



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS MATEMÁTICAS Y
DE LA ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA

**EL MOVIMIENTO BROWNIANO
EN LA ESFERA**

T E S I S

QUE PARA OPTAR POR EL GRADO DE:

MAESTRA EN CIENCIAS

P R E S E N T A:

XIAOPING MA

DIRECTOR DE TESIS:

DR. GERÓNIMO FRANCISCO URIBE BRAVO
(INSTITUTO DE MATEMÁTICAS, U.N.A.M.)

CIUDAD UNIVERSITARIA, Cd. Mx., Nov. 2021



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

Agradecimientos

Primeramente agradezco a toda mi familia, en especial a mis padres y mi tío por su comprensión, confianza y apoyo brindado por siempre.

Agradezco a mis docentes, sinodales y en especial a mi tutor, el Dr. Gerónimo Francisco Uribe Bravo, por haberme brindado su apoyo en toda la carrera del posgrado, su paciencia y dedicación de tiempo.

Le agradezco también a mis amigos por la confianza y el apoyo, en especial muchas gracias a la amiga Griselda Barón por su apoyo en la revisión gramática de la tesis.

Agradezco a CONACyT, por la beca de maestría que me brindó. Por último, la investigación fue realizada gracias al Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) de la UNAM, proyecto IN114720. Agradezco a la DGAPA-UNAM por la beca recibida.

Índice general

Agradecimientos	iii
Introducción	vii
1. El teorema de Arquímedes	1
1.1. Introducción	1
1.2. Distribución uniforme en la esfera \mathcal{S}^{d-1} y en la bola \mathcal{B}^d	3
1.2.1. Simulación de puntos uniformes en \mathcal{S}^{d-1} y en \mathcal{B}^d	6
1.3. Teorema de Arquímedes vía álgebra Beta-Gamma	7
2. El movimiento browniano en \mathbb{R}^d y en la esfera \mathcal{S}^{d-1}	11
2.1. El movimiento browniano en \mathbb{R}^d	11
2.2. La parte radial del movimiento browniano	20
2.2.1. Simulación exacta del proceso de Bessel	28
2.3. El movimiento browniano en la esfera \mathcal{S}^{d-1}	30
2.3.1. El operador de Laplace-Beltrami en \mathcal{S}^{d-1}	33
2.3.2. Representación skew-product del $\mathbf{MB}(\mathbb{R}^d)$	35
2.3.3. Representación skew-product del $\mathbf{MB}(\mathcal{S}^{d-1})$	36
3. La difusión de Wright-Fisher	40
3.1. El modelo genético de Wright-Fisher	41
3.1.1. El modelo de Wright-Fisher con mutación	43
3.2. La difusión de Wright-Fisher con mutación	43
3.3. Simulación exacta de la difusión de Wright-Fisher con mutación	46
4. Proyección del movimiento browniano $\mathbf{MB}(\mathcal{S}^{d-1})$	51
4.1. Proyección del movimiento browniano esférico	51
4.2. Representación skew-product del $\mathbf{MB}(\mathcal{S}^{d-1})$ y su relación con la difusión de Wright-Fisher	56
4.3. Simulación exacta del $\mathbf{MB}(\mathcal{S}^{d-1})$	60
Apéndice A. Implementación de algoritmos	65
Apéndice B. Sobre la simulación exacta del proceso ancestral	73
Bibliografía	76

Introducción

El movimiento browniano describe un movimiento aleatorio e irregular de las partículas pequeñas suspendidas en un medio fluido (líquido o gas), como resultado de choques contra las moléculas de dicho fluido. Este fenómeno fue observado primeramente por el botánico escocés Robert Brown en 1827, quien descubrió a través de su microscopio, el movimiento errático de partículas de polen en suspensión en el agua, sin embargo, no fue capaz de determinar la causa de este movimiento. Fue hasta el año 1905, que Albert Einstein publicó un artículo donde explicó que el movimiento observado por Brown era resultado de colisiones aleatorias con las moléculas de agua; y estudió el fenómeno browniano donde se dio cuenta, entre otras cosas, del efecto físico que tiene la temperatura, la viscosidad del fluido y el tamaño de la partícula. El estudio fue proseguido por los físicos de tanto renombre como Smoluchowski, Fokker, Planck, Ornstein y otros.

