UNIVERSIDAD DE LAS AMÉRICAS PUEBLA

Escuela de Ciencias

Departamento de Actuaría, Física y Matemáticas

Uso de Números Hiperbólicos para Medición Probabílistica y su aplicación en cadenas de Markov

Trabajo de investigación que presenta la estudiante

FRIDA ELIZABETH CARBAJAL ORDÓÑEZ

162723

Director

Marco Antonio Perez de la Rosa

San Andrés Cholula, Puebla. OTOÑO 2022

Trabajo de investigación que presenta la estudiante
FRIDA ELIZABETH CARBAJAL ORDÓÑEZ, 162723

Asesor del proyecto

MARCO ANTONIO PÉREZ DE LA ROSA

Presidente de Tesis

GERARDO ARIZMENDI ECHEGARAY

Secretario de Tesis

MIGUEL ÁNGEL REYES CORTÉS

DEDICATORIA

Sin duda alguna los primeros a los que este trabajo quiero dedicar es a mami y papi, porque gracias a ellos yo soy lo que soy y sin ellos no tendría la oportunidad de luchar cada día por mis sueños.
Va para mi hermanita Victoria, que siempre me apoya incondicionalmente y está en todas mis locuras. También para mi familia en Acala que, a pesar de la lejanía, los tengo siempre presentes en mi corazón.
Gracias infinitas también a mi mentor Marco por la paciencia y la pasión que me enseñó a poner en esta investigación, sin su guía y conocimiento, este trabajo no hubiera llegado a buen puerto.
Dios, gracias también por ser tan bueno en mi vida.

Índice general

Resumen

El presente escrito tiene como propósito principal tomar como base la idea de la probabilidad ya comúnmente utilizada para desarrollar una nueva teoría de probabilidad basada en las propiedades de los números hiperbólicos. Esto se desarrolla desde las bases recordando los axiomas de orden y tomando la teoría de la medida para la construcción de un espacio de medida probabilística que pueda ampliar la probabilidad tradicional y llevarla a un nivel bi-dimensional.
Es posible ver que la teoría de probabilidad tradicional es un caso especial dentro del espacio de probabilidad hiperbólica, por lo cual es evidente que tiene las bases necesarias para ser aplicada en cualquier área donde sean empleados cálculos probabilísticos. En el caso específico de este texto se aplicó la teoría nueva a las Cadenas de Markov y se encontró una importante conexión en esta área.

Capítulo 1

Introducción

La probabilidad es una de las ramas principales en las matemáticas. De ella depende todo análisis que considere la estimación de eventos futuros aleatorios y cambiantes. Por ello su estudio es amplio y las aplicaciones que derivan de este son útiles en muchos ámbitos dentro de cualquier área, tal como en la demografía, en estadística, física, economía, finanzas, e incluso en áreas de ciencias sociales y de la salud.
La capacidad de estimar eventos futuros posibles es útil en cualquier situación para todos los ámbitos de estudio. En las situaciones donde la incertidumbre suele ser una problemática, la probabilidad en apoyo de la estadística la reduce y da paso a resultados que ayuden a la toma de decisiones con mayor certeza.
Tener confianza en las decisiones que se toman y un futuro menos incierto es uno de los propósitos mejor logrados de la probabilidad.

1.1. Planteamiento del problema

Ahora bien, la probabilidad tradicional es bien conocida en todas las áreas por estimar siempre valores en el rango $[0,1]$ dentro la dimensión lineal de los números reales $\mathbb{R}$.
Del interés por los números no reales surge la pregunta principal de este trabajo de investigación: **¿Es posible desarrollar una nueva teoría de probabilidad dos-dimensional**.
En otras palabras, se busca probar que es posible considerar probabilidades en espacios bi-dimensionales que estimen medidas de certidumbre mediante el uso de números no reales, extendiendo así la teoría tradicional de probabilidad a una teoría dos-dimensional.

1.2. Objetivo General

Se busca responder a la pregunta de investigación al implementar un espacio de medida probabilística construida en el plano de los hiperbólicos que no solamente implemente nuevos métodos para trabajar con probabilidades bi-dimensionales, sino que de la misma manera complemente la teoría tradicional al incluirla como caso particular.

1.3. Objetivo Específico

Es importante que la nueva teoría desarrollada de un espacio de probabilidad hiperbólica sea aplicable en cualquiera de las áreas que necesite la estimación de probabilidades y reducción de incertidumbre.
En este trabajo específicamente nos centraremos en la implementación de la teoría desarrollada dentro de las Cadenas de Markov como un ejemplo de aplicación. Se busca comprobar la utilidad de las probabilidades bi-dimensionales en este tema puesto que es aplicable en diversos contextos de la vida cotidiana donde se requiera transiciones de estados estimados en determinados rangos de tiempo.

1.4. Justificación e Importancia

La teoría tradicional de probabilidad está bien adaptada, es conocida y aceptada por cualquier campo de estudio. Sin embargo, la curiosidad como matemáticos por seguir desarrollando y ampliar la teoría ha llevado a esta investigación por buscar más allá de lo ya entendido e indagar en la existencia de otros métodos que generen nuevas estimaciones considerando dos probabilidades en un sólo número (no real). La investigación de esta teoría es importante pues abre el estudio de los números hiperbólicos y su utilidad en áreas reales y aplicables

1.5. Delimitaciones

La teoría desarrollada está basada en un nivel de estudio considerablemente avanzado dentro de la probabilidad, pero más aún se basa en la comprensión amplia y bien fundamentada de los números complejos. Por lo tanto, es posible que para muchas áreas no sea posible una enseñanza tan profunda para su implementación y que el entendimiento de su funcionamiento se vea obstaculizado por la falta de bases matemáticas necesarias.
Es por ello que esta teoría está mayormente dirigida a las áreas de ciencias exactas.

Se espera que logren apropiar esta teoría, familiarizarla y adaptarla progresivamente a cualquier estudio que permita la doble dimensionalidad de las probabilidades.

1.6. Estructura y narrativa

En el Marco Teórico se establecen los conceptos necesarios clave para la construcción de la nueva teoría de probabilidad. Se consideran las relaciones de orden, el axioma del supremo y la teoría de la medida como bases para generar un buen espacio de probabilidad que cumpla con los requerimientos que misma manera que la teoría tradicional lo hace.
Mediante el capítulo de los Números Hiperbólicos se describen las propiedades y el funcionamiento de estos números complejos. De la misma manera que los reales, estos operan bajo sus respectivas operaciones básicas: suma, resta, multiplicación y división. Así como poseen diversas propiedades tales como: el inverso, su representación matricial, su conjugado y la idempotencia.
A lo largo de la sección de las bases para los hiperbólicos se retoma la propiedad de la idempotencia y se desarrolla un método para transformar el número hiperbólico en base normal a su forma en base idempotente, cuyas propiedades se exponen y desarrollan. Así también se desarrolla un orden parcial para la base idempotente, exponiendo sus propiedades por definición, y sus implicaciones para la suma y la multiplicación
En el capítulo de Espacio de probabilidad hiperbólica se desarrolla el punto máximo de la tesis, puesto que todos los elementos anteriores sirven para estructurar la teoría de probabilidad hiperbólica, desarrollándola al explicar sus propiedades y haciendo la comparativa con la probabilidad tradicional. De este modo también se describen algunas propiedades particulares tales como la condicionalidad y la independencia de eventos.
Se sigue con la aplicación de esta teoría en el capítulo siguiente mediante el caso especial de las Cadenas de Markov.

Capítulo 2

Marco Teórico

2.1. Relaciones de Orden

Para el desarrollo de nuestra teoría es necesario el conocimiento en las bases del álgebra superior, en específico en las relaciones binarias donde se establece un orden.

Sean X y Y conjuntos se entiende por relación binaria R de X en Y a la pareja ordenada de $(R, X \times Y)$, donde $R \subseteq X \times Y$. Siendo $x \in X, y \in Y$ decimos que x está relacionado con y si R es la relación de X en Y: $(x, y) \in R$

Definición 1. *Se entiende a un orden parcial en X si teniendo una relación S dentro del conjunto X denotada como $(S \subseteq X \times X)$ cumple con:*

1) $(x, x) \notin S\ \forall x \in S$.

2) (x, y) *y* $(y, z) \in S,$ *entonces* $(x, z) \in S$. *Véase [2]*

Si existe un orden parcial S en (x_1, x_2) se dice que x_1 es menor que x_2, y se escribe $x_1 < x_2$. Al decir que $x_1 \nless x_2$ interpretamos que $(x_1, x_2) \notin S$.

Así se tiene que

$$< parcial = \begin{cases} x \nless x,\ \forall x \in X \\ \text{si } x < y\ ,\ y < z \implies x < z \end{cases} \tag{2.1}$$

Con base en lo anterior se tiene también que:

a) $x \leq x,\ \forall x \in\ X$.
Por la definición de $\leq$ que establece que es menor o igual. Es evidente que para cualquier x dentro del conjunto debe cumplir únicamente con $x = x$

b) Se cumple la propiedad antisimétrica: $x = y$ cuando se satisface que: $x \leq y$ y $y \leq x$.

Es imposible tener una desigualdad total de $x \neq y$ puesto que esto nos llevaría asumir que $x < y < x$, por transitividad esto es $x < x$ lo que en un orden parcial no es permitido.

c) Si $x \leq y$ y $y \leq z \implies x \leq z$.
Se cumple por transitividad la segunda propiedad de (2.1) ahora en el caso específico menor o igual que.

Para que una relación de orden $<$ en X sea total, teniendo elementos $x, y \in X$ esta tiene que cumplir con que sus elementos sean comparables entre sí, como se muestra a continuación:

$$x < y,\ x = y,\ y < x$$

Se dice así que el $(X, <)$ es un conjunto totalmente ordenado y cada uno los elementos de sus pares ordenados deberá cumplir solamente una de las propiedades. Un ejemplo de un conjunto totalmente ordenado es el conocido conjunto de los números reales.

Lema 1. *Sea un conjunto ordenado $(X, <)$, cada pareja de elementos en X satisface una y solamente una de las propiedades, es decir todo x,y $\in X$ son comparables entre sí de una sola manera.*

Ejemplo Sea:

A= $\{\alpha, \beta, \gamma\}$

B= $\mathbf{P}(A) = \{\emptyset, \{\alpha\}, \{\beta\}, \{\gamma\}, \{\alpha, \beta\}, \{\alpha, \gamma\}, \{\beta, \gamma\}, \{\alpha, \beta, \gamma\}\}$

Teniendo una relación $\mathbf{R} = \{(a, b) \in A \times B\}$, podemos notar que el elemento que consiste en el conjunto vacío $\emptyset$ es comparable con todos los elementos. Sin embargo α, β, y γ no son comparables entre sí.
Por lo tanto podemos afirmar que, al no contar con la comparabilidad, no es un conjunto totalmente ordenado. Cumple únicamente con las características de un conjunto parcialmente ordenado.
Gracias a las propiedades de las relaciones de orden es posible establecer comparabilidad entre los elementos en un conjunto, y es así como se establecen los mínimos y máximos.

Definición 2. *Se le llama* **mínimo** *al elemento del conjunto A, cuando teniendo un conjunto parcialmente ordenado $(A, <)$, existe $a \in A$ para el cual $\forall b \in A$ cumple con $a < b$.*

Definición 3. *Se le llama* **máximo** *al elemento del conjunto A, cuando teniendo un conjunto parcialmente ordenado $(A, <)$, existe $a \in A$ para el cual $\forall b \in A$ cumple con $a > b$.*

Solo es posible contar con un elemento máximo y un elemento mínimo, además es importante remarcar que ambos son comparables con el resto de los elementos dentro del conjunto y que esta propiedad aplica asimismo entre ellos.

Para comprobar que existe un solo elemento mínimo y máximo basta con definir que teniendo dos elementos pertenecientes al conjunto A: a_1 y a_2, si ambos son mínimos se tiene que para cualquier $a \in A$, cumple con $a \geq a_1$, entonces se asume también que $a_2 \geq a_1$. Al a_2 ser un mínimo también asumimos que $a \geq a_2$ y $a_1 \geq a_2$. Teniendo esto, se resume a que $a_1 = a_2$.

Sucede lo mismo análogamente con un máximo.

Axioma 1. *Llamado axioma del supremo establece que :*

a) Todo conjunto no vacío acotado superiormente en los reales tiene un supremo, el cual es la menor de todas las cotas superiores.

b) Todo conjunto no vacío acotado inferiormente en los reales tiene un ínfimo, el cual es el mayor de todas las cotas inferiores.

Un orden parcial en un conjunto se entiende como un buen orden, cuando cada subconjunto no vacío tiene un elemento mínimo. A estos se les conoce como un conjunto con buen orden.

2.2. Axioma del Supremo

Posterior al entendimiento de las relacines de orden es necesario el entendimiento de uno de los teoremas más importantes en la teoría de los conjuntos.

Definición 4. *Sea* **S** *un subconjunto de* $\mathbb{R}$*, entonces* $\beta \in \mathbb{R}$ *es cota superior si cumple* $\beta \geq x, \forall x \in S$

Definición 5. *s es el supremo de S si* $\forall \beta$ *tal que* $\beta \geq x, \forall x \in \mathbf{S}$ *entonces* $s \leq \beta$*. Se podría considerar como el menor de todas las cotas superiores.*

Ejemplo Sea S el intervalo (-1,2), según la definición anterior sus cotas superiores son todos aquellos mayores que 2, tales como π, 3, 19 además del 2 por tratarse de un intervalo abierto. Así entonces el menor de todas las cotas superiores es el 2, por lo cual es el supremo.

Definición 6. *Sea* **S** *un subconjunto de* $\mathbb{R}$*, entonces* $\alpha \in \mathbb{R}$ *es cota inferior si cumple* $\alpha \leq x, \forall x \in S$

Definición 7. *ι es el ínfimo de S si* $\forall \alpha$ *tal que* $\alpha \leq x, \forall x \in \mathbf{S}$ *entonces* $\iota \geq \beta$

Ejemplo Sea S el intervalo (-1,3) según la definición de cota inferior, sus cotas inferiores todos aquellos menores que -1, tales como $-\pi$, -4, -9, incluyendo el -1. Así entonces el mayor de todas las cotas inferiores es el -1, siendo este el ínfimo del intervalo.

2.3. Teoría de la medida

Antes de establecer métodos para la medición de probabilidades, es necesario comprender las bases de la teoría de la medida, la cual se encuentra inmersa dentro del área de análisis matemático. En la enciclopedia McGraw-Hill de ciencia y tecnología se define medida de la siguiente manera.
Para explicar formalmente la teoría de la medida se establece a X como un conjunto universo arbitrario y a $\mathbf{S}$ como un conjunto de subconjuntos de X que cumple con las características siguientes:

1. Considerando a ϕ como el conjunto vacío, entonces $\phi \in \mathbf{S}$

2. Considerando el conjunto A, si $A \in \mathbf{S}$, entonces asimismo su complemento $A^c \in \mathbf{S}$

3. Considerando los conjuntos $A_1, A_2, ..., A_n \in \mathbf{S}$, la unión de todos estos sigue perteneciendo a $\mathbf{S}$, esto es
$$\bigcup_{n=1}^{\infty} A_n \in \mathbf{S}$$

4. Considerando los conjuntos $A_1, A_2, ..., A_n \in \mathbf{S}$, la intersección de todos estos sigue perteneciendo a $\mathbf{S}$, esto es
$$\bigcap_{n=1}^{\infty} A_n \in \mathbf{S}$$

Así pues, antes de proceder a establecer una $\sigma-$álgebra, es necesario primero establecer la definición de un Anillo Booleano. [6]

Definición 8. *Un anillo Booleano de conjuntos es una clase no vacía* $\mathbf{R}$ *de conjuntos tal que:*

$$E \in \mathbf{R} \qquad\qquad F \in \mathbf{R}$$

Entonces

$$E \cup F \in \mathbf{R} \qquad\qquad E - F \in \mathbf{R}$$

Lo cual significa que un anillo es una clase no vacía cerrada bajo la formación de uniones y diferencias.

Además

Definición 9. *Se entiende por $\sigma-$anillo a la clase de conjuntos no vacía* $\mathbf{C}$ *tal que*

* *Si $E \in \mathbf{C}$ y $F \in \mathbf{C}$, entonces $E - F \in \mathbf{C}$*

* *Si el conjunto numerable de $E_i \in \mathbf{C}$, tal que $i = 1, 2, 3, 4, ...,$ entonces*

$$\bigcup_{i=1}^{\infty} E_i \in \mathbf{C}$$

Definición 10. *Por lo tanto, cuando los puntos 1. 2. y 3. se satisfacen, entonces se genera lo que es conocido como una $\sigma-$álgebra, en la cual es posible realizar cualquier operación de conjuntos entre los subconjuntos$A_1, A_2, ..., A_n$ pertenecientes a* $\mathbf{S}$*, y el resultado de estas operaciones seguiría perteneciendo a* $\mathbf{S}$*. Entonces una $\sigma-$álgebra es un $\sigma-$anillo que contiene al universo. Así entonces, entender una $\sigma-$álgebra será de gran utilidad para comprender la teoría de la medida.*

Definición 11. *Se entiende por espacio medible a la pareja (X, σ) si X es un conjunto universo y σ es una $\sigma-$álgebra formada por una familia de subconjuntos $A_1, A_2, ..., An \in \mathbf{S}$ del conjunto universo X.*

Definición 12. *Entonces pues comprendemos por un espacio de medida a la tripleta (X, σ, μ), comprendiendo a X como un conjunto y σ como una $\sigma-$álgebra de los subconjuntos* $\mathbf{S}$ *en X, se dice que la medida μ será una función que asigna a cada conjunto $A_1, A_2, ..., An \in \mathbf{S}$ de σ un número bajo las siguientes condiciones:*

1. La medida de cualquier elemento de σ es un número real no negativo. Siendo una función con rango en $\mathbb{R} \in [0, \infty)$

2. La medida del vacio es cero.
$$\mu(\phi) = 0$$

3. Si A_1 y A_2 y pertenecen a $\mathbf{S}$ y $A_1 \cap A_2 = \phi$ o sea son disconjuntos, entonces la medida de la unión será igual a la medida de la suma esto es:
$$\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2)$$

De manera general, si $A_1, A_2, ..., A_n$ pertenecen a $\mathbf{S}$, entonces es cierto que la medidda de la unión de todos los subconjuntos de $\mathbf{S}$ será menor ó igual a la suma de las medidas de todos los subconjuntos de $\mathbf{S}$

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

Lo anterior se iguala cuando todos los conjuntos $A_1, A_2, ..., A_n$ son disjuntos entre sí, es decir, la intersección entre cualquiera de ellos resulta en el vacio ϕ. A ello se le conoce como la propiedad de aditividad contable de la medida.

Así entonces, si A_1 y $A_2 \in \mathbf{S}$ y $A_2 \subset A_1$, entonces la medida de A_1 será mayor o igual que la medida de A_2, esto es:

$$\mu\left(A_2\right) \leq \mu\left(A_1\right)$$

Ahora bien, si modificamos nuestro universo X a tener tamaño 1, estaríamos generando un espacio de probabilidad, donde las condiciones de espacio de medida se modifican y resultan en:

1. La probabilidad de cualquier evento de σ es un numero real no negativo, cuya función tiene un rango en $\in [0, 1]$

2. La probabilidad del vacio es cero.

$$\mu\left(\phi\right) = 0$$

3. Si A_1 y A_2 y pertenecen a $\mathbf{S}$ y $A_1 \cap A_2 = \phi$ o sea son disconjuntos, entonces la medida de la unión será igual a la medida de la suma esto es:

$$\mu\left(A_1 \cup A_2\right) = \mu\left(A_1\right) + \mu\left(A_2\right)$$

Capítulo 3

Números Hiperbólicos

Un número hiperbólico $\mathbb{D}$ se define como:

$$\mathbb{D} = \{x + ky \mid x, y \in \mathbb{R}\} \tag{3.1}$$

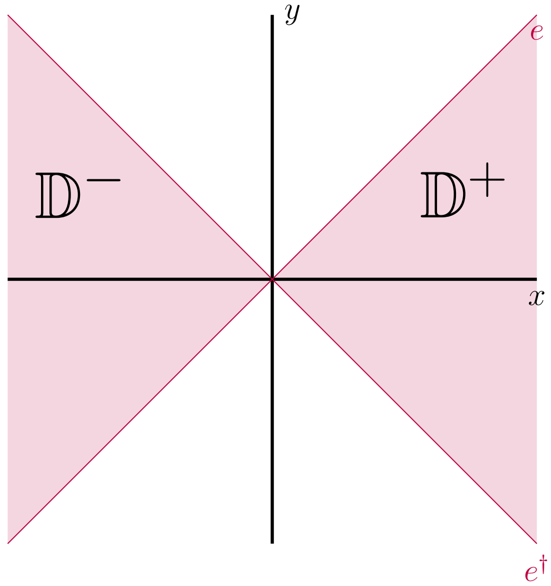

Figura 3.1: Plano Hiperbólico

Al igual que en el plano complejo $\mathbb{C}$, en el plano hiperbólico un número se compone de partes reales y de una unidad imaginaria k, la cual satisface que $k^2 = 1$. De la definición de $\mathbb{D}$ (3.1), podemos asumir que la multiplicación y suma de dos números hiperbólicos funciona de la misma manera que en los números reales. En caso de resultar una k^2 esta será sustituida por 1.

