplicando el caso $\alpha = 0$ sobre la función $f - \alpha$,

$$0 \leq \int_A (f - \alpha) d\mu = \int_A f d\mu - \alpha \mu(A),$$

por tanto

$$\int_A f d\mu \geq \alpha \mu(A).$$

Para el apartado b) empleamos que $\inf_{n \geq 1} x_n = -\sup_{n \geq 1} -x_n$. Así

$$\begin{aligned} \inf_{n \geq 1} A_n(f)(x) < \alpha &\iff -\sup_{n \geq 1} -A_n(f)(x) < \alpha \iff -\sup_{n \geq 1} A_n(-f)(x) < \alpha \\ &\iff \sup_{n \geq 1} A_n(-f)(x) > -\alpha. \end{aligned}$$

Aplicando el apartado a),

$$\int_A -f d\mu \geq -\alpha \mu(A) \iff \int_A f d\mu \leq \alpha \mu(A).$$

■

Ya nos encontramos en disposición de probar el teorema ergódico de Birkhoff 1.48. La prueba aquí presente es una variación de la que puede encontrarse en [12, cap 2, teo 2.3].

Demostración. (Prueba del teorema 1.48). En primer lugar, para cada $\alpha, \beta \in \mathbb{R}$ con $\alpha < \beta$ definimos los conjuntos

$$E_{\alpha, \beta} = \left\{ x \in \mathbb{X} : \liminf_{n \rightarrow +\infty} A_n(f)(x) < \alpha < \beta < \limsup_{n \rightarrow +\infty} A_n(f)(x) \right\}.$$

Vamos a probar que $\mu(E_{\alpha, \beta}) = 0$ para cada α y β . Así, la unión sobre cualquier par de racionales, $\alpha < \beta$, también tendrá medida nula de donde se deducirá que $\liminf_{n \rightarrow +\infty} A_n(f)(x) = \limsup_{n \rightarrow +\infty} A_n(f)(x)$ $\mu - c.t.p.$ y por tanto existe el límite $f^*(x) = \lim_{n \rightarrow +\infty} A_n(f)(x)$.

Para comenzar, vamos a ver que $\liminf_{n \rightarrow +\infty} A_n(f)(x)$ y $\limsup_{n \rightarrow +\infty} A_n(f)(x)$ son funciones τ -invariantes. Comenzamos con el límite inferior:

$$\begin{aligned} \liminf_{n \rightarrow +\infty} A_n(f)(\tau(x)) &= \liminf_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(\tau^{k+1}(x)) = \liminf_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=1}^n f(\tau^i(x)) \\ &= \liminf_{n \rightarrow +\infty} -\frac{1}{n} f(x) + \left(\frac{n+1}{n} \right) \frac{1}{n+1} \sum_{i=0}^n f(\tau^i(x)) \\ &= \liminf_{n \rightarrow +\infty} A_{n+1}(f)(x) = \liminf_{n \rightarrow +\infty} A_n(f)(x). \end{aligned}$$

El límite superior es análogo, sustituyendo \liminf por \limsup en la cuenta anterior. De aquí podemos deducir

$$\begin{aligned} \tau^{-1}(E_{\alpha, \beta}) &= \{x : \tau(x) \in E_{\alpha, \beta}\} \\ &= \left\{ x : \liminf_{n \rightarrow +\infty} A_n(f)(\tau(x)) < \alpha < \beta < \limsup_{n \rightarrow +\infty} A_n(f)(\tau(x)) \right\} \\ &= \left\{ x : \liminf_{n \rightarrow +\infty} A_n(f)(x) < \alpha < \beta < \limsup_{n \rightarrow +\infty} A_n(f)(x) \right\} = E_{\alpha, \beta}, \end{aligned}$$

por consiguiente $E_{\alpha, \beta}$ es τ -invariante. Además, tanto el límite superior como inferior son funciones medibles, al serlo f , por lo tanto el conjunto $E_{\alpha, \beta}$ es medible.

Por otro lado,

$$E_{\alpha, \beta} \subset \left\{ x : \inf_{n \geq 1} A_n(f)(x) < \alpha < \beta < \sup_{n \geq 1} A_n(f)(x) \right\},$$

por lo que estamos en las condiciones para aplicar el lema 1.52, obteniéndose

$$\int_{E_{\alpha, \beta}} f d\mu \leq \alpha \mu(E_{\alpha, \beta}) \leq \beta \mu(E_{\alpha, \beta}) \leq \int_{E_{\alpha, \beta}} f d\mu.$$

Como $\alpha < \beta$, necesariamente $\mu(E_{\alpha,\beta}) = 0$. Así, $\mu(E_{\alpha,\beta}) = 0$ para cada $\alpha < \beta$. De esta forma, llamando E al conjunto

$$\left\{ x : \liminf_{n \rightarrow +\infty} A_n(f)(x) < \limsup_{n \rightarrow +\infty} A_n(f)(x) \right\},$$

resulta

$$E = \bigcup_{\alpha < \beta \in \mathbb{Q}} \left\{ x : \liminf_{n \rightarrow +\infty} A_n(f)(x) < \alpha < \beta < \limsup_{n \rightarrow +\infty} A_n(f)(x) \right\}$$

y $\mu(E) = 0$. Así, $\liminf_{n \rightarrow +\infty} A_n(f)(x) = \limsup_{n \rightarrow +\infty} A_n(f)(x)$ $\mu - c.t.p.$ y por tanto existe el límite $f^*(x) = \lim_{n \rightarrow +\infty} A_n(f)(x)$.

Por otro lado, claramente $f^* \circ \tau = f^*$ $\mu - c.t.p.$, ya que el límite inferior y superior son τ -invariantes e iguales a f^* .

Veamos ahora que $f^* \in \mathcal{L}^1(\mathbb{X})$ y $\|f^*\|_1 \leq \|f\|_1$. Es claro que $|A_n(f)| \leq A_n(|f|)$, $n = 1, 2, \dots$, ya que

$$\left| \frac{1}{n} \sum_{i=0}^{n-1} f(\tau^i) \right| \leq \frac{1}{n} \sum_{i=0}^{n-1} |f(\tau^i)| = \frac{1}{n} \sum_{i=0}^{n-1} |f|(\tau^i),$$

por consiguiente

$$\int_{\mathbb{X}} |f^*| d\mu = \int_{\mathbb{X}} \left| \lim_{n \rightarrow +\infty} A_n(f)(x) \right| d\mu \leq \int_{\mathbb{X}} \lim_{n \rightarrow +\infty} A_n(|f|)(x) d\mu.$$

Por el lema de Fatou,

$$\int_{\mathbb{X}} \lim_{n \rightarrow +\infty} A_n(|f|)(x) d\mu \leq \liminf_{n \rightarrow +\infty} \int_{\mathbb{X}} A_n(|f|)(x) d\mu. \quad (1.4)$$

Además, $\|A_n(|f|)\|_1 = \| |f| \|_1$, ya que

$$\int_{\mathbb{X}} \frac{1}{n} \sum_{i=0}^{n-1} |f \circ \tau^i| d\mu = \frac{1}{n} \sum_{i=0}^{n-1} \int_{\mathbb{X}} |f \circ \tau^i| d\mu = \frac{1}{n} \sum_{i=0}^{n-1} \int_{\mathbb{X}} |f| \circ \tau^i d\mu$$

y aplicando el cambio de variable de la observación 4 en cada integral, se obtienen n integrales iguales, resultando

$$\frac{1}{n} \sum_{i=0}^{n-1} \int_{\mathbb{X}} |f| \circ \tau^i d\mu = \int_{\mathbb{X}} |f| d\mu.$$

Sustituyendo en 1.4 se tiene

$$\int_{\mathbb{X}} |f^*| d\mu \leq \int_{\mathbb{X}} |f| d\mu$$

y como $f \in \mathcal{L}^1(\mathbb{X})$ entonces $f^* \in \mathcal{L}^1(\mathbb{X})$ y $\|f^*\|_1 \leq \|f\|_1$.

Por último, tenemos que ver que $\int_{\mathbb{X}} f^* d\mu = \int_{\mathbb{X}} f d\mu$. No es restrictivo suponer, redefiniendo si es necesario f^* en un conjunto de medida nula, que solo toma valores finitos, es decir, $f^* : \mathbb{X} \rightarrow \mathbb{R}$.

Para ello vamos a recuperar los conjuntos $X(n, k)$ del teorema 1.47. Así pues, para cada $k \in \mathbb{Z}$, $n = 0, 1, 2, \dots$, definimos

$$X(n, k) = \left\{ x \in \mathbb{X} : \frac{k}{2^n} \leq f^*(x) < \frac{k+1}{2^n} \right\}.$$

Como se vio en 1.47 estos son conjuntos casi τ -invariantes al serlo la función f^* , y para cada n fijo, $\mathbb{X} = \bigcup_{k \in \mathbb{Z}} X(n, k)$.

Fijamos $\varepsilon > 0$ tal que $\sup_{n \geq 1} A_n(f) > \frac{k}{2^n} - \varepsilon$ en $X(n, k)$. Así, por el lema 1.52, resulta

$$\int_{X(n, k)} f d\mu \geq \left(\frac{k}{2^n} - \varepsilon \right) \mu(X(n, k)).$$

De igual forma, $\inf_{n \geq 1} A_n(f) < \left(\frac{k+1}{2^n} \right)$ en $X(n, k)$, y por el lema 1.52 se obtiene

$$\int_{X(n, k)} f d\mu \leq \left(\frac{k+1}{2^n} \right) \mu(X(n, k)).$$

Entonces

$$\left(\frac{k}{2^n} - \varepsilon \right) \mu(X(n, k)) \leq \int_{X(n, k)} f d\mu \leq \left(\frac{k+1}{2^n} \right) \mu(X(n, k))$$

y tomando el límite cuando ε tiende a 0

$$\left(\frac{k}{2^n} \right) \mu(X(n, k)) \leq \int_{X(n, k)} f d\mu \leq \left(\frac{k+1}{2^n} \right) \mu(X(n, k)).$$

Por otro lado, por la definición de $X(n, k)$, se tiene

$$\left(\frac{k}{2^n} \right) \mu(X(n, k)) \leq \int_{X(n, k)} f^* d\mu \leq \left(\frac{k+1}{2^n} \right) \mu(X(n, k))$$

y restando ambas expresiones y tomando valores absolutos

$$\left| \int_{X(n, k)} f d\mu - \int_{X(n, k)} f^* d\mu \right| \leq \frac{1}{2^n} \mu(X(n, k)).$$

Sumando sobre k en la expresión anterior, como $\mathbb{X} = \bigcup_{k \in \mathbb{Z}} X(n, k)$, se obtiene

$$\left| \int_{\mathbb{X}} f d\mu - \int_{\mathbb{X}} f^* d\mu \right| \leq \frac{1}{2^n} \mu(\mathbb{X})$$

y tomando límites cuando n tiende a infinito se obtiene

$$\int_{\mathbb{X}} f d\mu = \int_{\mathbb{X}} f^* d\mu.$$

■

El siguiente corolario es el principal resultado que vamos a emplear en este trabajo, ya que lo vamos a aplicar en las transformaciones de Markov que serán estudiadas más adelante, para

aclarar la dinámica en el sistema dinámico en el que se trabajará. Este corolario nos asegura que si tenemos una transformación ergódica τ , la frecuencia relativa con la que los elementos $\{\tau^n(x)\}_{n=0}^{+\infty}$, con $x \in \mathbb{X}$, caen en E , es igual a la medida de E , $\mu(E)$. La demostración es de elaboración propia y el resultado puede encontrarse en [1, cap 3, cor 3.3.1].

Corolario 1.53. *Sea $(\mathbb{X}, \mathcal{A}, \mu, \tau)$ un sistema dinámico. Si τ es ergódica entonces*

$$f^* = \int_{\mathbb{X}} f d\mu, \quad \mu - c.t.p.$$

Más aún, si $E \in \mathcal{A}$ entonces

$$\frac{1}{n} \sum_{i=0}^{n-1} \chi_E(\tau^i(x)) \longrightarrow \mu(E), \quad \mu - c.t.p.$$

Demostración. Por el teorema ergódico de Birkhoff 1.48, $f^* \circ \tau = f^*$ $\mu - c.t.p.$ Como τ es ergódica, el teorema 1.47 garantiza que f^* es constante $\mu - c.t.p.$

Por otro lado, el teorema ergódico 1.48 garantiza que $\int_{\mathbb{X}} f^* d\mu = \int_{\mathbb{X}} f d\mu$. Como f^* es constante, sustituyendo en la expresión anterior se tiene

$$\int_{\mathbb{X}} f d\mu = \int_{\mathbb{X}} f^* d\mu = f^* \int_{\mathbb{X}} d\mu = f^*.$$

Si $f = \chi_E$, sustituyendo,

$$\chi_E^* = \int_{\mathbb{X}} \chi_E d\mu = \mu(E).$$

Por lo tanto, por el teorema ergódico 1.48 se obtiene lo que buscábamos:

$$\frac{1}{n} \sum_{i=0}^{n-1} \chi_E(\tau^i(x)) \longrightarrow \mu(E), \quad \mu - c.t.p.$$

■

El operador de Perron-Frobenius

En este capítulo se van a estudiar en profundidad el operador de Perron-Frobenius y sus características, que serán necesarias para poder llevar a cabo el correcto desarrollo de los diferentes resultados del trabajo. Como ya hemos comentado anteriormente, este operador va a ser una de las principales herramientas del trabajo y puede verse como una extensión natural de los teoremas de Perron y Perron-Frobenius que estudiamos en el trabajo fin de grado [10], aunque será en el capítulo cuatro donde veremos más clara esta relación, a través de la representación matricial del operador.

Entre las ventajas más importantes de este operador destacaremos que permite probar la existencia y unicidad de medidas invariantes absolutamente continuas y establecer diferentes propiedades sobre ellas y sus densidades. En el último capítulo las veremos con detalle.

Este operador ya fue introducido informalmente, en el marco de los espacios de medida abstractos, ahora concretamos al caso $(I = [a, b], \mathcal{B}, \lambda)$, que es el que realmente nos interesa.

Supongamos que la transformación medible $\tau : I \rightarrow I$ es no singular, es decir, $\lambda(A) = 0$ implica $\lambda(\tau^{-1}(A)) = 0$ para todo $A \in \mathcal{B}$. Sea $f \in \mathcal{L}^1(I)$. Entonces (recuérdese la observación 1)

$$\mu_f(A) = \int_A f d\lambda$$

es una medida real y $\mu_f \ll \lambda$. Por la proposición 1.14, $\tau_*\mu_f \ll \lambda$, donde la medida real $\tau_*\mu_f$ está dada por $\tau_*\mu_f(A) = \mu_f(\tau^{-1}(A))$. Por el teorema de Radon-Nikodym 1.21, existe una única función $P_\tau f \in \mathcal{L}^1(I)$ tal que

$$\int_{\tau^{-1}(A)} f d\lambda = \mu_f(\tau^{-1}(A)) = \tau_*\mu_f(A) = \int_A P_\tau f d\lambda$$

para todo $A \in \mathcal{B}$. Este es el operador de Perron-Frobenius:

Definición 2.1. Sea $\tau : I \rightarrow I$ una transformación no singular en $(I, \mathcal{B}, \lambda)$. Se define el *operador de Perron-Frobenius* $P_\tau : \mathcal{L}^1(I) \rightarrow \mathcal{L}^1(I)$ como sigue: para cualquier $f \in \mathcal{L}^1(I)$, $P_\tau f$ es la única función de $\mathcal{L}^1(I)$ que cumple

$$\int_A P_\tau f d\lambda = \int_{\tau^{-1}(A)} f d\lambda$$

para cualquier $A \in \mathcal{B}$.

Notemos que si $f \in \mathcal{D}(I)$ entonces μ_f y $\tau_*\mu_f$ son medidas de probabilidad, por lo que también $P_\tau f \in \mathcal{D}(I)$. Por tanto, según nos convenga, veremos P_τ como $P_\tau : \mathcal{L}^1(I) \rightarrow \mathcal{L}^1(I)$ o $P_\tau : \mathcal{D}(I) \rightarrow \mathcal{D}(I)$.

2.1. Propiedades del operador de Perron-Frobenius

En esta sección se presentan diferentes propiedades del operador de Perron-Frobenius, que serán necesarias para probar una buena parte de los resultados de los siguientes capítulos, ya que el operador estará prácticamente involucrado en todos ellos. Pueden encontrarse en [1, cap 4].

Proposición 2.2 (Linealidad). $P_\tau : \mathcal{L}^1(I) \rightarrow \mathcal{L}^1(I)$ es un operador lineal.