Desde el punto de vista matemático, el primer modelo del movimiento browniano fue desarrollado por Norbert Wiener, quien inició en 1918 el estudio del movimiento browniano, puso la teoría en una base matemática firme, y originó el desarrollo de la teoría de los procesos estocásticos. Bajo la influencia de Wiener, Andrei Kolmogórov, Paul Lévy y la de Kiyosi Itô, la teoría se desarrolló rápidamente.

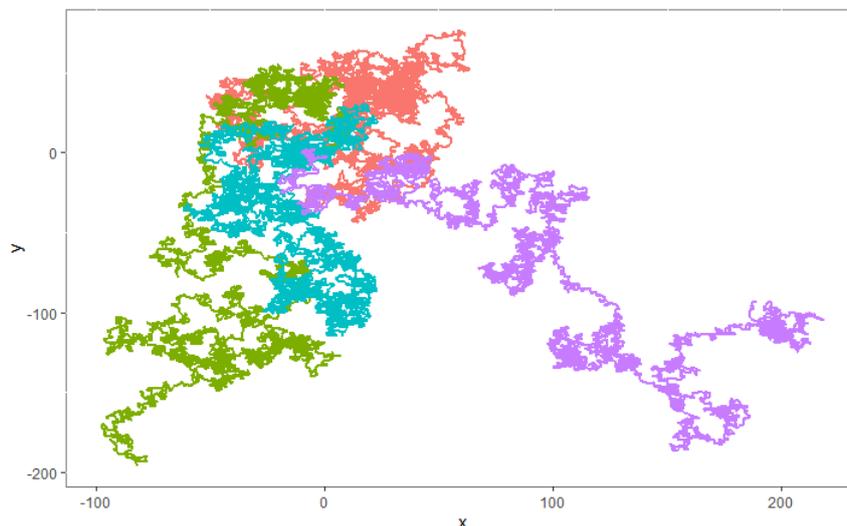


Figura 1: Ejemplo de de 4 partículas en movimiento aleatorio bi-dimensional.

Hasta hoy en día, el movimiento browniano ya se ha estudiado durante más de cien años, y ha adquirido gran influencia en numerosos áreas de las ciencias puras y aplicadas como la física, la química, la biología, la economía y la sociología; entre ellos muchos fenómenos pueden describirse por el movimiento browniano: tales como el movimiento aleatorio de bacterias pequeñas, las fluctuaciones de los mercados de acciones, las partículas de humo suspendidas en el aire, etc. Aunque se lleva más de cien años de investigación con el movimiento browniano (lo abreviaremos por MB), el estudio de los movimientos aleatorios representado por el MB sigue siendo de interés. Por ejemplo, se interesan los movimientos

aleatorios en el plano bidimensional, en las superficies curvadas o en las variedades de alta dimensión ([Hsu02]). En particular, nos interesa el movimiento browniano en la superficie de una esfera, que sea de dimensión 2 o de dimensión mayor.

¿Por qué el movimiento en la esfera? Por ejemplo, como los elefantes marinos, que recorren a grandes distancias y son rastreados, debido a que los biólogos se interesan en describir sus rutas. Los animales buscarán alimentos a lo largo del camino, es decir, se apartarán de la ruta directa desde el origen hasta el destino, y esto puede modelarse como una fluctuación estocástica ([Bri97]). Además, la proteína de una membrana celular que siempre hace movimientos irregulares sobre la membrana; y el movimiento de una bacteria en el líquido giratorio, etc. En la naturaleza no existen solamente superficies suaves, sino también otros tipos de superficies curvadas, por lo que, el estudio del movimiento browniano en las superficies tiene su importancia en la aplicación a la vida real.

El objetivo principal de la presente tesis es presentar las bases necesarias para entender el movimiento browniano en la esfera unitaria \mathcal{S}^{d-1} del espacio euclidiano \mathbb{R}^d de dimensión $d \geq 2$, y tomando como base el artículo de [MMUB20] para la simulación computacional de muestras exactas del incremento de un MB en la esfera.