Asimismo también es posible categorizar los números hiperbólicos como se muestra a continuación:

$$\mathbb{D}^{+} = \{x + ky \mid x^2 - y^2 \geqslant 0,\ x \geqslant 0\} \tag{3.2}$$

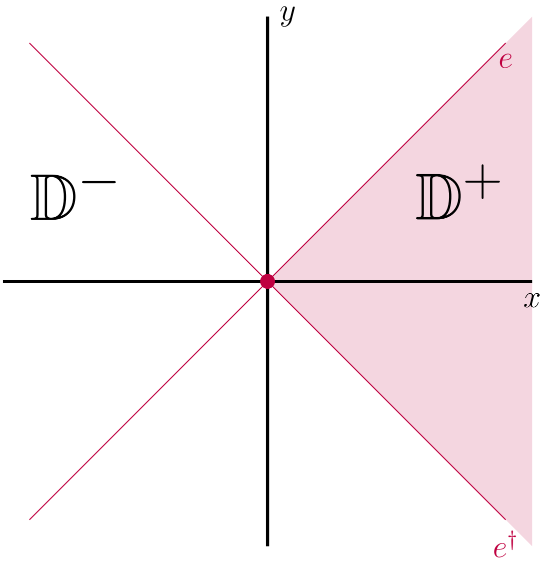

Figura 3.2: Hiperbólicos no negativos

La ecuación (3.2) describe aquellos números hiperbólicos no negativos. Podemos observar que el cero está incluido.

$$\mathbb{D}^{+} \setminus \{0\} = \{x + ky \mid x^2 - y^2 \geqslant 0,\ x > 0\} \tag{3.3}$$

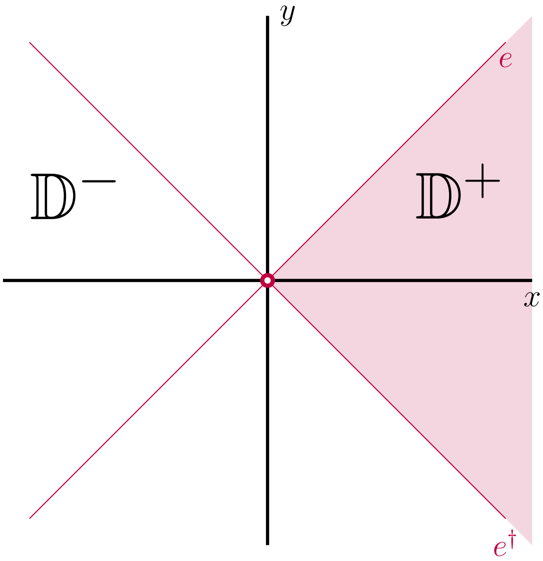

Figura 3.3: Hiperbólicos positivos

La ecuación (3.3) describe el subconjunto que engloba los números hiperbólicos positivos, podemos notar que el cero no esta incluído.

$$\mathbb{D}^{-} = \{x + ky \mid x^2 - y^2 \geqslant 0,\ x \leqslant 0\} \tag{3.4}$$

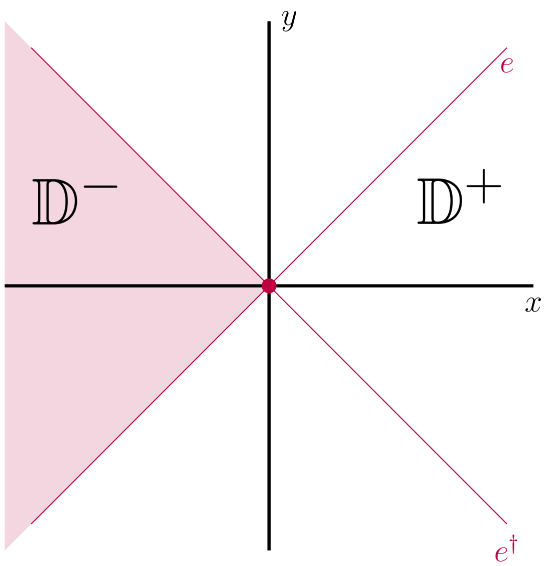

Figura 3.4: Hiperbólicos no positivos

La ecuación (3.4) es llamado subconjunto de números no positivos, aquí el cero forma parte de este.

$$\mathbb{D}^{-} \diagdown \{0\} = \{x + ky \mid x^2 - y^2 \geqslant 0,\ x < 0\} \tag{3.5}$$

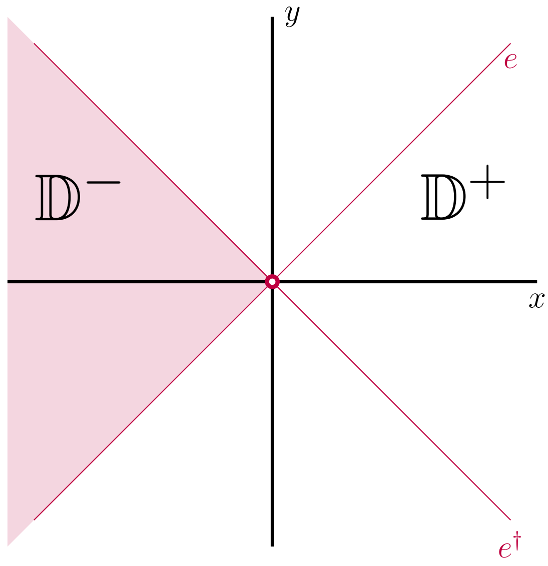

Figura 3.5: Hiperbólicos negativos

En (3.5) se describe el subconjunto de los números hiperbólicos negativos, al excluir el cero

3.1. Operaciones básicas y propiedades

Es necesario conocer cómo interactúan los números hiperbólicos entre sí con las operaciones básicas, así como entender de sus propiedades individuales y la fácil implementación de estas.

3.1.1. Suma, multiplicación y división

Considerando dos números hiperbólicos $z_1 = x_1 + \mathbf{k}y_1$ y $z_2 = x_2 + \mathbf{k}y_2$, las operaciones entre estos se plantean a continuación

- Suma

$$z_1 + z_2 = (x_1 + x_2) + \mathbf{k}(y_1 + y_2) \tag{3.6}$$

Demostración:
Se agrupan las partes reales de los números y las que contienen el factor imaginario

$$z_1 + z_2 = x_1 + \mathbf{k}y_1 + x_2 + \mathbf{k}y_2 = x_1 + x_2 + \mathbf{k}y_1 + \mathbf{k}y_2 = (x_1 + x_2) + \mathbf{k}(y_1 + y_2)$$

- Multiplicación
$$z_1 z_2 = (x_1 x_2 + y_1 y_2) + \mathbf{k}(x_1 y_2 + x_2 y_1) \tag{3.7}$$
Demostración:
$$\begin{aligned} z_1 z_2 &= (x_1 + \mathbf{k}y_1)(x_2 + \mathbf{k}y_2) \\ &= x_1x_2 + x_1y_2\mathbf{k} + y_1x_2\mathbf{k} + y_1y_2\mathbf{k}^2 \qquad \text{sabiendo que } k^2 = 1 \\ &= x_1x_2 + y_1y_2 + x_1y_2\mathbf{k} + x_2y_1\mathbf{k} \\ &= (x_1x_2 + y_1y_2) + \mathbf{k}(x_1y_2 + x_2y_1) \end{aligned}$$

- División:
$$\frac{z_1}{z_2} = \frac{x_1x_2 - y_1y_2}{x_2^2 - y_2^2} + \mathbf{k}\frac{x_1y_1 - x_1y_2}{x_2^2 - y_2^2}$$
Demostración:
Es necesario multiplicar la división por $1 = \frac{x_2 - \mathbf{k}y_2}{x_2 - \mathbf{k}y_2}$
$$\begin{aligned} \frac{z_1}{z_2} &= \frac{x_1 + \mathbf{k}y_1}{x_2 + \mathbf{k}y_2} = \frac{x_1 + \mathbf{k}y_1}{x_2 + \mathbf{k}y_2}\left[\frac{x_2 - \mathbf{k}y_2}{x_2 - \mathbf{k}y_2}\right] \\ &= \frac{(x_1 + \mathbf{k}y_1)(x_2 - \mathbf{k}y_2)}{x_2^2 - \mathbf{k}^2y_2^2} \\ &= \frac{(x_1x_2 + y_1y_2) + \mathbf{k}(x_1y_2 + x_2y_1)}{x_2^2 - y_2^2} \\ &= \frac{x_1x_2 - y_1y_2}{x_2^2 - y_2^2} + \mathbf{k}\frac{x_1y_1 - x_1y_2}{x_2^2 - y_2^2} \end{aligned}$$

Ejemplo
Sea $z_1 = 6 + 10\mathbf{k}$ y $z_2 = 3 - 15\mathbf{k}$, entonces se tiene que $x_1 = 6, y_1 = 10, x_2 = 3, y_2 = -15$

- Suma
$$z_1 + z_2 = (6 + 3) + \mathbf{k}(10 + (-15)) = 9 - 5\mathbf{k}$$

- Multiplicación
$$z_1z_2 = (6 * 3 + 10 * (-15)) + \mathbf{k}(6 * (-15) + 3 * 10) = -132 - 60\mathbf{k}$$

- División:
$$\frac{z_1}{z_2} = \frac{6 * 3 - 10 * (-15)}{3^2 - (-15)^2} + \mathbf{k}\frac{6 * 10 - 6 * (-15)}{3^2 - (-15)^2} = \frac{168}{-216} + \mathbf{k}\frac{150}{-216}$$

3.1.2. Inverso multiplicativo de un hiperbólico

Considerando un número hiperbólico definido como $z = x + \mathbf{k}y$, entonces se tiene que

Su inverso es

$$z^{-1} = \frac{x}{x^2 - y^2} - \mathbf{k}\frac{y}{x^2 - y^2}$$

Demostración:
Se sabe que el inverso o recíproco de un número multiplicado por este satisface la igualdad con la unidad, lo que significa que $zz^{-1} = 1$, así entonces se busca el número $z^{-1} = a + \mathbf{k}b$ que cumpla con esta ecuación.

$$\begin{aligned} zz^{-1} &= 1 \\ (x + \mathbf{k}y)(a + \mathbf{k}b) &= 1 + 0\mathbf{k} \\ xa + xb\mathbf{k} + ya\mathbf{k} + by\mathbf{k}^2 &= 1 + 0\mathbf{k} \\ (xa + by) + (xb + ya)\mathbf{k} &= 1 + 0\mathbf{k} \end{aligned}$$

A lo cual, se genera un sistema de ecuaciones haciendo coincidir las partes reales con las que contienen el factor $\mathbf{k}$:

$$\begin{cases} xa + by = 1 \\ xb + ya = 0 \end{cases}$$

Despejando en la primera ecuación se obtiene que $a = \frac{1-by}{x}$ y sustituyendo en la segunda ecuación sigue que:

$$\begin{aligned} xb + y\left[\frac{1 - by}{x}\right] &= 0 \\ xb + \frac{y - by^2}{x} &= 0 \\ xb &= \frac{by^2 - y}{x} \\ x^2b &= by^2 - y \\ x^2b - by^2 &= -y \\ b(x^2 - y^2) &= -y \\ b &= \frac{-y}{x^2 - y^2} \end{aligned}$$

Sabiendo el valor de b, lo usamos para encontrar el de a

$$\begin{aligned} a = \frac{1 - by}{x} &= \frac{1 - \left[\frac{-y}{x^2-y^2}\right]y}{x} = \frac{1 + \frac{y^2}{x^2-y^2}}{x} = \frac{\frac{x^2-y^2+y^2}{x^2-y^2}}{x} = \frac{\frac{x^2}{x^2-y^2}}{x} \\ &= \frac{x^2}{x(x^2 - y^2)} = \frac{x}{x^2 - y^2} \end{aligned}$$

Por lo tanto $z^{-1} = a + \mathbf{k}b = \frac{x}{x^2-y^2} - \mathbf{k}\frac{y}{x^2-y^2}$

Ejemplo

Sea el número hiperbólico $z = 7 + 6\mathbf{k}$, entonces se tiene que $x = 7$ y $y = 6$ su inverso es

$$z^{-1} = \frac{7}{7^2 - 6^2} - \mathbf{k}\frac{6}{7^2 - 6^2} = \frac{7}{13} - \mathbf{k}\frac{6}{13}$$

3.1.3. Representación Matricial de un Hiperbólico

Su representación matricial es

$$x + \mathbf{k}y \leftrightarrow \begin{pmatrix} x & y \\ y & x \end{pmatrix} \tag{3.8}$$

Demostración:

Considerando la matriz $M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ y elevándola al cuadrado se obtiene la matriz identidad, lo cual satisface con la propiedad del factor imaginario en los hiperbólicos $k^2 = 1$.

$$\mathbf{M}^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0*0+1*1 & 0*1+1*0 \\ 1*0+0*1 & 1*1+0*0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$$

Por lo tanto, se puede generar una matriz a partir de un número hiperbólico considerando la matriz identidad I_2 para la parte real y la de propiedad hiperbólica establecida previamente para la parte imaginaria. Teniendo un número hiperbólico $z = x + \mathbf{k}y$ este se puede reescribir como sigue:

$$\begin{aligned} xI_2 + y\mathbf{M} &= x\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + y\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} + \begin{pmatrix} 0 & y \\ y & 0 \end{pmatrix} \\ &= \begin{pmatrix} x & y \\ y & x \end{pmatrix} \end{aligned}$$

Ejemplo

Sea el número hiperbólico $z = 9 + 3.5\mathbf{k}$, entonces se tiene que $x = 9$ y $y = 3.5$ su representación matricial es

$$\begin{pmatrix} 9 & 3.5 \\ 3.5 & 9 \end{pmatrix}$$

Una de las mayores ventajas de esta representación matricial es la facilidad a la que se reducen las operaciones entre dos números hiperbólicos, puesto a que se reduce a la suma, multiplicación o inversa de matrices.

3.1.4. Conjugado

Definición 13. *Siendo un número hiperbólico de la forma $z = x+\mathbf{k}y$, su conjugado estará denotado por $z^\dagger = x - \mathbf{k}y$.*

Ejemplo
Sea $z = 7 + 5\mathbf{k}$, su conjugado esta denotado por $z^\dagger = 7 - 5\mathbf{k}$.

Observación 1. *Así pues, nótese que la multiplicación de un número por su conjugado estará definido de la siguiente manera:*

$$\begin{aligned} zz^\dagger &= (x + \mathbf{k}y)(x - \mathbf{k}y) \\ &= x^2 - x\mathbf{k}y + x\mathbf{k}y - \mathbf{k}^2y^2 \\ &= x^2 - \mathbf{k}^2y^2,\ \mathbf{k}^2 = 1 \\ &= x^2 - y^2 \end{aligned}$$

Se puede comprobar que se llega a lo mismo usando la regla de multiplicación entre dos números usada previamente. Así pues, podemos denotar este resultado como el valor absoluto o magnitud de un número hiperbólico, como se muestra a continuación:

$$|z|^2_{hyp} = x^2 - y^2 \in \mathbb{R} \tag{3.9}$$

Ejemplo
Sea $z = 2 + 4\mathbf{k}$ y su conjugado $z^\dagger = 2 - 4\mathbf{k}$. La multiplicación se desarrolla a contiuación:

$$\begin{aligned} zz^\dagger &= (2 + 4\mathbf{k})(2 - 4\mathbf{k}) \\ &= 2^2 - (2)4\mathbf{k} + (2)4\mathbf{k} - \mathbf{k}^2 4^2 \\ &= 4 - 16\mathbf{k}^2,\ \mathbf{k}^2 = 1 \\ &= 4 - 16 \\ &= -12 \end{aligned}$$

Así entonces podemos denotar que cuando el valor de y sea mayor que el de x, tendremos como resultado una magnitud negativa. Los valores absolutos negativos son posibles dentro de los números hiperbólicos.

$$|z|^2_{hyp} = -12 \in \mathbb{R}$$

De esta forma podemos deducir que cuando $x > y$ se tiene que $|z|^2_{hyp} \in \mathbb{R}^+$. El caso del primer ejemplo.

Sin embargo, el caso interesante a analizar es cuando x = y, puesto que de aquí se deriva la teoría siguiente:

Definición 14. *Un número hiperbólico es divisor de cero si satisface que:*

1) $x, y \in \mathbb{R}$ *y ambos son distintos de cero.*

2) $x^2 - y^2 = 0$*, esto se cumple cuando* $x = \pm y$*.*

Así, si $z = x + \mathbf{k}y$ es un hiperbólico divisor de cero, y al mismo tiempo su conjugado $z^\dagger = x - \mathbf{k}y$ tiene los coeficientes reales distintos de cero. Entonces su multiplicación satisface que $zz^\dagger = 0$

Ejemplo: Sea $z = 4 + 4\mathbf{k}$, su conjugado esta denotado por $z^\dagger = 4 - 4\mathbf{k}$. Podemos comprobar que z y $z^\dagger$ son divisores de cero puesto que cumplen con:

1) $x = 4,\ y = 4$. Ambos en el conjunto $\mathbb{R}$ y distintos de cero.

2) $4^2 - 4^2 = 0$, esto se cumple puesto que x y y son ambos 4

Asi pues se satisface:

$$\begin{aligned}
zz^\dagger &= (4 + 4\mathbf{k})(4 - 4\mathbf{k}) \\
&= 4^2 - (4)4\mathbf{k} + (4)4\mathbf{k} - \mathbf{k}^2 4^2 \\
&= 16 - 16\mathbf{k} + 16\mathbf{k} + 16\mathbf{k}^2,\ \mathbf{k}^2 = 1 \\
&= 16 - 16 \\
&= 0
\end{aligned}$$

Por lo tanto cuando $x^2 = y^2$ se puede generalizar a

$$z = \lambda(1 \pm \mathbf{k}),\ \lambda \in \mathbb{R} \tag{3.10}$$

3.1.5. Idempotencia

Definición 15. *Un número hiperbólico idempotente es aquel que, retomando (14) se establece* $z = e$ *y* $\lambda = \frac{1}{2}$ *Así:*

$$e = \frac{1}{2}(1 + \mathbf{k}) \qquad\qquad e^\dagger = \frac{1}{2}(1 - \mathbf{k}) \tag{3.11}$$

Entonces cumple con:

$$\begin{aligned}
ee^{\dagger} &= \left(\frac{1}{2}+\frac{k}{2}\right)\left(\frac{1}{2}-\frac{k}{2}\right)\\
&= \left(\frac{1}{2}\right)^{2}-\left(\frac{1}{2}\right)\left(\frac{k}{2}\right)+\left(\frac{1}{2}\right)\left(\frac{k}{2}\right)-\left(\frac{k}{2}\right)^{2}\\
&= \frac{1}{4}-\frac{k^2}{4}\\
&= \frac{1}{4}-\frac{1}{4}(k^2)\\
&= \frac{1}{4}-\frac{1}{4}\\
&= 0
\end{aligned}$$

Análogamente se concluye que $e^{\dagger}e = 0$

Observación 2. *Nótese que cualquier idempotente es divisor de cero.*

Así pues los idempotentes satisfacen también:

$$\begin{aligned}
e^2 = ee &= \left(\frac{1}{2}+\frac{k}{2}\right)\left(\frac{1}{2}+\frac{k}{2}\right)\\
&= \left(\frac{1}{2}\right)^{2}+2\left(\frac{1}{2}\right)\left(\frac{k}{2}\right)+\left(\frac{k}{2}\right)^{2}\\
&= \frac{1}{4}+\frac{k}{2}+\left(\frac{1}{4}\right)k^2\\
&= \frac{1}{4}+\frac{k}{2}+\left(\frac{1}{4}\right)\\
&= \frac{1}{2}+\frac{k}{2}\\
&= e
\end{aligned}$$

Con base en las demostraciones anteriores se pueden obtener las siguientes afirmaciones, entre otras:

- $e^n = e,\ n \in \mathbb{R}$
- $(e^{\dagger})^2 = e^{\dagger}$
- $e^2 - e = 0$
- $e(e-1) = 0$

3.1.6. Fórmula de Euler para hiperbólicos

De la misma manera en la que se desarrolla la fórmula de Euler para los números complejos, es posible su desarrollo para los números hiperbólicos.
Es necesario considerar para ello los senos y cosenos hiperbólicos, sabiendo que, mediante las series de Maclaurin estos se expresan:

$$\cosh x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^4}{4!} + \dots \qquad \operatorname{senh} x = 1 + \frac{x}{1!} + \frac{x^3}{3!} + \frac{x^5}{5!} + \dots$$

Mediante las series de Taylor se sabe que para todo $\phi \in \mathbb{R}$

$$e^{\mathbf{k}\phi} = 1 + \frac{\mathbf{k}\phi}{1!} + \frac{\mathbf{k}^2\phi^2}{2!} + \frac{\mathbf{k}^3\phi^3}{3!} + \dots$$

Y usando las propiedades de la unidad imaginaria $\mathbf{k}$ se tiene que cuando es elevada a una potencia par $\mathbf{k}^{2,4,6,8,\cdots} = 1$ y cuando se eleva a una potencia impar $\mathbf{k}^{1,3,5,7,\cdots} = \mathbf{k}$, entonces es posible reescribirlo como

$$e^{\mathbf{k}\phi} = \left(1 + \frac{\phi^2}{2!} + \frac{\phi^4}{4!} + \frac{\phi^6}{6!} \dots\right) + \mathbf{k}\left(1 + \frac{\phi}{1!} + \frac{\phi^3}{3!} + \frac{\phi^5}{5!} + \dots\right)$$

Lo cual hace referencia directa con las series previamente mencionadas de $\cosh x$ y $\sinh x$ cuando $x = \phi$ para todo $\phi \in \mathbb{R}$

$$e^{\mathbf{k}\phi} = \cosh\phi + \mathbf{k}\sinh\phi$$

Capítulo 4

Bases para Números Hiperbólicos

Definición 16. *Para comprender las bases es necesario aclarar que desde el principio de este escrito se ha empleado la* ***base común*** *que denota a un número hiperbólico en su forma en la ecuación (3.1)*

Definición 17. *Tomando en cuenta los números idempotentes, se obtiene la* ***base idempotente****, la cual se define con base en la base común de la forma (3.1) de la siguiente manera*

$$\begin{aligned} Z &= z_1 e + z_2 e^{\dagger} \\ donde: z_1 &= x + y \\ z_2 &= x - y \end{aligned} \tag{4.1}$$

Así entonces, es posible la transformación de una base a otra, A continuación se muestra el cambio de base idempotente a la base común:

$$\begin{aligned} Z &= z_1 e + z_2 e^{\dagger} \\ &= (x+y)\mathbf{e} + (x-y)\mathbf{e}^{\dagger} \\ &= (x+y)\left(\frac{1}{2} + \frac{k}{2}\right) + (x-y)\left(\frac{1}{2} - \frac{k}{2}\right) \\ &= x + ky \end{aligned}$$

■

Ejemplo: Teniendo en base común

$$z = 2 + 6k,\ x = 2,\ y = 6$$

Entonces transformando de base común a base idempotente el número quedaría como

$$\begin{aligned} Z =&(2+6)\mathbf{e}+(2-6)\mathbf{e}^{\dagger} \\ =&8\mathbf{e}-4\mathbf{e}^{\dagger} \end{aligned} \tag{4.2}$$

Para ir de la base idempotente a la base común.

$$\begin{aligned} Z &= 8\mathbf{e}-4\mathbf{e}^{\dagger} \\ &= 8\left(\frac{1}{2}+\frac{k}{2}\right)-4\left(\frac{1}{2}-\frac{k}{2}\right) \\ &= \left(\frac{8}{2}\right)+\left(\frac{8k}{2}\right)-\left(\frac{4}{2}\right)+\left(\frac{4k}{2}\right) \\ &= \frac{4}{2}+\frac{12k}{2} \\ &= 2+6k \end{aligned}$$

4.1. Operaciones con base idempotente

Así como se definen las operaciones para los números hiperbólicos en su base normal, se definen las operaciones básicas cuando el hiperbólico está en la base idempotente.
Considerando dos números hiperbólicos

$$z_1 = x_1 + y_1 k \qquad\qquad z_2 = x_2 + y_2 k$$

cuya transformación en base idempotente es

$$\begin{aligned} Z_1 &= \alpha_1\mathbf{e}+\alpha_2\mathbf{e}^{\dagger} \\ Z_2 &= \omega_1\mathbf{e}+\omega_2\mathbf{e}^{\dagger} \end{aligned}$$

Considerando $\alpha_1 = x_1+y_1, \alpha_2 = x_1-y_1$ y $\omega_1 = x_2+y_2, \omega_2 = x_2-y_2$, las operaciones entre Z_1 y Z_2 se describen a continuación

- Suma

$$Z_3 = Z_1 + Z_2 = (\alpha_1+\omega_1)\mathbf{e}+(\alpha_2+\omega_2)\mathbf{e}^{\dagger} \tag{4.3}$$

  Demostración:
  Se debe comprobar que la transformación de Z_3 a su base normal resulta en

la ecuación [3.6]

$$
\begin{aligned}
&Z_3 = Z_1 + Z_2\\
&= (\alpha_1 + \omega_1)\mathbf{e} + (\alpha_2 + \omega_2)\mathbf{e}^\dagger\\
&= (x_1 + y_1 + x_2 + y_2)\left(\frac{1}{2} + \frac{k}{2}\right) + (x_1 - y_1 + x_2 - y_2)\left(\frac{1}{2} - \frac{k}{2}\right)\\
&= (x_1 + y_1 + x_2 + y_2)\left(\frac{1}{2} + \frac{k}{2}\right) + (x_1 - y_1 + x_2 - y_2)\left(\frac{1}{2} - \frac{k}{2}\right)\\
&= \frac{x_1 + y_1 + x_2 + y_2}{2} + \frac{kx_1 + ky_1 + kx_2 + ky_2}{2} +\\
&\frac{x_1 - y_1 + x_2 - y_2}{2} + \frac{-kx_1 + ky_1 - kx_2 + ky_2}{2}\\
&= (x_1 + x_2) + (y_1 + y_2)\,k
\end{aligned}
$$

■

- Resta

$$Z_3 = Z_1 - Z_2 = (\alpha_1 - \omega_1)\mathbf{e} + (\alpha_2 - \omega_2)\mathbf{e}^\dagger \tag{4.4}$$

La demostración de la resta es análoga a la de la suma.