Demostración. Sea $A \subset I$ un conjunto de Borel y sean $\alpha, \beta \in \mathbb{R}$. Entonces, si $f, g \in \mathcal{L}^1(I)$, se tiene

$$\begin{aligned} \int_A P_\tau(\alpha f + \beta g)d\lambda &= \int_{\tau^{-1}(A)} (\alpha f + \beta g)d\lambda = \alpha \int_{\tau^{-1}(A)} f d\lambda + \beta \int_{\tau^{-1}(A)} g d\lambda \\ &= \alpha \int_A P_\tau f d\lambda + \beta \int_A P_\tau g d\lambda = \int_A (\alpha P_\tau f + \beta P_\tau g)d\lambda. \end{aligned}$$

Como esto es cierto para cualquier conjunto $A \in \mathcal{B}$, se obtiene

$$P_\tau(\alpha f + \beta g) = \alpha P_\tau f + \beta P_\tau g. \quad \blacksquare$$

Seguimos con la positividad del operador.

Proposición 2.3 (Positividad). Sea $f \in \mathcal{L}^1(I)$ con $f \geq 0$. Entonces $P_\tau f \geq 0$.

Demostración. Sea $A = \{x \in I : (P_\tau f)(x) < 0\}$. Entonces

$$0 \geq \int_A P_\tau f d\lambda = \int_{\tau^{-1}(A)} f d\lambda \geq 0 \implies \int_A P_\tau f d\lambda = 0,$$

pero como $P_\tau f$ es negativo en A , necesariamente $\lambda(A) = 0$, por lo tanto $P_\tau f \geq 0$. \blacksquare

Proposición 2.4 (Preservación de integrales). Sea $f \in \mathcal{L}^1(I)$. Entonces

$$\int_I P_\tau f d\lambda = \int_I f d\lambda.$$

Demostración. Por definición

$$\int_I P_\tau f d\lambda = \int_{\tau^{-1}(I)} f d\lambda.$$

Como $\tau^{-1}(I) = I$, se tiene

$$\int_I P_\tau f d\lambda = \int_{\tau^{-1}(I)} f d\lambda = \int_I f d\lambda.$$

El operador de Perron-Frobenius es asimismo una aplicación contractiva: ■

Proposición 2.5 (Contractiva). $P_\tau : \mathcal{L}^1(I) \longrightarrow \mathcal{L}^1(I)$ es una contracción, es decir, $\|P_\tau f\|_1 \leq \|f\|_1$ para toda $f \in \mathcal{L}^1(I)$, dándose la igualdad cuando $f \geq 0$.

Demostración. Dada $f \in \mathcal{L}^1(I)$, sean f^+ y f^- su parte positiva y negativa, respectivamente. Entonces $f^+, f^- \in \mathcal{L}^1(I)$ y $f^+, f^- \geq 0$, además $f = f^+ - f^-$ y $|f| = f^+ + f^-$. Por la linealidad de $P_\tau f$, proposición 2.2, se cumple

$$P_\tau f = P_\tau(f^+ - f^-) = P_\tau f^+ - P_\tau f^-.$$

Por lo tanto, tomando valor absoluto y usando las proposiciones 2.2 y 2.3,

$$|P_\tau f| = |P_\tau f^+ - P_\tau f^-| \leq |P_\tau f^+| + |P_\tau f^-| = P_\tau f^+ + P_\tau f^- = P_\tau(f^+ + f^-) = P_\tau|f|.$$

Tomando normas y aplicando la proposición 2.4,

$$\|P_\tau f\|_1 = \int_I |P_\tau f| d\lambda \leq \int_I P_\tau|f| d\lambda = \int_I |f| d\lambda = \|f\|_1.$$

Por lo tanto P_τ es contractiva.

Por otro lado, si $f \geq 0$, por 2.3, $P_\tau f \geq 0$ y aplicando la proposición 2.4,

$$\|P_\tau f\|_1 = \int_I |P_\tau f| d\lambda = \int_I P_\tau f d\lambda = \int_I f d\lambda = \int_I |f| d\lambda = \|f\|_1,$$

obteniéndose la igualdad. ■

De este resultado se puede deducir que $P_\tau : \mathcal{L}^1(I) \longrightarrow \mathcal{L}^1(I)$ es continuo respecto a la norma \mathcal{L}^1 , ya que

$$\|P_\tau f - P_\tau g\|_1 \leq \|f - g\|_1.$$

La siguiente propiedad es importante por su gran utilidad: establece el valor del operador para la composición de transformaciones.

Proposición 2.6. Dadas $\tau : I \longrightarrow I$ y $\sigma : I \longrightarrow I$ no singulares, entonces $P_{\tau \circ \sigma} f = (P_\tau \circ P_\sigma) f$. En particular, $P_{\tau^n} f = P_\tau^n f$.

Demostración. Sea $f \in \mathcal{L}^1(I)$. Como τ y σ son no singulares, $\lambda(A) = 0$ implica $\lambda(\tau^{-1}(A)) = 0$ y $\lambda((\tau \circ \sigma)^{-1}(A)) = \lambda(\sigma^{-1}(\tau^{-1}(A))) = 0$, es decir, $\tau \circ \sigma$ es no singular y podemos calcular el operador de Perron-Frobenius asociado, $P_{\tau \circ \sigma}$, teniéndose

$$\int_A P_{\tau \circ \sigma} f d\lambda = \int_{(\tau \circ \sigma)^{-1}(A)} f d\lambda.$$

Por otro lado, usando la definición del operador de Perron-Frobenius, se tiene la siguiente igualdad:

$$\int_A P_\tau(P_\sigma f) d\lambda = \int_{\tau^{-1}(A)} P_\sigma f d\lambda = \int_{\sigma^{-1}(\tau^{-1}(A))} f d\lambda.$$

Igualando ambas expresiones se obtiene

$$\int_A P_{\tau \circ \sigma} f d\lambda = \int_{(\tau \circ \sigma)^{-1}(A)} f d\lambda = \int_A P_\tau(P_\sigma f) d\lambda,$$

por tanto $P_{\tau \circ \sigma} f = (P_\tau \circ P_\sigma) f$. Procediendo por inducción, claramente se obtiene que $P_{\tau^n} f = P_\tau^n f$. ■

En este punto debemos hacernos la pregunta: ¿cuándo una función f será un punto fijo del operador P_τ ? Pues bien, esto ocurrirá si y solo si la medida real definida por f es τ -invariante, en particular, cuando f sea una densidad obtenemos una medida de probabilidad μ τ -invariante absolutamente continua respecto a la medida de Lebesgue λ (en este contexto, diremos que f es *la densidad de la medida* μ). Este resultado es de vital importancia para probar el principal resultado de este trabajo. Lo concretamos en la siguiente proposición que podemos encontrar en [1, cap 4].

Proposición 2.7. Sean $\tau : I \rightarrow I$ no singular y $f^* \in \mathcal{L}^1(I)$. Entonces $P_\tau f^* = f^*$ si y solo si $\mu(A) = \mu_{f^*}(A) = \int_A f^* d\lambda$ es τ -invariante, es decir, si y solo si $\mu(\tau^{-1}(A)) = \mu(A)$ para todo conjunto $A \in \mathcal{B}$.

Demostración. Supongamos que $\mu(\tau^{-1}(A)) = \mu(A)$ para cualquier conjunto medible $A \in \mathcal{B}$. Entonces por la definición de μ

$$\int_{\tau^{-1}(A)} f^* d\lambda = \mu(\tau^{-1}(A)) = \mu(A) = \int_A f^* d\lambda$$

y por tanto

$$\int_A P_\tau f^* d\lambda = \int_{\tau^{-1}(A)} f^* d\lambda = \int_A f^* d\lambda.$$

Como $A \in \mathcal{B}$ es arbitrario, se garantiza $P_\tau f^* = f^*$.

Veamos la otra implicación.

Supongamos que $P_\tau f^* = f^*$. Entonces

$$\int_A P_\tau f^* d\lambda = \int_A f^* d\lambda = \mu(A).$$

Por definición

$$\int_A P_\tau f^* d\lambda = \int_{\tau^{-1}(A)} f^* d\lambda = \mu(\tau^{-1}(A)),$$

combinando ambas expresiones, se obtiene que $\mu(A) = \mu(\tau^{-1}(A))$ para todo A medible. ■

2.2. Representación del operador de Perron-Frobenius

En esta sección del trabajo se muestra una representación del operador de Perron-Frobenius para la clase de funciones monótonas a trozos del intervalo en sí mismo. Esta familia incluye muchas transformaciones de interés en el análisis y modelado de dinámicas unidimensionales. En particular, nosotros trabajaremos con transformaciones expansivas y de Markov (estas últimas las presentamos en el capítulo final).

Comenzamos con la definición de dichas transformaciones monótonas a trozos. Puede verse en [1, cap 4].

Definición 2.8. Dado $I = [a, b]$, la transformación $\tau : I \rightarrow I$ se dice *monótona a trozos* si existe una partición (minimal) de I , $a = a_0 < a_1 < \dots < a_q = b$ (su *partición asociada*) tal que

- a) $\tau|_{(a_{i-1}, a_i)} = \tau_i$ es una función que puede extenderse a una función de clase C^1 en $[a_{i-1}, a_i]$, $i = 1, \dots, q$;
- b) $|\tau'(x)| > 0$ en (a_{i-1}, a_i) , $i = 1, \dots, q$.

Si junto con b) se tiene $|\tau'(x)| \geq \alpha > 1$ en todos los lugares en los que existe la derivada, τ se llama *expansiva a trozos*. Por b) τ es estrictamente monótona en cada (a_{i-1}, a_i) .

Un ejemplo de una transformación monótona a trozos puede verse en la figura 2.1.

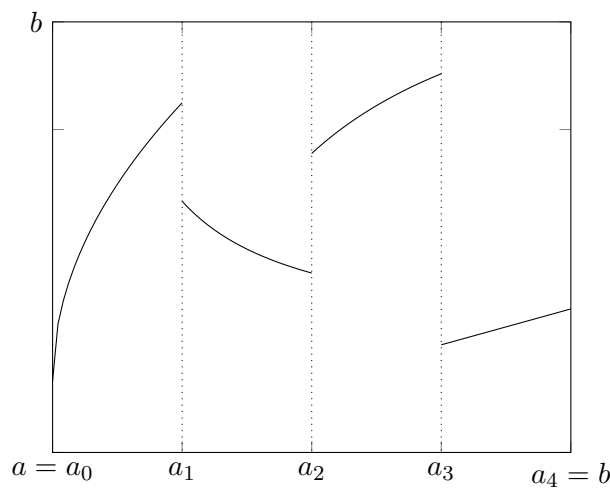


Figura 2.1: Ejemplo de transformación monótona a trozos.

Veamos ahora cuál es la forma del operador de Perron-Frobenius para este tipo de transformaciones. La descripción se puede encontrar en [1, cap 4]. Sea τ una transformación monótona a trozos. Por la definición de P_τ se tiene

$$\int_A P_\tau f d\lambda = \int_{\tau^{-1}(A)} f d\lambda \quad (2.1)$$

para cualquier conjunto de Borel $A \subset I$.

Como τ es monótona en cada $J_i = (a_{i-1}, a_i)$, $i = 1, \dots, q$, se puede definir la inversa de $\tau_i = \tau|_{J_i}$ en cada uno de los intervalos J_i . Sean

$$\phi_i = \tau_{|J_i}^{-1} = \tau_i^{-1},$$

con $B_i = \tau(J_i)$, entonces $\phi_i : B_i \rightarrow J_i$ y los conjuntos $\tau^{-1}(A)$ y $\cup_{i=1}^q \phi_i(B_i \cap A)$ se diferencian en a lo sumo un número finito de puntos, donde los conjuntos $\{\phi_i(B_i \cap A)\}_{i=1}^q$ son mutuamente disjuntos y dependiendo de A pueden ser vacíos.

Sustituyendo lo anterior en la ecuación 2.1 se obtiene

$$\begin{aligned} \int_A P_\tau f d\lambda &= \int_{\tau^{-1}(A)} f d\lambda = \int_{\cup_{i=1}^q \phi_i(B_i \cap A)} f d\lambda = \sum_{i=1}^q \int_{\phi_i(B_i \cap A)} f d\lambda \\ &= \sum_{i=1}^q \int_{B_i \cap A} f(\phi_i(x)) |\phi_i'(x)| d\lambda \end{aligned}$$

usando un cambio de variable en cada i . Entonces

$$\begin{aligned} \int_A P_\tau f d\lambda &= \sum_{i=1}^q \int_A f(\phi_i(x)) |\phi_i'(x)| \chi_{B_i}(x) d\lambda \\ &= \int_A \sum_{i=1}^q \frac{f(\tau_i^{-1}(x))}{|\tau'(\tau_i^{-1}(x))|} \chi_{\tau(J_i)}(x) d\lambda. \end{aligned}$$

Al ser A arbitrario se tiene que

$$P_\tau f(x) = \sum_{i=1}^q \frac{f(\tau_i^{-1}(x))}{|\tau'(\tau_i^{-1}(x))|} \chi_{\tau(J_i)}(x) \quad (2.2)$$

y escribiéndolo de una forma más compacta

$$P_\tau f(x) = \sum_{z \in \{\tau^{-1}(x)\}} \frac{f(z)}{|\tau'(z)|},$$

donde para cada x el conjunto $\{\tau^{-1}(x)\}$ consiste en como mucho q puntos. Y se entiende que $P_\tau f(x) = 0$ si $x \notin \cup_{i=1}^q \tau(J_i)$.

Veamos dos ejemplos de cálculo del operador de Perron-Frobenius siguiendo la descomposición anterior. Ambos son ejercicios propuestos en el libro [1, cap 4].

Ejemplo 1. Sea $\tau : [0, 1] \rightarrow [0, 1]$ definida por

$$\tau(x) = \begin{cases} 2x, & 0 \leq x \leq \frac{1}{2}, \\ -x + \frac{3}{2}, & \frac{1}{2} \leq x \leq 1. \end{cases}$$

Vamos a encontrar P_τ y mostrar que si

$$\tilde{f}(x) = \begin{cases} 0, & 0 \leq x < \frac{1}{2}, \\ g(x), & \frac{1}{2} \leq x \leq 1, \end{cases}$$

donde g es simétrica con respecto a la recta $x = \frac{3}{4}$ (es decir, $g(\frac{3}{2} - x) = g(x)$ para todo $x \in [\frac{1}{2}, 1]$), entonces $P_\tau \tilde{f} = \tilde{f}$.

En primer lugar, vamos a representar gráficamente la función τ , véase la figura 2.2.

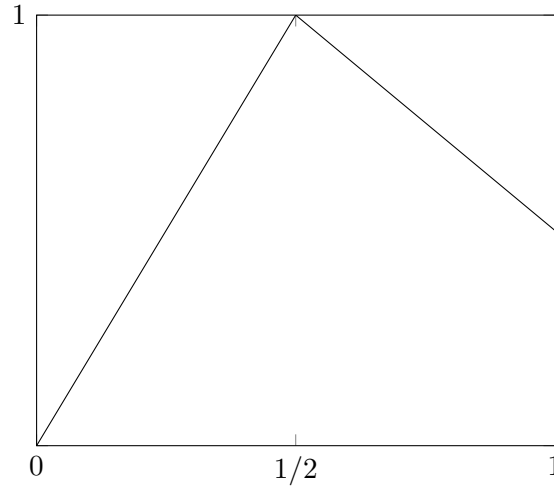


Figura 2.2: Gráfica de τ .

Atendiendo a la forma de la función, podemos observar dos componentes $\tau_1(x) = 2x$ si $x \in (0, \frac{1}{2})$ y $\tau_2(x) = -x + \frac{3}{2}$ si $x \in (\frac{1}{2}, 1)$, de forma que estamos trabajando en la partición $\mathcal{P} = \{0, \frac{1}{2}, 1\}$. Cuyas derivadas son $\tau_1'(x) = 2$ si $x \in (0, \frac{1}{2})$ y $\tau_2'(x) = -1$ si $x \in (\frac{1}{2}, 1)$.

Necesitamos saber el valor de sus inversas, calculémoslas. Para eso despejamos x en cada una de las expresiones anteriores, obteniendo:

$$\phi_1(y) = \tau_1^{-1}(y) = \frac{y}{2}, \quad \phi_2(y) = \tau_2^{-1}(y) = \frac{3}{2} - y.$$

Además, se tiene que $\tau_1((0, \frac{1}{2})) = (0, 1)$ y $\tau_2((\frac{1}{2}, 1)) = (\frac{1}{2}, 1)$, con lo que ya tenemos todos los ingredientes necesarios para calcular P_τ , usando la representación 2.2. Así, dada f se tiene:

$$P_\tau f(x) = \frac{1}{2} f\left(\frac{x}{2}\right) \chi_{(0,1)} + f\left(\frac{3}{2} - x\right) \chi_{(\frac{1}{2},1)} = \frac{1}{2} f\left(\frac{x}{2}\right) + f\left(\frac{3}{2} - x\right) \chi_{(\frac{1}{2},1)}.$$

En el caso particular de \tilde{f} se tendrá

$$P_\tau \tilde{f}(x) = \begin{cases} 0, & 0 < x < \frac{1}{2}, \\ g\left(\frac{3}{2} - x\right), & \frac{1}{2} < x < 1. \end{cases}$$

Como g es simétrica, $g(\frac{3}{2} - x) = g(x)$ y $P_\tau \tilde{f} = \tilde{f}$.