En el Capítulo 1, introducimos el teorema de Arquímedes en la teoría de la probabilidad, donde se establece que la proyección ortogonal de la esfera unitaria $\mathcal{S}^2 \subset \mathbb{R}^3$ en cualesquiera de sus coordenadas se preserva la distribución uniforme, y en la sección 1.2 se estudia la distribución uniforme en la esfera \mathcal{S}^{d-1} y en la bola unitaria \mathcal{B}^d , en donde vemos que la distribución uniforme en la esfera es la única medida invariante bajo rotaciones. Así, con la definición de distribuciones uniformes en la esfera y en la bola, se dan los algoritmos para la simulación de muestras de puntos uniformes en \mathcal{S}^{d-1} y en \mathcal{B}^d , y la implementación de algoritmos (ver Apéndice A) nos permite visualizar el teorema de Arquímedes en términos probabilísticos y su generalización a la esfera \mathcal{S}^{d-1} de dimensión $d \geq 3$ ([AM13], [CDH17]). También se presentan algunas propiedades probabilísticas de la proyección ortogonal de una distribución uniforme en la esfera \mathcal{S}^{d-1} sobre la bola \mathcal{B}^m para $m < d$, para posteriormente en el Capítulo 4 verificar que la caracterización de la distribución uniforme como la medida invariante para el movimiento browniano en la esfera, se preserva bajo proyección ortogonal.

En el Capítulo 2, revisamos la definición y las propiedades importantes del movimiento browniano en el espacio euclidiano, y se estudia su parte radial como un proceso de Bessel en la sección 2.2, con su función de densidad de transición dada explícitamente ([RY99]). Además, se obtiene un algoritmo para simular muestras exactas de una trayectoria del proceso de Bessel. Por otro lado, al considerar la esfera \mathcal{S}^{d-1} como una variedad rotacionalmente simétrica e inmersa en el espacio ambiente \mathbb{R}^d , que el movimiento browniano en la variedad es un proceso típico de Markov a tiempo continuo, y la simple definición es: Sea \mathcal{M} una variedad riemanniana, entonces el movimiento browniano en \mathcal{M} es un proceso de difusión \mathbf{Z} con espacio de estado el espacio de trayectorias $\mathbb{W}(\mathcal{M})$ en la variedad \mathcal{M} , y con generador $\frac{1}{2}\Delta_{\mathcal{M}}$, donde $\Delta_{\mathcal{M}}$ es el operador de Laplace-Beltrami en \mathcal{M} ([Hsu02]). Concretamente, en el Capítulo 2, se define al movimiento browniano en la esfera como solución única a una ecuación diferencial estocástica llamada *la representación de Stroock para el movimiento browniano esférico*. Finalmente, se estudiará la descomposición skew-product del movimiento browniano en \mathbb{R}^d , y análogamente, el movimiento browniano \mathbf{Z} en la esfera también tiene una descomposición skew-product en coordenadas polares $(r, \boldsymbol{\theta})$ (cf. [IM65], [RW00], [Hsu02]). Presentamos la demostración completa de estos dos resultados usando la técnica del artículo de [MMUB18] para la descomposición skew-product de una difusión que vive en la bola unitaria.

Sin embargo, como la transformación de coordenadas polares $\mathbf{z} = (r, \boldsymbol{\theta}) \in \mathcal{S}^{d-1}$ en

coordenadas cartesianas está dada por $\mathbf{z} = ((\sin r)\boldsymbol{\theta}^\top, \cos r)^\top \in \mathbb{R}^d$, entonces, podemos expresar el movimiento browniano \mathbf{Z} en términos de las coordenadas polares por $\mathbf{Z} = ((\sin r)\boldsymbol{\theta}^\top, \cos r)^\top$. Sea $Z^d = \cos r$ la última coordenada de \mathbf{Z} , bajo una transformación lineal $g(z) = \frac{1-z}{2}$ para Z^d , se lleva a un proceso de difusión de Wright-Fisher neutral con mutación ([MMUB20]); esto es, el proceso que introducimos su definición en el Capítulo 3, como un modelo genético que describe la evolución de frecuencia alélica en una población a lo largo de generaciones, vemos que está definida como una cadena de Markov, y por la aproximación de difusión, se entiende que un proceso de difusión de Wright-Fisher con mutación es la única solución fuerte a una ecuación diferencial estocástica y con sus valores en el intervalo $[0, 1]$. Como la difusión de Wright-Fisher neutral con mutación tiene la densidad de transición explícita ([GS10]), y el artículo de [JS17] presentó un algoritmo para la simulación de muestras exactas de la densidad de transición de una difusión de Wright-Fisher con mutación, así también el algoritmo para la simulación de muestras de la distribución del proceso puro de muerte A_∞^θ . Revisaremos estos algoritmos en la sección 3.3, y se realiza la implementación del algoritmo para generar muestra de trayectoria de una difusión de Wright-Fisher, con lo que se puede analizar y visualizar las propiedades de sus fronteras y su distribución estacionaria.