- Multiplicación

$$Z_3 = Z_1 Z_2 = (\alpha_1 \omega_1)\mathbf{e} + (\alpha_2 \omega_2)\mathbf{e}^\dagger \tag{4.5}$$

Demostración:
Se busca probar que la transformación de Z_3 a su base normal resulta en la ecuación [3.7].
Teniendo que

$$\alpha_1 \omega_1 = (x_1 + y_1)(x_2 + y_2) = x_1 x_2 + x_1 y_2 + x_2 y_1 + y_1 y_2$$

$$\alpha_2 \omega_2 = (x_1 - y_1)(x_2 - y_2) = x_1 x_2 - x_1 y_2 - x_2 y_1 + y_1 y_2$$

Se desglosa

$$\begin{aligned}
&Z_3 = Z_1 Z_2 \\
&= (\alpha_1\omega_1)\mathbf{e} + (\alpha_2\omega_2)\mathbf{e}^\dagger \\
&= (x_1x_2 + x_1y_2 + x_2y_1 + y_1y_2)\left(\frac{1}{2} + \frac{k}{2}\right) \\
&+ (x_1x_2 - x_1y_2 - x_2y_1 + y_1y_2)\left(\frac{1}{2} - \frac{k}{2}\right) \\
&= \frac{x_1x_2 + x_1y_2 + x_2y_1 + y_1y_2}{2} + \frac{kx_1x_2 + kx_1y_2 + kx_2y_1 + ky_1y_2}{2} \\
&+ \frac{x_1x_2 - x_1y_2 - x_2y_1 + y_1y_2}{2} + \frac{-kx_1x_2 + kx_1y_2 + kx_2y_1 - ky_1y_2}{2} \\
&= (x_1x_2 + y_1y_2) + k(x_1y_2 + x_2y_1)
\end{aligned}$$

■

- División

$$Z_3 = \frac{Z_1}{Z_2} = \left(\frac{\alpha_1}{\omega_1}\right)\mathbf{e} + \left(\frac{\alpha_2}{\omega_2}\right)\mathbf{e}^\dagger \tag{4.6}$$

La demostración de la división es análoga a la de la multiplicación.

- Potencia
A partir de la multiplicación, podemos derivar la potencia de cualquier número, asumiendo el caso particular donde $Z_1 = Z_2$ resultaría en

$$Z_1^2 = \alpha_1^2\mathbf{e} + \alpha_2^2\mathbf{e}^\dagger$$

Y generalizando la potencia a cualquier exponente n, se tiene que

$$Z_1^n = \alpha_1^n\mathbf{e} + \alpha_2^n\mathbf{e}^\dagger \tag{4.7}$$

Demostración:
Es posible demostrar la ecuación [4.7] si consideramos que $A = Z_1^2$ y $B = Z_1$, así entonces el producto entre A y B es

$$\begin{aligned}
AB &= Z_1^2 Z_1 = Z_1^3 \\
&= \left(\alpha_1^2\alpha_1\right)\mathbf{e} + \left(\alpha_2^2\alpha_2\right)\mathbf{e}^\dagger \\
&= \alpha_1^3\mathbf{e} + \alpha_2^3\mathbf{e}^\dagger
\end{aligned}$$

Y se cumple la ecuación [4.7] para $n = 3$. Se puede seguir entonces que para $n = 4$, sea $C = Z_1^3$, el producto de B y C resulta en

$$\begin{aligned} CB &= Z_1^3 Z_1 = Z_1^4 \\ &= \left(\alpha_1^3 \alpha_1\right) \mathbf{e} + \left(\alpha_2^3 \alpha_2\right) \mathbf{e}^\dagger \\ &= \alpha_1^4 \mathbf{e} + \alpha_2^4 \mathbf{e}^\dagger \end{aligned}$$

Por lo que para $n = 4$ se cumple la regla de potencia. De esta forma se puede seguir la demostración por inducción matemática para todo $n \in \mathbb{R}$.

4.2. Forma matricial para base idempotente

De la misma manera en la que cualquier hiperbólico en su base normal se puede expresar en forma matricial, es posible expresar la base idempotente también. Considerando a $\mathbf{e}$ entonces se tiene que su expresión matricial es

$$\mathbf{E} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \tag{4.8}$$

Y para $\mathbf{e}^\dagger$ si expresión matricial estaría dada por

$$\mathbf{E}^\dagger = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \tag{4.9}$$

Así entonces un número hiperbólico en base idempotente $Z = z_1 e + z_2 e^\dagger$ en la ecuación (4.1) con ayuda de las matrices $\mathbf{E}$ y $\mathbf{E}^\dagger$ se transformaría a

$$Z = z_1 \mathbf{E} + z_2 \mathbf{E}^\dagger \tag{4.10}$$

Demostración:
De la misma manera, es posible transformar de la base idempotente y llegar a la

representación matricial en la base normal

$$\begin{aligned}
Z =& z_1\mathbf{E} + z_2\mathbf{E}^\dagger \\
=& (x+y)\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} + (x-y)\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \\
=& \begin{pmatrix} \frac{x+y}{2} & \frac{x+y}{2} \\ \frac{x+y}{2} & \frac{x+y}{2} \end{pmatrix} + \begin{pmatrix} \frac{x-y}{2} & -\frac{x-y}{2} \\ -\frac{x-y}{2} & \frac{x+y}{2} \end{pmatrix} \\
=& \begin{pmatrix} \frac{(x+y)+(x-y)}{2} & \frac{(x+y)-(x-y)}{2} \\ \frac{(x+y)-(x-y)}{2} & \frac{(x+y)+(x-y)}{2} \end{pmatrix} \\
=& \begin{pmatrix} \frac{2x}{2} & \frac{2y}{2} \\ \frac{2y}{2} & \frac{2x}{2} \end{pmatrix} \\
=& \begin{pmatrix} x & y \\ y & x \end{pmatrix}
\end{aligned}$$

■

La principal ventaja de la notación idempotente en las matrices recae principalmente en la multiplicación entre matrices 2x2, puesto que el proceso tradicional se simplifica simplemente a multiplicación de los coeficientes z_1 y z_2 de la matriz en base idempotente. Si posteriormente se desea saber los valores en base normal de la matriz resultante, esto es posible simplemente dividiendo entre dos la suma de los coeficientes para los valores en la diagonal principal y dividiendo entre dos la resta para los coeficientes en la diagonal secundaria. Dicho de otra manera, sea

$$M_1 = \begin{pmatrix} x_1 & y_1 \\ y_1 & x_1 \end{pmatrix} \qquad M_2 = \begin{pmatrix} x_2 & y_2 \\ y_2 & x_2 \end{pmatrix}$$

Entonces sus representaciones hiperbólicas matriciales en base idempotente son

$$\begin{aligned}
m_1 &= (x_1+y_1)\mathbf{E} + (x_1-y_1)\mathbf{E}^\dagger \\
m_2 &= (x_2+y_2)\mathbf{E} + (x_2-y_2)\mathbf{E}^\dagger
\end{aligned}$$

Y la multiplicación entre m_1 y m_2 sería

$$m_3 = m_1m_2 = (x_1+y_1)(x_2+y_2)\mathbf{E} + (x_1-y_1)(x_2-y_2)\mathbf{E}^\dagger$$

Para simplicidad consideramos $a = (x_1+y_1)(x_2+y_2)$ y $b = (x_1-y_1)(x_2-y_2)$ Ahora, si se desea encontrar la forma matricial de la matriz producto se tiene que

$$x_3 = \frac{a+b}{2} \qquad y_3 = \frac{a-b}{2}$$

Y por lo tanto se tendría finalmente $M_3 = \begin{pmatrix} x_3 & y_3 \\ y_3 & x_3 \end{pmatrix}$

A continuación se demuestra que el proceso tradicional de productos matriciales concuerda con el desarrollo previamente mencionado. Considerando que

$$M_3 = \begin{pmatrix} x_1 & y_1 \\ y_1 & x_1 \end{pmatrix} \begin{pmatrix} x_2 & y_2 \\ y_2 & x_2 \end{pmatrix} = \begin{pmatrix} x_3 & y_3 \\ y_3 & x_3 \end{pmatrix}$$

Entonces se tiene que en la diagonal principal

$$\begin{aligned} x_3 = \frac{a+b}{2} &= \frac{(x_1+y_1)(x_2+y_2)+(x_1-y_1)(x_2-y_2)}{2} \\ &= \frac{x_1x_2+x_1y_2+y_1x_2+y_1y_2+x_1x_2-x_1x_2-y_1x_2+y_1y_2}{2} \\ &= \frac{2x_1x_2+2y_1y_2}{2} \\ &= x_1x_2+y_1y_2 \end{aligned}$$

Y para los valores en la diagonal secundaria

$$\begin{aligned} y_3 = \frac{a-b}{2} &= \frac{(x_1+y_1)(x_2+y_2)-(x_1-y_1)(x_2-y_2)}{2} \\ &= \frac{x_1x_2+x_1y_2+y_1x_2+y_1y_2+x_1x_2-x_1x_2-y_1x_2+y_1y_2}{2} \\ &= \frac{2x_1y_2+2y_1x_2}{2} \\ &= x_1y_2+y_1x_2 \end{aligned}$$

Por lo tanto se llega al mismo resultado al que se llegaría con la multiplicación tradicional de dos matrices 2x2.
Mediante un ejemplo se puede apreciar mejor la simplicidad del método de representación hiperbólica matricial en base idempotente

Ejemplo:
Sean $M_1 = \begin{pmatrix} 7 & 9.5 \\ 9.5 & 7 \end{pmatrix}$ y $M_2 = \begin{pmatrix} 5 & 0.8 \\ 0.8 & 5 \end{pmatrix}$, entonces su representación hiperbólica matricial en base idempotente es

$$m_1 = 16.5\mathbf{E} - 2.5\mathbf{E}^\dagger \qquad\qquad m_2 = 5.8\mathbf{E} + 4.2\mathbf{E}^\dagger$$

Y por lo tanto la multiplicación resultaría en

$$m_3 = 16.5 * 5.8\mathbf{E} - 2.5 * 4.2\mathbf{E}^\dagger = 95.7\mathbf{E} - 10.5\mathbf{E}^\dagger$$

Y volviendo a la base normal de la matriz se tiene que $x_3 = \frac{95.7+10.5}{2} = 53.1$ y $y_3 = \frac{95.7-10.5}{2} = 42.6$, por lo cual $M_3 = \begin{pmatrix} 53.1 & 42.6 \\ 42.6 & 53.1 \end{pmatrix}$

4.3. Axioma del supremo en hiperbólicos

De la misma manera en a que previamente se expuso este axioma aplicado a los números reales, es importante entender sus implicaciones dentro de los números hiperbólicos, ello servirá para demostrar que de la misma manera son un conjunto ininterrumpido de números donde no hay huecos.

Si se tiene un intervalo $\mathbf{A} \subset \mathbb{D}$ acotado superior e inferiormente en $\mathbb{D}$, entonces encontramos que existen para este conjuntos de cotas $\mathbb{D}$-superiores y $\mathbb{D}$- inferiores cuando, considerando un número $\tau \in \mathbf{A}$, se compara τ de forma que τ es mayor que todas las cotas inferiores, y al mismo tiempo es menor que todas las cotas superiores. Conforme a lo anterior encontramos las siguientes definiciones:

Definición 18. *Si $\mathbb{A} \in \mathbb{D}$ está acotado superiormente, se le entiende por supremo a aquel número menor de todas las cotas $\mathbb{D}$-superiores. Se le denota como* $sup_{\mathbb{D}}\mathbf{A} = sup\mathbf{A}_1 e + sup\mathbf{A}_2\mathbf{e}^{\dagger}$

Definición 19. *Si $\mathbb{A} \in \mathbb{D}$ está acotado inferiormente, se le entiende por ínfimo a aquel número mayor de todas las cotas $\mathbb{D}$-inferiores. Se le denota como* $inf_{\mathbb{D}}\mathbf{A} = inf\mathbf{A}_1 e + inf\mathbf{A}_2\mathbf{e}^{\dagger}$.

4.4. Orden parcial en los números hiperbólicos

Es posible encontrar una similitud entre los numeros no negativos dentro de $\mathbb{D}$ y los $\mathbb{R}$, y esta es una buena razón para asumir que existe cierto orden en $\mathbb{D}$ similar al que encontramos en $\mathbb{R}$.

Considerando las características de orden expuestas en alguna sección previa, es importante estabecer la existencia de un orden parcial dentro de lo números hiperbólicos, mediante la comparabilidad de sus elementos. Esto es, si consideramos dos números en $\mathbb{D}$, siendo estos:

$$Z_1 = \beta_1 e + \beta_2 e^{\dagger} \qquad\qquad Z_2 = \gamma_1 e + \gamma_2 e^{\dagger}$$

Donde $\beta_1,\ \beta_2,\ \gamma_1,\ \gamma_2 \in \mathbb{R}$.

Entonces decimos que existe una relación de orden parcial entre estos dos números cuando se cumple que $Z_2 - Z_1 \in \mathbb{D}^+$. En otras palabras se dice que $Z_2 \succeq Z_1$ ó $Z_1 \preceq Z_2$ cuando se tiene que:

$$\gamma_1 \geq \beta_1, \qquad\qquad \gamma_2 \geq \beta_2$$

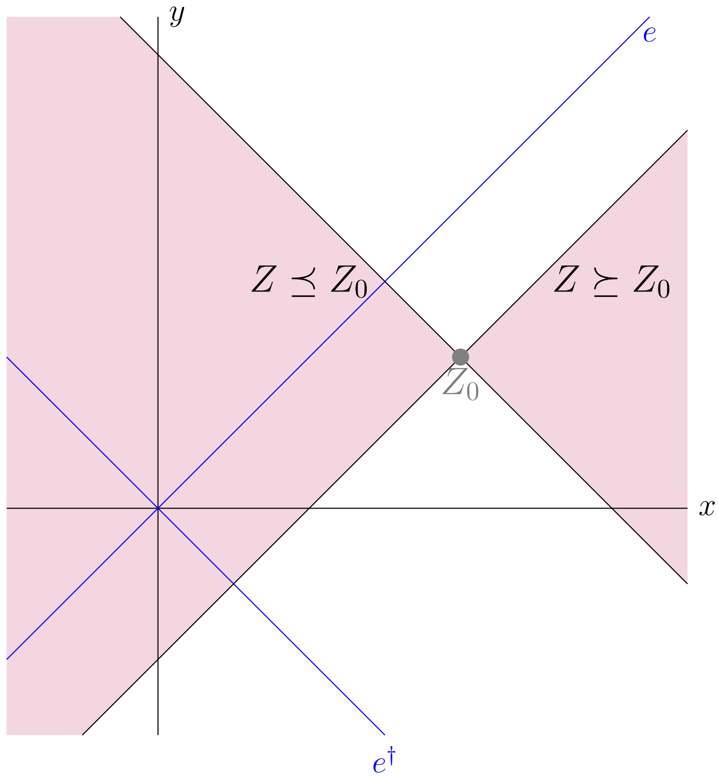

Figura 4.1: Orden parcial de un hiperbólico

En la figura previa se observa una relación de orden parcial del número Z_0 respecto al resto del plano hiperbólico.

Ejemplo: Siendo

$$Z_1 = 7e + 9e^{\dagger} \qquad\qquad Z_2 = 9e + 11e^{\dagger}$$

donde $\beta_1 = 7,\ \beta_2 = 9,\ \gamma_{1=9},\ \gamma_2 = 11, \in \mathbb{R}$. Afirmamos que existe una relación de orden pacial entre Z_1 y Z_2 puesto que se cumple:

$$\begin{aligned} Z_2 - Z_1 &= 9e + 11e^{\dagger} - \left(7e + 9e^{\dagger}\right) \\ &= (9-7)e + (11-9)e^{\dagger} \\ &= 2e + 2e^{\dagger} \in \mathbb{D}^{+} \end{aligned}$$

Puesto que

$$9 \geq 7 \qquad\qquad 11 \geq 9$$

Por lo tanto se concluye que $Z_2 \succeq Z_1$

Así pues considerando cualquier otro número $z_0 = a+\mathbf{k}b \in \mathbb{D}$, y tomando como base la estructura de los números hiperbólicos desarrollados en la base idempotente es posible encontrar que el plano hiperbólico se secciona en cuatro partes, siendo el nuevo eje de referencia el número z. Este plano estará basado en la comparabilidad de dos números hiperbólicos como sigue:

- Si $Z_0 \geq Z$, entonces se dice que este cuadrante será el de todo número hiperbólico que sea mayor o igual que Z_0
- Si $Z_0 \leq Z$, entonces se dice que este cuadrante será el de todo número hiperbólico que sea menor o igual que Z_0
- Aquellos que no sean comparables con Z en $\mathbb{D}$ corresponden a los dos cuadrantes restantes.

4.4.1. Propiedades del orden parcial

La comparabilidad de los números hiperbólicos involucra ciertas propiedades del orden parcial $\preceq$. Tomando α, β, y γ como números hiperbólicos se muestran estas a continuación.

- Si α y β con comparables en el orden parcial $\preceq$, entonces se cumple una de las siguientes condiciones:

$$\begin{aligned} \alpha \prec \beta \\ \alpha \succ \beta \\ \alpha = \beta \end{aligned} \tag{4.11}$$

- Si $\alpha \prec \beta$ y $\beta \preceq \gamma$, entonces $\alpha \prec \gamma$
- Si $\alpha \preceq \beta$ y $\beta \prec \gamma$, entonces $\alpha \prec \gamma$

4.4.2. Conexiones de suma

- $\alpha \prec \gamma$ implica que $\alpha + \beta \prec \gamma + \beta$
- $\alpha \preceq \gamma$ implica que $\alpha + \beta \preceq \gamma + \beta$
- $0 \prec \alpha$ implica que $-\alpha \prec 0$
- $\alpha_1 \preceq \alpha_2$ y $\beta \preceq \gamma$ implica que $\alpha_1 + \beta \preceq \alpha_2 + \gamma$
- $\alpha_1 \preceq \alpha_2$ y $\beta \prec \gamma$ implica que $\alpha_1 + \beta \prec \alpha_2 + \gamma$

4.4.3. Conexiones de multiplicación

- Si α y β son hiperbólicos no negativos de la forma 3.2 su producto:

$$\alpha \cdot \beta \in \mathbb{D}^+$$

- Si α y β son hiperbólicos positivos de la forma 3.3 su producto:

$$\alpha \cdot \beta \in \mathbb{D}^+ \setminus \{0\}$$

- Si α y β son hiperbólicos negativos de la forma 3.5 su producto será estrictamente positivo :

$$\alpha \cdot \beta \succ 0$$

- Si α es un número hiperbólico positivo y β es negativo su producto será estrictamente negativo:

$$\alpha \cdot \beta \prec 0 \tag{4.12}$$

- Si $\alpha \prec \beta$ y $\gamma \succ 0$ entonces $\alpha \cdot \gamma \prec \beta \cdot \gamma$

- Si $\alpha \prec \beta$ y $\gamma \prec 0$ entonces $\alpha \cdot \gamma \prec \beta \cdot \gamma$

- Si $\alpha \succ 0$ y $\alpha \succ \beta$ entonces $\beta^{-1} \succ 0$ y $\beta^{-1} \prec \alpha^{-1}$. Esto dice que si α es un numero positivo , entonces también lo son su invertible e inverso

Capítulo 5

Espacio de probabilidad hiperbólica

En la sección de Teoría de la Medida, se trataron las definiciones para espacios medibles y espacios de medida, así pues, esto es base principal para definir un espacio de probabilidad.

Definición 20. *Un espacio de probabilidad se define por la tripleta* $(\Omega, \sigma, \mathbb{P})$*.*

A continuación se explica cada elemento de la tripleta:

Ω Considerado como el universo, en el cual se toma en cuenta todos los resultados posibles de un experimento. Este universo es llamado espacio muestral.