Ejemplo 2. Sea $\tau : [0, 1] \rightarrow [0, 1]$ definida por

$$\tau(x) = \begin{cases} -2x + 1, & 0 \leq x \leq \frac{1}{2}, \\ 2x - 1, & \frac{1}{2} \leq x \leq 1. \end{cases}$$

Calculamos P_τ .

En primer lugar, vamos a representar gráficamente la función τ , véase la figura 2.3.

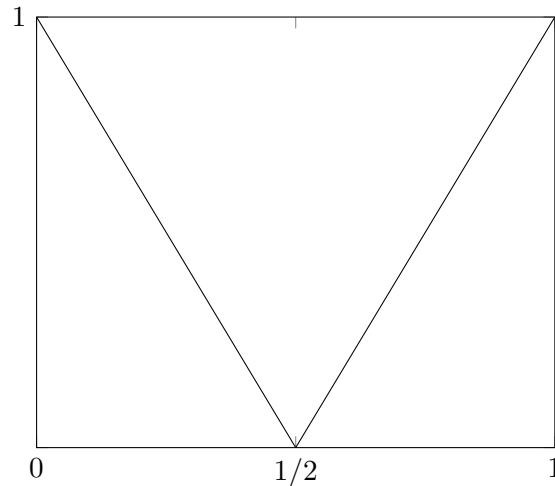


Figura 2.3: Gráfica de τ .

Atendiendo a la forma de la función, podemos observar dos componentes $\tau_1(x) = -2x + 1$ si $x \in (0, \frac{1}{2})$ y $\tau_2(x) = 2x - 1$ si $x \in (\frac{1}{2}, 1)$, de forma que estamos trabajando en la partición $\mathcal{P} = \{0, \frac{1}{2}, 1\}$, siendo las derivadas $\tau_1'(x) = -2$ si $x \in (0, \frac{1}{2})$ y $\tau_2'(x) = 2$ si $x \in (\frac{1}{2}, 1)$.

Despejando las respectivas inversas, obtenemos

$$\tau_1^{-1}(y) = \frac{1-y}{2}, \quad \tau_2^{-1}(y) = \frac{y+1}{2}.$$

Además, se tiene que $\tau_1((0, \frac{1}{2})) = (0, 1)$ y $\tau_2((\frac{1}{2}, 1)) = (0, 1)$, con lo que ya tenemos todos los ingredientes necesarios para calcular P_τ usando la representación 2.2. En concreto, dada f , se tiene

$$P_\tau f(x) = \frac{1}{2} f\left(\frac{1-x}{2}\right) + \frac{1}{2} f\left(\frac{x+1}{2}\right).$$

En particular, $\tilde{f} = 1$ es un punto fijo de P_τ , es decir, la propia medida de Lebesgue λ es τ -invariante.

Transformaciones expansivas a trozos

El objetivo de este capítulo es mostrar la existencia de medidas invariantes absolutamente continuas (respecto a la medida de Lebesgue), en ocasiones las denotaremos por *acim*, para transformaciones expansivas a trozos usando funciones de variación acotada. Además, probaremos que las funciones de densidad de dichas medidas serán de variación acotada.

Para llevar a cabo este desarrollo es necesario introducir algo de notación.

Seguimos trabajando en el intervalo $I = [a, b]$ y con la medida de Lebesgue λ . Denotamos por $\mathcal{T}(I)$ a la clase de las transformaciones $\tau : I \rightarrow I$ que cumplen las siguientes condiciones:

- τ es expansiva a trozos, es decir, existe una partición asociada $\mathcal{P} = \{J_i = (a_{i-1}, a_i), i = 1, \dots, q\}$ de I tal que $\tau_i = \tau|_{J_i}$ es de clase C^1 , y $|\tau'_i(x)| \geq \alpha > 1$ para cada $i = 1, \dots, q$ y para todo $x \in J_i$ y puede extenderse a una función de clase C^1 en $I_i = [a_{i-1}, a_i]$.
- $g(x) = \frac{1}{|\tau'(x)|}$ es una función de variación acotada. Aunque esto no es relevante, podemos suponer para que nuestras funciones estén definidas de un modo uniforme que $\tau(a_i) = a$ y $g(a_i) = 0$ en los puntos de la partición.

Para cada $n \geq 1$, definimos la partición $\mathcal{P}^{(n)}$ como

$$\mathcal{P}^{(n)} \equiv \bigvee_{k=0}^{n-1} \tau^{-k}(\mathcal{P}) \equiv \{J_{i_0} \cap \tau^{-1}(J_{i_1}) \cap \dots \cap \tau^{-n+1}(J_{i_{n-1}}) : J_{i_j} \in \mathcal{P}\}.$$

Claramente, τ^n es expansiva a trozos con partición asociada $\mathcal{P}^{(n)}$ si \mathcal{P} es la partición asociada a τ .

Veamos a continuación una serie de lemas que serán necesarios para probar que estas funciones tienen una *acim* cuya densidad es de variación acotada. Nuestro primer lema establece una desigualdad sobre la variación del operador de Perron-Frobenius de una función de variación acotada. Puede verse en [1, cap 5].

Lema 3.1. Sean $\tau \in \mathcal{T}(I)$ y $\delta = \min_{i=1, \dots, q} \lambda(J_i)$. Entonces para cada $f \in BV(I)$

$$V_I(P_\tau f) \leq AV_I f + B \int_I |f| d\lambda,$$

donde $A = \frac{3}{\alpha} + \max_{1 \leq i \leq q} V_{I_i} g$ y $B = \frac{2}{\alpha \delta} + \frac{1}{\delta} \max_{1 \leq i \leq q} V_{I_i} g$.

Demostración. Sea $a = x_0 < x_1 < \dots < x_r = b$ una partición de I independiente de la partición \mathcal{P} . Definimos $\phi_i = \tau_i^{-1}$ y vamos a usar la expresión 2.2. De esta forma

$$P_\tau f(x) = \sum_{i=1}^q \frac{f(\tau_i^{-1}(x))}{|\tau_i'(\tau_i^{-1}(x))|} \chi_{\tau(J_i)}(x) = \sum_{i=1}^q f(\phi_i(x))g(\phi_i(x))\chi_{\tau(J_i)}(x).$$

Aprovechando la expresión anterior podemos realizar el siguiente desarrollo sobre el sumatorio $\sum_{j=1}^r |P_\tau f(x_j) - P_\tau f(x_{j-1})|$:

$$\begin{aligned} & \sum_{j=1}^r |P_\tau f(x_j) - P_\tau f(x_{j-1})| \\ &= \sum_{j=1}^r \left| \sum_{i=1}^q f(\phi_i(x_j))g(\phi_i(x_j))\chi_{\tau(J_i)}(x_j) - \sum_{i=1}^q f(\phi_i(x_{j-1}))g(\phi_i(x_{j-1}))\chi_{\tau(J_i)}(x_{j-1}) \right| \\ &= \sum_{j=1}^r \sum_{i=1}^q |f(\phi_i(x_j))g(\phi_i(x_j))\chi_{\tau(J_i)}(x_j) - f(\phi_i(x_{j-1}))g(\phi_i(x_{j-1}))\chi_{\tau(J_i)}(x_{j-1})|, \end{aligned}$$

donde en el último paso hemos usado la desigualdad triangular.

Descartados los términos nulos, este doble sumatorio lo vamos a descomponer en dos partes

$$\sum_{(i,j) \in A} \quad \text{y} \quad \sum_{(i,j) \in B},$$

donde $A = \{(i, j) : \text{se cumple la condición 1}\}$ y $B = \{(i, j) : \text{se cumple la condición 2 o 3}\}$, y las condiciones son:

1. Cuando $\chi_{\tau(J_i)}(x_j) = \chi_{\tau(J_i)}(x_{j-1}) = 1$.
2. Cuando $\chi_{\tau(J_i)}(x_j) = 1$ y $\chi_{\tau(J_i)}(x_{j-1}) = 0$.
3. Cuando $\chi_{\tau(J_i)}(x_j) = 0$ y $\chi_{\tau(J_i)}(x_{j-1}) = 1$.

Comenzamos con la primera opción, $\chi_{\tau(J_i)}(x_j) = \chi_{\tau(J_i)}(x_{j-1}) = 1$, obteniendo el siguiente sumando:

$$\begin{aligned} & \sum_{(i,j) \in A} |f(\phi_i(x_j))g(\phi_i(x_j)) - f(\phi_i(x_{j-1}))g(\phi_i(x_{j-1}))| \\ &= \sum_{(i,j) \in A} |f(\phi_i(x_j))g(\phi_i(x_j)) - f(\phi_i(x_j))g(\phi_i(x_{j-1})) \\ & \quad + f(\phi_i(x_j))g(\phi_i(x_{j-1})) - f(\phi_i(x_{j-1}))g(\phi_i(x_{j-1}))| \\ &\leq \sum_{(i,j) \in A} |f(\phi_i(x_j))[g(\phi_i(x_j)) - g(\phi_i(x_{j-1}))]| + \sum_{(i,j) \in A} |g(\phi_i(x_{j-1})) [f(\phi_i(x_j)) - f(\phi_i(x_{j-1}))]| \\ &\leq \sum_{i=1}^q (\sup_{I_i} |f|) V_{I_i} g + (\sup_I |g|) \sum_{i=1}^q V_{I_i} f, \end{aligned}$$

esto último por pertenecer $\phi_i(x_j), \phi_i(x_{j-1})$ a $J_i \subset I_i$.

Usando la proposición 1.28, $|g| = g$, $\sup_I g \leq \frac{1}{\alpha}$, el valor de δ y la observación 3 al teorema 1.29, lo anterior estaría acotado por

$$\begin{aligned} \max_{1 \leq i \leq q} V_{I_i} g \sum_{i=1}^q \left(V_{I_i} f + \frac{1}{\lambda(I_i)} \int_{I_i} |f| d\lambda \right) + \sup_I g \sum_{i=1}^q V_{I_i} f \\ \leq \max_{1 \leq i \leq q} V_{I_i} g \left(V_I f + \frac{1}{\delta} \int_I |f| d\lambda \right) + \frac{1}{\alpha} V_I f. \end{aligned} \quad (3.1)$$

Pasamos ahora a estudiar los otros dos casos de manera conjunta. En primer lugar, notar que $\chi_{\tau(J_i)}(x_j) = 1$ y $\chi_{\tau(J_i)}(x_{j-1}) = 0$ ocurre solo si $x_j \in \tau(J_i)$ y $x_{j-1} \notin \tau(J_i)$, por lo tanto x_j y x_{j-1} están en lados opuestos de un extremo de $\tau(J_i)$. Lo mismo ocurre en el otro caso.

Por tanto, para cada J_i , se tiene como mucho un par x_j, x_{j-1} con $x_j \in \tau(J_i)$ y $x_{j-1} \notin \tau(J_i)$ y otro par $x_{j'}, x_{j'-1}$ con $x_{j'} \notin \tau(J_i)$ y $x_{j'-1} \in \tau(J_i)$. Así, el otro sumando quedaría acotado por:

$$\sum_{(i,j) \in B} (|f(\phi_i(x_j))g(\phi_i(x_j))| + |f(\phi_i(x_{j'-1}))g(\phi_i(x_{j'-1}))|) \leq \frac{1}{\alpha} \sum_{i=1}^q (|f(\phi_i(x_j))| + |f(\phi_i(x_{j'-1}))|), \quad (3.2)$$

debido a $\sup_I g \leq \frac{1}{\alpha}$, y entendiendo que los números $f(\phi_i(x_j))$ y $f(\phi_i(x_{j'-1}))$ pueden ser cero.

Si llamamos $s_i = \phi_i(x_j)$ y $r_i = \phi_i(x_{j'-1})$, que son dos puntos de $J_i \subset I_i$, 3.2 se transforma en

$$\frac{1}{\alpha} \sum_{i=1}^q (|f(s_i)| + |f(r_i)|) \leq \sum_{i=1}^q (2|f(v_i)| + |f(r_i) - f(v_i)| + |f(s_i) - f(v_i)|),$$

donde $v_i \in J_i$ es un punto tal que $|f(v_i)| \leq \frac{1}{\lambda(I_i)} \int_{I_i} |f| d\lambda$. De esta forma, 3.2 resulta acotada por

$$\frac{1}{\alpha} \sum_{i=1}^q \left(2V_{I_i} f + \frac{2}{\lambda(I_i)} \int_{I_i} |f| d\lambda \right) \leq \frac{2}{\alpha} V_I f + \frac{2}{\alpha\delta} \int_I |f| d\lambda. \quad (3.3)$$

Combinando las cotas 3.1 y 3.3 se obtiene

$$\sum_{j=1}^r |P_\tau f(x_j) - P_\tau f(x_{j-1})| \leq \left(\frac{3}{\alpha} + \max_{1 \leq i \leq q} V_{I_i} g \right) V_I f + \left(\frac{2}{\alpha\delta} + \frac{1}{\delta} \max_{1 \leq i \leq q} V_{I_i} g \right) \int_I |f| d\lambda.$$

Por la definición de $V_I(P_\tau f)$, tomando supremos en la expresión anterior, se obtiene el resultado buscado:

$$V_I(P_\tau f) \leq \left(\frac{3}{\alpha} + \max_{1 \leq i \leq q} V_{I_i} g \right) V_I f + \left(\frac{2}{\alpha\delta} + \frac{1}{\delta} \max_{1 \leq i \leq q} V_{I_i} g \right) \int_I |f| d\lambda. \quad \blacksquare$$

En el siguiente lema se usa una variación de la prueba realizada en [1, cap 5].

Lema 3.2. *Sea $\tau \in \mathcal{T}(I)$. Entonces para cada $n \geq 1$*

$$W_n \equiv \max_{J \in \mathcal{P}^{(n)}} V_J g_n \leq \frac{n}{\alpha^{n-1}} W_1, \quad (3.4)$$

donde $g_n = |(g \circ \tau^{n-1}) \cdot (g \circ \tau^{n-2}) \cdots (g \circ \tau) \cdot g|$, en particular $g_1 = g$.

Demostración. Notemos que, por la regla de la cadena, g_n coincide con $|(\tau^n)'|^{-1}$ en los intervalos de la partición $\mathcal{P}^{(n)}$ y se anula en los puntos de la partición.

Para llevar a cabo la prueba vamos a proceder por inducción sobre n .

En primer lugar, el caso $n = 1$ es directo, ya que $W_1 \leq W_1$. Supongamos que 3.4 es cierto para $n = 1, 2, 3, \dots, k$ con $k \geq 1$, es decir

$$W_k \leq \frac{k}{\alpha^{k-1}} W_1.$$

Veamos si 3.4 sigue siendo cierto para $n = k + 1$. Para ello tomamos $J \in \mathcal{P}^{(k+1)}$ y $x_0 < x_1 < \dots < x_m$ una partición de \bar{J} . Entonces

$$\sum_{j=1}^m |g_{k+1}(x_j) - g_{k+1}(x_{j-1})| = \sum_{j=1}^m |g_k(\tau(x_j))g_1(x_j) - g_k(\tau(x_{j-1}))g_1(x_{j-1})|,$$

donde se ha utilizado que

$$\begin{aligned} g_{k+1}(x_j) &= |g(\tau^k(x_j)) \cdot g(\tau^{k-1}(x_j)) \cdots g(\tau(x_j))g(x_j)| \\ &= |g(\tau^{k-1}(\tau(x_j))) \cdot g(\tau^{k-2}(\tau(x_j))) \cdots g(\tau(x_j))g(x_j)| \\ &= |g(\tau^{k-1}(\tau(x_j))) \cdot g(\tau^{k-2}(\tau(x_j))) \cdots g(\tau(x_j))| |g(x_j)| = g_k(\tau(x_j))g_1(x_j). \end{aligned}$$

Continuando con la suma anterior,

$$\begin{aligned} &\sum_{j=1}^m |g_k(\tau(x_j))g_1(x_j) - g_k(\tau(x_{j-1}))g_1(x_{j-1})| \\ &= \sum_{j=1}^m |g_k(\tau(x_j))g_1(x_j) - g_k(\tau(x_{j-1}))g_1(x_j) + g_k(\tau(x_{j-1}))g_1(x_j) - g_k(\tau(x_{j-1}))g_1(x_{j-1})| \\ &\leq \sum_{j=1}^m (|g_k(\tau(x_j)) - g_k(\tau(x_{j-1}))|g_1(x_j)| + |g_k(\tau(x_{j-1}))|(g_1(x_j) - g_1(x_{j-1}))) \\ &\leq \frac{1}{\alpha} W_k + \frac{1}{\alpha^k} W_1 \leq \frac{k}{\alpha \cdot \alpha^{k-1}} W_1 + \frac{1}{\alpha^k} W_1 = \frac{k+1}{\alpha^k} W_1, \end{aligned}$$

donde en la última parte se ha empleado la hipótesis de inducción para $n = k$ y que $g(x) \leq \frac{1}{\alpha}$ en todo punto x . Notemos que los puntos $\tau(x_1), \dots, \tau(x_{m-1})$ están en dentro de uno de los intervalos J' de $\mathcal{P}^{(k)}$. Los puntos $\tau(x_0)$ y $\tau(x_m)$ pueden ser, o no, extremos de \bar{J}' . Si alguno de ellos no lo es, digamos $\tau(x_0)$, es porque $x_0 \in \mathcal{P}$ y entonces $\tau(x_0) = a$ y $g_k(\tau(x_0)) = 0$, que es el mismo valor que g_k toma en los extremos de \bar{J}' .