σ Se refiere a una $\sigma-$álgebra de la colección de subconjuntos de ω

$\mathbb{P}$ Se define como la medida de probabilidad, la cual es una función que asigna un número no negativo a cualquier elemento de σ

Entendiendo mejor los números hiperbólicos y sus propiedades, nos adentramos en una de sus aplicaciones de interés: la probabilidad. Consideremos entonces un espacio (Ω, Σ) donde la función:

$$\mathbf{P}_{\mathbb{D}} : \Sigma \longrightarrow \mathbb{D}$$

es llamada una probabilidad $\mathbb{D}$-valuada si cumple con los siguientes puntos:

(i) $\mathbf{P}_{\mathbb{D}(\mathbf{A})} \succeq 0$

(ii) $\mathbf{P}_{\mathbb{D}(\mathbf{\Omega})} = p$ donde p toma tres posibles valores: $1, e$ ó $e^{\dagger}$

(iii) Dada una secuencia $\mathbf{A}_n \subset \Sigma$ de eventos mutuamente excluyentes:

$$\mathbf{P}_{\mathbb{D}}\left(\bigcup_{n=1}^{\infty}\mathbf{A}_n\right)=\sum_{n=1}^{\infty}\mathbf{P}_{\mathbb{D}}\left(\mathbf{A}_n\right)$$

Así pues, considerando que $\mathbf{P}_{\mathbb{D}}$ es un número hiérbólico, esto se denota como $\mathbf{P}_{\mathbb{D}}\left(A\right)\in\mathbb{D}$. El espacio de probabilidad $\mathbb{D}$ está definido por $(\Omega,\sigma,\mathbf{P}_{\mathbb{D}})$. Donde una variable aleatoria hiperbólica o variable $\mathbb{D}$-aleatoria es una función:

$$\mathbf{X}_{\mathbb{D}}:\Omega\longrightarrow\mathbb{D} \tag{5.1}$$

Retomando las propiedades de cambio de base que se establecieron previamente, se establece que las probabilidades $\mathbb{D}$-valuadas podrán ser escritas tanto como en la base normal como en la idempotente.

$$\mathbf{P}_{\mathbb{D}}\left(A\right)=p_1\left(A\right)+p_2\left(A\right)\mathbf{k}=P_1\left(A\right)\mathbf{e}+P_2\left(A\right)\mathbf{e}^{\dagger} \tag{5.2}$$

Conforme a la conversión de bases:

$$P_1\left(A\right)=p_1\left(A\right)+p_2\left(A\right) \qquad P_2\left(A\right)=p_1\left(A\right)-p_2\left(A\right) \tag{5.3}$$

Y considerando la propiedad (**i**), establecemos que $P_1\left(A\right)$ y $P_2\left(A\right)$ son ambos mayores o iguales a cero.
Así pues, la propiedad (**ii**) nos da como resultado:

$$\mathbf{P}_{\mathbb{D}}\left(\Omega\right)=p=P_1\left(\Omega\right)\mathbf{e}+P_2\left(\Omega\right)\mathbf{e}^{\dagger} \tag{5.4}$$

Asumiendo que:

* Cuando $p=1$, entonces $P_1=1, P_2=1$

* Cuando $p=\mathbf{e}$, entonces $P_1=1, P_2=0$

* Cuando $p=\mathbf{e}^{\dagger}$ entonces $P_1=0, P_2=1$

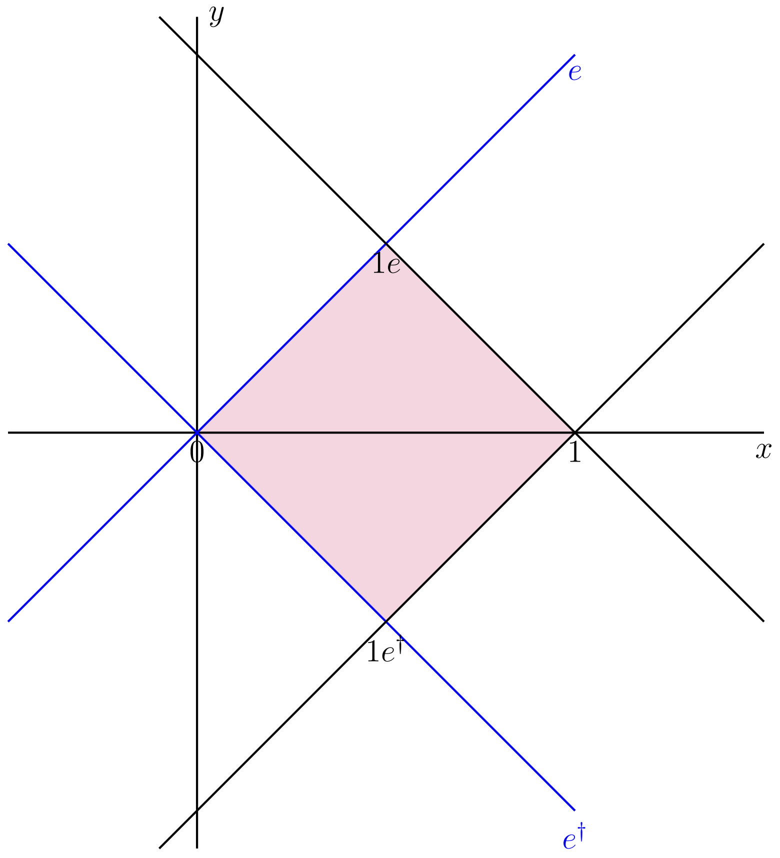

Figura 5.1: Espacio de probabilidad hiperbólica

En la figura previa es posible observar al intervalo que se genera en el plano hiperbólico considerando la base idempotente, el cual genera nuestro espacio de probabilidad.

Ejemplo:
Supongamos que tenemos la probabilidad hiperbólica $\mathbf{P}_{\mathbb{D}}(A) = 0.53 + 0.15\mathbf{k}$ Así entonces se tendría que $p_1(A) = 0.53$ y $p_2(A) = 0.15$, por lo tanto conforme a la conversión de base mostrada en 5.6

$$P_1(A) = 0.53 + 0.15 = 0.68 \qquad P_2(A) = 0.53 - 0.15 = 0.38 \tag{5.5}$$

Entonces la probabilidad original puede ser rescrita como $\mathbf{P}_{\mathbb{D}}(A) = 0.68e + 0.38e^{\dagger}$

5.1. En relación con la probabilidad real

Además, en relación con la probabilidad tradicional nótese que cuando en la ecuación (15) se tiene que $p_1(A) = \alpha$ y $p_2 = 0$, obtenemos caso específico de la probabilidad real tradicional, dónde α es un número no negativo entre cero y uno. Podemos confirmar que la probabilidad hiperbólica es una generalización de la teoría tradicional de probabilidad real.

Observación 3. *Toda probabilidad real es una probabilidad hiperbólica, pero no toda hiperbólica es una real.*

Además, basándonos en las ecuaciones (16) para la transformación a la base idempotente, observamos que

$$P_1(A) = \alpha + 0 \qquad\qquad P_2(A) = \alpha - 0 \tag{5.6}$$

Entonces, en particular podemos remarcar que las probabilidades reales transformadas a la base idempotente son de la forma.

$$\mathbf{P}_{\mathbb{D}}(A) = \alpha\mathbf{e} + \alpha\mathbf{e}^{\dagger}$$

Ejemplo:
Considerando la probabilidad real $\mathbf{P}(A) = 0.42 = 0.42 + 0\mathbf{k}$, y por lo tanto se tendría que $P_1(A) = 0.42{+}0$ y $P_2(A) = 0.42{-}0$. Esto resulta en que la probabilidad real puede ser transformada a una hiperbólica con base idempotente de la forma $\mathbf{P}_{\mathbb{D}}(A) = 0.42e + 0.42e^{\dagger}$

5.2. Propiedades de medidas de probabilidad hiperbólicas

Entre otras propiedades importantes de las medidas de probabilidad se encuentran:

1. Dado un evento y su complemento tal que $A, A^c \in \sigma$, se tiene que la suma de probabilidades hiperbólicas de dicho evento y su complemento es igual a p Esto es
$$\mathbf{P}_{\mathbb{D}}(A) + \mathbf{P}_{\mathbb{D}}(A^c) = p \tag{5.7}$$

2. La probabilidad hiperbólica del vacio es igual a cero
$$\mathbf{P}_{\mathbb{D}}(\emptyset) = 0 \tag{5.8}$$

3. Teniendo dos eventos $A, B \in \sigma$ tal que $A \subset B$, entonces sus probabilidades son comparables respecto a un orden parcial, lo que es
$$\mathbf{P}_{\mathbb{D}}(A) \preceq \mathbf{P}_{\mathbb{D}}(B) \tag{5.9}$$

4. El principio de inclusión y exclusión el cual plantea que, dada una colección de eventos $A_1, A_2, ..., A_n$, se tiene que

$$\mathbf{P}_{\mathbb{D}}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbf{P}_{\mathbb{D}}\left(A_i\right) - \sum_{1\leq i\leq j\leq n} \mathbf{P}_{\mathbb{D}}\left(A_i \cap A_j\right) \tag{5.10}$$

$$+ \sum_{1\leq i\leq j\leq k\leq n} \mathbf{P}_{\mathbb{D}}\left(A_i \cap A_j \cap A_k\right) + ... \tag{5.11}$$

$$+(-1)^{n-1}\mathbf{P}_{\mathbb{D}}\left(A_i \cap ... \cap A_n\right).$$

Demostración: Por inducción tenemos que para $n = 2$ y sabiendo que $A_1 \cup A_2 = A_1 \cup (A_1^c \cap A_2)$ tenemos en términos de probabilidad que:

$$\mathbf{P}_{\mathbb{D}}\left(A_1 \cup A_2\right) = \mathbf{P}_{\mathbb{D}}\left(A_1\right) + \mathbf{P}_{\mathbb{D}}\left(A_1^c \cap A_2\right)$$

Además $A_2 = (A_1 \cap A_2) \cup (A_1^c \cap A_2)$, por lo tanto
$\mathbf{P}_{\mathbb{D}}\left(A_1^c \cap A_2\right) = \mathbf{P}_{\mathbb{D}}\left(A_2\right) - \mathbf{P}_{\mathbb{D}}\left(A_1 \cap A_2\right)$, entonces:

$$\begin{aligned}\mathbf{P}_{\mathbb{D}}\left(A_1 \cup A_2\right) &= \mathbf{P}_{\mathbb{D}}\left(A_1\right) + \mathbf{P}_{\mathbb{D}}\left(A_1^c \cap A_2\right)\\ &= \mathbf{P}_{\mathbb{D}}\left(A_1\right) + \mathbf{P}_{\mathbb{D}}\left(A_2\right) - \mathbf{P}_{\mathbb{D}}\left(A_1 \cap A_2\right)\end{aligned} \tag{5.12}$$

Ahora, haciendo $n = m$ en la ecuación (5.10) tenemos

$$\mathbf{P}_{\mathbb{D}}\left(\bigcup_{i=1}^{m} A_i\right) = \sum_{i=1}^{m} \mathbf{P}_{\mathbb{D}}\left(A_i\right) - \sum_{1\leq i\leq j\leq m} \mathbf{P}_{\mathbb{D}}\left(A_i \cap A_j\right) \tag{5.13}$$

$$+ \sum_{1\leq i\leq j\leq k\leq m} \mathbf{P}_{\mathbb{D}}\left(A_i \cap A_j \cap A_k\right) + ... \tag{5.14}$$

$$+(-1)^{m-1}\mathbf{P}_{\mathbb{D}}\left(A_i \cap ... \cap A_m\right).$$

Asumimos entonces, que al ser válido para $n = m$, es válido para $n = m + 1$
Usando la demostración para $n = 2$ asumimos que $A_1 = A_1 \cup A_2 \cup ... \cup A_m$ y que $A_2 = A_{m+1}$, por lo tanto:

$$\mathbf{P}_{\mathbb{D}}\left(A_1 \cup A_2\right) = \mathbf{P}_{\mathbb{D}}\left(A_1 \cup A_2 \cup ... \cup A_m \cup A_{m+1}\right) = \mathbf{P}_{\mathbb{D}}\left(\bigcup_{i=1}^{m+1} A_i\right)$$

Aplicando lo obtenido en (5.12) tenemos que:

$$\begin{aligned}
\mathbf{P}_{\mathbb{D}}\left(\bigcup_{i=1}^{m+1} A_i\right) &= \mathbf{P}_{\mathbb{D}}\left(A_1 \cup A_2 \cup ... \cup A_m\right) + \mathbf{P}_{\mathbb{D}}\left(A_{m+1}\right) \\
&\quad -\mathbf{P}_{\mathbb{D}}\left(\left(A_1 \cup A_2 \cup ... \cup A_m\right) \cap A_{m+1}\right) \\
&= \mathbf{P}_{\mathbb{D}}\left(A_1 \cup A_2 \cup ... \cup A_m\right) + \mathbf{P}_{\mathbb{D}}\left(A_{m+1}\right) \\
&\quad -\mathbf{P}_{\mathbb{D}}\left(\left(A_1 \cup A_2 \cup ... \cup A_m\right) \cap A_{m+1}\right) \\
&= \mathbf{P}_{\mathbb{D}}\left(A_1 \cup A_2 \cup ... \cup A_m\right) + \mathbf{P}_{\mathbb{D}}\left(A_{m+1}\right) \\
&\quad -\mathbf{P}_{\mathbb{D}}\left(\left(A_1 \cap A_{m+1}\right) \cup \left(A_2 \cap A_{m+1}\right) \cup ... \right. \\
&\quad \left. \cup \left(A_m \cap A_{m+1}\right)\right) \\
&= \mathbf{P}_{\mathbb{D}}\left(\bigcup_{i=1}^{m} A_i\right) + \mathbf{P}_{\mathbb{D}}\left(A_{m+1}\right) - \mathbf{P}_{\mathbb{D}}\left(\bigcup_{i=1}^{m}\left(A_i \cap A_{m+1}\right)\right)
\end{aligned}$$

Retomando la ecuación (5.13) y sustituyendo para las uniones de conjuntos, tenemos que:

$$\begin{aligned}
\mathbf{P}_{\mathbb{D}}\left(\bigcup_{i=1}^{m+1} A_i\right) &= \sum_{i=1}^{m} \mathbf{P}_{\mathbb{D}}\left(A_i\right) && \text{(a)} \\
&- \sum_{1 \leq i \leq j \leq m} \mathbf{P}_{\mathbb{D}}\left(A_i \cap A_j\right) && \text{(b)} \\
&+ \sum_{1 \leq i \leq j \leq k \leq m} \mathbf{P}_{\mathbb{D}}\left(A_i \cap A_j \cap A_k\right) + \dots && \text{(c)} \\
&+ (-1)^{m-1} \mathbf{P}_{\mathbb{D}}\left(A_i \cap \cdots \cap A_m\right) && \text{(d)} \\
&+ \mathbf{P}_{\mathbb{D}}\left(A_{m+1}\right) && \text{(e)} \\
&- \sum_{i=1}^{m} \mathbf{P}_{\mathbb{D}}\left(\left(A_i \cap A_{m+1}\right)\right) && \text{(f)} \\
&+ \sum_{1 \leq i \leq j \leq m} \mathbf{P}_{\mathbb{D}}\left(\left(A_i \cap A_{m+1}\right) \cap A_j\right) && \text{(g)} \\
&- \sum_{1 \leq i \leq j \leq k \leq m} \mathbf{P}_{\mathbb{D}}\left(\left(A_i \cap A_{m+1}\right) \cap A_j \cap A_k\right) - \dots && \text{(h)} \\
&- (-1)^{m-1} \mathbf{P}_{\mathbb{D}}\left(\left(A_i \cap A_{m+1}\right) \cap \cdots \cap A_m\right)
\end{aligned}$$

A continuación, se procede a sumar los términos semejantes del desarrollo anterior:

- $\star$ La pareja de (a) y (b) se reduce a

$$\sum_{i=1}^{m+1} \mathbf{P}_{\mathbb{D}}\left(A_i\right)$$

- $\star$ La suma (b) contiene todas las intersecciones entre los conjuntos desde A_1 hasta A_m y la suma (f) contiene las intersecciones de todos los conjuntos con A_{m+1}. Así entonces, al juntar estas sumas se están considerando todas las intersecciones desde A_1 hasta A_{m+1}. Esto es:

$$-\sum_{1\leq i\leq j\leq m+1} \mathbf{P}_{\mathbb{D}}\left(A_i \cap A_j\right)$$

- $\star$ De la misma forma sucede para las sumas de tres intersecciones (c) y (g), las cuales juntas contienen todas las intersecciones de tres desde A_1 hasta A_{m+1}, resultando en la siguiente forma:

$$+\sum_{1\leq i\leq j\leq k\leq m+1} \mathbf{P}_{\mathbb{D}}\left(A_i \cap A_j \cap A_k\right) + ...$$

  Y así se continúa para las intersecciones de 4, 5, 6, ... conjuntos desde A_1 hasta A_{m+1}

- $\star$ Finalmente, para el caso de las sumas de intersecciones de m conjuntos desde A_1 hasta A_{m+1} conjuntos se unen (d) y (h). Quedando:

$$+(-1)^m \mathbf{P}_{\mathbb{D}}\left(A_i \cap ... \cap A_{m+1}\right)$$

Por lo tanto, al unir todo lo anterior, queda demostrado que para $n = m+1$:

$$\begin{aligned} \mathbf{P}_{\mathbb{D}}\left(\bigcup_{i=1}^{m+1} A_i\right) = & \sum_{i=1}^{m+1} \mathbf{P}_{\mathbb{D}}\left(A_i\right) - \sum_{1\leq i\leq j\leq m+1} \mathbf{P}_{\mathbb{D}}\left(A_i \cap A_j\right) \\ & + \sum_{1\leq i\leq j\leq k\leq m+1} \mathbf{P}_{\mathbb{D}}\left(A_i \cap A_j \cap A_k\right) + ... \\ & + (-1)^m \mathbf{P}_{\mathbb{D}}\left(A_i \cap ... \cap A_{m+1}\right). \end{aligned}$$

■

5. Generalizando lo anterior a una inequidad tenemos que dados n eventos $A_1, A_2, ..., A_n$, la probabilidad de su unión es comparable en un orden parcial con la suma de sus probabilidades, lo que es:

$$\mathbf{P}_{\mathbb{D}}\left(\bigcup_{i=1}^{n} A_i\right) \preceq \sum_{i=1}^{n} \mathbf{P}_{\mathbb{D}}\left(A_i\right) \tag{5.15}$$

6. Así mismo también se establece un Teorema de continuidad de probabilidad hiperbólica.
Sea $\{A_n : n \in \mathbb{N}\}$ una sucesión no decreciente de eventos $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots \subseteq A_n \subseteq \ldots$, entonces se tiene que

$$\mathbf{P}_{\mathbb{D}}\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mathbf{P}_{\mathbb{D}}\left(A_n\right) \tag{5.16}$$

Demostración: Sabemos que $\mathbb{P}\left(A_n\right) \leq \mathbb{P}\left(A_{n+1}\right)$ puesto que $A_n \subseteq An+1$. Entonces la sucesión numérica de las probabilidades de A_n es no decreciente y acotada superiormente por uno, por lo tanto el límite si existe. Se define

$$\begin{aligned} B_1 &= A_1 \\ B_m &= A_m - A_{m-1} \end{aligned}$$

Así entonces la sucesión $\{B_n : n \in \mathbb{N}\}$ es un conjunto de eventos disjuntos por pares, y entonces se tiene

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{m=1}^{\infty} B_m$$

Aplicando la probabilidad

$$\begin{aligned} \mathbf{P}_{\mathbb{D}}\left(\bigcup_{n=1}^{\infty} A_n\right) &= \mathbf{P}_{\mathbb{D}}\left(\bigcup_{n=1}^{\infty} B_m\right) \\ &= \sum_{m=1}^{\infty} \mathbf{P}_{\mathbb{D}}\left(B_m\right) \\ &= \mathbf{P}_{\mathbb{D}}\left(B_1\right) + \sum_{m=2}^{\infty} \mathbf{P}_{\mathbb{D}}\left(B_m\right) \\ &= \mathbf{P}_{\mathbb{D}}\left(A_1\right) + \sum_{m=2}^{\infty} \mathbf{P}_{\mathbb{D}}\left(A_m - A_{m-1}\right) \\ &= \mathbf{P}_{\mathbb{D}}\left(A_1\right) + \sum_{m=2}^{\infty} \left[\mathbf{P}_{\mathbb{D}}\left(A_m\right) - \mathbf{P}_{\mathbb{D}}\left(A_{m-1}\right)\right] \\ &= \mathbf{P}_{\mathbb{D}}\left(A_1\right) + \lim_{n\to\infty} \sum_{m=2}^{n} \left[\mathbf{P}_{\mathbb{D}}\left(A_m\right) - \mathbf{P}_{\mathbb{D}}\left(A_{m-1}\right)\right] \end{aligned}$$

Enfocándonos en desarrollar la suma se obtiene

$$\begin{aligned}
&\lim_{n\to\infty}\sum_{m=2}^{n}\left[\mathbf{P}_{\mathbb{D}}\left(A_m\right)-\mathbf{P}_{\mathbb{D}}\left(A_{m-1}\right)\right]\\
&=\lim_{n\to\infty}\left(\left[\mathbf{P}_{\mathbb{D}}\left(A_2\right)-\mathbf{P}_{\mathbb{D}}\left(A_1\right)\right]+\left[\mathbf{P}_{\mathbb{D}}\left(A_3\right)-\mathbf{P}_{\mathbb{D}}\left(A_2\right)\right]+\dots\right.\\
&\left.+\left[\mathbf{P}_{\mathbb{D}}\left(A_{n-1}\right)-\mathbf{P}_{\mathbb{D}}\left(A_{n-2}\right)\right]+\left[\mathbf{P}_{\mathbb{D}}\left(A_n\right)-\mathbf{P}_{\mathbb{D}}\left(A_{n-1}\right)\right]\right)\\
&=\lim_{n\to\infty}\left(-\mathbf{P}_{\mathbb{D}}\left(A_1\right)+\left[\mathbf{P}_{\mathbb{D}}\left(A_2\right)-\mathbf{P}_{\mathbb{D}}\left(A_2\right)\right]+\left[\mathbf{P}_{\mathbb{D}}\left(A_3\right)-\mathbf{P}_{\mathbb{D}}\left(A_3\right)\right]+\dots\right.\\
&\left.+\left[\mathbf{P}_{\mathbb{D}}\left(A_{n-1}\right)-\mathbf{P}_{\mathbb{D}}\left(A_{n-1}\right)\right]+\mathbf{P}_{\mathbb{D}}\left(A_n\right)\right)\\
&=\lim_{n\to\infty}\left(\mathbf{P}_{\mathbb{D}}\left(A_n\right)-\mathbf{P}_{\mathbb{D}}\left(A_1\right)\right)\\
&=\lim_{n\to\infty}\left(\mathbf{P}_{\mathbb{D}}\left(A_n\right)\right)-\mathbf{P}_{\mathbb{D}}\left(A_1\right)
\end{aligned}$$

Sustituyendo lo obtenido en la ecuación que se venía desarrollando previamente, obtenemos

$$\begin{aligned}
\mathbf{P}_{\mathbb{D}}\left(\bigcup_{n=1}^{\infty}A_n\right)&=\mathbf{P}_{\mathbb{D}}\left(A_1\right)+\lim_{n\to\infty}\left(\mathbf{P}_{\mathbb{D}}\left(A_n\right)\right)-\mathbf{P}_{\mathbb{D}}\left(A_1\right)\\
&=\lim_{n\to\infty}\mathbf{P}_{\mathbb{D}}\left(A_n\right)
\end{aligned}$$

■

Análogamente sucede para el caso $A_1 \supseteq A_2 \supseteq \cdots \supseteq A_n \supseteq \dots$, se obtiene que

$$\mathbf{P}_{\mathbb{D}}\left(\bigcap_{n=1}^{\infty}A_n\right)=\lim_{n\to\infty}\mathbf{P}_{\mathbb{D}}\left(A_n\right)$$

La demostración se realiza tomando los complementos y siguiendo el mismo método anterior.