De esta forma, como $V_J g_n \leq \frac{n}{\alpha^{n-1}} W_1$ para cualquier $J \in \mathcal{P}^{(n)}$, en particular

$$W_n \equiv \max_{J \in \mathcal{P}^{(n)}} V_J g_n \leq \frac{n}{\alpha^{n-1}} W_1.$$

El siguiente lema también lo podemos encontrar en [1, cap 5].

Lema 3.3. *Sea $\tau \in \mathcal{T}(I)$. Entonces existen constantes $0 < r < 1$, $C > 0$ y $R > 0$ tales que para cualesquiera $f \in BV(I)$ y $n \geq 1$ se cumple*

$$\|P_\tau^n f\|_{BV} \leq Cr^n \|f\|_{BV} + R\|f\|_1.$$

Demostración. Definimos $\alpha_n = \inf_{x \in I} (|(\tau^n)'(x)|)$ y $\delta_n = \min_{J \in \mathcal{P}(n)} \lambda(J)$ y, como en el lema anterior, $W_n = \max_{J \in \mathcal{P}(n)} V_{\bar{J}} g_n$. Entonces $\alpha_n \geq \alpha^n$. Por el lema 3.2 sabemos que $W_n \leq n\alpha^{1-n}W_1$. Entonces como $\alpha > 1$ y $\alpha_n \geq \alpha^n$ podemos encontrar un $k \geq 1$ tal que

$$r_k = \frac{3}{\alpha_k} + W_k < 1.$$

Fijamos este k y establecemos $R_k = \frac{r_k}{\delta_k}$, $C_1 = \max\{1, r_1, r_2, r_3, \dots, r_{k-1}\}$, $C_2 = \max\{\frac{r_1}{\delta_1}, \frac{r_2}{\delta_2}, \dots, \frac{r_k}{\delta_k}\}$ y $n = jk + i$, con $0 \leq i \leq k - 1$. Por la proposición 2.6, se tiene $P_{\tau^n} = (P_{\tau^k})^j P_{\tau^i}$, y usando el lema 3.1 y que P_τ es una contracción (proposición 2.5), resulta

$$\begin{aligned} V_I P_{\tau^n} f &= V_I P_{\tau^k}^j (P_{\tau^i} f) \leq r_k [V_I P_{\tau^k}^{j-1} (P_{\tau^i} f)] + R_k \|f\|_1 \\ &\leq r_k [r_k V_I P_{\tau^k}^{j-2} (P_{\tau^i} f) + R_k \|f\|_1] + R_k \|f\|_1 \\ &\vdots \\ &\leq r_k^j V_I (P_{\tau^i} f) + (r_k^{j-1} + r_k^{j-2} + \dots + r_k + 1) R_k \|f\|_1 \\ &\leq C_1 r_k^j V_I f + (r_k^{j-1} + r_k^{j-2} + \dots + r_k + 1) R_k \|f\|_1 + r_k^j C_2 \|f\|_1 \\ &\leq C_1 r_k^j V_I f + \left(\frac{1}{1 - r_k} \right) C_2 \|f\|_1. \end{aligned}$$

Entonces por la definición de la norma $\|\cdot\|_{BV}$,

$$\begin{aligned} \|P_{\tau^n} f\|_{BV} &= \|P_{\tau^n} f\|_1 + \inf_{h=P_{\tau^n} f \text{ c.t.p.}} V_I h \leq \|P_{\tau^n} f\|_1 + \inf_{\bar{f}=f \text{ c.t.p.}} V_I P_{\tau^n} \bar{f} \\ &\leq \|f\|_1 + \inf_{\bar{f}=f \text{ c.t.p.}} C_1 r_k^j V_I \bar{f} + \left(\frac{1}{1 - r_k} \right) C_2 \|f\|_1 \\ &\leq \|f\|_1 + C_1 r_k^j \|f\|_{BV} + \left(\frac{1}{1 - r_k} \right) C_2 \|f\|_1. \end{aligned}$$

Tomando $r = r_k^{\frac{1}{k}}$, $C = C_1 r^{-(k-1)}$ y $R = \frac{C_2}{1 - r_k} + 1$ se tiene

$$\|P_{\tau^n} f\|_{BV} \leq Cr^n \|f\|_{BV} + R\|f\|_1.$$

Con todos estos resultados ya somos capaces de probar el teorema principal de este capítulo, que nos da la existencia de una medida invariante absolutamente continua con función

de densidad de variación acotada. De hecho, las funciones de densidad de medidas invariantes absolutamente continuas serán siempre de variación acotada. La prueba de la existencia puede encontrarse en [1, cap 5, teo 5.2.1] y la segunda parte es de elaboración propia.

Teorema 3.4. *Sea $\tau \in \mathcal{T}(I)$. Entonces τ admite una medida invariante absolutamente continua. Más aún, las funciones de densidad de las medidas τ -invariantes absolutamente continuas están en $BV([a, b])$.*

Demostración. Denotamos por $\mathbf{1} = \chi_I$ a la función característica del intervalo I , cuya variación es 0 al ser constante y su norma $\|\mathbf{1}\|_{BV} = b - a$. Consideramos la familia de funciones $\{P_{\tau^n} \mathbf{1}\}_{n=0}^{+\infty}$. Por el lema anterior 3.3, se tiene

$$\|P_{\tau^n} \mathbf{1}\|_{BV} \leq Cr^n \|\mathbf{1}\|_{BV} + R\|\mathbf{1}\|_1 = Cr^n(b - a) + R(b - a) \leq (C + R)(b - a),$$

para cada n . Entonces $\{P_{\tau^n} \mathbf{1}\}_{n=0}^{+\infty}$ es un subconjunto acotado de $BV(I)$. Lo mismo se aplica al conjunto $\{f_n = \frac{1}{n} \sum_{j=0}^{n-1} P_{\tau^j} \mathbf{1}\}_{n=1}^{+\infty}$, ya que

$$\begin{aligned} \|f_n\|_{BV} &= \left\| \frac{1}{n} \sum_{j=0}^{n-1} P_{\tau^j} \mathbf{1} \right\|_{BV} \leq \frac{1}{n} \sum_{j=0}^{n-1} \|P_{\tau^j} \mathbf{1}\|_{BV} \\ &= \frac{1}{n} \sum_{j=0}^{n-1} \|P_{\tau^j} \mathbf{1}\|_{BV} \leq \frac{1}{n} \sum_{j=0}^{n-1} (C + R)(b - a) = (C + R)(b - a), \end{aligned}$$

para cada n (por la proposición 2.6). Por la proposición 1.38, existe una subsucesión $(f_{n_k})_{k=1}^{+\infty}$ que converge en $\mathcal{L}^1(I)$ a una función f^* . Veamos que en efecto $P_{\tau} f^* = f^*$, de donde deduciremos la primera afirmación del teorema. Por la desigualdad triangular,

$$\|P_{\tau} f^* - f^*\|_1 \leq \|P_{\tau} f^* - P_{\tau} f_{n_k}\|_1 + \|P_{\tau} f_{n_k} - f_{n_k}\|_1 + \|f_{n_k} - f^*\|_1.$$

Como f_{n_k} tiende a f^* en $\mathcal{L}^1(I)$, dado $\varepsilon > 0$ arbitrario, cuando k tiende a infinito, el primer y el tercer sumando serán menores a $\varepsilon/3$, el primero porque el operador es una contracción lineal, proposiciones 2.5 y 2.2. Para el segundo sumando

$$\begin{aligned} \|P_{\tau} f_{n_k} - f_{n_k}\|_1 &= \left\| \frac{1}{n_k} (P_{\tau} \mathbf{1} + P_{\tau}^2 \mathbf{1} + \dots + P_{\tau}^{n_k-1} \mathbf{1} + P_{\tau}^{n_k} \mathbf{1}) - \frac{1}{n_k} (\mathbf{1} + P_{\tau} \mathbf{1} + \dots + P_{\tau}^{n_k-1} \mathbf{1}) \right\|_1 \\ &= \frac{1}{n_k} \|P_{\tau}^{n_k} \mathbf{1} - \mathbf{1}\|_1 \leq \frac{1}{n_k} (\|P_{\tau}^{n_k} \mathbf{1}\|_1 + \|\mathbf{1}\|_1) = \frac{1}{n_k} (\|\mathbf{1}\|_1 + \|\mathbf{1}\|_1) = \frac{2(b - a)}{n_k} < \frac{\varepsilon}{3} \end{aligned}$$

cuando k tiende a infinito, también gracias a las proposiciones 2.5 y 2.2. Por lo tanto, $\|P_{\tau} f^* - f^*\|_1 \leq \varepsilon$ con ε arbitrario mayor que 0, de donde se deduce que $P_{\tau} f^* = f^*$. Como $f^* \geq 0$, por la proposición 2.3, $f^*/\|f^*\|_1$ es densidad de una medida τ -invariante absolutamente continua. Notemos que las proposiciones 2.3 y 2.4 implican que $\|f_{n_k}\|_1 = b - a$ para todo n_k por lo que $\|f^*\|_1 = b - a \neq 0$.

Veamos la segunda parte. Ya sabemos que las funciones de densidad de las medidas invariantes absolutamente continuas se corresponden con los puntos fijos del operador de Perron-Frobenius

positivos y de norma \mathcal{L}^1 unitaria. Pues bien, ahora vamos a ver que dichos puntos fijos, en particular las densidades, están en $BV(I)$. Sea f^* uno de dichos puntos fijos. Entonces, por la densidad de $BV(I)$ en $\mathcal{L}^1(I)$, proposición 1.36, puede encontrarse una sucesión de funciones $f_m \in BV(I)$ convergentes a f^* en $\mathcal{L}^1(I)$. Además, como $f^* \in \mathcal{L}^1(I)$ y $\|f_m\|_1 \leq \|f_m - f^*\|_1 + \|f^*\|_1$, existe $K > 0$ tal que $\|f_m\|_1 < K$, para todo m . Definimos las funciones $g_{m,n} = Q_n(f_m)$, donde Q_n viene dado por

$$Q_n(f) = \frac{1}{n} \sum_{j=0}^{n-1} P_\tau^j f.$$

Notemos que las funciones $g_{m,n}$ están en $BV([a, b])$ por el lema 3.3. Más aún,

$$\begin{aligned} \|g_{m,n}\|_{BV} &= \left\| \frac{1}{n} \sum_{j=0}^{n-1} P_\tau^j(f_m) \right\|_{BV} \leq \frac{1}{n} \sum_{j=0}^{n-1} \|P_\tau^j(f_m)\|_{BV} \leq \frac{1}{n} \sum_{j=0}^{n-1} (Cr^j \|f_m\|_{BV} + R\|f_m\|_1) \\ &= C\|f_m\|_{BV} \frac{1}{n} \sum_{j=0}^{n-1} r^j + R\|f_m\|_1. \end{aligned}$$

Tomando límite cuando n tiende a infinito, como $\sum_{j=0}^{n-1} r^j$ es una progresión geométrica con valor finito al ser $0 < r < 1$, resulta

$$\lim_{n \rightarrow \infty} \|g_{m,n}\|_{BV} \leq R\|f_m\|_1 < RK.$$

En definitiva, para todo m existe un cierto n_m tal que $\|g_{m,n_m}\|_{BV} < RK$, con RK independiente de m .

Por otro lado, al ser el operador de Perron-Frobenius una contracción (proposición 2.5), lineal (proposición 2.2) y f^* un punto fijo del mismo, se obtiene

$$\|g_{m,n} - f^*\|_1 = \|Q_n(f_m) - f^*\|_1 = \|Q_n(f_m) - Q_n(f^*)\|_1 = \|Q_n(f_m - f^*)\|_1 \leq \|f_m - f^*\|_1.$$

En particular, $g_{m,n_m} \rightarrow f^*$ en $\mathcal{L}^1(I)$.

Combinando los dos resultados anteriores, es decir, $\|g_{m,n_m}\|_{BV} < RK$ para todo m y $g_{m,n_m} \rightarrow f^*$ en $\mathcal{L}^1([a, b])$ y aplicando la proposición 1.39, se obtiene $f^* \in BV(I)$ como queríamos demostrar. ■

Transformaciones de Markov

El objetivo de este capítulo es demostrar la existencia de una única medida invariante absolutamente continua y equivalente con la medida de Lebesgue, respecto la cual una transformación de Markov lineal es ergódica. Y con este resultado deduciremos que en el sistema dinámico $(I, \mathcal{B}, \mu, \tau)$, la frecuencia relativa con la que una órbita cae en un subconjunto de Borel $A \in \mathcal{B}$, converge a la medida del conjunto A , $\mu(A)$ y no solo en casi todo punto respecto de la medida μ , sino también respecto de la medida de Lebesgue λ . Para ello, haremos uso de buena parte de los resultados que hemos probado en capítulos anteriores, junto con nuevos resultados de este capítulo. Además, ilustraremos el teorema con algún ejemplo.

En primer lugar, vamos a ponernos en situación de las transformaciones con las que vamos a trabajar en este capítulo. Seguimos, como en todo el trabajo, con el intervalo $I = [a, b]$ y transformaciones $\tau : I \rightarrow I$ monótonas a trozos con partición asociada $a = a_0 < a_1 < \dots < a_q = b$. Para $i = 1, \dots, q$, llamamos $J_i = (a_{i-1}, a_i)$ y la restricción de τ a J_i como τ_i . Ahora, si τ_i es un difeomorfismo de J_i a una unión de intervalos J_j conectados, es decir de la forma $(a_{j(i)}, a_{k(i)})$, se dice que τ es de *Markov* y la partición $\mathcal{P} = \{J_i\}_{i=1}^q$ se llama *partición de Markov con respecto a τ* y se dice que τ está en la clase $\mathcal{T}_M(I)$. Si además cada τ_i es lineal en cada J_i se dice que τ es una *transformación de Markov lineal a trozos*.

Comenzamos con la definición de la matriz de incidencia de una transformación.

Definición 4.1. Sean $\tau : I \rightarrow I$ una transformación monótona a trozos y $\mathcal{P} = \{J_i\}_{i=1}^q$ su partición asociada. Se define la *matriz de incidencia* de τ como la matriz $A_\tau = (a_{ij})_{1 \leq i, j \leq q}$ donde

$$a_{ij} = \begin{cases} 1, & \text{si } J_i \subset \tau(J_j), \\ 0, & \text{en otro caso.} \end{cases}$$

Esta noción de matriz de incidencia resulta ser de gran utilidad cuando τ es una transformación de Markov, ya que si $a_{ij} = 0$ entonces $J_i \cap \tau(J_j) = \emptyset$. Más aún, una entrada b_{ij} positiva en A_τ^k equivale a decir que $\tau^k(J_j) \supset J_i$, mientras $b_{ij} = 0$ equivale a $J_i \cap \tau^k(J_j) = \emptyset$.

Veamos una serie de definiciones del álgebra lineal y teoría de grafos presentados en su mayoría en [10, caps 2,6]. Serán necesarios para realizar una interpretación matricial del operador de Perron-Frobenius.

Definición 4.2. Se dice que una matriz $A = (a_{ij}) \in M_{q \times q}(\mathbb{R})$ no negativa (es decir, con todas sus entradas no negativas) es *primitiva* si existe un entero positivo k tal que $A^k > 0$ (es decir, todas sus entradas son positivas).

Definición 4.3. Se dice que una matriz $A = (a_{ij}) \in M_{q \times q}(\mathbb{R})$ es *de permutación* si todas sus entradas son nulas excepto una cualquiera por cada fila y columna, que debe ser igual a 1.

Definición 4.4. Una matriz $A = (a_{ij}) \in M_{q \times q}(\mathbb{R})$ no negativa se llama *matriz estocástica* cuando sus filas o sus columnas suman 1. En particular, se llama *estocástica por columnas* si son ellas las que suman 1 y *estocástica por filas* si ellas suman 1.

Definición 4.5. Un *grafo dirigido* G es un par $G = (V, E)$ donde V es un conjunto finito (de vértices) $V = \{v_1, v_2, \dots, v_q\}$ y E un conjunto (de aristas) contenido en $V \times V$, donde la arista dirigida (v_i, v_j) va del vértice v_i al vértice v_j .

Definición 4.6. Dado un grafo $G = (V, E)$, un *camino* desde el vértice v_i al vértice v_j es una secuencia de vértices $v_i, v_{i_1}, v_{i_2}, \dots, v_{i_k}, v_j$, que empieza en v_i y acaba en v_j , donde para cada vértice v (excepto el último) de la secuencia existe una arista $e \in E$ que lo une con el siguiente de la secuencia. Se llama *longitud del camino* al número de aristas que lo forman. Se llama *ciclo* a aquel camino cuyo vértice inicial y final coinciden sin repetirse ningún vértice intermedio del camino.

Por otro lado, una matriz también puede ser asociada a un grafo dirigido. Notemos la permutación de los índices i, j en la definición siguiente:

Definición 4.7. El *grafo asociado a una matriz no negativa* $A = (a_{ij}) \in M_{q \times q}(\mathbb{R})$ se llama $G(A)$ y tiene por conjunto de vértices $V = \{v_1, v_2, \dots, v_q\}$ y la arista $(v_j, v_i) \in E$ cuando $a_{ij} \neq 0$.

Definición 4.8. Un grafo dirigido $G = (V, E)$ se llama *fuertemente conexo* si para cada par de vértices $v_j, v_i \in V$ existe un camino de v_i a v_j .

Definición 4.9. Una matriz no negativa A es *irreducible* si su grafo asociado $G(A)$ es fuertemente conexo.

El siguiente resultado se puede encontrar en [10, cap 6], no presentamos una prueba porque ya la realizamos en el trabajo fin de grado.

Teorema 4.10. *Una matriz no negativa A es primitiva si es irreducible y su grafo asociado tiene dos ciclos cuyas longitudes son enteros coprimos.*

4.1. Transformaciones de Markov lineales a trozos

El objetivo en esta sección es representar el operador de Perron-Frobenius, P_τ , de una transformación de Markov lineal a trozos, a través de una matriz.

Comenzamos fijando una partición del intervalo, \mathcal{P} , y definiendo la clase de todas las funciones constantes en los intervalos de la partición anterior, denotada por \mathcal{S} . En consonancia con la definición 1.34, estas funciones reciben el nombre de *funciones de paso en \mathcal{P}* . De esta forma se tiene la siguiente relación:

$$f \in \mathcal{S} \text{ si y solo si } f = \sum_{j=1}^q \pi_j \chi_{J_j},$$

para ciertas constantes π_1, \dots, π_q . Identificamos a f con el vector columna $\pi^f = (\pi_1, \dots, \pi_q)^T$.

Veamos el siguiente teorema que caracteriza el operador de Perron-Frobenius a través de una matriz M_τ . La siguiente prueba es una variación de la realizada en [1, cap 4, teo 9.2.1], en el sentido de que la matriz empleada es diferente, al dar a π^f una interpretación de vector columna y no vector fila.

Teorema 4.11. *Sea $\tau \in \mathcal{T}_{\mathcal{M}}(I)$ un transformación lineal de Markov a trozos con partición asociada $\mathcal{P} = \{J_i\}_{i=1}^q$ de I . Entonces existe una matriz $q \times q$, M_τ , tal que $P_\tau f = M_\tau \pi^f$ para toda $f \in \mathcal{S}$, con π^f el vector columna obtenido de f .*

La matriz M_τ es de la forma $M_\tau = (m_{ij})$, donde

$$m_{ij} = \frac{a_{ij}}{|\tau'_j|} = \frac{\lambda(J_j \cap \tau^{-1}(J_i))}{\lambda(J_i)}, \quad 1 \leq i, j \leq q,$$

y $A_\tau = (a_{ij})$ es la matriz de incidencia de τ y τ'_j el valor de τ' sobre J_j .

Demostración. Por la fórmula 2.2, obtenida en la representación del operador de Perron-Frobenius, se tiene

$$P_\tau f(x) = \sum_{i=1}^q f(\tau_i^{-1}(x)) |\tau'_i(\tau_i^{-1}(x))|^{-1} \chi_{\tau(J_i)}(x).$$

Vamos a realizar la prueba distinguiendo dos casos: un primer caso particular y un segundo más general.

En el primer caso, vamos a suponer que $f = \chi_{J_k} \in \mathcal{S}$ para algún $1 \leq k \leq q$, es decir $\pi^f = (0, \dots, 1, \dots, 0)^T$. Entonces, sustituyendo en la expresión anterior,

$$P_\tau f(x) = P_\tau \chi_{J_k}(x) = \sum_{i=1}^q \chi_{J_k}(\tau_i^{-1}(x)) |\tau'_i(\tau_i^{-1}(x))|^{-1} \chi_{\tau(J_i)}(x).$$

Como $\tau_i^{-1}(I) = J_i$, $\chi_{J_k}(\tau_i^{-1}(x))$ será 0 para todo $i \neq k$ y 1 cuando $i = k$. Sustituyendo

$$P_\tau f(x) = |\tau'_k(\tau_k^{-1}(x))|^{-1} \chi_{\tau(J_k)}(x).$$

Además, al ser τ lineal en J_k , $\tau' = \tau'_k$ es constante en J_k . Así, como $\tau_k^{-1}(x) \in J_k$ se tiene

$$P_\tau f(x) = |\tau'_k|^{-1} \chi_{\tau(J_k)}(x).$$

Por lo tanto $P_\tau f = M_\tau \pi^f$.

Veamos el caso general. Sea $f \in \mathcal{S}$, es decir,

$$f = \sum_{j=1}^q \pi_j \chi_{J_j} = (\pi_1, \dots, \pi_q)^T.$$

Por la proposición 2.2 sabemos que el operador P_τ es lineal, por tanto

$$P_\tau f = P_\tau \left(\sum_{j=1}^q \pi_j \chi_{J_j} \right) = \sum_{j=1}^q \pi_j P_\tau(\chi_{J_j})$$

y por el caso anterior

$$P_\tau f(x) = \sum_{j=1}^q \pi_j |\tau'_j|^{-1} \chi_{\tau(J_j)}(x). \quad (4.1)$$

Esto garantiza que $P_\tau f \in \mathcal{S}$ y se podrá escribir $P_\tau f = (d_1, \dots, d_q)^T$.

Estudiamos ahora la parte derecha de la expresión 4.1. Dado $x \in J_i$, el j -ésimo sumando será de la forma $\pi_j |\tau'_j|^{-1}$ si y solo si $x \in \tau(J_j)$, es decir si y solo si $J_i \subset \tau(J_j)$. Pero recordemos que $a_{ij} = 1$ si $J_i \subset \tau(J_j)$ y 0 en el resto en la matriz de incidencia $A_\tau = (a_{ij})$, y que

$$M_\tau = (m_{ij}) = a_{ij} |\tau'_j|^{-1}.$$

Entonces

$$d_i = \sum_{j=1}^q \pi_j m_{ij}$$

y

$$P_\tau f = (d_1, \dots, d_q)^T = M_\tau (\pi_1, \dots, \pi_q)^T = M_\tau \pi^f.$$

■

A la matriz M_τ se le suele llamar *matriz inducida por τ* . Notemos que es una matriz no negativa y cada entrada no nula en una columna j es igual a $|\tau'_j|^{-1}$. Además, τ determina M_τ de manera única, pero no al revés porque no importa si τ_j es creciente o decreciente en J_j .

4.2. Funciones propias asociadas a las matrices inducidas por transformaciones de Markov lineales a trozos

El objetivo en esta sección es probar que una transformación $\tau : I \rightarrow I$ de Markov lineal a trozos cumple que su matriz inducida M_τ tiene 1 como valor propio. De este hecho se puede deducir que el vector propio positivo asociado a 1, visto como una función de paso f en \mathcal{P} , es la densidad de una acim para τ (cuando su norma \mathcal{L}^1 sea 1). En efecto, en la sección anterior hemos probado que $P_\tau f = M_\tau \pi^f$; por tanto, si es un vector propio se tiene que

$$P_\tau f = M_\tau \pi^f = \pi^f = f$$

y f es un punto fijo de P_τ , y la proposición 2.7 garantiza que f es la densidad de una acim para τ .

Si \mathcal{P} es una partición con intervalos de igual longitud, entonces M_τ es estocástica por columnas y 1 es valor propio, [10]. Además, si M_τ es primitiva, por el teorema de Perron-Frobenius estudiado en [10], sus multiplicidades algebraica y geométrica son 1 y existe un vector propio positivo.

Vamos a probar que lo mismo es cierto sin necesidad de que la matriz sea estocástica. El siguiente teorema puede encontrarse en [1, cap 9, teo 9.3.1].

Teorema 4.12. *Sea $\tau : I \rightarrow I$ una transformación de Markov lineal a trozos con matriz inducida M_τ . Entonces M_τ tiene 1 como valor propio. Si además M_τ es primitiva, las multiplicidades algebraicas y geométrica del valor propio 1 son 1 y el espacio propio está generado por un vector positivo.*

Demostración. Recordemos que $\{a_0, a_1, \dots, a_q\}$ define la partición de \mathcal{P} donde τ es de Markov. Definimos

$$\delta = \prod_{j=1}^q (a_j - a_{j-1})$$

y

$$\delta_i = \frac{\delta}{a_i - a_{i-1}} = \prod_{\substack{j=1 \\ j \neq i}}^q (a_j - a_{j-1}).$$

Consideramos la matriz diagonal $D = (d_{ij})$ dada por $d_{ii} = \delta_i, i = 1, \dots, q$ y 0 en el resto. De esta forma, su matriz inversa $E = D^{-1}$ viene dada por $e_{ii} = \delta_i^{-1}, i = 1, \dots, q$ y cero en el resto.

Como τ es de Markov la imagen de cada intervalo de la partición es unión de otros intervalos de la partición. Supongamos que τ lleva J_j a $J_i \cup J_{i+1} \cup \dots \cup J_{i+k}$. Entonces, como τ'_j es la pendiente de una recta, al ser lineal, que une los puntos $(a_j, a_{i+k}), (a_{j-1}, a_{i-1})$ o $(a_j, a_{i-1}), (a_{j-1}, a_{i+k})$, su valor absoluto vale $|\tau'_j| = (a_{i+k} - a_{i-1}) / (a_j - a_{j-1})$. Por la definición de M_τ , se tiene que la j -ésima columna de M_τ tiene entradas $(a_j - a_{j-1}) / (a_{i+k} - a_{i-1})$ en las filas i hasta $i+k$ y cero en el resto de filas. Definimos la matriz $B = D^{-1}M_\tau D$, que cumple $b_{rs} = \delta_r^{-1}m_{rs}\delta_s$. Esta matriz es estocástica por columnas, veámoslo. Para ello, vamos a calcular la suma de una columna arbitraria j de B :

$$\begin{aligned} \sum_{r=1}^q b_{rj} &= \sum_{r=1}^q \delta_r^{-1}m_{rj}\delta_j = \sum_{r=i}^{i+k} \delta_r^{-1} \frac{(a_j - a_{j-1})}{(a_{i+k} - a_{i-1})} \delta_j \\ &= \sum_{r=i}^{i+k} \delta_r^{-1} \frac{(a_j - a_{j-1})}{(a_{i+k} - a_{i-1})} \frac{\delta}{(a_j - a_{j-1})} = \frac{\delta}{(a_{i+k} - a_{i-1})} \left[\frac{1}{\delta_i} + \frac{1}{\delta_{i+1}} + \dots + \frac{1}{\delta_{i+k}} \right] \\ &= \frac{\delta}{(a_{i+k} - a_{i-1})} \left[\frac{a_i - a_{i-1}}{\delta} + \frac{a_{i+1} - a_i}{\delta} + \dots + \frac{a_{i+k} - a_{i+k-1}}{\delta} \right] = 1. \end{aligned}$$

Por lo tanto B es estocástica por columnas y por [10, cap 6, prop 6.8] 1 es valor propio de B y además es de módulo máximo (es decir, el radio espectral de B es 1). Por [10, cap 2, prop 2.22] M_τ tiene a 1 como valor propio de módulo máximo.

Además, si M_τ es primitiva, por el teorema de Perron-Frobenius para matrices primitivas, [10, cap 5, teo 5.1], 1 (por ser de módulo máximo) posee multiplicidades algebraica y geométrica 1 y un vector propio positivo. ■

Veamos un ejemplo de cómo calcular la matriz inducida M_τ y hallar los vectores propios del operador correspondientes al valor propio 1. Es un ejercicio propuesto en [1].

Ejemplo 3. Sea $\tau : [0, 1] \rightarrow [0, 1]$ definida por

$$\tau(x) = \begin{cases} 2x + \frac{1}{2}, & 0 \leq x \leq \frac{1}{4}, \\ -x + \frac{5}{4}, & \frac{1}{4} \leq x \leq \frac{1}{2}, \\ -2x + \frac{7}{4}, & \frac{1}{2} \leq x \leq \frac{3}{4}, \\ -x + 1, & \frac{3}{4} \leq x \leq 1. \end{cases}$$

En primer lugar, vamos a representar gráficamente la función τ , véase la figura 4.1.

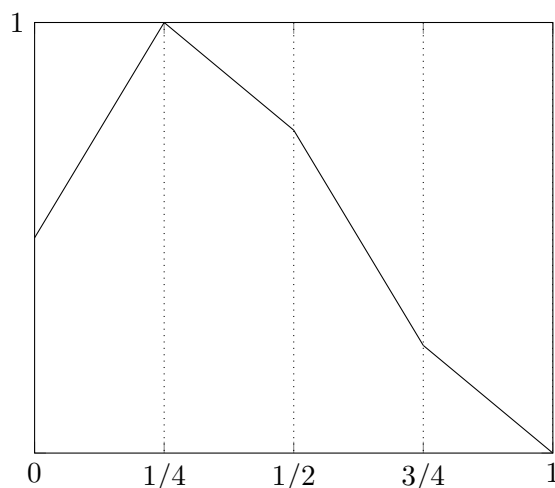


Figura 4.1: Gráfica de τ .

Vamos a considerar la partición $\mathcal{P} = \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$ con intervalos asociados $J_1 = (0, \frac{1}{4})$, $J_2 = (\frac{1}{4}, \frac{1}{2})$, $J_3 = (\frac{1}{2}, \frac{3}{4})$ y $J_4 = (\frac{3}{4}, 1)$. Como $\tau(J_1) = (\frac{1}{2}, 1)$, $\tau(J_2) = (\frac{3}{4}, 1)$, $\tau(J_3) = (\frac{1}{4}, \frac{3}{4})$ y $\tau(J_4) = (0, \frac{1}{4})$, la imagen de cada uno de esos intervalos se corresponde con uniones de los J_i y por tanto τ es una transformación de Markov y además lineal. Por lo que podemos calcular su matriz inducida M_τ respecto a esta partición.

Para calcular la matriz inducida, lo vamos a hacer de dos formas distintas de acuerdo al teorema 4.11. Primero a través de la matriz de incidencia, para ello calculamos los valores a_{ij} usando su definición, comprobando las relaciones $J_i \subset \tau(J_j)$. Así, obtenemos $a_{11} = 0$, $a_{12} = 0$, $a_{13} = 0$, $a_{14} = 1$, $a_{21} = 0$, $a_{22} = 0$, $a_{23} = 1$, $a_{24} = 0$, $a_{31} = 1$, $a_{32} = 0$, $a_{33} = 1$, $a_{34} = 0$, $a_{41} = 1$, $a_{42} = 1$, $a_{43} = 0$ y $a_{44} = 0$. Por otro lado, si calculamos el valor de la derivada de τ en cada J_i , denotándola por τ'_i , obtenemos $|\tau'_1| = 2$, $|\tau'_2| = 1$, $|\tau'_3| = 2$ y $|\tau'_4| = 1$. Sustituyendo en M_τ resulta:

$$M_\tau = \left(\frac{a_{ij}}{|\tau'_j|} \right) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 1 & 0 & 0 \end{pmatrix}.$$

Notemos que la matriz es estocástica por columnas por tener todos los intervalos de la partición igual longitud.