Tomando en cuenta todo lo anterior, es posible encontrar las notables similitudes entre la teoría de probabilidad tradicional y la teoría de probabilidad hiperbólica. Lo que si se tiene en común es que para ambas teorías existe una variable en la cual existe el factor de incertidumbre.
La probabilidad tradicional esta centrada en tener un método de medida en el cual, si un evento tiene oportunidad nula de ocurrir, a este se le asignará un cero de probabilidad, sin embargo si se tiene una certeza total de la oportunidad de que el evento ocurra, entonces se tiene un uno. Así pues, esta siempre oscila entre 0 y 1,únicamente en el plano real.
Así pues, la probabilidad hiperbólica sigue con este mismo propósito, el de asignar valores y medidas para cuantificar qué tan cierto o incierto es una afirmación de

un evento en el futuro, sin embargo, a diferencia dela probabilidad tradicional, se contempla la estimación de diferentes variaciones a la vez, lo que genera una ampliación en la dimensión contemplada para la medición de las probabilidades.
Ambas teorías de probabilidad establecen sus tres axiomas principales, en los cuales basan su funcionalidad. Tanto el evento $\mathbf{P}(A)$ como el evento $\mathbf{P}_{\mathbb{D}}(A)$ tienen una medida probabilidad que es positiva, mayor o igual a cero, pero siempre menor o igual a 1, en el caso particular de $\mathbf{P}_{\mathbb{D}}(A)$ considerando también $\mathbf{e}$ y †.
De la misma manera, para ambas teorías existen determinadas álgebras que apoyan las uniones e intersecciones de los eventos, para generar probabilidades más acertadas.

5.3. Probabilidad Condicional

Ahora bien, una vez establecida gran parte de la teoría base respecto a los números hipérbólicos y su relación con la probabilidad, es importante enriquecer esta teoría de la misma manera en la que la probabilidad clásica, de esta manera aseguramos que la aplicación de la probabilidad hiperbólica sea práctica para dar soluciones en aplicaciones reales.
Al aprender de probabilidad, una de las temáticas principales es la condicionalidad entre eventos, por lo cual este capítulo estará dedicado al desarrollo de ello dentro de esta teoría.

Definición 21. *Sea $(\Omega, \sigma, \mathbf{P}_{\mathbb{D}})$.la tripleta que describe a un espacio hiperbólico probabilístico y siendo A y B dos eventos, la probabilidad del evento A bajo la condición de que el evento B suceda se denota como $\mathbf{P}_{\mathbb{D}}(A|B)$*

Sin embargo, a diferencia de la probabilidad clásica, en este caso esta condicionalidad estará sujeta a ciertas restricciones, las cuales se describen a continuación:
Para cuando $\mathbf{P}_{\mathbb{D}}(B)$ no es un divisor de cero [I]

(1) $\mathbf{P}_{\mathbb{D}}(A|B) := \frac{\mathbf{P}_{\mathbb{D}}(A\cap B)}{\mathbf{P}_{\mathbb{D}}(B)}$
Esto si $\mathbf{P}_{\mathbb{D}}(B) \succ 0$ y no es divisor de cero, de esta manera se garantiza que existe el inverso multiplicativo y no se indefine la ecuación.

(2) $\mathbf{P}_{\mathbb{D}}(A|B) := \mathbf{P}_{\mathbb{D}}(A)$
Cuando $\mathbf{P}_{\mathbb{D}}(B) = 0$

Cuando es el caso de que $\mathbf{P}_{\mathbb{D}}(B)$ es un divisor de cero, y tomando en cuenta que $P_1(B) = \lambda_1$ y $P_2(B) = \lambda_2$[II]

[I]Vease definición 14 y ecuación (3.10)
[II]Véase ecuaciones (5.6)

(3) $\mathbf{P}_{\mathbb{D}}\left(A|B\right) := \frac{\mathbf{P}_{\mathbb{D}}(A\cap B)}{\lambda_1}e + \mathbf{P}_{\mathbb{D}}\left(A\right)e^{\dagger}$
Si $\mathbf{P}_{\mathbb{D}}\left(B\right) = \lambda_1 e$, $\lambda_1 > 0$. Se contempla evidentemente que $\lambda_2 = 0$

(4) $\mathbf{P}_{\mathbb{D}}\left(A|B\right) := \mathbf{P}_{\mathbb{D}}\left(A\right)e + \frac{\mathbf{P}_{\mathbb{D}}(A\cap B)}{\lambda_2}e^{\dagger}$
Si $\mathbf{P}_{\mathbb{D}}\left(B\right) = \lambda_2 e^{\dagger}$, $\lambda_2 > 0$. Se contempla evidentemente que $\lambda_1 = 0$

Al mismo tiempo es posible comprobar que el caso (3) y (4) concuerdan con lo descrito en (1) gracias a las propiedades de la base idempotente, como se puede ver a continuación.
Sea

$$\mathbf{P}_{\mathbb{D}}(A) = P_1(A)e + P_2(A)e^{\dagger}$$

entonces es cierto que

$$\mathbf{P}_{\mathbb{D}}(A|B) = P_1(A|B)e + P_2(A|B)e^{\dagger}$$

Desarrollando esto con (1) obtenemos que

$$\begin{aligned}\mathbf{P}_{\mathbb{D}}(A|B) =& P_1(A|B)e + P_2(A|B)e^{\dagger}\\ =& \frac{\mathbf{P}_1\left(A\cap B\right)}{\mathbf{P}_1\left(B\right)}e + \frac{\mathbf{P}_2\left(A\cap B\right)}{\mathbf{P}_2\left(B\right)}e^{\dagger}\end{aligned}$$

Considerando $\mathbf{P}_{\mathbb{D}}(B) = P_1(B)e + P_2(B)e^{\dagger}$, donde $P_1(B) = \lambda_1$ y $P_2(B) = 0$, basándonos en (2), entonces se tiene (3)

$$\mathbf{P}_{\mathbb{D}}(A|B) = \frac{\mathbf{P}_1\left(A\cap B\right)}{\lambda_1}e + \mathbf{P}_{\mathbb{D}}(A)e^{\dagger}$$

Análogamente sucede cuando $P_1(B) = 0$ y $P_2(B) = \lambda_2$, resultando en (4)

$$\mathbf{P}_{\mathbb{D}}(A|B) = \mathbf{P}_{\mathbb{D}}(A)e + \frac{\mathbf{P}_1\left(A\cap B\right)}{\lambda_2}e^{\dagger}$$

Así entonces es posible ver que los 4 casos están estrechamente relacionados entre ellos. Para entender los casos (3) y (4) de mejor manera se plantea un ejemplo cuando $\mathbf{P}_{\mathbb{D}}(B)$ es un divisor de cero, dónde $x^2 = y^2$ en términos de la base normal.
EJEMPLO:

* Sea $x = .3, y = .3$, se tiene que

$$\begin{aligned}\mathbf{P}_{\mathbb{D}}(B) =& .3 + .3\mathbf{k}\\ =& (.3 + .3)e + (.3 - .3)e^{\dagger}\\ =& .6e + 0e^{\dagger} = .6e\end{aligned}$$

De esta forma, dentro de una probabilidad condicional corresponde al caso (3), donde $P_1(B) = \lambda_1 = .6$ y $P_2(B) = 0$

* Sea $x = .3, y = -.3$, se tiene que

$$\begin{aligned}\mathbf{P}_{\mathbb{D}}(B) =& .3 - .3\mathbf{k}\\ =&(.3 + (-.3))e + (.3 - (-.3))e^{\dagger}\\ =&0e + .6e^{\dagger} = .6e^{\dagger}\end{aligned}$$

  De esta forma, dentro de una probabilidad condicional corresponde al caso (4), donde$P_1(B) = 0$ y $P_2(B) = \lambda_2 = .6$

Así pues también se puede definir la probabilidad condicional de un evento consigo mismo como $\mathbf{P}_{\mathbb{D}}(B|B) = p$, donde p puede tomar 3 valores $1, e, e^{\dagger}$ como se muestra a continuación.

1. Si $\mathbf{P}_{\mathbb{D}}(B)$ no es un divisor de cero

$$\mathbf{P}_{\mathbb{D}}(B|B) = \frac{\mathbf{P}_{\mathbb{D}}(B \cap B)}{\mathbf{P}_{\mathbb{D}}(B)} = \frac{\mathbf{P}_{\mathbb{D}}(B)}{\mathbf{P}_{\mathbb{D}}(B)} = 1$$

2. Si es un divisor de cero de la forma $\mathbf{P}_{\mathbb{D}}(B) = \lambda_1 e$ entonces aplicando el caso (3) de la probabilidad condicional

$$\begin{aligned}\mathbf{P}_{\mathbb{D}}(B|B) =& \frac{\mathbf{P}_{\mathbb{D}}(B \cap B)}{\lambda_1}e + \mathbf{P}_{\mathbb{D}}(B)e^{\dagger} = \frac{\mathbf{P}_{\mathbb{D}}(B)}{\lambda_1}e + \lambda_1 e e^{\dagger}\\ =& \frac{\lambda_1 e}{\lambda_1}e + 0 = ee = e\end{aligned}$$

3. Si es un divisor de cero de la forma $\mathbf{P}_{\mathbb{D}}(B) = \lambda_2 e^{\dagger}$ entonces aplicando el caso (4) de la probabilidad condicional

$$\begin{aligned}\mathbf{P}_{\mathbb{D}}(B|B) =& \mathbf{P}_{\mathbb{D}}(B)e + \frac{\mathbf{P}_{\mathbb{D}}(B \cap B)}{\lambda_2}e^{\dagger} = \lambda_2 e^{\dagger} e + \frac{\mathbf{P}_{\mathbb{D}}(B)}{\lambda_2}e^{\dagger}\\ =& 0 + \frac{\lambda_2 e^{\dagger}}{\lambda_2 e^{\dagger}} = e^{\dagger}e^{\dagger} = e^{\dagger}\end{aligned}$$

De la misma forma se aplica que para conjuntos de eventos independientes entre ellos, si denominamos a una unión de conjuntos infinitos como $C = \bigcup_{k=1}^{\infty} C_k$ donde para cualesquiera $i \neq j$ se tiene que $C_i \cap C_j = \emptyset$, entonces:

1. Si $\mathbf{P}_{\mathbb{D}}(B)$ no es un divisor de cero

$$\begin{aligned}\mathbf{P}_{\mathbb{D}}(C|B) =& \frac{\mathbf{P}_{\mathbb{D}}(C \cap B)}{\mathbf{P}_{\mathbb{D}}(B)} = \frac{\mathbf{P}_{\mathbb{D}}\left(\bigcup_{k=1}^{\infty} C_k \cap B\right)}{\mathbf{P}_{\mathbb{D}}(B)} = \frac{\sum_{k=1}^{\infty}\mathbf{P}_{\mathbb{D}}(C_k \cap B)}{\mathbf{P}_{\mathbb{D}}(B)}\\ =& \sum_{k=1}^{\infty}\frac{\mathbf{P}_{\mathbb{D}}(C_k \cap B)}{\mathbf{P}_{\mathbb{D}}(B)} = \sum_{k=1}^{\infty}\mathbf{P}_{\mathbb{D}}(C_k|B)\end{aligned}$$

Para el caso de los divisores de los divisores de cero se llega a la misma conclusión para cualquiera de los dos casos: $\mathbf{P}_{\mathbb{D}}(B) = \lambda_1 e$ ó $\mathbf{P}_{\mathbb{D}}(B) = \lambda_2 e^{\dagger}$, tal como se muestra a continuación.

2. Si $\mathbf{P}_{\mathbb{D}}(B) = \lambda_1 e$, entonces es un divisor de cero, por lo tanto es necesario aplicar el caso (3) de probabilidad condicional. Lo que resultaría en

$$\begin{aligned}
\mathbf{P}_{\mathbb{D}}\left(C|B\right) =& \frac{\mathbf{P}_{\mathbb{D}}\left(C \cap B\right)}{\lambda_1} e + \mathbf{P}_{\mathbb{D}}\left(C\right) e^{\dagger} \\
=& \mathbf{P}_{\mathbb{D}}\left(\bigcup_{k=1}^{\infty} C_k \cap B\right) \frac{e}{\lambda_1} + \mathbf{P}_{\mathbb{D}}\left(\bigcup_{k=1}^{\infty} C_k\right) e^{\dagger} \\
=& \sum_{k=1}^{\infty} \mathbf{P}_{\mathbb{D}}\left(C_k \cap B\right) \frac{e}{\lambda_1} + \sum_{k=1}^{\infty} \mathbf{P}_{\mathbb{D}}(C_k) e^{\dagger} \\
=& \sum_{k=1}^{\infty} \left[\frac{\mathbf{P}_{\mathbb{D}}\left(C_k \cap B\right)}{\lambda_1} e + \mathbf{P}_{\mathbb{D}}\left(C_k\right) e^{\dagger}\right] \\
=& \sum_{k=1}^{\infty} \mathbf{P}_{\mathbb{D}}(C_k|B)
\end{aligned}$$

3. Si $\mathbf{P}_{\mathbb{D}}(B) = \lambda_2 e^{\dagger}$, entonces es un divisor de cero, por lo tanto es necesario aplicar el caso (4) de probabilidad condicional. Lo que resultaría en

$$\begin{aligned}
\mathbf{P}_{\mathbb{D}}\left(C|B\right) =& \mathbf{P}_{\mathbb{D}}\left(C\right) e + \frac{\mathbf{P}_{\mathbb{D}}\left(C \cap B\right)}{\lambda_2} e^{\dagger} \\
=& \mathbf{P}_{\mathbb{D}}\left(\bigcup_{k=1}^{\infty} C_k\right) e + \mathbf{P}_{\mathbb{D}}\left(\bigcup_{k=1}^{\infty} C_k \cap B\right) \frac{e^{\dagger}}{\lambda_2} \\
=& \sum_{k=1}^{\infty} \mathbf{P}_{\mathbb{D}}(C_k) e + \sum_{k=1}^{\infty} \mathbf{P}_{\mathbb{D}}\left(C_k \cap B\right) \frac{e^{\dagger}}{\lambda_2} \\
=& \sum_{k=1}^{\infty} \left[\mathbf{P}_{\mathbb{D}}\left(C_k\right) e + \frac{\mathbf{P}_{\mathbb{D}}\left(C_k \cap B\right)}{\lambda_2} e^{\dagger}\right] \\
=& \sum_{k=1}^{\infty} \mathbf{P}_{\mathbb{D}}(C_k|B)
\end{aligned}$$

La demostración para los casos de divisores de cero son independientes a las que se muestran en el artículo base

Teorema 1. *El teorema de multiplicación establece que sea* $(\Omega, \sigma, \mathbf{P}_{\mathbb{D}})$ *la tripleta que describe a un espacio hiperbólico probabilístico y siendo* A *y* B *dos eventos, se* *tiene que la probabilidad de la intersección entre ambos eventos, es igual a la probabilidad del evento B multiplicada por la probabilidad del evento A condicionado a B. Así entonces:*

$$\mathbf{P}_{\mathbb{D}}\left(A \cap B\right) = \mathbf{P}_{\mathbb{D}}(B)\mathbf{P}_{\mathbb{D}}(A|B)$$

Para demostrar este teorema es necesario dividir en 4 casos diferentes.
Los primeros dos, cuando $\mathbf{P}_{\mathbb{D}}(B)$ no es divisor de cero

1. Cuando $\mathbf{P}_{\mathbb{D}}(B) > 0$, la demostración sigue conforme al caso (1) de probabilidad condicional:

   $$\mathbf{P}_{\mathbb{D}}\left(A|B\right) = \frac{\mathbf{P}_{\mathbb{D}}\left(A \cap B\right)}{\mathbf{P}_{\mathbb{D}}\left(B\right)}$$

   Haciendo un fácil despeje obtenemos

   $$\mathbf{P}_{\mathbb{D}}\left(A \cap B\right) = \mathbf{P}_{\mathbb{D}}(B)\mathbf{P}_{\mathbb{D}}(A|B)$$

2. Cuando $\mathbf{P}_{\mathbb{D}}(B) = 0$, se demuestra basándose en el caso (2) de la probabilidad condicional, como sigue:

   $$\mathbf{P}_{\mathbb{D}}\left(A|B\right) = \mathbf{P}_{\mathbb{D}}\left(A\right)$$

   Por lo tanto se cumple el teorema puesto que

   $$\begin{aligned}\mathbf{P}_{\mathbb{D}}\left(A \cap B\right) =& \mathbf{P}_{\mathbb{D}}(B)\mathbf{P}_{\mathbb{D}}(A|B)\\ =& \mathbf{P}_{\mathbb{D}}(B)\mathbf{P}_{\mathbb{D}}(A)\\ =& 0 \cdot \mathbf{P}_{\mathbb{D}}(A) = 0\end{aligned}$$

Los otros dos casos se establecen cuando $\mathbf{P}_{\mathbb{D}}(B)$ es divisor de cero

3. Si $\mathbf{P}_{\mathbb{D}}(B) = \lambda_1 e$ con $\lambda_1 > 0$, entonces se ubica en el caso (3) de la probabilidad condicional

   $$\mathbf{P}_{\mathbb{D}}\left(A|B\right) = \frac{\mathbf{P}_{\mathbb{D}}\left(A \cap B\right)}{\lambda_1} e + \mathbf{P}_{\mathbb{D}}\left(A\right) e^{\dagger}$$

   Multiplicando esta ecuación por $\mathbf{P}_{\mathbb{D}}(B)$ en ambos lados, procede de la siguiente manera

$$\begin{aligned}
\mathbf{P}_{\mathbb{D}}(B)\mathbf{P}_{\mathbb{D}}\left(A|B\right) =& \mathbf{P}_{\mathbb{D}}(B)\left[\frac{\mathbf{P}_{\mathbb{D}}\left(A\cap B\right)}{\lambda_1}e+\mathbf{P}_{\mathbb{D}}\left(A\right)e^{\dagger}\right]\\
=&\frac{\mathbf{P}_{\mathbb{D}}(B)\mathbf{P}_{\mathbb{D}}\left(A\cap B\right)}{\lambda_1}e+\mathbf{P}_{\mathbb{D}}(B)\mathbf{P}_{\mathbb{D}}\left(A\right)e^{\dagger}\\
=&\frac{\lambda_1 e\cdot\mathbf{P}_{\mathbb{D}}\left(A\cap B\right)}{\lambda_1}e+\lambda_1 e\cdot\mathbf{P}_{\mathbb{D}}\left(A\right)e^{\dagger}\\
=&\mathbf{P}_{\mathbb{D}}\left(A\cap B\right)e+0\\
=&\left[\mathbf{P}_1\left(A\cap B\right)e+\mathbf{P}_2\left(A\cap B\right)e^{\dagger}\right]e\\
=&\mathbf{P}_1\left(A\cap B\right)e\\
=&\mathbf{P}_{\mathbb{D}}\left(A\cap B\right)
\end{aligned}$$

3. Si $\mathbf{P}_{\mathbb{D}}(B)=\lambda_2 e^{\dagger}$ con $\lambda_2>0$, entonces se ubica en el caso (4) de la probabilidad condicional

$$\mathbf{P}_{\mathbb{D}}\left(A|B\right)=\mathbf{P}_{\mathbb{D}}\left(A\right)e+\frac{\mathbf{P}_{\mathbb{D}}\left(A\cap B\right)}{\lambda_2}e^{\dagger}$$

Multiplicando ambos lados de la ecuación por $\mathbf{P}_{\mathbb{D}}(B)$, queda

$$\begin{aligned}
\mathbf{P}_{\mathbb{D}}(B)\mathbf{P}_{\mathbb{D}}\left(A|B\right) =& \mathbf{P}_{\mathbb{D}}(B)\left[\mathbf{P}_{\mathbb{D}}\left(A\right)e+\frac{\mathbf{P}_{\mathbb{D}}\left(A\cap B\right)}{\lambda_2}e^{\dagger}\right]\\
=&\mathbf{P}_{\mathbb{D}}(B)\mathbf{P}_{\mathbb{D}}\left(A\right)e+\frac{\mathbf{P}_{\mathbb{D}}(B)\mathbf{P}_{\mathbb{D}}\left(A\cap B\right)}{\lambda_2}e^{\dagger}\\
=&\lambda_2 e^{\dagger}\cdot\mathbf{P}_{\mathbb{D}}\left(A\right)e+\frac{\lambda_2 e^{\dagger}\cdot\mathbf{P}_{\mathbb{D}}\left(A\cap B\right)}{\lambda_2}e^{\dagger}\\
=&0+\mathbf{P}_{\mathbb{D}}\left(A\cap B\right)e^{\dagger}\\
=&\left[\mathbf{P}_1\left(A\cap B\right)e+\mathbf{P}_2\left(A\cap B\right)e^{\dagger}\right]e^{\dagger}\\
=&\mathbf{P}_2\left(A\cap B\right)e^{\dagger}\\
=&\mathbf{P}_{\mathbb{D}}\left(A\cap B\right)
\end{aligned}$$

■

Lo anterior se refiere a cuando tenemos exclusivamente dos eventos, sin embargo es posible generalizar a n eventos.Así entonces, definimos $A_1, A_2, \ldots, A_n$ como n eventos aleatorios

De la misma manera en la que se desarrolló previamente, se tiene que considerar el caso para aquellas probabilidades que son divisores de cero y aquellas que no.