Vamos a calcularla ahora teniendo en cuenta la condición $m_{ij} = \frac{\lambda(J_j \cap \tau^{-1}(J_i))}{\lambda(J_i)}$. Para ello, debemos calcular primero el valor de τ_i^{-1} para cada i . De esta forma, despejando, obtenemos

$$\tau_1^{-1}(y) = \frac{2y-1}{4}, \quad \tau_2^{-1}(y) = \frac{5}{4} - y, \quad \tau_3^{-1}(y) = \frac{7-4y}{8} \quad \text{y} \quad \tau_4^{-1}(y) = 1 - y.$$

Atendiendo a la gráfica de la función y al valor de las inversas calculado, deducimos que la imagen inversa de cada intervalo de la partición es

$$\tau^{-1}(J_1) = \tau_4^{-1}(J_1) = \left(\frac{3}{4}, 1 \right),$$

$$\tau^{-1}(J_2) = \tau_3^{-1}(J_2) = \left(\frac{5}{8}, \frac{3}{4} \right),$$

$$\tau^{-1}(J_3) = \tau_1^{-1}(J_3) \cup \tau_3^{-1}(J_3) = \left(0, \frac{1}{8} \right) \cup \left(\frac{1}{2}, \frac{5}{8} \right)$$

y

$$\tau^{-1}(J_4) = \tau_1^{-1}(J_4) \cup \tau_2^{-1}(J_4) = \left(\frac{1}{8}, \frac{1}{4} \right) \cup \left(\frac{1}{4}, \frac{1}{2} \right).$$

Calculemos los coeficientes m_{ij} , para ello tenemos en cuenta que $\lambda(J_i) = \frac{1}{4}$ para cada i :

$$m_{11} = \frac{\lambda\left(\left(0, \frac{1}{4}\right) \cap \left(\frac{3}{4}, 1\right)\right)}{\frac{1}{4}} = 0, \quad m_{12} = \frac{\lambda\left(\left(\frac{1}{4}, \frac{1}{2}\right) \cap \left(\frac{3}{4}, 1\right)\right)}{\frac{1}{4}} = 0, \quad m_{13} = \frac{\lambda\left(\left(\frac{1}{2}, \frac{3}{4}\right) \cap \left(\frac{3}{4}, 1\right)\right)}{\frac{1}{4}} = 0,$$

$$m_{14} = \frac{\lambda\left(\left(\frac{3}{4}, 1\right) \cap \left(\frac{3}{4}, 1\right)\right)}{\frac{1}{4}} = 1, \quad m_{21} = \frac{\lambda\left(\left(0, \frac{1}{4}\right) \cap \left(\frac{5}{8}, \frac{3}{4}\right)\right)}{\frac{1}{4}} = 0, \quad m_{22} = \frac{\lambda\left(\left(\frac{1}{4}, \frac{1}{2}\right) \cap \left(\frac{5}{8}, \frac{3}{4}\right)\right)}{\frac{1}{4}} = 0,$$

$$m_{23} = \frac{\lambda\left(\left(\frac{1}{2}, \frac{3}{4}\right) \cap \left(\frac{5}{8}, \frac{3}{4}\right)\right)}{\frac{1}{4}} = \frac{1}{2}, \quad m_{24} = \frac{\lambda\left(\left(\frac{3}{4}, 1\right) \cap \left(\frac{5}{8}, \frac{3}{4}\right)\right)}{\frac{1}{4}} = 0,$$

$$m_{31} = \frac{\lambda\left(\left(0, \frac{1}{4}\right) \cap \left[\left(0, \frac{1}{8}\right) \cup \left(\frac{1}{2}, \frac{5}{8}\right)\right]\right)}{\frac{1}{4}} = \frac{1}{2}, \quad m_{32} = \frac{\lambda\left(\left(\frac{1}{4}, \frac{1}{2}\right) \cap \left[\left(0, \frac{1}{8}\right) \cup \left(\frac{1}{2}, \frac{5}{8}\right)\right]\right)}{\frac{1}{4}} = 0,$$

$$m_{33} = \frac{\lambda\left(\left(\frac{1}{2}, \frac{3}{4}\right) \cap \left[\left(0, \frac{1}{8}\right) \cup \left(\frac{1}{2}, \frac{5}{8}\right)\right]\right)}{\frac{1}{4}} = \frac{1}{2}, \quad m_{34} = \frac{\lambda\left(\left(\frac{3}{4}, 1\right) \cap \left[\left(0, \frac{1}{8}\right) \cup \left(\frac{1}{2}, \frac{5}{8}\right)\right]\right)}{\frac{1}{4}} = 0,$$

$$m_{41} = \frac{\lambda\left(\left(0, \frac{1}{4}\right) \cap \left[\left(\frac{1}{8}, \frac{1}{4}\right) \cup \left(\frac{1}{4}, \frac{1}{2}\right)\right]\right)}{\frac{1}{4}} = \frac{1}{2}, \quad m_{42} = \frac{\lambda\left(\left(\frac{1}{4}, \frac{1}{2}\right) \cap \left[\left(\frac{1}{8}, \frac{1}{4}\right) \cup \left(\frac{1}{4}, \frac{1}{2}\right)\right]\right)}{\frac{1}{4}} = 1,$$

$$m_{43} = \frac{\lambda\left(\left(\frac{1}{2}, \frac{3}{4}\right) \cap \left[\left(\frac{1}{8}, \frac{1}{4}\right) \cup \left(\frac{1}{4}, \frac{1}{2}\right)\right]\right)}{\frac{1}{4}} = 0 \quad \text{y} \quad m_{44} = \frac{\lambda\left(\left(\frac{3}{4}, 1\right) \cap \left[\left(\frac{1}{8}, \frac{1}{4}\right) \cup \left(\frac{1}{4}, \frac{1}{2}\right)\right]\right)}{\frac{1}{4}} = 0,$$

quedando de nuevo la misma matriz M_τ de antes.

Vamos a calcular los puntos fijos de P_τ , pero antes de eso vamos a comprobar si la matriz M_τ es primitiva. En efecto, el grafo asociado a la matriz es fuertemente conexo, ya que podemos encontrar un ciclo que conecta a todos los nodos, en concreto $1 \rightarrow 3 \rightarrow 2 \rightarrow 4 \rightarrow 1$. Además, podemos encontrar dos ciclos de longitudes coprimas, el anterior y $3 \rightarrow 3$ (longitud 1), por lo que el teorema 4.10 nos garantiza la primitividad. Así, por el teorema 4.12 sabemos que las multiplicidades algebraicas y geométricas del valor propio 1 son 1. Para terminar, calculemos los puntos fijos del operador de Perron-Frobenius que son funciones de paso, es decir, los vectores π^f tales que $M_\tau \pi^f = \pi^f$. Supongamos que $\pi^f = (x, y, z, t)^T$. Entonces

$$M_\tau \pi^f = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \pi^f$$

proporciona las relaciones

$$x = t, \quad z = 2y, \quad x = z \quad \text{y} \quad \frac{1}{2}x + y = t,$$

por lo que π^f es de la forma

$$\pi^f = (2y, y, 2y, 2y)^T$$

y todos los puntos fijos de P_τ constantes a trozos se corresponden con los múltiplos de $(2, 1, 2, 2)^T$.

El siguiente teorema nos garantiza que si tenemos una transformación de Markov lineal a trozos e $\inf |\tau'| > 1$ donde exista la derivada, entonces toda función de densidad invariante por τ es constante a trozos. La prueba puede verse en [1, cap 9, teo 9.4.1].

Teorema 4.13. *Sea $\tau \in \mathcal{T}_M(I)$ lineal a trozos tal que $\alpha = \inf |\tau'| > 1$ donde la derivada exista. Entonces la densidad de toda acim para τ , es más, todo punto fijo de P_τ , es una función de paso en la partición asociada a τ .*

Demostración. Sea $f \in \mathcal{L}^1(I)$ tal que $P_\tau f = f$. Como $\inf |\tau'| > 1$ y $\frac{1}{|\tau'(x)|}$ es una función constante a trozos (por tanto de variación acotada) se tiene que $\tau \in \mathcal{T}(I)$. Por la segunda parte del teorema 3.4 sabemos que f es de variación acotada en I . Como es un punto fijo de P_τ resulta que

$$P_\tau f(x) = \sum_{j=1}^q f(\tau_j^{-1}(x)) \frac{1}{|\tau_j'|} \chi_{\tau(J_j)}(x) = f(x).$$

Sea J_k un intervalo de la partición \mathcal{P} , y sean $x, y \in J_k$ fijos con $x \neq y$. Entonces $\chi_{\tau(J_j)}(x) = \chi_{\tau(J_j)}(y)$ para todo j . Por lo tanto

$$\begin{aligned} f(x) - f(y) &= P_\tau f(x) - P_\tau f(y) = \sum_{j=1}^q \frac{1}{|\tau'_j|} [f(\tau_j^{-1}(x)) - f(\tau_j^{-1}(y))] \chi_{\tau(J_j)}(x) \\ &= \sum_{i_1} \frac{1}{|\tau'_{i_1}|} [f(\tau_{i_1}^{-1}(x)) - f(\tau_{i_1}^{-1}(y))], \end{aligned}$$

donde el índice i_1 recorre los enteros $j \in \{1, \dots, q\}$ tales que $x \in \tau(J_j)$. Repitiendo el proceso para cada índice i_1 ,

$$f(\tau_{i_1}^{-1}(x)) - f(\tau_{i_1}^{-1}(y)) = \sum_{i_2} \frac{1}{|\tau'_{i_2}|} [f(\tau_{i_2}^{-1}\tau_{i_1}^{-1}(x)) - f(\tau_{i_2}^{-1}\tau_{i_1}^{-1}(y))].$$

Repitiendo el proceso teniendo en cuenta que cada $\frac{1}{\alpha} \geq \frac{1}{|\tau'_{i_l}|}$ y tomando valores absolutos se obtiene

$$\begin{aligned} |f(x) - f(y)| &\leq \frac{1}{\alpha} \sum_{i_1} |f(\tau_{i_1}^{-1}(x)) - f(\tau_{i_1}^{-1}(y))| \leq \frac{1}{\alpha^2} \sum_{i_1} \sum_{i_2} |f(\tau_{i_2}^{-1}\tau_{i_1}^{-1}(x)) - f(\tau_{i_2}^{-1}\tau_{i_1}^{-1}(y))| \\ &\vdots \\ &\leq \frac{1}{\alpha^n} \sum_{i_1} \cdots \sum_{i_n} |f(\tau_{i_n}^{-1} \cdots \tau_{i_1}^{-1}(x)) - f(\tau_{i_n}^{-1} \cdots \tau_{i_1}^{-1}(y))|. \end{aligned}$$

Como τ es monótona a trozos, el conjunto

$$\{(\tau_{i_n}^{-1} \cdots \tau_{i_1}^{-1}(x), \tau_{i_n}^{-1} \cdots \tau_{i_1}^{-1}(y))\}_{i_1, i_2, \dots, i_n}$$

es una colección finita de como mucho q^n intervalos que no se solapan.

El sumatorio anterior está claramente acotado por la variación total de la función en el intervalo I , y como f es de variación acotada, $V_I f \leq K$. Así, sustituyendo en esa expresión se obtiene

$$|f(x) - f(y)| \leq \frac{1}{\alpha^n} V_I f < \varepsilon$$

para un n suficientemente grande al ser $V_I f \leq K$. Por lo tanto, $f(x) = f(y)$ y f es constante en J_k . Repitiendo el proceso en cada intervalo queda probado el resultado. ■

Ejemplo 4. Con este ejemplo buscamos mostrar lo expresado en el teorema anterior. Recordemos la función τ del ejemplo 2. Sea $\tau : [0, 1] \rightarrow [0, 1]$ definida por

$$\tau(x) = \begin{cases} -2x + 1, & 0 \leq x \leq \frac{1}{2}, \\ 2x - 1, & \frac{1}{2} < x \leq 1, \end{cases}$$

donde habíamos calculado que el operador de Perron-Frobenius era

$$P_\tau f(x) = \frac{1}{2} f\left(\frac{1-x}{2}\right) + \frac{1}{2} f\left(\frac{x+1}{2}\right).$$

Atendiendo a la forma de la función, donde $\mathcal{P} = \{0, \frac{1}{2}, 1\}$, la matriz de incidencia es claramente

$$A_\tau = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

y la matriz asociada

$$M_\tau = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Como $\inf |\tau'| > 1$, por el teorema 4.13, todo punto fijo de P_τ es una función de paso y por tanto los únicos puntos fijos serán las soluciones del sistema $M_\tau \pi^f = \pi^f$, por el teorema 4.11. Calculemoslos

$$M_\tau \pi^f = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

obteniendo la relación

$$x = y,$$

por lo tanto, todos los puntos fijos de P_τ son de la forma $(x, x)^T$, es decir múltiplos de $(1, 1)^T$. Por lo tanto, la única acim para τ tiene densidad $f = 1$, es decir, es la propia medida de Lebesgue.

Observación 5. La condición de que $\inf |\tau'| > 1$ es esencial para que los puntos fijos de P_τ sean funciones de paso, véase el ejemplo 1. En él teníamos la transformación $\tau : [0, 1] \rightarrow [0, 1]$ definida por

$$\tau(x) = \begin{cases} 2x, & 0 \leq x \leq \frac{1}{2}, \\ -x + \frac{3}{2}, & \frac{1}{2} \leq x \leq 1, \end{cases}$$

donde claramente $\inf |\tau'| \leq 1$ (en particular vale 1). Y como se vio, puede tener puntos fijos de la forma

$$f(x) = \begin{cases} 0, & 0 < x < \frac{1}{2}, \\ g(x), & \frac{1}{2} < x < 1, \end{cases}$$

donde g es simétrica con respecto a la recta $x = \frac{3}{4}$.

Así, si $g(x) = 3 - 4x$ para $x \in (\frac{1}{2}, \frac{3}{4})$ y $g(x) = 4x - 3$ para $x \in [\frac{3}{4}, 1)$, por ejemplo, se tiene que f no es constante a trozos. En particular, si se toma normalizada la función, se tendrá una función de densidad que no es constante a trozos y no sería cierto el resultado del teorema.

4.3. Existencia y unicidad de medidas invariantes absolutamente continuas

Los resultados de esta sección pueden encontrarse en su mayoría en [5]. Nos disponemos ya a probar el resultado que teníamos como objetivo en este trabajo, el cual garantiza que si tenemos una transformación τ de Markov lineal a trozos con matriz de incidencia primitiva entonces existe una única medida invariante absolutamente continua equivalente a la medida de Lebesgue, respecto a la cual τ es ergódica. Para ello, es necesario introducir una serie de resultados que trabajan sobre la primitividad de la matriz inducida y el concepto de transformación completa. Vamos a trabajar constantemente con una partición de Markov asociada a $\tau \in \mathcal{T}_M(I)$, $\mathcal{P} = \{J_i\}_{i=1}^q$ con matriz de incidencia A_τ . Recordamos la notación $\tau_i = \tau|_{J_i}$, y sea l_i la recta que une los puntos extremos de $\tau_i(J_i)$ donde l'_i es su pendiente. Definimos

$$\beta_i = \inf_{x \in J_i} |\tau'_i(x)|$$

y

$$\delta_i = \frac{\beta_i}{|l'_i|},$$

donde $\delta_i = 1$ si τ_i es lineal, ya que tanto β_i como $|l'_i|$ representa la pendiente de τ_i en valor absoluto. Además $\delta_i \leq 1$ al tomar el ínfimo en β_i , por el teorema del valor medio.

Asimismo, definimos

$$\rho = \min_{i=1, \dots, q} \lambda(J_i)$$

y

$$\gamma = \frac{\rho}{\max_{i=1, \dots, q} \lambda(J_i)},$$

y si K es un intervalo contenido en algún intervalo J_p de \mathcal{P} , escribiremos

$$\eta(K) = \lambda(K)/\lambda(J_p).$$

Definición 4.14. Sea $\tau : I \rightarrow I$ una transformación medible. Decimos que τ es *completa* si para cualquier intervalo $A \subset I$ existe k (dependiente de A) tal que

$$\overline{\tau^k(A)} = I.$$

Definición 4.15. Sea $\tau \in \mathcal{T}_M(I)$. Se dice que τ cumple la *condición (A)* si para cada intervalo $A \subset I$ existen l y J_j (dependientes de A) tales que

$$\tau^l(A) \supset J_j.$$

Lema 4.16. Sea $\tau \in \mathcal{T}_M(I)$ y supongamos que A_τ es primitiva. Si τ cumple la condición (A), entonces τ es completa.

Demostración. Al ser A_τ primitiva, para un cierto entero positivo k se cumple que A_τ^k tiene todas sus entradas positivas. Así, por la definición de matriz de incidencia, se cumple que $J_i \subset \tau^k(J_j)$ para todo j e i , por tanto $\overline{\tau^k(J_j)} = I$ para todo j . Ahora, si $A \subset I$ y $\tau^l(A) \supset J_j$, tendremos

$$\overline{\tau^{l+k}(A)} = I.$$

Por tanto τ es completa. ■

Lema 4.17. Sean $\tau \in \mathcal{T}_M(I)$ y $K \subset J_p$ un intervalo, con $\overline{\tau(J_p)} = \overline{\cup_{j=0}^m J_{p'+j}}$ y $m \geq 0$. Entonces

$$\frac{\lambda(\tau(K))}{\sum_{j=0}^m \lambda(J_{p'+j})} \geq \delta_p \eta(K),$$

En particular, si $\tau(J_p) = J_{p'}$ entonces

$$\eta(\tau(K)) \geq \delta_p \eta(K).$$

Demostración. Por el teorema del valor medio y $K \subset J_p$ se tiene que $\lambda(\tau(K)) \geq \beta_p \lambda(K)$, al ser β_p el ínfimo de las derivadas en J_p . Además,

$$\delta_p = \frac{\beta_p}{|l'_p|} = \frac{\beta_p \lambda(J_p)}{\sum_{j=0}^m \lambda(J_{p'+j})}.$$

Combinando ambos resultados,

$$\lambda(\tau(K)) \geq \beta_p \lambda(K) = \delta_p \frac{\sum_{j=0}^m \lambda(J_{p'+j}) \lambda(K)}{\lambda(J_p)},$$

y despejando

$$\frac{\lambda(\tau(K))}{\sum_{j=0}^m \lambda(J_{p'+j})} \geq \delta_p \frac{\lambda(K)}{\lambda(J_p)} = \delta_p \eta(K).$$

La segunda afirmación es inmediata. ■

Lema 4.18. Sea $\tau \in \mathcal{T}_M(I)$. Si A_τ es irreducible y no es una matriz de permutación, entonces la condición (A) se satisface si

$$\delta_1 \delta_2 \cdots \delta_q (1 + \gamma) > 1.$$

Demostración. Sea A un intervalo. Definimos una sucesión de intervalos $K^{(1)}, K^{(2)}, \dots, K^{(j)}, \dots$ contenidos en intervalos de la partición \mathcal{P} con $K^{(i+1)} \subset \tau(K^{(i)})$ y $A \supset K^{(1)}$. El objetivo es mostrar que la sucesión $(\eta(K^{(i)}))_i$ alcanza el valor 1 en un número finito de pasos: de esta forma, para unos ciertos i y p se tendrá $\lambda(K^{(i)})/\lambda(J_p) = 1$, es decir $K^{(i)} = J_p$, cumpliéndose la condición (A).

Sea $J' \in \mathcal{P}$ tal que $A \cap J' \neq \emptyset$. Definimos $K^{(1)} = A \cap J'$ y asumimos que $K^{(i)} \subset J_p$ (si $K^{(i)} = J_p$ ya habríamos terminado). Definimos $K^{(i+1)}$ como sigue:

1. Si $\tau : J_p \longrightarrow J_{p'}$, entonces $K^{(i+1)} = \tau(K^{(i)}) \subset J_{p'}$.

2. Si $\tau : J_p \rightarrow \bigcup_{j=0}^m J_{p'+j}$, $m \geq 1$ tenemos tres posibilidades:

- a) Si $\tau(K^{(i)}) \subset J_{p'+l}$ para algún $0 \leq l \leq m$, entonces definimos $K^{(i+1)} = \tau(K^{(i)}) \subset J_{p'+l}$. En este caso, o bien $K^{(i+1)} = J_{p'+l}$ y ya terminaría la sucesión con $\eta(K^{(i+1)}) = 1$ o bien $K^{(i+1)} \subset J_{p'+l}$.
- b) Si $\tau(K^{(i)}) \subset \overline{J_{p'+l} \cup J_{p'+l+1}}$ y $\tau(K^{(i)})$ se interseca con ambos intervalos, sea $H^{(k)} = \tau(K^{(i)}) \cap J_{p'+l+k}$ con $k = 0$ o 1 . Si para algún $k = 0, 1$ se da $H^{(k)} = J_{p'+l+k}$ tomamos $K^{(i+1)} = H^{(k)}$ y terminaríamos la sucesión con $\eta(K^{(i+1)}) = 1$ (en caso de que se dé la igualdad en los dos valores de k es indiferente el $K^{(i+1)}$ que se tome). En caso contrario, elegimos como $K^{(i+1)}$ cualquiera de los intervalos $H^{(k)}$, es indiferente el que se elija.
- c) Si $\tau(K^{(i)}) \cap J_{p'+l-1} \neq \emptyset$ y $\tau(K^{(i)}) \cap J_{p'+l+1} \neq \emptyset$ para algún l , entonces, por la continuidad a trozos de τ , $J_{p'+l} \subset \tau(K^{(i)})$. Tomando $K^{(i+1)} = J_{p'+l}$ la secuencia termina con $\eta(K^{(i+1)})$.

Así, podemos obtener una serie de desigualdades para los casos anteriores. Para el caso 1), por el lema anterior 4.17, se tiene

$$\eta(K^{(i+1)}) \geq \delta_p \eta(K^{(i)}).$$

Para el caso 2)a), de nuevo por el lema anterior 4.17, resulta

$$\delta_p \eta(K^{(i)}) \leq \frac{\lambda(K^{(i+1)})}{\sum_{j=0}^m \lambda(J_{p'+j})} \leq \frac{\lambda(K^{(i+1)})}{\lambda(J_{p'+l}) + \rho} \leq \frac{\lambda(K^{(i+1)})/\lambda(J_{p'+l})}{1 + \frac{\rho}{\lambda(J_{p'+l})}} \leq \frac{\eta(K^{(i+1)})}{1 + \gamma},$$

por lo tanto

$$\eta(K^{(i+1)}) \geq \delta_p (1 + \gamma) \eta(K^{(i)}).$$

Con la construcción anterior, se puede observar que, en cada caso, la sucesión $(K^{(i)})_i$ terminará cuando, para una cierta iteración i , $K^{(i)}$ pertenezca a la partición \mathcal{P} , es decir $K^{(i)} = J_p$, con $1 \leq p \leq q$. En esta situación, claramente se satisface la condición (A) por lo que solo faltaría ver que en efecto la sucesión $(K^{(i)})_i$ termina.

En primer lugar, vamos a ver que si la sucesión fuera infinita entonces el caso 2)b) no puede ocurrir más de una vez, para ello vamos a proceder por reducción al absurdo. Supongamos que ocurre dos veces, siendo $K^{(i)}$ y $K^{(j)}$, con $i < j$, un par de intervalos obtenidos tras aplicar la construcción del caso 2)b). Entonces,

$$\overline{\tau(K^{(i-1)})} = [c, x_{p'+l}] \cup [x_{p'+l}, d], \text{ con } x_{p'+l-1} < c \text{ y } x_{p'+l+1} > d.$$

Más aún, por la definición de $K^{(i)}$, que es igual a $H^{(0)}$ o a $H^{(1)}$, $x_{p'+l}$ es un punto extremo de $K^{(i)}$. Por ser τ de Markov y monótona en cada intervalo de su partición asociada, $K^{(j-1)}$ debe ser un intervalo con uno de sus puntos extremos perteneciente a la partición \mathcal{P} , digamos x_p . Suponiendo que

$$\overline{\tau(K^{(j-1)})} = [c', x_r] \cup [x_r, d'],$$

x_p se ha transformado en c' o d' , por lo que uno de estos puntos estaría en \mathcal{P} y la sucesión terminaría, lo que es una contradicción con la infinitud de la sucesión.

Sea s el índice tal que $K^{(s)}$ es obtenido de $K^{(s-1)}$ mediante la construcción del caso 2)b) (si este caso no llegara a suceder se toma $s = 1$, es decir sería el primero). Consideramos la sucesión de números $\eta(K^{(s)}), \eta(K^{(s+1)}), \dots$. Si esta sucesión fuera infinita, entonces $\eta(K^{(s+i)}) < 1$ para todo $i \geq 0$.

Ahora, para todo $p \geq s$, $K^{(p+1)}$ solo puede obtenerse de las construcciones realizadas en los casos 1) y 2)a). Para cierto p , sea k el índice más pequeño tal que $K^{(p+k)}$ se obtiene aplicando el caso 2)a), y digamos $K^{(p+j)} \subseteq J_{i_j} \in \mathcal{P}$ para todo $0 \leq j < k$. Como la matriz de incidencia de \mathcal{P} , A_τ , es irreducible (su grafo asociado es fuertemente conexo, es decir, existe un camino entre cualquiera de sus nodos) y no es de permutación, y estamos en la situación del caso 1), los elementos de la sucesión $(J_{i_j})_{j=0}^{k-1}$ no se pueden repetir, con lo que $k \leq q$. Aplicando la desigualdad del caso 1) sucesivamente resulta

$$\eta(K^{(p+k-1)}) \geq \delta_{i_{k-2}} \eta(K^{(p+k-2)}) \geq \dots \geq \delta_{i_0} \dots \delta_{i_{k-2}} \eta(K^{(p)}).$$

Como $K^{(p+k)}$ se formaría a partir del caso 2)a), aplicando la desigualdad de dicho caso se obtiene

$$\eta(K^{(p+k)}) \geq \delta_{i_{k-1}} (1 + \gamma) \eta(K^{(p+k-1)}) \geq \delta_{i_0} \dots \delta_{i_{k-2}} \delta_{i_{k-1}} (1 + \gamma) \eta(K^{(p)}),$$

por lo tanto

$$\eta(K^{(p+k)}) \geq \delta_1 \dots \delta_{q-1} \delta_q (1 + \gamma) \eta(K^{(p)}),$$

ya que todos los δ_{i_j} son distintos y los $\delta_i \leq 1$.

Para terminar sean $k_i > 0$ los números tales que $K^{(s+\sum_{i=1}^t k_i)}$ son los intervalos que se van obteniendo mediante la construcción del caso 2)a). Entonces

$$\begin{aligned} \eta(K^{(s+\sum_{i=1}^t k_i)}) &\geq \delta_1 \dots \delta_q (1 + \gamma) \eta(K^{(s+\sum_{i=1}^{t-1} k_i)}) \geq [\delta_1 \dots \delta_q (1 + \gamma)]^2 \eta(K^{(s+\sum_{i=1}^{t-2} k_i)}) \\ &\geq \dots \geq [\delta_1 \dots \delta_q (1 + \gamma)]^t \eta(K^{(s)}) > 1 \end{aligned}$$

para un t suficientemente grande, ya que $\delta_1 \dots \delta_q (1 + \gamma) > 1$ por hipótesis. Esto contradice la suposición de que la sucesión de números $\eta(K^{(s)}), \eta(K^{(s+1)}), \dots$, cumpla $\eta(K^{(s+i)}) < 1$ para todo $i \geq 0$ por ser infinita. Por lo tanto es finita y la sucesión de conjuntos $K^{(i)}$ también lo es, cumpliéndose la condición (A). ■

Observación 6. Si la transformación τ es lineal a trozos entonces $\delta_i = 1$ para todo $i = 1, 2, \dots, q$ y la condición del lema anterior 4.18 se transforma en $1 + \gamma > 1$ que siempre es cierta, por lo tanto la condición (A) siempre se cumplirá bajo el resto de condiciones.

En el siguiente lema interpretamos \mathcal{P} como la familia de puntos de la partición, $\mathcal{P} = \{a_i\}_{i=1}^q$.

Lema 4.19. *Supongamos que $\tau \in \mathcal{T}_{\mathcal{M}}(I)$ cumple la condición (A). Entonces*

$$\bigcup_{j=1}^{+\infty} \tau^{-j}(\mathcal{P})$$

es denso en I .

Demostración. Vamos a proceder por reducción al absurdo. Supongamos que el conjunto no es denso en I . Entonces existe un intervalo $A = (c, d)$ tal que $\tau^j(A) \cap \mathcal{P} = \emptyset$ para todo $j \geq 1$, y por tanto cada iterada de A está contenida en algún intervalo de \mathcal{P} y todas las iteradas de τ son estrictamente monótonas en A . Si A' es un intervalo estrictamente contenido en A , todas las iteradas $\tau^j(A')$ estarán estrictamente contenidas en intervalos de la partición, contradiciéndose la condición (A). ■

Observación 7. Como vimos al inicio del tercer capítulo, si τ es expansiva a trozos en \mathcal{P} claramente τ^n es expansiva a trozos en la partición $\mathcal{P}^{(n)}$. Análogamente, si $\tau \in \mathcal{T}_{\mathcal{M}}(I)$ es lineal a trozos en \mathcal{P} , $\tau^n \in \mathcal{T}_{\mathcal{M}}(I)$ lo será en $\mathcal{P}^{(n)}$.

Lema 4.20. *Sea $\tau \in \mathcal{T}_{\mathcal{M}}(I)$ lineal a trozos y cumpliendo la condición (A). Entonces existen m_0 y $\kappa > 1$ tales que para todo $m \geq m_0$*

$$\inf_x |(\tau^m(x))'| \geq \kappa.$$

Demostración. Por la observación anterior 7, sabemos que τ^l es una transformación de Markov lineal a trozos en la partición $\mathcal{P}^{(l)}$. Entonces, por el lema 4.19, podemos escoger un m_0 suficientemente grande tal que $\mathcal{P}^{(m_0)}$ tenga la propiedad de que el intervalo más grande de ella tenga medida de Lebesgue σ menor que ρ , es decir que tenga menor medida que el intervalo más pequeño de la partición original \mathcal{P} . Ahora, en cada intervalo $J_i^{(m_0)} \in \mathcal{P}^{(m_0)}$, τ^{m_0} es lineal y $\tau^{m_0}(J_i^{(m_0)}) \supset J_j$, para algún intervalo de la partición \mathcal{P} . Entonces

$$|(\tau^{m_0}(x))'| \geq \frac{\lambda(J_j)}{\lambda(J_i^{(m_0)})} \geq \frac{\rho}{\sigma} = \kappa > 1.$$

Por lo tanto,

$$\inf_x |(\tau^{m_0}(x))'| \geq \kappa.$$

De hecho, la misma afirmación funciona para todo $m \geq m_0$, porque los intervalos de $\mathcal{P}^{(m)}$ son más pequeños que los de $\mathcal{P}^{(m_0)}$. ■

Con todos estos resultados, ya estamos en condiciones de probar el resultado principal de este trabajo. Antes de comenzar la prueba, usando el teorema 4.10 y la definición de matriz irreducible, observemos que, por ejemplo, una matriz será primitiva si es irreducible y además algún elemento de su diagonal es estrictamente positivo. Esto es debido a que al tener un elemento de la diagonal estrictamente positivo, su grafo asociado posee un ciclo de longitud 1 que será coprimo con cualquier otro ciclo formado y como el grafo es fuertemente conexo se deduce que

la matriz sería primitiva. Si traducimos esto al lenguaje de las transformaciones y sus matrices asociadas es equivalente a decir que su matriz de incidencia sea irreducible y la transformación cumpla que $J_i \subset \tau(J_i)$ para algún i , es decir, que algún intervalo de la partición tenga un punto fijo. La prueba del teorema es de elaboración propia.

Teorema 4.21. *Si $\tau \in \mathcal{T}_{\mathcal{M}}(I)$ es lineal a trozos y su matriz de incidencia es primitiva, entonces τ admite una única acim μ . Más aún, μ es equivalente a λ y τ es ergódica respecto a μ .*

Demostración. Como $A_\tau^k > 0$ para un cierto k , esto implica que $\tau^k(J_i) \supset J_j$ para todo par de intervalos J_i, J_j de \mathcal{P} , por lo que A_τ es irreducible y no es de permutación. Por lo tanto estamos en las condiciones para aplicar el lema 4.18 (recuérdese también la observación 6) y deducir que τ cumple la condición (A). Más aún, podemos aplicar el lema 4.20 que garantiza que para un cierto m se cumple

$$\inf_x |(\tau^m(x))'| \geq k > 1.$$

Así pues, la transformación $\phi = \tau^m$ será expansiva y de Markov lineal a trozos por la observación 7.

Bajo dichas premisas estamos en disposición de aplicar el teorema 4.13 a ϕ , lo que implica que toda función de densidad ϕ -invariante es una función de paso en la partición $\mathcal{Q} = \mathcal{P}^{(m)} = \{K_l\}_{l=1}^t$ de ϕ , es decir, toda función que cumple $P_\phi f = f$, con $f \in \mathcal{D}(I)$, será una función de paso y densidad de una acim, llamémosla μ , con $\mu(A) = \int_A f d\lambda$.

Lo siguiente es notar que la matriz de incidencia de ϕ es también primitiva. Observemos que A_ϕ no es, ni está conectada directamente con, la iterada A_τ^m , porque las dimensiones de las matrices A_ϕ y A_τ son distintas (la partición asociada a ϕ tiene muchos más intervalos que la partición asociada a τ). Ahora bien, el lema 4.16 garantiza que τ es completa, y por tanto ϕ también es completa, y si una transformación de Markov es completa entonces es claro que su matriz de incidencia es primitiva. Por lo tanto, la matriz inducida por ϕ , M_ϕ , también será primitiva.