Teorema 2. *Cuando se cumple que la probabilidad hiperbólica de la intersección de estos n eventos no genera un divisor de cero, se tiene que esta es:*

$$\mathbf{P}_{\mathbb{D}}\left(\bigcap_{i=1}^{n} A_i\right) = \mathbf{P}_{\mathbb{D}}(A_1)\mathbf{P}_{\mathbb{D}}\left(A_2|A_1\right)\ldots\mathbf{P}_{\mathbb{D}}\left(A_n|A_1\cap\cdots\cap A_{n-1}\right)$$

Demostración:
Sea $B=\bigcap_{i=1}^{n-1} A_i$, se tiene que

$$\mathbf{P}_{\mathbb{D}}\left(\bigcap_{i=1}^{n} A_i\right) = \mathbf{P}_{\mathbb{D}}\left(\bigcap_{i=1}^{n-1} A_i\cap A_n\right) = \mathbf{P}_{\mathbb{D}}\left(B\cap A_n\right) = \mathbf{P}_{\mathbb{D}}\left(A_n\cap B\right)$$

Usando el teorema 1, para el caso 1 previamente demostrado, sigue que

$$\mathbf{P}_{\mathbb{D}}\left(A_n\cap B\right) = \mathbf{P}_{\mathbb{D}}(B)\mathbf{P}_{\mathbb{D}}(A_n|B) = \mathbf{P}_{\mathbb{D}}\left(\bigcap_{i=1}^{n-1} A_i\right)\mathbf{P}_{\mathbb{D}}\left(A_n\middle|\bigcap_{i=1}^{n-1} A_i\right)$$

Sea $C=\bigcap_{i=1}^{n-2} A_i$ y considerando que

$$\begin{aligned}\mathbf{P}_{\mathbb{D}}(B) =& \mathbf{P}_{\mathbb{D}}\left(\bigcap_{i=1}^{n-1} A_i\right) = \mathbf{P}_{\mathbb{D}}\left(C\cap A_{n-1}\right) = \mathbf{P}_{\mathbb{D}}\left(A_{n-1}\cap C\right)\\ =& \mathbf{P}_{\mathbb{D}}(C)\mathbf{P}_{\mathbb{D}}(A_{n-1}|C) = \mathbf{P}_{\mathbb{D}}\left(\bigcap_{i=1}^{n-2} A_i\right)\mathbf{P}_{\mathbb{D}}\left(A_{n-1}|\bigcap_{i=1}^{n-2} A_i\right)\end{aligned}$$

Sea $D=\bigcap_{i=1}^{n-3} A_i$, y considerando que

$$\begin{aligned}\mathbf{P}_{\mathbb{D}}(C) =& \mathbf{P}_{\mathbb{D}}\left(\bigcap_{i=1}^{n-2} A_i\right) = \mathbf{P}_{\mathbb{D}}\left(D\cap A_{n-2}\right) = \mathbf{P}_{\mathbb{D}}\left(A_{n-2}\cap D\right)\\ =& \mathbf{P}_{\mathbb{D}}(D)\mathbf{P}_{\mathbb{D}}(A_{n-2}|D) = \mathbf{P}_{\mathbb{D}}\left(\bigcap_{i=1}^{n-3} A_i\right)\mathbf{P}_{\mathbb{D}}\left(A_{n-2}|\bigcap_{i=1}^{n-3} A_i\right)\end{aligned}$$

De esta manera, se siguen tomando los grupos de intersecciones como las intersecciones de los primeros elementos en intersección con último evento, lo que genera

$$\mathbf{P}_{\mathbb{D}}\left(\bigcap_{i=1}^{n} A_i\right) = \mathbf{P}_{\mathbb{D}}\left(A_1\cap\cdots\cap A_n\right)$$

$$= \mathbf{P}_{\mathbb{D}}\left(\bigcap_{i=1}^{n-3} A_i\right)\mathbf{P}_{\mathbb{D}}\left(A_{n-2}|\bigcap_{i=1}^{n-3} A_i\right)\mathbf{P}_{\mathbb{D}}\left(A_{n-1}|\bigcap_{i=1}^{n-2} A_i\right)\mathbf{P}_{\mathbb{D}}\left(A_n|\bigcap_{i=1}^{n-1} A_i\right)$$

Por lo tanto, aplicando todo lo anterior hasta llegar a los primeros términos, se llega a que

$$\mathbf{P}_{\mathbb{D}}\left(\bigcap_{i=1}^{n} A_i\right) = \mathbf{P}_{\mathbb{D}}\left(A_1 \cap \cdots \cap A_n\right)$$

$$= \mathbf{P}_{\mathbb{D}}(A_1)\mathbf{P}_{\mathbb{D}}\left(A_2|A_1\right)\dots\mathbf{P}_{\mathbb{D}}\left(A_{n-1}|\bigcap_{i=1}^{n-2} A_i\right)\mathbf{P}_{\mathbb{D}}\left(A_n|A_1 \cap \cdots \cap A_{n-1}\right)$$

■

Teorema 3. *Sea $A_1, A_2, \dots A_n$ un conjunto de n eventos, donde al menos para algún $A_i, i \in 1, 2, \dots n$ se tiene que su probabilidad es un divisor de cero de la forma $\mathbf{P}_{\mathbb{D}}(A_i) = \lambda_1 e, \lambda_1 > 0$, entonces $\mathbf{P}_{\mathbb{D}}\left(A_1 \cap A_2 \cap \cdots \cap A_i \cap \dots A_n\right)$ también es un divisor de cero, y se cumple que*

$$\mathbf{P}_{\mathbb{D}}\left(\bigcap_{i=1}^{n} A_i\right) = \mathbf{P}_{\mathbb{D}}(A_1)\mathbf{P}_{\mathbb{D}}\left(A_2|A_1\right)\dots\mathbf{P}_{\mathbb{D}}\left(A_n|A_1 \cap \cdots \cap A_{n-1}\right)$$

Demostración:
Sabemos por hipótesis que al menos para un evento su probabilidad resulta en un divisor de cero, haciendo a $\mathbf{P}_{\mathbb{D}}(A_1) = \lambda_1 e, \lambda_1 > 0$, entonces tenemos que sea $B = \bigcap_{i=1}^{n-1} A_i$

$$\mathbf{P}_{\mathbb{D}}\left(A_1 \cap \cdots \cap A_n\right) = \mathbf{P}_{\mathbb{D}}\left(\bigcap_{i=1}^{n-1} A_i \cap A_n\right) = \mathbf{P}_{\mathbb{D}}\left(B \cap A_n\right) =$$

Usando el teorema 1, en el caso 3, tenemos que

$$\mathbf{P}_{\mathbb{D}}\left(A_n \cap B\right) = \mathbf{P}_{\mathbb{D}}(B)\mathbf{P}_{\mathbb{D}}(A_n|B) = \mathbf{P}_{\mathbb{D}}\left(\bigcap_{i=1}^{n-1} A_i\right)\mathbf{P}_{\mathbb{D}}\left(A_n|\bigcap_{i=1}^{n-1} A_i\right)$$

La demostración se sigue como en el caso anterior. ■

Teorema 4. *Sea $A_1, A_2, \dots A_n$ un conjunto de n eventos, donde al menos para algún $A_i, i \in 1, 2, \dots n$ se tiene que su probabilidad es un divisor de cero de la forma $\mathbf{P}_{\mathbb{D}}(A_i) = \lambda_2 e^{\dagger}, \lambda_2 > 0$, entonces $\mathbf{P}_{\mathbb{D}}\left(A_1 \cap A_2 \cap \cdots \cap A_i \cap \dots A_n\right)$ también es un divisor de cero, y se cumple que*

$$\mathbf{P}_{\mathbb{D}}\left(\bigcap_{i=1}^{n} A_i\right) = \mathbf{P}_{\mathbb{D}}(A_1)\mathbf{P}_{\mathbb{D}}\left(A_2|A_1\right)\dots\mathbf{P}_{\mathbb{D}}\left(A_n|A_1 \cap \cdots \cap A_{n-1}\right)$$

Demostración:
Sabemos por hipótesis que al menos para un evento su probabilidad resulta en un divisor de cero, haciendo a $\mathbf{P}_{\mathbb{D}}(A_1) = \lambda_1 e, \lambda_1 > 0$, entonces tenemos que sea $B = \bigcap_{i=1}^{n-1} A_i$

$$\mathbf{P}_{\mathbb{D}}\left(A_1 \cap \cdots \cap A_n\right) = \mathbf{P}_{\mathbb{D}}\left(\bigcap_{i=1}^{n-1} A_i \cap A_n\right) = \mathbf{P}_{\mathbb{D}}\left(B \cap A_n\right) =$$

Usando el teorema 1, en el caso 4, tenemos que

$$\mathbf{P}_{\mathbb{D}}\left(A_n \cap B\right) = \mathbf{P}_{\mathbb{D}}(B)\mathbf{P}_{\mathbb{D}}(A_n|B) = \mathbf{P}_{\mathbb{D}}\left(\bigcap_{i=1}^{n-1} A_i\right)\mathbf{P}_{\mathbb{D}}\left(A_n|\bigcap_{i=1}^{n-1} A_i\right)$$

La demostración se sigue como en el caso anterior. ■

5.4. Independencia de eventos

Por otro lado, necesitamos un concepto que nos ayude a comprender si un evento aleatorio tiene o no efecto alguno en otro evento aleatorio. Para ello se define la independencia entre dos eventos.

Definición 22. *Sea A y B dos eventos aleatorios.*

1. *El evento A es independiente del evento B si cumple que la probabilidad de A dado B, es en sí la probabilidad de A. Esto es:*

$$\mathbf{P}_{\mathbb{D}}(A|B) = \mathbf{P}_{\mathbb{D}}(A)$$

2. *El evento B es independiente del evento A si cumple que la probabilidad de B dado A, es en sí la probabilidad de B. Esto es:*

$$\mathbf{P}_{\mathbb{D}}(B|A) = \mathbf{P}_{\mathbb{D}}(B)$$

3. *Los eventos A y B son mutuamente independientes si se cumplen los puntos 1 y 2 simultáneamente.*

De estas definiciones podemos encontrar varios escenarios para ambos eventos.

(a) **Cualesquiera cero**
Considerando que las probabilidades para cualquiera de los dos eventos sean cero: $\mathbf{P}_{\mathbb{D}}(A) = \mathbf{P}_{\mathbb{D}}(B) = 0$, entonces si definimos que

$$\mathbf{P}_{\mathbb{D}}(A|B) = \mathbf{P}_{\mathbb{D}}(A), \qquad \mathbf{P}_{\mathbb{D}}(B|A) = \mathbf{P}_{\mathbb{D}}(B)$$

se concluye que A y B son mutuamente independientes.
Incluso si consideramos solo a un evento con propabilidad de cero, se llegaría a la misma conclusión, véase como:

$$\begin{aligned}
&\mathbf{P}_{\mathbb{D}}(A) = 0 \Rightarrow \mathbf{P}_{\mathbb{D}}(A \cap B) = \mathbf{P}_{\mathbb{D}}(A)\mathbf{P}_{\mathbb{D}}(B) = 0 \\
&\mathbf{P}_{\mathbb{D}}(A|B) = \frac{\mathbf{P}_{\mathbb{D}}(A \cap B)}{\mathbf{P}_{\mathbb{D}}(B)} = \frac{\mathbf{P}_{\mathbb{D}}(A)\mathbf{P}_{\mathbb{D}}(B)}{\mathbf{P}_{\mathbb{D}}(B)} = \mathbf{P}_{\mathbb{D}}(A) = 0 \\
&\mathbf{P}_{\mathbb{D}}(B|A) = \frac{\mathbf{P}_{\mathbb{D}}(B \cap A)}{\mathbf{P}_{\mathbb{D}}(A)} = \frac{\mathbf{P}_{\mathbb{D}}(B)\mathbf{P}_{\mathbb{D}}(A)}{\mathbf{P}_{\mathbb{D}}(A)} = \mathbf{P}_{\mathbb{D}}(B)
\end{aligned}$$

Entonces A y B son mutuamente independientes

b) **No divisores de cero**
Considerando que $\mathbf{P}_{\mathbb{D}}(B)$ y $\mathbf{P}_{\mathbb{D}}(B)$ no estén en los divisores de cero la idependencia de A y B es equivalente a asumir que

$$\mathbf{P}_{\mathbb{D}}(A \cap B) = \mathbf{P}_{\mathbb{D}}(A)\mathbf{P}_{\mathbb{D}}(B) \tag{5.17}$$

Entonces se diría que

$$\mathbf{P}_{\mathbb{D}}(A|B) = \frac{\mathbf{P}_{\mathbb{D}}(A \cap B)}{\mathbf{P}_{\mathbb{D}}(B)} = \frac{\mathbf{P}_{\mathbb{D}}(A)\mathbf{P}_{\mathbb{D}}(B)}{\mathbf{P}_{\mathbb{D}}(B)} = \mathbf{P}_{\mathbb{D}}(A)$$

Similarmente para

$$\mathbf{P}_{\mathbb{D}}(B|A) = \frac{\mathbf{P}_{\mathbb{D}}(B \cap A)}{\mathbf{P}_{\mathbb{D}}(A)} = \frac{\mathbf{P}_{\mathbb{D}}(B)\mathbf{P}_{\mathbb{D}}(A)}{\mathbf{P}_{\mathbb{D}}(A)} = \mathbf{P}_{\mathbb{D}}(B)$$

Así entonces decimos que A y B son mutuamente excluyentes.

c) **Ambos divisores de cero sobre el mismo eje idempotente**
•Sobre el eje e
Asumiendo que $\mathbf{P}_{\mathbb{D}}(A)$ y $\mathbf{P}_{\mathbb{D}}(B)$ son ambos divisores de cero de la forma $\mathbf{P}_{\mathbb{D}}(A) = \lambda e$ y $\mathbf{P}_{\mathbb{D}}(B) = \mu e$, entonces $\mathbf{P}_1(A) = \lambda$ y $\mathbf{P}_1(B) = \mu$ donde λ y μ son valores positivos. De esta manera $\mathbf{P}_2(A) = \mathbf{P}_2(B) = 0$ Así entonces $\mathbf{P}_{\mathbb{D}}(A \cap B) = \theta e$ puesto que $\mathbf{P}_1(A \cap B) = \theta$ y $\mathbf{P}_2(A \cap B) = 0$ Asumiendo que A es independiente de B, entonces es cierto que

$$\mathbf{P}_{\mathbb{D}}(A|B) = \mathbf{P}_{\mathbb{D}}(A) = \lambda e$$

Desarrollando según el tercer caso de probabilidad condicional, tenemos que

$$\mathbf{P}_{\mathbb{D}}\left(A|B\right) = \frac{\mathbf{P}_{\mathbb{D}}\left(A \cap B\right)}{\mu}e + \mathbf{P}_{\mathbb{D}}\left(A\right)e^{\dagger} = \frac{\mathbf{P}_1\left(A \cap B\right)}{\mu}e = \frac{\theta}{\mu}e$$

Así entonces se llega a la igualdad $\lambda e = \frac{\theta}{\mu} e$, estableciendo que si y sólo si $\lambda = \frac{\theta}{\mu}$, entonces A es independiente de B.
Ahora asumiendo que B es independiente de A, se tiene que

$$\mathbf{P}_{\mathbb{D}}(B|A) = \mathbf{P}_{\mathbb{D}}(B) = \mu e$$

Desarrollando según el tercer caso de probabilidad condicional, tenemos que

$$\mathbf{P}_{\mathbb{D}}\left(B|A\right) = \frac{\mathbf{P}_{\mathbb{D}}\left(B \cap A\right)}{\lambda} e + \mathbf{P}_{\mathbb{D}}\left(B\right) e^{\dagger} = \frac{\mathbf{P}_{1}\left(B \cap A\right)}{\lambda} e = \frac{\theta}{\lambda} e$$

De la misma manera se obtiene una igualdad $\mu e = \frac{\theta}{\lambda} e$, estableciendo que si y sólo si $\mu = \frac{\theta}{\lambda}$, entonces B es independiente de A.
Véase que la igualdad para la independencia de A con B y de B con A es la misma. Así entonces concluimos que cuando se cumple que $\theta = \lambda\mu$, entonces A y B son mutuamente excluyentes.

$$\mathbf{P}_{\mathbb{D}}(A \cap B) = \mathbf{P}_{1}(A \cap B)e = \theta e = \lambda\mu e = \lambda e \cdot \mu e = \mathbf{P}_{\mathbb{D}}(A)\mathbf{P}_{\mathbb{D}}(B)$$

•**Sobre** $e^{\dagger}$
Funciona de la misma manera para cuando el caso en el que $\mathbf{P}_{\mathbb{D}}(A)$ y $\mathbf{P}_{\mathbb{D}}(B)$ son ambos divisores de cero de la forma $\mathbf{P}_{\mathbb{D}}(A) = \lambda e^{\dagger}$ y $\mathbf{P}_{\mathbb{D}}(B) = \mu e^{\dagger}$entonces se tiene que $\mathbf{P}_{1}(A) = \mathbf{P}_{1}(B) = 0$ y $\mathbf{P}_{2}(A) = \lambda$ y $\mathbf{P}_{2}(B) = \mu$ donde λ y μ son valores positivos. Así entonces $\mathbf{P}_{\mathbb{D}}(A \cap B) = \theta e^{\dagger}$ puesto que $\mathbf{P}_{2}(A \cap B) = \theta$ y $\mathbf{P}_{1}(A \cap B) = 0$ Asumiendo que A es independiente de B, entonces es cierto que

$$\mathbf{P}_{\mathbb{D}}(A|B) = \mathbf{P}_{\mathbb{D}}(A) = \lambda e^{\dagger}$$

Desarrollando según el cuarto caso de probabilidad condicional, tenemos que

$$\mathbf{P}_{\mathbb{D}}(A|B) = \mathbf{P}_{\mathbb{D}}(A)e + \frac{\mathbf{P}_{\mathbb{D}}(A \cap B)}{\mu} e^{\dagger} = \frac{\mathbf{P}_{2}\left(A \cap B\right)}{\mu} e^{\dagger} = \frac{\theta}{\mu} e^{\dagger}$$

Así entonces se llega a la igualdad $\lambda e^{\dagger} = \frac{\theta}{\mu} e^{\dagger}$, estableciendo que si y sólo si $\lambda = \frac{\theta}{\mu}$, entonces A es independiente de B.
Ahora asumiendo que B es independiente de A, se tiene que

$$\mathbf{P}_{\mathbb{D}}(B|A) = \mathbf{P}_{\mathbb{D}}(B) = \mu e^{\dagger}$$

Desarrollando según el tercer caso de probabilidad condicional, tenemos que

$$\mathbf{P}_{\mathbb{D}}(B|A) = \mathbf{P}_{\mathbb{D}}(B)e + \frac{\mathbf{P}_{\mathbb{D}}(B \cap A)}{\lambda} e^{\dagger} = \frac{\mathbf{P}_{2}\left(B \cap A\right)}{\lambda} e^{\dagger} = \frac{\theta}{\lambda} e^{\dagger}$$

De la misma manera se obtiene una igualdad $\mu e^{\dagger} = \frac{\theta}{\lambda}e^{\dagger}$, estableciendo que si y sólo si $\mu = \frac{\theta}{\lambda}$, entonces B es independiente de A.
Véase que la igualdad para la independencia de A con B y de B con A es la misma. Así entonces concluimos que cuando se cumple que $\theta = \lambda\mu$, entonces A y B son mutuamente excluyentes.

$$\mathbf{P}_{\mathbb{D}}(A \cap B) = \mathbf{P}_2(A \cap B)e^{\dagger} = \theta e^{\dagger} = \lambda\mu e^{\dagger} = \lambda e^{\dagger} \cdot \mu e^{\dagger} = \mathbf{P}_{\mathbb{D}}(A)\mathbf{P}_{\mathbb{D}}(B)$$

d) **Ambos divisores de cero sobre el distinto eje idempotente**
Ahora considerando que las probabilidades de los dos eventos son ambos divisores de cero, pero cada uno sobre uno de los dos ejes idempotentes, esto es $\mathbf{P}_{\mathbb{D}}(A) = \lambda e$ y $\mathbf{P}_{\mathbb{D}}(B) = \mu e^{\dagger}$, ambos diferentes de cero. Por lo tanto, se tendría que $\mathbf{P}_{\mathbb{D}}(A \cap B) = 0$, pues $\mathbf{P}_1(A \cap B) = \mathbf{P}_2(A \cap B) = 0$.
Así pues considerando el caso 4 de probabilidad condicional para probar que A es independiente de B, tendríamos que:

$$\mathbf{P}_{\mathbb{D}}(A|B) = \mathbf{P}_{\mathbb{D}}(A)e + \frac{\mathbf{P}_{\mathbb{D}}(A \cap B)}{\mu}e^{\dagger} = \mathbf{P}_{\mathbb{D}}(A)e = \lambda ee = \lambda e = \mathbf{P}_{\mathbb{D}}(A)$$

Igualmente para probar que B es independiente de A, ahora usando el caso 3 de probabilidad condicional:

$$\mathbf{P}_{\mathbb{D}}(B|A) = \frac{\mathbf{P}_{\mathbb{D}}(A \cap B)}{\mu}e + \mathbf{P}_{\mathbb{D}}(A)e^{\dagger} = \mathbf{P}_{\mathbb{D}}(B)e^{\dagger} = \mu e^{\dagger}e^{\dagger} = \mu e^{\dagger} = \mathbf{P}_{\mathbb{D}}(B)$$

Por consiguiente A y B son eventos mutuamente independientes, y por lo tanto

$$\mathbf{P}_{\mathbb{D}}(A \cap B) = \mathbf{P}_{\mathbb{D}}(A)\mathbf{P}_{\mathbb{D}}(B) = \lambda e \cdot \mu e^{\dagger} = 0$$

e) **Un divisor de cero y un número hiperbólico con inverso**
Consideramos para este caso a $\mathbf{P}_{\mathbb{D}}(A) = \lambda e$ y a $\mathbf{P}_{\mathbb{D}}(B) = \mu_1 e + \mu_2 e^{\dagger}$, ambos distintos de cero. Se tendría que $\mathbf{P}_1(A) = \lambda$, $\mathbf{P}_2(A) = 0$, $\mathbf{P}_1(B) = \mu_1$ y $\mathbf{P}_2(B) = \mu_2$.
Puesto que solo hay intersección en el eje e, entonces $\mathbf{P}_1(A \cap B) = \theta \geq 0$ y $\mathbf{P}_2(A \cap B) = 0$ esto se traduciría en

$$\mathbf{P}_{\mathbb{D}}(A \cap B) = \mathbf{P}_1(A \cap B)e + \mathbf{P}_2(A \cap B)e^{\dagger} = \theta e$$

Suponiendo que es A es independiente de B, entonces decimos que

$$\mathbf{P}_{\mathbb{D}}(A|B) = \mathbf{P}_{\mathbb{D}}(A) = \lambda e$$

Por otro lado, usando el primer caso de probabilidad condicional como sigue

$$\mathbf{P}_{\mathbb{D}}(A|B) = \frac{\mathbf{P}_{\mathbb{D}}(A \cap B)}{\mathbf{P}_{\mathbb{D}}(B)} = \frac{\theta e}{\mu_1 e + \mu_2 e^{\dagger}} = \frac{\theta}{\mu_1}e$$

Por consiguiente se obtiene la igualdad $\lambda e = \frac{\theta}{\mu_1} e$, esto es $\lambda = \frac{\theta}{\mu_1}$.

De la misma manera asumiendo que B es independiente de A, se establece que

$$\mathbf{P}_{\mathbb{D}}(B|A) = \mathbf{P}_{\mathbb{D}}(B) = \mu_1 e + \mu_2 e^{\dagger}$$

Y usando el tercer caso de probabilidad condicional se desarrolla lo siguiente

$$\begin{aligned}
\mathbf{P}_{\mathbb{D}}(B|A) &= \frac{\mathbf{P}_{\mathbb{D}}(B \cap A)}{\lambda} e + \mathbf{P}_{\mathbb{D}}(B) e^{\dagger} = \frac{\mathbf{P}_1 (B \cap A)}{\lambda} e + [\mu_1 e + \mu_2 e^{\dagger}] e^{\dagger} \\
&= \frac{\mathbf{P}_1 (B \cap A)}{\lambda} e + \mu_2 e^{\dagger} = \frac{\theta}{\lambda} e + \mu_2 e^{\dagger} = \mu_1 e + \mu_2 e^{\dagger} = \mathbf{P}_{\mathbb{D}}(B)
\end{aligned}$$

Para cuando se cumple la igualdad de $\mu_1 = \frac{\theta}{\lambda}$.

Por lo tanto, se dice que los eventos A y B son mutuamente independientes cuando se tiene $\theta = \lambda\mu_1$ Entonces se concluye que

$$\mathbf{P}_{\mathbb{D}}(A \cap B) = \theta e = \lambda\mu_1 e = \lambda e \cdot [\mu_1 e + \mu_2 e^{\dagger}] = \mathbf{P}_{\mathbb{D}}(A)\mathbf{P}_{\mathbb{D}}(B)$$

Corolario 1. *Dado dos eventos Ay B, entonces A es independiente de B si y sólo si B es independiente de B, es decir, deben ser mutuamente independientes, y por lo tanto se tiene que:*

$$\mathbf{P}_{\mathbb{D}}(A \cap B) = \mathbf{P}_{\mathbb{D}}(A)\mathbf{P}_{\mathbb{D}}(B)$$

Teorema 5. *Si A y B son mutuamente independientes, entonces existe independencia entre estos eventos y sus complementos. Esto es entre A y B^C, A^C y B, A^C y B^C*

Demostración:

Probando la independencia para los eventos A y B^C, si se considera que el evento $A = (A \cap B) \cup (A \cap B^C)$, entonces su probabilidad está dada por $\mathbf{P}_{\mathbb{D}}(A) = \mathbf{P}_{\mathbb{D}}(A \cap B) + \mathbf{P}_{\mathbb{D}}(A \cap B^C)$, y despejando para $\mathbf{P}_{\mathbb{D}}(A \cap B^C)$, se tiene:

$$\mathbf{P}_{\mathbb{D}}(A \cap B^C) = \mathbf{P}_{\mathbb{D}}(A) - \mathbf{P}_{\mathbb{D}}(A \cap B)$$

Puesto que A y B son eventos mutuamente independientes, sabemos por el colorario anterior que $\mathbf{P}_{\mathbb{D}}(A \cap B) = \mathbf{P}_{\mathbb{D}}(A)\mathbf{P}_{\mathbb{D}}(B)$. Así que

$$\mathbf{P}_{\mathbb{D}}(A \cap B^C) = \mathbf{P}_{\mathbb{D}}(A) - \mathbf{P}_{\mathbb{D}}(A)\mathbf{P}_{\mathbb{D}}(B) = \mathbf{P}_{\mathbb{D}}(A)(1 - \mathbf{P}_{\mathbb{D}}(B)) \tag{5.18}$$

El siguiente paso depende depende directamente del valor de $\mathbf{P}_{\mathbb{D}}(A)$ puesto que de ahí se define el valor que se tenga para la probabilidad del universo. Recordando que $\mathbf{P}_{\mathbb{D}}(\Omega)$ es de la forma dada en la ecuación (5.4) y que la p puede tomar valores para $1, e$ y $e^{\dagger}$ dependiendo de la probabilidad dada de A. Se explican los casos a continuación.

Caso 1 $\mathbf{P}_{\mathbb{D}}(A)$ no es un divisor de cero, por lo tanto $\mathbf{P}_{\mathbb{D}}(\Omega)$ tampoco lo es, y la probabilidad podría ser unicamente $p = 1$ (nótese que no es posible considerar los casos de $p = e$ ó $p = e^{\dagger}$, puesto que nuestro universo no está en los divisores de cero) Entonces se tendría que:

$$\mathbf{P}_{\mathbb{D}}(A \cap B^C) = \mathbf{P}_{\mathbb{D}}(A)(1 - \mathbf{P}_{\mathbb{D}}(B)) = \mathbf{P}_{\mathbb{D}}(A)\mathbf{P}_{\mathbb{D}}(B^C)$$

Caso 2 $\mathbf{P}_{\mathbb{D}}(A)$ es un divisor de cero de la forma $\mathbf{P}_{\mathbb{D}}(A) = \lambda e$ por lo tanto $\mathbf{P}_{\mathbb{D}}(\Omega)$ puede tomar valores para:

$p = 1$: Entonces $\mathbf{P}_{\mathbb{D}}(\Omega) = 1$ no está en los divisores de cero y se puede considerar a la probabilidad de B como $\mathbf{P}_{\mathbb{D}}(B) = \mu_1 e + \mu_2 e^{\dagger}$ y a $\mathbf{P}_{\mathbb{D}}(B^C) = 1 - \mathbf{P}_{\mathbb{D}}(B) = 1 - (\mu_1 e + \mu_2 e^{\dagger})$ Recordando que $1 = 1e + 1e^{\dagger}$, entonces $\mathbf{P}_{\mathbb{D}}(B^C) = (1 - \mu_1)e + (1 - \mu_2)e^{\dagger}$ Con esto establecido y siguiendo lo obtenido en (5.18) se desarrolla que:

$$\begin{aligned}
\mathbf{P}_{\mathbb{D}}(A \cap B^C) =& \mathbf{P}_{\mathbb{D}}(A) - \mathbf{P}_{\mathbb{D}}(A)\mathbf{P}_{\mathbb{D}}(B) \\
=& \lambda e - \lambda e(\mu_1 e + \mu_2 e^{\dagger}) \\
=& \lambda e - \lambda e(\mu_1 e) \\
=& \lambda e(1 - \mu_1 e) \\
=& \lambda e((1 - \mu_1)e + (1 - \mu_2)e^{\dagger}) \\
=& \mathbf{P}_{\mathbb{D}}(A)\mathbf{P}_{\mathbb{D}}(B^C)
\end{aligned}$$

$p = e$: Entonces $\mathbf{P}_{\mathbb{D}}(\Omega) = e$ está en los divisores de cero y se considera entonces $\mathbf{P}_{\mathbb{D}}(B) = \mu_1 e$, por lo que $\mathbf{P}_{\mathbb{D}}(B^C) = (1 - \mu_1)e$ Se tiene que:

$$\begin{aligned}
\mathbf{P}_{\mathbb{D}}(A \cap B^C) =& \mathbf{P}_{\mathbb{D}}(A) - \mathbf{P}_{\mathbb{D}}(A)\mathbf{P}_{\mathbb{D}}(B) \\
=& \lambda e - \lambda e(\mu_1 e) \\
=& \lambda e - \lambda e(\mu_1 e) \\
=& \lambda e(1 - \mu_1 e) \\
=& \lambda e(1 - \mu_1)e \\
=& \mathbf{P}_{\mathbb{D}}(A)\mathbf{P}_{\mathbb{D}}(B^C)
\end{aligned}$$

$p = e^{\dagger}$: Puesto que $\mathbf{P}_{\mathbb{D}}(A)$ está definida en el eje idempotente e no es posible que $\mathbf{P}_{\mathbb{D}}(\Omega) = e^{\dagger}$, y este caso no se desarrolla.

Caso 3 $\mathbf{P}_{\mathbb{D}}(A)$ es un divisor de cero de la forma $\mathbf{P}_{\mathbb{D}}(A) = \lambda e^{\dagger}$ por lo tanto $\mathbf{P}_{\mathbb{D}}(\Omega)$ puede tomar valores para:

$p = 1$: Entonces $\mathbf{P}_{\mathbb{D}}(\Omega) = 1$ no está en los divisores de cero y se puede considerar a la probabilidad de B como $\mathbf{P}_{\mathbb{D}}(B) = \mu_1 e + \mu_2 e^\dagger$ y a $\mathbf{P}_{\mathbb{D}}(B^C) = 1 - \mathbf{P}_{\mathbb{D}}(B) = 1 - (\mu_1 e + \mu_2 e^\dagger)$ Recordando que $1 = 1e + 1e^\dagger$, entonces $\mathbf{P}_{\mathbb{D}}(B^C) = (1-\mu_1)e + (1-\mu_2)e^\dagger$ Con esto establecido y siguiendo lo obtenido en (5.18) se desarrolla que:

$$\begin{aligned}
\mathbf{P}_{\mathbb{D}}(A \cap B^C) =& \mathbf{P}_{\mathbb{D}}(A) - \mathbf{P}_{\mathbb{D}}(A)\mathbf{P}_{\mathbb{D}}(B) \\
=& \lambda e^\dagger - \lambda e^\dagger(\mu_1 e + \mu_2 e^\dagger) \\
=& \lambda e^\dagger - \lambda e^\dagger(\mu_2 e^\dagger) \\
=& \lambda e^\dagger(1 - \mu_2 e^\dagger) \\
=& \lambda e^\dagger((1-\mu_1)e + (1-\mu_2)e^\dagger) \\
=& \mathbf{P}_{\mathbb{D}}(A)\mathbf{P}_{\mathbb{D}}(B^C)
\end{aligned}$$

$p = e$: Puesto que $\mathbf{P}_{\mathbb{D}}(A)$ está definida en el eje idempotente $e^\dagger$ no es posible que $\mathbf{P}_{\mathbb{D}}(\Omega) = e$, y este caso no se desarrolla.

$p = e^\dagger$: Entonces $\mathbf{P}_{\mathbb{D}}(\Omega) = e^\dagger$ está en los divisores de cero y se considera entonces $\mathbf{P}_{\mathbb{D}}(B) = \mu_2 e^\dagger$, por lo que $\mathbf{P}_{\mathbb{D}}(B^C) = (1-\mu_2)e^\dagger$ Se tiene que:

$$\begin{aligned}
\mathbf{P}_{\mathbb{D}}(A \cap B^C) =& \mathbf{P}_{\mathbb{D}}(A) - \mathbf{P}_{\mathbb{D}}(A)\mathbf{P}_{\mathbb{D}}(B) \\
=& \lambda e^\dagger - \lambda e^\dagger(\mu_2 e^\dagger) \\
=& \lambda e^\dagger - \lambda e^\dagger(\mu_2 e^\dagger) \\
=& \lambda e^\dagger(1 - \mu_2 e^\dagger) \\
=& \lambda e^\dagger(1-\mu_2)e^\dagger \\
=& \mathbf{P}_{\mathbb{D}}(A)\mathbf{P}_{\mathbb{D}}(B^C)
\end{aligned}$$

■

Definición 23. *Teniendo n eventos aleatorios: $A_1, A_2, \ldots A_n$ decimos que son conjuntamente independientes si para cualquier subconjunto con $r \in \{2, \ldots, n$, se tiene que:*

$$\mathbf{P}_{\mathbb{D}}(A_{i_1} \cap A_{i_2} \cdots \cap A_{i_r)} = \mathbf{P}_{\mathbb{D}}(A_{i_1})\mathbf{P}_{\mathbb{D}}(A_{i_2}) \ldots \mathbf{P}_{\mathbb{D}}(A_{i_r})$$

Donde los subíndices cumplen que $1 \leq i_1 \leq i_2 \leq \ldots \leq i_r \leq n$

Así entonces, para el caso cuando $r = 2$, se dice que los n eventos son independientes a pares, y cuando los n eventos son mutuante independientes, es fácil generalizar que

$$\mathbf{P}_{\mathbb{D}}(A_1 \cap A_2 \cdots \cap A_n) = \mathbf{P}_{\mathbb{D}}(A_1)\mathbf{P}_{\mathbb{D}}(A_2) \ldots \mathbf{P}_{\mathbb{D}}(A_n) \tag{5.19}$$

Definición 24. *Sea espacio de probabilidad hiperbólica* $(\Omega, \sigma, \mathbb{P})$ *La colección de eventos* $\{H_1, H_2, \ldots, H_n\}$ *es un sistema fundamental de eventos si los* n *eventos eventos son mutuamente independientes con probabilidades no negativas, para los cuales se tiene que* $H_1 \cup H_2 \cup \cdots \cup H_n = \Omega$.

Teorema 6. *El teorema de probabilidad total establece que dado un espacio de probabilidad* $\mathbb{D}$ *definido por* $(\Omega, \sigma, \mathbf{P}_{\mathbb{D}})$*, un evento aleatorio* A *y un Sistema Fundamental de Eventos* $\{H_1, H_2, \ldots, H_n\}$*, entonces se cumple que:*

$$\mathbf{P}_{\mathbb{D}}(A) = \sum_{i=1}^{n} \mathbf{P}_{\mathbb{D}}(H_i)\mathbf{P}_{\mathbb{D}}(A|H_i) \tag{5.20}$$

Demostración: Sabemos que $A = A \cap \Omega = A \cap (H_1 \cup H_2 \cup \cdots \cup H_n) = (A \cap H_1) \cup (A \cap H_2) \cup \cdots \cup (A \cap H_i) = \bigcup_{i=1}^{n} A \cap H_i$, por lo tanto la probabilidad de A está dada por

$$\mathbf{P}_{\mathbb{D}}(A) = \sum_{i=1}^{n} \mathbf{P}_{\mathbb{D}}(A \cap H_i)$$

Multiplicando por $1 = \frac{\mathbf{P}_{\mathbb{D}}(H_i)}{\mathbf{P}_{\mathbb{D}}(H_i)}$ y usando el teorema de multiplicación donde $\mathbf{P}_{\mathbb{D}}(A|H_i) = \frac{\mathbf{P}_{\mathbb{D}}(A \cap H_i)}{\mathbf{P}_{\mathbb{D}}(H_i)}$, se tiene que:

$$\begin{aligned}
\mathbf{P}_{\mathbb{D}}(A) &= \sum_{i=1}^{n} \mathbf{P}_{\mathbb{D}}(A \cap H_i) \\
&= \sum_{i=1}^{n} \mathbf{P}_{\mathbb{D}}(A \cap H_i)\frac{\mathbf{P}_{\mathbb{D}}(H_i)}{\mathbf{P}_{\mathbb{D}}(H_i)} \\
&= \sum_{i=1}^{n} \mathbf{P}_{\mathbb{D}}(H_i)\frac{\mathbf{P}_{\mathbb{D}}(A \cap H_i)}{\mathbf{P}_{\mathbb{D}}(H_i)} \\
&= \sum_{i=1}^{n} \mathbf{P}_{\mathbb{D}}(H_i)\mathbf{P}_{\mathbb{D}}(A|H_i)
\end{aligned}$$

■

Teorema 7. *El teorema de Bayes establece que dado un espacio de probabilidad* $\mathbb{D}$ *definido por* $(\Omega, \sigma, \mathbf{P}_{\mathbb{D}})$*, un evento aleatorio* A *y un Sistema Fundamental de Eventos* $\{H_1, H_2, \ldots, H_n\}$*, entonces se cumple que:*

1. *Cuando* $\mathbf{P}_{\mathbb{D}}(A)$ *no es divisor de cero*

$$\mathbf{P}_{\mathbb{D}}(H_j|A) = \frac{\mathbf{P}_{\mathbb{D}}(H_j)\mathbf{P}_{\mathbb{D}}(A|H_j)}{\mathbf{P}_{\mathbb{D}}(A)}$$

2. *Cuando* $\mathbf{P}_{\mathbb{D}}(A)$ *es un divisor de cero de la forma* $\mathbf{P}_{\mathbb{D}}(A) = \lambda e$

$$\mathbf{P}_{\mathbb{D}}(H_j|A)e = \frac{\mathbf{P}_{\mathbb{D}}(H_j)\mathbf{P}_{\mathbb{D}}(A|H_j)}{\mathbf{P}_{\mathbb{D}}(A)}e$$

3. *Cuando* $\mathbf{P}_{\mathbb{D}}(A)$ *es un divisor de cero de la forma* $\mathbf{P}_{\mathbb{D}}(A) = \lambda e^{\dagger}$

$$\mathbf{P}_{\mathbb{D}}(H_j|A)e^{\dagger} = \frac{\mathbf{P}_{\mathbb{D}}(H_j)\mathbf{P}_{\mathbb{D}}(A|H_j)}{\mathbf{P}_{\mathbb{D}}(A)}e^{\dagger}$$

Donde $\mathbf{P}_{\mathbb{D}}(A)$ *se puede obtener mediante el teorema de probabilidad anterior descrito previamente.*

Demostración:

1. Usando nuevamente el teorema 1 de multiplicación y recordado que se cumple $\mathbf{P}_{\mathbb{D}}(A \cap H_j) = \mathbf{P}_{\mathbb{D}}(H_j \cap A)$, se tiene que

$$\mathbf{P}_{\mathbb{D}}(A \cap H_j) = \mathbf{P}_{\mathbb{D}}(H_j)\mathbf{P}_{\mathbb{D}}(A|H_j) = \mathbf{P}_{\mathbb{D}}(A)\mathbf{P}_{\mathbb{D}}(H_j|A)$$

Despejando para $\mathbf{P}_{\mathbb{D}}(H_j|A)e$ se encuentra que:

$$\mathbf{P}_{\mathbb{D}}(H_j|A) = \frac{\mathbf{P}_{\mathbb{D}}(H_j)\mathbf{P}_{\mathbb{D}}(A|H_j)}{\mathbf{P}_{\mathbb{D}}(A)}$$

2. Puesto que $\mathbf{P}_{\mathbb{D}}(A) = \lambda e$, se utiliza el caso 3 de la definición de Probabilidad Condicional:

$$\mathbf{P}_{\mathbb{D}}(H_j|A) = \frac{\mathbf{P}_{\mathbb{D}}(H_j \cap A)}{\lambda}e + \mathbf{P}_{\mathbb{D}}(H_j)e^{\dagger}$$

Multiplicando $\mathbf{P}_{\mathbb{D}}(A)e$ de ambos lados de la ecuación, sigue:

$$\begin{aligned}
\mathbf{P}_{\mathbb{D}}(H_j|A)\mathbf{P}_{\mathbb{D}}(A)e =& \left[\frac{\mathbf{P}_{\mathbb{D}}(H_j \cap A)}{\lambda}e + \mathbf{P}_{\mathbb{D}}(H_j)e^{\dagger}\right]\mathbf{P}_{\mathbb{D}}(A)e \\
=& \frac{\mathbf{P}_{\mathbb{D}}(H_j \cap A)\mathbf{P}_{\mathbb{D}}(A)}{\lambda}e + \mathbf{P}_{\mathbb{D}}(H_j)e^{\dagger}\mathbf{P}_{\mathbb{D}}(A)e \\
=& \frac{\mathbf{P}_{\mathbb{D}}(H_j \cap A)\lambda e}{\lambda}e \\
=& \mathbf{P}_{\mathbb{D}}(H_j \cap A)e = \mathbf{P}_{\mathbb{D}}(A \cap H_j)e \\
=& \mathbf{P}_{\mathbb{D}}(H_j)\mathbf{P}_{\mathbb{D}}(A|H_j)e
\end{aligned}$$

Despejando para $\mathbf{P}_{\mathbb{D}}(H_j|A)$ se llega a lo establecido en el punto 2 del teorema.

3. Puesto que $\mathbf{P}_{\mathbb{D}}(A) = \lambda e^{\dagger}$, se utiliza el caso 4 de la definición de Probabilidad Condicional:
$$\mathbf{P}_{\mathbb{D}}(H_j|A) = \mathbf{P}_{\mathbb{D}}(H_j)e + \frac{\mathbf{P}_{\mathbb{D}}(H_j \cap A)}{\lambda}e^{\dagger}$$
Multiplicando $\mathbf{P}_{\mathbb{D}}(A)e^{\dagger}$ de ambos lados de la ecuación, sigue:
$$\begin{aligned}\mathbf{P}_{\mathbb{D}}(H_j|A)\mathbf{P}_{\mathbb{D}}(A)e^{\dagger} &= \left[\mathbf{P}_{\mathbb{D}}(H_j)e + \frac{\mathbf{P}_{\mathbb{D}}(H_j \cap A)}{\lambda}e^{\dagger}\right]\mathbf{P}_{\mathbb{D}}(A)e^{\dagger} \\ &= \mathbf{P}_{\mathbb{D}}(H_j)e\mathbf{P}_{\mathbb{D}}(A)e^{\dagger} + \frac{\mathbf{P}_{\mathbb{D}}(H_j \cap A)\mathbf{P}_{\mathbb{D}}(A)}{\lambda}e^{\dagger} \\ &= \frac{\mathbf{P}_{\mathbb{D}}(H_j \cap A)\lambda e^{\dagger}}{\lambda}e^{\dagger} \\ &= \mathbf{P}_{\mathbb{D}}(H_j \cap A)e^{\dagger} = \mathbf{P}_{\mathbb{D}}(A \cap H_j)e^{\dagger} \\ &= \mathbf{P}_{\mathbb{D}}(H_j)\mathbf{P}_{\mathbb{D}}(A|H_j)e^{\dagger}\end{aligned}$$
Despejando para $\mathbf{P}_{\mathbb{D}}(H_j|A)e^{\dagger}$ se llega a lo establecido en el punto 3 del teorema.

■

Capítulo 6

Aplicación en Cadenas de Markov

Ahora bien, teniendo las bases teóricas en los números, probabilidades y propiedades hiperbólicas, es importante encontrar una manera eficaz de implementar estos métodos en las áreas ya conocidas de la probabilidad, no sólo para enriquecer esta teoría, sino también para hacer de ella una herramienta útil y práctica en la solución de problemáticas reales. Por ello, el desarrollo de la aplicación encajó ideal en un área de la probabilidad tradicional en la que es posible implementar gran parte de la teoría previamente establecida: las Cadenas de Markov.
Algunos principales ejemplos de las aplicaciones de estas cadenas de Markov son:

- **Procesos de ramificación:** describe un crecimiento poblacional, donde cada miembro de la n-ésima generación genera nuevos miembros independientemente a las generaciones pasadas.
- **Caminatas aleatorias:** se registran las saltos en los recorridos a lo largo del tiempo de un individuo, los movimientos futuros son independientes de los pasados.
- **Procesos de Poisson:** se usa para estudiar procesos que consideran continuidad en el tiempo, por lo general la frecuencia de algún evento dentro o la variación en los tiempos de llegada de este.
- **Cadenas de Markov Monte Carlo:** su uso es importante en estadística cuando se estima un parámetro desconocido y los resultados estadísticos de este resultan en una distribución.

Sabiendo que un proceso aleatorio cuya variación depende del tiempo es conocido como proceso estocástico, este tiene la **Propiedad Markoviana** si la variable depende únicamente del valor en el tiempo presente, es decir su futuro es independiente de su comportamiento en tiempos pasados.