Como $P_\phi f = f$ es equivalente a $M_\phi \pi^f = \pi^f$ (equivalencia que se puede emplear al ser f una función de paso), dichas funciones de densidad se corresponden con vectores propios no negativos (v_1, v_2, \dots, v_t) de M_ϕ asociados al valor propio 1 tales que $\sum_{l=1}^t v_l \lambda(K_l) = 1$, lo que, al ser M_ϕ primitiva, garantiza la unicidad de la función de densidad f y por tanto de μ para ϕ (teorema 4.12). Esto garantiza la unicidad de la medida μ para τ , ya que las medidas invariantes absolutamente continuas de τ lo son también para ϕ y como τ posee al menos una, la correspondiente al vector propio positivo de valor propio 1 de la matriz inducida que visto como función de paso tiene norma \mathcal{L}^1 igual a 1 (es primitiva, luego existe), necesariamente debe ser esa misma medida μ .

Veamos la ergodicidad respecto a la medida μ . Para ello vamos a proceder por reducción al absurdo. Supongamos que no es ergódica. Entonces existen conjuntos medibles disjuntos A y B , con $0 < \mu(A), \mu(B) < 1$, tales que $\tau^{-1}(A) = A$ y $\tau^{-1}(B) = B$ y $A \cup B = I$. Entonces podemos

construir infinitas medidas, llamémoslas μ_δ con $0 \leq \delta \leq 1$, tales que

$$\mu_\delta(C) = \delta \frac{\mu(C \cap A)}{\mu(A)} + (1 - \delta) \frac{\mu(C \cap B)}{\mu(B)}.$$

Pues bien estas medidas son invariantes y absolutamente continuas respecto τ . En efecto, si $\lambda(C) = 0$ entonces $\mu(C) = 0$ y como $C \cap A \subset C$, $C \cap B \subset C$, resulta $\mu(C \cap A) = 0$ y $\mu(C \cap B) = 0$, obteniéndose

$$\mu_\delta(C) = \delta \frac{\mu(C \cap A)}{\mu(A)} + (1 - \delta) \frac{\mu(C \cap B)}{\mu(B)} = 0,$$

luego $\mu_\delta \ll \lambda$. Al ser μ τ -invariante, se cumple que para cualquier conjunto medible D , $\mu(\tau^{-1}(D)) = \mu(D)$, por tanto, dado C ,

$$\begin{aligned} \mu_\delta(\tau^{-1}(C)) &= \delta \frac{\mu(\tau^{-1}(C) \cap A)}{\mu(A)} + (1 - \delta) \frac{\mu(\tau^{-1}(C) \cap B)}{\mu(B)} \\ &= \delta \frac{\mu(\tau^{-1}(C) \cap \tau^{-1}(A))}{\mu(A)} + (1 - \delta) \frac{\mu(\tau^{-1}(C) \cap \tau^{-1}(B))}{\mu(B)} \\ &= \delta \frac{\mu(\tau^{-1}(C \cap A))}{\mu(A)} + (1 - \delta) \frac{\mu(\tau^{-1}(C \cap B))}{\mu(B)} = \delta \frac{\mu(C \cap A)}{\mu(A)} + (1 - \delta) \frac{\mu(C \cap B)}{\mu(B)} \\ &= \mu_\delta(C). \end{aligned}$$

Al haber infinitas medidas invariantes absolutamente continuas se contradice la unicidad de μ , por lo tanto τ es ergódica.

Por último, veamos la equivalencia entre λ y μ . Como $\mu \ll \lambda$, falta ver $\lambda \ll \mu$. Para ello tomamos un conjunto medible $A \subset I$ tal que $\mu(A) = 0$ y vamos a ver que $\lambda(A) = 0$. No es restrictivo suponer que A está contenido en un intervalo J_{j_0} de la partición asociada, ya que los puntos tienen medida de Lebesgue nula.

$$\text{Como } f = \sum_{j=1}^q \pi_j \chi_{J_j},$$

$$0 = \mu(A) = \int_A f d\lambda = \pi_{j_0} \lambda(A),$$

y como todos los coeficientes π_j son positivos (teorema 4.12), concluimos $\lambda(A) = 0$, por tanto λ y μ son equivalentes. ■

De este teorema podemos extraer importantes consecuencias. Del corolario del teorema ergódico de Birkhoff 1.53, se deduce que en el sistema dinámico $(I, \mathcal{B}, \mu, \tau)$, la frecuencia relativa con la que la órbita de un punto $x \in I$, $\{\tau^n(x)\}_{n=0}^{+\infty}$, cae dentro de un intervalo $A \subset I$ converge a su medida, es decir,

$$\frac{1}{n} \sum_{i=0}^{n-1} \chi_A(\tau^i(x)) \longrightarrow \mu(A), \quad \mu - c.t.p.$$

y por la equivalencia de las medidas

$$\frac{1}{n} \sum_{i=0}^{n-1} \chi_A(\tau^i(x)) \longrightarrow \mu(A), \quad \lambda - c.t.p.$$

Además,

$$\mu(A) = \sum_{j=1}^q \pi_j \lambda(A \cap J_j),$$

siendo $(\pi_1, \pi_2, \dots, \pi_q)^T$ el vector propio de valor propio 1 de M_τ que cumple

$$\sum_{j=1}^q \pi_j \lambda(J_j) = 1.$$

En definitiva, hemos descrito la dinámica de τ en $\lambda - c.t.p.$

Para terminar el trabajo, veamos un ejemplo donde poner de manifiesto su teorema final. Vamos a calcular la única acim para una transformación bajo las condiciones del teorema anterior. Este ejemplo puede verse en [13, cap 2].

Ejemplo 5. Sea $\tau : [0, 1] \rightarrow [0, 1]$ definida por

$$\tau(x) = \begin{cases} 3x + \frac{1}{4}, & 0 \leq x < \frac{1}{4}, \\ 2(x - \frac{1}{4}), & \frac{1}{4} \leq x < \frac{1}{2}, \\ 2 - 2(x + \frac{1}{4}), & \frac{1}{2} \leq x < \frac{3}{4}, \\ -3x + \frac{13}{4}, & \frac{3}{4} \leq x \leq 1. \end{cases}$$

En primer lugar, vamos a representar gráficamente la función τ , véase la figura 4.2.

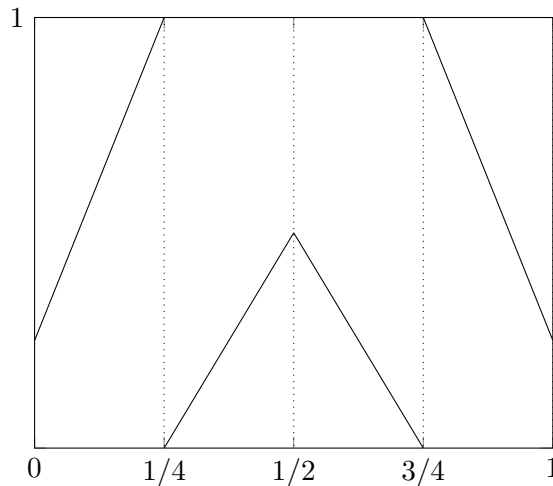


Figura 4.2: Gráfica de τ .

Vamos a considerar la partición $\mathcal{P} = \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$ con intervalos asociados $J_1 = (0, \frac{1}{4})$, $J_2 = (\frac{1}{4}, \frac{1}{2})$, $J_3 = (\frac{1}{2}, \frac{3}{4})$ y $J_4 = (\frac{3}{4}, 1)$. Como $\tau(J_1) = (\frac{1}{4}, 1)$, $\tau(J_2) = (0, \frac{1}{2})$, $\tau(J_3) = (0, \frac{1}{2})$ y $\tau(J_4) = (\frac{1}{4}, 1)$, la imagen de cada uno de esos intervalos se corresponde con uniones de los J_i y por tanto τ es una transformación de Markov y además lineal. Por lo que podemos calcular su matriz inducida M_τ respecto a esta partición.

Para calcular la matriz inducida, lo vamos a hacer a partir de la matriz de incidencia. Para ello calculamos los valores a_{ij} usando su definición, comprobando las relaciones $J_i \subset (J_j)$. Así, obtenemos $a_{11} = 0$, $a_{12} = 1$, $a_{13} = 1$, $a_{14} = 0$, $a_{21} = 1$, $a_{22} = 1$, $a_{23} = 1$, $a_{24} = 1$, $a_{31} = 1$, $a_{32} = 0$, $a_{33} = 0$, $a_{34} = 1$, $a_{41} = 1$, $a_{42} = 0$, $a_{43} = 0$ y $a_{44} = 1$. Por otro lado, si calculamos la derivada de τ en cada J_i , denotándola por τ'_i , obtenemos $|\tau'_1| = 3$, $|\tau'_2| = 2$, $|\tau'_3| = 2$ y $|\tau'_4| = 3$. Sustituyendo en M_τ resulta:

$$M_\tau = \left(\frac{a_{ij}}{|\tau'_j|} \right) = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} \end{pmatrix}.$$

Vamos a calcular los puntos fijos de P_τ , pero antes de eso vamos a comprobar si la matriz M_τ es primitiva. En efecto, el grafo asociado a la matriz es fuertemente conexo, ya que podemos encontrar un ciclo que conecta a todos los nodos, en concreto $1 \rightarrow 4 \rightarrow 3 \rightarrow 2 \rightarrow 1$. Además, podemos encontrar dos ciclos de longitudes coprimas, el anterior y $2 \rightarrow 2$ (longitud 1), por lo que el teorema 4.10 nos garantiza la primitividad. Además, como $\inf |\tau'| > 1$ el teorema 4.13 garantiza que todos los puntos fijos de P_τ son funciones de paso.

Por lo tanto, para calcular los puntos fijos del operador de Perron-Frobenius, debemos encontrar las soluciones de $M_\tau \pi^f = \pi^f$, por el teorema 4.11. Supongamos que $\pi^f = (x, y, z, t)^T$. Entonces

$$M_\tau = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}$$

proporciona las relaciones

$$\frac{y}{2} + \frac{z}{2} = x, \quad \frac{x}{3} + \frac{y}{2} + \frac{z}{2} + \frac{t}{3} = y, \quad \frac{x}{3} + \frac{t}{3} = z \quad \text{y} \quad \frac{x}{3} + \frac{t}{3} = t,$$

de las que se deducen

$$x = 2t, \quad y = 3t, \quad \text{y} \quad z = t.$$

por lo que π^f es de la forma

$$\pi^f = (2t, 3t, t, t)^T$$

y todos los puntos fijos de P_τ son múltiplos de $(2, 3, 1, 1)^T$.

Para terminar, necesitamos calcular el vector normalizado, es decir, aquel vector cuya función asociada tenga norma 1 en la norma \mathcal{L}^1 y a partir de dicha función obtener la forma de la medida invariante absolutamente continua para la que es una densidad. Para ello, consideramos la función

dada por

$$f(x) = \begin{cases} 2t, & 0 \leq x < \frac{1}{4}, \\ 3t, & \frac{1}{4} \leq x < \frac{1}{2}, \\ t, & \frac{1}{2} \leq x < \frac{3}{4}, \\ t, & \frac{3}{4} \leq x \leq 1, \end{cases}$$

y vamos a ver que valor de t hace que sea unitaria.

Calculamos su norma:

$$\begin{aligned} \|f\|_1 &= \int_{[0,1]} f(x)dx = \int_{[0, \frac{1}{4})} 2tdx + \int_{[\frac{1}{4}, \frac{1}{2})} 3tdx + \int_{[\frac{1}{2}, \frac{3}{4})} tdx + \int_{[\frac{3}{4}, 1]} tdx \\ &= \frac{2t}{4} + \frac{3t}{4} + \frac{t}{4} + \frac{t}{4} = \frac{7t}{4}. \end{aligned}$$

Como esta debe ser 1, necesariamente $t = \frac{4}{7}$ y la función que buscamos es

$$f^*(x) = \begin{cases} \frac{8}{7}, & 0 < x < \frac{1}{4}, \\ \frac{12}{7}, & \frac{1}{4} < x < \frac{1}{2}, \\ \frac{4}{7}, & \frac{1}{2} < x < \frac{3}{4}, \\ \frac{4}{7}, & \frac{3}{4} < x < 1. \end{cases}$$

Si, por ejemplo, $A = [\frac{1}{3}, \frac{2}{3}]$, entonces

$$\mu(A) = \frac{12}{7} \cdot \frac{1}{6} + \frac{4}{7} \cdot \frac{1}{6} = \frac{8}{21}.$$

Así pues, la órbita de un punto $x \in [0, 1]$ elegido al azar acabará visitando A casi un tercio de las veces.

Bibliografía

- [1] A. Boyarsky y P. Góra. *Laws of chaos: invariant measures and dynamical systems in one dimension*, Birkhäuser, Boston, 1997.
- [2] A. Boyarsky y G. Haddad, *All invariant densities of piecewise linear Markov maps are piecewise constant*, Adv. Appl. Math. 2 (1981), 284-289.
- [3] A. Boyarsky y M. Scarowsky, *On a class of transformations which have unique absolutely continuous invariant measures*, Trans. Amer. Math. Soc. 255 (1979), 243-262.
- [4] B. D. Craven, *Lebesgue measure & integral*, Pitman Books Limited, Londres, 1982.
- [5] N. Friedman y A. Boyarsky, *Irreducibility and primitivity using Markov maps*, Linear Algebra and Appl. 37 (1981), 103-117.
- [6] A. García Nogales, *Teorías de la medida y de la probabilidad*, Servicio de publicaciones Universidad de Extremadura, Cáceres, 2008.
- [7] A. A. Kosjakin y E. A. Sandier, *Propiedades ergódicas de una cierta clase de transformaciones derivables a trozos en un segmento*, Izv. Vysš. Učebn. Zaved. Matematika 1972, núm. 3(118), 32-40. (En ruso)
- [8] A. Lasota y J. A. Yorke, *On the existence of invariant measures for piecewise monotonic transformations*, Trans. Amer. Math. Soc. 186 (1973), 481-488.
- [9] T.-Y. Li y J. A. Yorke, *Ergodic transformations from an interval into itself*, Trans. Amer. Math. Soc. 235 (1978), 183-192.
- [10] A. López Molina. *El teorema de Perron-Frobenius, Google y la demografía*, Trabajo Fin de Grado, Universidad de Murcia, 2019.
- [11] I. P. Natanson, *Theory of functions of a real variable*, Frederick Ungar Publ. Co., Nueva York, 1955.
- [12] K. Petersen, *Ergodic theory*, Cambridge University Press, Nueva York, 1989.
- [13] A. Swishchuk y S. Islam. *Random dynamical systems in finance*, CRC Press, Boca Raton, 2013.
- [14] P. Walters, *An introduction to ergodic theory*, Springer, Nueva York, 1981.

Índice alfabético

- $\mathcal{T}(I)$, 43
- σ -álgebra de Borel, 4
- σ -álgebra, 4
- clase de todas las funciones constantes en los intervalos de \mathcal{P} , \mathcal{S} , 53
- camino en un grafo, 52
- ciclo, 52
- condición (A), 61
- conjunto casi invariante, 25
- conjunto invariante, 25
- conjunto negativo para μ , 8
- conjunto positivo para μ , 8
- conjuntos de Borel, 4
- conjuntos medibles, 4
- densidad de la medida μ , 38
- derivada de Radon-Nikodym, 12
- descomposición de Hahn, 10
- ergódica, 25
- espacio \mathcal{L}^1 , 5
- espacio de funciones de variación acotada en $[a, b]$, $BV([a, b])$, 13
- espacio de las funciones de densidad, 5
- espacio de medida, 4
- espacio de probabilidad, 4
- espacio medible, 4
- familia uniformemente acotada, 17
- función casi invariante, 25
- función constante a trozos, 21
- función de densidad, 5
- función de variación acotada, 13
- función integrable, 5
- función invariante, 25
- función simple, 7
- funciones de paso en \mathcal{P} , 53
- grafo asociado a una matriz no negativa $G(A)$, 52
- grafo dirigido, 52
- grafo dirigido fuertemente conexo, 52
- longitud camino, 52
- matriz de incidencia de τ , 51
- matriz de permutación, 52
- matriz estocástica, 52
 - por columnas, 52
 - por filas, 52
- matriz inducida por τ , M_τ , 54
- matriz irreducible, 52
- matriz primitiva, 52
- medible, 5
- medida, 4
 - absolutamente continua, 6
 - de probabilidad, 4
 - finita, 4
 - real, 6
- medidas equivalentes, 6
- norma, 5
- norma \mathcal{L}^1 de f , 5
- operador de Perron-Frobenius, 12
- partición de Markov con respecto a τ , 51
- precompacto en $\mathcal{L}^1([a, b])$, 23
- puntos finales de una partición, 12

sistema dinámico, 25

transformación completa, 61

transformación de Markov, 51

transformación de Markov lineal a trozos, 51

transformación expansiva a trozos, 39

transformación medible, 4

transformación monótona a trozos, 39

transformación no singular, 6

variación de f , 13