Sea X_1 la variable aleatoria en el tiempo inicial y X_i la que define en el tiempo $i = \{2, 3, \ldots, n, \ldots\}$, entonces considerando los posibles estados que puede tomar la variable desde el tiempo inicial como el conjunto de $S = \{s_1, s_2, \ldots, s_n, \ldots\}$ un proceso estocástico se define como

$$\mathbf{P}(X_{n+1} = s_{n+1} | X_1 = s_1, X_2 = s_2, \ldots, X_n = s_n)$$

Donde el estado futuro de la variable aleatoria depende del estado presente y de los pasados, así pues, un proceso estocástico con la propiedad Markoviana donde sólo se depende del estado presente n, el estado futuro $n+1$ se define como

$$\mathbf{P}(X_{n+1} = s_{n+1} | X_n = s_n) \tag{6.1}$$

Definición 25. *La secuencia de estas variables aleatorias X_i es llamada una Cadena de Markov si cumplen con la propiedad descrita en la ecuación anterior (6.1). Además esta es llamada homogénea si las probabilidades condicionales no dependen del valor de n, esto es: $\forall i, j \in S$ se tiene que*

$$p_{i,j} = \mathbf{P}(X_{n+1} = j | X_n = i)$$

Definición 26. *La probabilidad condicional de que X_{n+1} llegue al estado j dado que la variable aleatoria X_n se encuentra en el estado i es llamada* ***probabilidad de transición de un paso***

Así entonces asumiendo para este trabajo que existe la homogeneidad en las cadenas de Markov, para proceder con los cálculos de probabilidades es necesario considerar

1. Matriz de transición P
   Contiene en una matriz las probabilidades de transición de paso (o tiempo) al siguiente. Esto es $P = (p_{i,j} : i, j \in S)$, dado que:

   $$p_{i,j} = \mathbf{P}(X_{n+1} = j | X_n = i) \geq 0$$

   Se le dice que es **estocástica** puesto que cada renglón de la matriz debe contener la probabilidad total de los eventos, ó sea sumar 1:

   $$\sum_{j \in S} p_{i,j} = 1$$

   Demostración: Puesto que todo $p_{i,j}$ es un número no negativo entre cero y uno, entonces

   $$\sum_{j \in S} p_{i,j} = \sum_{j \in S} \mathbf{P}(X_{n+1} = j | X_n = i) = 1$$

2. Distribución inicial λ
   Plantea las probabilidades para todos posibles estados iniciales que X_0 pueda tomar. Esto es $\lambda = (\lambda_i : i \in S)$ dado que:

$$\lambda_i = \mathbf{P}(X_0 = i) \geq 0$$

   De la misma manera la suma de todas estas probabilidades deben resultar en 1:

$$\sum_{i \in S} \lambda_i = 1$$

   Demostración: Puesto que todo λ_i es un número no negativo entre cero y uno, entonces

$$\sum_{i \in S} \lambda_i = \sum_{i \in S} \mathbf{P}(X_0 = i) = 1$$

Así entonces, teniendo las probabilidades de la matriz de transición entre dos tiempos y la distribución inicial para los estados, es posible encontrar la **distribución conjunta**, la cual genera la probabilidad en el tiempo n que considera de cada uno de las probabilidades de los estados previos.

Definición 27. *Sea i_k el estado que tomó X_k para $k \in \{0, 1, 2, \ldots, n\}$, entonces la **distribución conjunta** de X_n estará dada por*

$$\mathbf{P}(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \lambda_{i_0} p_{i_0,i_1} p_{i_1,i_2} \ldots p_{i_{n-1},i_n} \tag{6.2}$$

Teorema 8. *Sea la distribución inicial λ y la matriz de transición P. La secuencia de variables aleatorias $X = (X_n : n \geq 0)$ es una cadena Markov con distribución inicial λ y matriz de transición P, si y solo si cumple con la distribución conjunta descrita previamente en la ecuación (6.2)*

- **Demostración**
  Asumiendo a X como una cadena de Markov con distribución inicial λ y matriz de transición P y considerando A_k como el evento $X_k = i_k$, entonces la ecuación (6.2) puede ser reescrita como

$$\mathbf{P}\left(A_0 \cap A_1 \cap A_2 \cap \cdots \cap A_n\right) = \lambda_{i_0} p_{i_0,i_1} p_{i_1,i_2} \ldots p_{i_{n-1},i_n} \tag{6.3}$$

  Retomando propiedades de multiplicación condicional

$$\mathbf{P}\left(A_0 \cap A_1 \cap A_2 \cap \cdots \cap A_n\right) =$$
$$\mathbf{P}\left(A_0 \cap A_1 \cap A_2 \cap \cdots \cap A_{n-1}\right) \mathbf{P}\left(A_n | A_0 \cap A_1 \cap A_2 \cap \cdots \cap A_{n-1}\right)$$

Dada la propiedad markoviana entonces

$$\mathbf{P}\left(A_0 \cap A_1 \cap A_2 \cap \cdots \cap A_n\right) = \\ \mathbf{P}\left(A_0 \cap A_1 \cap A_2 \cap \cdots \cap A_{n-1}\right) \mathbf{P}\left(A_n | A_{n-1}\right)$$

Asumiendo homogeneidad en el proceso y retomando la notación establecida previamente $\mathbf{P}\left(A_n|A_{n-1}\right) = \mathbf{P}(X_n = i_n | X_{n-1} = i_{n-1}) = p_{i_{n-1},i_n}$ Se reescribe

$$\mathbf{P}\left(A_0 \cap A_1 \cap A_2 \cap \cdots \cap A_n\right) = \\ \mathbf{P}\left(A_0 \cap A_1 \cap A_2 \cap \cdots \cap A_{n-1}\right) p_{i_{n-1},i_n}$$

De esta forma, se sigue el mismo método, por ejemplo: $\mathbf{P}\left(A_0 \cap A_1 \cap \cdots \cap A_{n-1}\right) = \mathbf{P}\left(A_0 \cap A_1 \cap \cdots \cap A_{n-2}\right) p_{i_{n-2},i_n-1}$. Se sigue este proceso hasta desglosar en todas las probabilidades de transición y la del estado inicial mostrada en el lado derecho de la ecuación (6.2)
Ahora, sabiendo que es cierta la ecuación 6.3 para todo n, cuando $n = 0$ encontramos la distribución inicial $\mathbf{P}(A_0) = \lambda_{i_0}$ y teniendo

$$\mathbf{P}\left(A_{n+1} | A_0 \cap A_1 \cap \cdots \cap A_n\right) = \frac{\mathbf{P}\left(A_0 \cap A_1 \cap \cdots \cap A_{n-1}\right)}{\mathbf{P}\left(A_0 \cap A_1 \cap \cdots \cap A_n\right))}$$

Entonces reescribiéndolo

$$\begin{aligned} \mathbf{P}\left(A_{n+1} | A_0 \cap A_1 \cap \cdots \cap A_n\right) =& \frac{\mathbf{P}\left(A_0 \cap A_1 \cap \cdots \cap A_{n-1}\right)}{\mathbf{P}\left(A_0 \cap A_1 \cap \cdots \cap A_n\right))} \\ =& \frac{\lambda_{i_0} p_{i_0,i_1} p_{i_1,i_2} \dots p_{i_{n-1},i_n} p_{i_n,i_{n+1}}}{\lambda_{i_0} p_{i_0,i_1} p_{i_1,i_2} \dots p_{i_{n-1},i_n}} \end{aligned}$$

Lo cual se reduce a la expresión

$$\mathbf{P}\left(A_{n+1} | A_0 \cap A_1 \cap \cdots \cap A_n\right) = p_{i_n,i_{n+1}}$$

Lo cual indica que el estado i_{n+1} depende únicamente del estado i_n y no de los estados previos $i_0, i_1, \dots, i_{n-1}$, por lo cual se dice que X es una cadena homogénea de Markov con cadena de transición P y distribución inicial λ ■

Además de ser homogénea, una cadena de Markov puede contar con una distribución estacionaria

Definición 28. *Suponiendo una distribución τ en X, se dice que τ es estacionaria si en el tiempo inicial 0 y en el tiempo siguiente 1 se tiene que $\tau_0 = \tau_1 = \tau$.*
Por lo tanto una distribución es estacionaria si satisface que

$$\tau = \tau P$$

Teorema 9. *Si una cadena de Markov empieza con una distribución estacionaria τ, se mantiene para siempre estacionaria.*

Demostración:
Teniendo una cadena de Markov donde $\tau_n(i)$ denota la distribución en el tiempo n, esto es $\tau_n(i) = \mathbf{P}[X_n = i]$ y considerando un espacio de estados finito $S = [1, \ldots, N]$, entonces esta cadena tiene una matriz de transición cuadrada de $N \times N$ que denotamos previamente como $p_{i,j}$
Sabiendo que

$$p_{i,j} = \mathbf{P}(X_{n+1} = j | X_n = i)$$

Por la ley de la probabilidad total se tiene que

$$\begin{aligned}
\tau_{n+1}(j) &= \mathbf{P}[X_{n+1} = j] \\
&= \sum_{i=1}^{N} \mathbf{P}[X_n = i]\mathbf{P}[X_{n+1} = j | X_n = i] \\
&= \sum_{i=1}^{N} \tau_n(i) p_{i,j}
\end{aligned}$$

Considerando que $\tau_n = [\tau_n(1), \ldots, \tau_n(N)]$ entonces es cierto que

$$\tau_{n+1} = \tau_n P \tag{6.4}$$

Por lo tanto, evaluando n en los diferentes estados se desarrolla

$$\begin{aligned}
\tau_1 &= \tau_0 P \\
\tau_2 &= \tau_1 P = \tau_0 P^2 \\
\tau_3 &= \tau_2 P = \tau_1 P^2 = \tau_0 P^3
\end{aligned}$$

Generalizando esto a $\tau_n = \tau_0 P^n$

Así entonces, para una distribución estacionaria es cierto que

$$\tau = \tau P$$

Por lo tanto, si una cadena de Markov empieza con una distribución estacionaria τ, se mantiene para siempre estacionaria.

Corolario 2. *Teniendo una matriz de probabilidad P simétrica NxN, entonces la distribución uniforme tal que $\lambda_i = \frac{1}{N}$, para todo $i \in [1, 2, ..., N]$ es estacionaria.*

Definición 29. *Se dice que una matriz de transición P es* ***doblemente estocástica*** *si la suma de las columnas y las filas resulta en 1.*

Ahora bien, podemos afirmar de la definición anterior de la representación matricial para los hiperbólicos señalada en capítulos anteriores [3.8], tiene la característica de ser una matriz de transición doblemente estocástica cuando se cumple que $x \in [0,1]$, y la variable $y = 1 - x$.
Ejemplo
Sea el número hiperbólico $P = .75 + .25\mathbf{k}$, entonces se tiene que $x = .75$ y se cumple que $y = 1 - .75 = .25$ su representación matricial es

$$P = \begin{pmatrix} .75 & .25 \\ .25 & .75 \end{pmatrix}$$

Se cumple que es una matriz simétrica donde la suma de sus filas y sus columnas resultan en 1 respectivamente, por lo tanto P es una matriz de probabilidad doblemente estocástica.

6.1. Potencias de matrices doblemente estocástica

Es posible obtener la matriz de probabilidades de transición de n pasos al elevar a la n potencia la matriz de probabilidades de transición de un paso.
Ahora bien si consideramos una matriz dimensión 2×2 doblemente estocástica, es posible transformar esta a su versión hiperbólica, puesto que las entradas de las diagonales en la matriz cumplen con la forma matricial de un hiperbólico.
Así entonces, si se busca encontrar la potencia n de dicha matriz, es posible hacerlo al potenciar su transformación de número hiperbólico en base idempotente.
Esto es, sea la matriz doblemente estocástica

$$P = \begin{pmatrix} x & y \\ y & x \end{pmatrix}$$

Esta se puede reescribir en número hiperbólico como

$$p = x + y\mathbf{k}$$

Cuya transformación a base idempotente, considerando que $x + y = 1$ por ser P una matriz doblemente estocástica, resultaría en

$$p = e + (x - y)e^{\dagger}$$

Recordando la propiedades de potencias para números hiperbólicos idempotentes [4.7], la n-ésima potencia de Z estaría dada por

$$p^n = e + (x - y)^n e^{\dagger}$$

Ahora bien, al transformar nuevamente a su forma matricial, es importante comprobar que P^n cumple las condiciones de una matriz doblemente estocástica. Así entonces, considerando las formas matriciales descritas en las ecuaciones (4.8) y (4.9) para e como la y $e^\dagger$ se tendría

$$\begin{aligned} P^n =& \mathbf{E} + (x-y)^n \mathbf{E}^\dagger \\ =& \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} + (x-y)^n \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \\ =& \begin{pmatrix} \frac{1}{2} + \frac{(x-y)^n}{2} & \frac{1}{2} - \frac{(x-y)^n}{2} \\ \frac{1}{2} - \frac{(x-y)^n}{2} & \frac{1}{2} + \frac{(x-y)^n}{2} \end{pmatrix} \end{aligned}$$

Para probar la doble estocásticidad de la matriz P^n es necesario probar que

1. $\frac{1}{2} + \frac{(x-y)^n}{2} + \frac{1}{2} - \frac{(x-y)^n}{2} = 1$
2. $0 \leq \frac{1}{2} + \frac{(x-y)^n}{2} \leq 1$
3. $0 \leq \frac{1}{2} - \frac{(x-y)^n}{2} \leq 1$

La primera se demuestra por sí misma. Para probar la segunda y la tercera es necesario tomar en cuenta que $x + y = 1$, entonces $y = 1 - x$ y por lo tanto $x - y = x - (1 - x) = 2x - 1$.
Así transformando la desigualdad principal de restricción para los valores de X, se tiene

$$\begin{aligned} 0 \leq x \leq 1 & \\ 0 \leq 2x \leq 2 & \qquad \textit{Duplicando la inecuación} \\ -1 \leq 2x - 1 \leq 1 & \qquad \textit{Restando uno} \\ -1 \leq 2x - 1 \leq 1 & \qquad \textit{Restando uno} \\ -1 \leq x - y \leq 1 & \qquad \textit{Usando la igualdad previa} \end{aligned}$$

Ahora bien, para llegar a las inecuaciones en los puntos 2 y 3, se eleva a la n-ésima potencia

$$-1^n \leq (x-y)^n \leq 1^n \tag{6.5}$$

Considerando diferentes casos para n par e impar y $x - y$ positivo o negativo encontramos dos resultados para esta inecuación.
Para los casos donde:

- n es par y $x - y$ es positivo
- n es impar y $x - y$ es positivo

- n es par y $x-y$ es negativo

$$0 \leq (x-y)^n \leq 1$$

- n es impar y $x-y$ es negativo

$$-1 \leq (x-y)^n \leq 0$$

Ahora bien, transformando ambos casos para buscar la inecuación 2.

$$\begin{aligned} 0 \leq (x-y)^n \leq 1 \qquad & \qquad -1 \leq (x-y)^n \leq 0 \\ 0 \leq \frac{(x-y)^n}{2} \leq \frac{1}{2} \qquad & \qquad -\frac{1}{2} \leq \frac{(x-y)^n}{2} \leq 0 \qquad (6.6) \\ \frac{1}{2} \leq \frac{1}{2} + \frac{(x-y)^n}{2} \leq 1 \qquad & \qquad 0 \leq \frac{1}{2} + \frac{(x-y)^n}{2} \leq \frac{1}{2} \qquad (6.7) \end{aligned}$$

Se puede observar en (6.6) que las ecuaciones se dividen entre dos, y posteriormente en (6.7) se les suma un medio para llegar a la inecuación deseada. Así, podemos ver que para cualquier caso se cumple el requisito 2. para las entradas en la diagonal principal de la matriz potencia.

Ahora bien, buscando demostrar el requisito 3. se sigue el mismo procedimiento transformando las inecuaciones.

$$\begin{aligned} 0 \leq (x-y)^n \leq 1 \qquad & \qquad -1 \leq (x-y)^n \leq 0 \\ -1 \leq -(x-y)^n \leq 0 \qquad & \qquad 0 \leq -(x-y)^n \leq 1 \qquad (6.8) \\ -\frac{1}{2} \leq -\frac{(x-y)^n}{2} \leq 0 \qquad & \qquad 0 \leq -\frac{(x-y)^n}{2} \leq \frac{1}{2} \qquad (6.9) \\ 0 \leq \frac{1}{2} - \frac{(x-y)^n}{2} \leq \frac{1}{2} \qquad & \qquad \frac{1}{2} \leq \frac{1}{2} - \frac{(x-y)^n}{2} \leq 1 \qquad (6.10) \end{aligned}$$

Mediante las ecuaciones en (6.8) se multiplica por uno negativo y se invierte la inecuación, posteriormente con (6.9) se divide entre dos, para concluir la transformación al sumar un medio en las ecuaciones (6.10). Con ello encontramos que las entradas en la diagonal secundaria para la matriz potencia de igual manera cumple con la restricción 3. Por lo tanto cumple con ser una matriz doblemente estocástica al recurrir a su transformación matricial.

Ejemplo Supóngase un país con una población en edad productiva tiene una tasa de desempleo del 3.7 % durante un trimestre del año, asumimos así que la probabilidad de estar empleado en un trimestre es del .963 contra la probabilidad de desempleo de .037. Si consideramos los estados como *empleado* y *desempleado*, podemos armar una matriz de transición para representar la probabilidad de pasar de estar empleado a estar desempleado de un trimestre a otro. Así entonces hacemos $x = .963, y = .037$ y obtenemos la matriz de transición de un trimestre como

$$P = \begin{pmatrix} .963 & .037 \\ .037 & .963 \end{pmatrix}$$

Si deseáramos saber cuál es la probabilidad de alguien empleado de quedar desempleado el año siguiente, entonces se busca la transición en el paso 4, por lo que se eleva la matriz de probabilidad a la cuarta potencia. Mediante lo descrito anteriormente, es transformamos esta matriz al número hiperbólico $p = .963 + .037\mathbf{k}$, y llevándolo a la base idempotente esto es $p = e + .926e^{\dagger}$

Capítulo 7

Conclusión

A lo largo de este trabajo de investigación, se utilizaron las bases teóricas ya conocidas de álgebra, teoría de la medida y probabilidad para desarrollar una nueva teoría de probabilidad que fuera funcional en una doble dimensionalidad. Se buscó probar la construcción de probabilidades bi-dimensionalidades mediante números no reales, en específico a través del uso de hiperbólicos.
Así pues, entendiendo bien el funcionamiento de los números hiperbólicos, sus propiedades y comprendiendo las tranformaciones a la base idempotente, se estableció una relación de orden que creara un álgebra funcional dentro de el espacio hiperbólico.
De esta manera, fue posible adentrarse a desarrollar mediante la teoría de la medida un espacio de probabilidad hiperbólica que cumpliera con los axiomas necesarios para la correcta funcionalidad de las probabilidades considerando números hiperbólicos en su base idempotente. Y se comprobó que funciona a la perfección considerando las propiedades de complemento, unión, intersección, independencia y condicionalidad de eventos.

Considerando lo anterior, se logra contestar la pregunta inicial de investigación planteada en la introducción, afirmando que si es posible desarrollar una nueva teoría de probabilidad considerando un espacio bidimensional.
Respecto al objetivo general podemos afirmar que esta nueva construcción de la probabilidad considerando números no reales brinda nuevos métodos que amplían el estudio de la probabilidad tradicional conservando sus propiedades. De esta manera, es importante remarcar que la teoría de probabilidad comúnmente conocida, en especial se encuentra como caso particular de la nueva teoría desarrollada.
Ahora bien, recordando los objetivos específicos, se logró hacer aplicable lo desarrollado mediante la rama de probabilidad de Cadenas de Markov, puesto que el espacio de probabilidad hiperbólico creado, mediante la propiedad de la transformación matricial de los hiperbólicos en base idempotente, permitió generar matrices doble-

mente estocásticas. la investigación y el desarrollo de varias demostraciones llevó al descubrimiento en el que la matriz potencia obtenida de la transformación de la potencia de un hiperbólico en base idempotente derivado de una matriz doblemente estocástica, asimismo es doblemente estocástica.
No sólo se lograron los objetivos generales y específicos, sino que se encontró una conexión entre los números hiperbólicos y las cadenas de Markov que no se había planteado previamente, y se considera un resultado de suma importancia dentro de esta investigación .
Por otro lado es importante hacer hincapié en que la escasez de información de este tema fue un factor importante para que la teoría se tuviera que desarrollar desde lo básico. Muchas demostraciones fueron elaboradas completamente desde el inicio para la correcta construcción de la teoría, con base en lo aprendido en varias áreas de mi formación universitaria, tales como Álgebra Superior, Teoría de Matrices, Modelos Estocásticos, Análisis Matemático, entre otras.
Se afirma que esta investigación está mayormente dirigida a las áreas de ciencias exactas, puesto que la comprensión del contenido en este trabajo requiere de un amplio conocimiento en la probabilidad y buen entendimiento de los números no reales.
Con esto, se concluye afirmando que considerando que se ha desarrollado un espacio de probabilidad hiperbólica que permite la bidimensional, es posible la aplicación de esta nueva teoría de probabilidad en varios ámbitos, en este trabajo se demostró para el caso en particular de Cadenas de Markov.
Sabiendo que la teoría de probabilidad tradicional es útil para facilitar la predicción de eventos y la estimación de inertidumbre, y entendiendo que ésta es un caso particular de la nueva probabilidad hiperbólica, es posible encontrar más aplicaciones que solucionen problemas en cualquier área de estudio donde se requiera, ampliando la probabilidad a un espacio bidimensional.

Bibliografía

[1] Luna-Elizarrarás, M. E., Shapiro, M., Struppa, D. C., Vajiac, A. (2015). Bicomplex holomorphic functions: the algebra, geometry and analysis of bicomplex numbers. Birkhäuser.

[2] Gómez Laveaga, C. (2014). Álgebra Superior Curso Completo. Distrito Federal, México: Universidad Nacional Autónoma de México.

[3] Alpay, D., Luna-Elizarrarás, M. E., & Shapiro, M. (2016, 14 julio). Kolmogorov's Axioms for Probabilities with Values in Hyperbolic Numbers. Recuperado de `https://doi.org/10.1007/s00006-016-0706-6`

[4] Mathai, A. M., & Rathie, P. N. (1997). Probability and Statistics (1.a ed.). London, United Kingdom: Palgrave Macmillan. Recuperado de `https://doi.org/10.1007/978-1-349-02767-5`

[5] Parker, S. P. (1984). McGraw-Hill concise encyclopedia of science & technology. McGraw-Hill.

[6] Halmos, P. R. (2013). Measure theory (Vol. 18). Springer.

[7] Tsitsiklis, J. (2005). 6.436 J/15.085 J Fundamentals of Probability, Fall 2005.

[8] Kulyabov, D. S., Korolkova, A. V., & Gevorkyan, M. N. (2020). Hyperbolic numbers as Einstein numbers. Journal of Physics: Conference Series, 1557(1), 012027. `https://doi.org/10.1088/1742-6596/1557/1/012027`

[9] Nagwa. (2022). Lesson Explainer: Matrix Representation of Complex Numbers. Recuperado 10 de febrero de 2022, de `https://www.nagwa.com/en/explainers/152196980513/`