Por último, en el Capítulo 4, primero revisamos los resultados principales del artículo de [MMUB18] sobre la proyección del movimiento browniano esférico. Donde se entiende que la caracterización de la distribución uniforme en \mathcal{S}^{d-1} como, la única medida invariante para el movimiento browniano en la esfera se preserva bajo proyección ortogonal; y se establece que el proceso \mathbf{X} obtenido de la proyección del movimiento browniano \mathbf{Z} en la esfera \mathcal{S}^{n+k-1} sobre la bola \mathcal{B}^n tiene una descomposición skew-product análoga a la del MB en \mathbb{R}^n , y su norma al cuadrado es una difusión de Wright-Fisher con mutación. Con estos resultados, como se presenta en la sección 4.2, se puede expresar a \mathbf{Z} en sus primeras $(d-1)$ coordenadas \mathbf{X} junto con su última coordenada Z^d , donde \mathbf{X} tiene una descomposición skew-product con su parte radial como transformación de una difusión de Wright-Fisher con mutación \mathcal{Y} , y su parte angular es un movimiento browniano $\boldsymbol{\varphi}$ en la esfera \mathcal{S}^{d-2} con escala de tiempo cambiado en función de la difusión de Wright-Fisher \mathcal{Y} , que es estacionario (i.e., $\boldsymbol{\varphi}_t$ distribuye uniformemente en la esfera \mathcal{S}^{d-2} para todo $t > 0$) cuando \mathbf{Z} tiene valor inicial en el polo norte $\mathbf{o} = (0, \dots, 0, 1)^\top \in \mathcal{S}^{d-1}$ o en el antipodal $-\mathbf{o}$, el polo sur, de acuerdo a las propiedades de fronteras de la difusión de Wright-Fisher con mutación. Esto es, otra presentación de la descomposición de skew-product del movimiento browniano \mathbf{Z} en la esfera \mathcal{S}^{d-1} ([MMUB20]). Y retomamos la descomposición skew-product del movimiento browniano $\mathbf{Z} = (r, \boldsymbol{\theta})$ en \mathcal{S}^{d-1} estudiado en la sección 2.3.3, con la transformación sobre su componente radial: $g(r) = \frac{1-\cos r}{2}$, y transformada en coordenadas cartesianas, podemos llegar a la misma representación skew-product como dada en [MMUB20]. Con esta representación, revisamos el algoritmo para la simulación exacta de incrementos del movimiento browniano \mathbf{Z} que se presentó en [MMUB20], además, realizamos la implementación del algoritmo.

Capítulo 1

El teorema de Arquímedes

1.1. Introducción

El matemático E. T. Bell en su libro *Men of Mathematics* ([Bel37]) dice lo siguiente sobre Arquímedes: “Cualquier enumeración de los tres matemáticos ‘más grandes’ de la historia, incluiría el nombre de Arquímedes. Los otros dos que se asocian con él de ordinario son Newton y Gauss. Considerando la relativa riqueza o pobreza de la ciencia matemática y la física en las respectivas edades en que estos gigantes vivieron, y estimando sus logros en el contexto de su época, pondría a Arquímedes en el primer lugar.”

Arquímedes (cerca de 287-212 a.C.) fue un físico, matemático, ingeniero y astrónomo griego, nació en Siracusa, Sicilia de la Antigua Grecia; y estudió en Alejandría, Egipto cuando era joven. El principio de Arquímedes y sus numerosas aplicaciones prácticas son muy conocidos actualmente, y uno de sus trabajos importantes en la mecánica es la ley de la palanca escrita en la obra *On the equilibrium of plans*, en donde exclamó: “¡Denme un punto de apoyo y moveré el mundo!”, una exclamación de Arquímedes que se ha conservado a través de los siglos.

Arquímedes también se conoce como “el gran geómetra” en la historia de la ciencia, estudió la obra *Elementos* de Euclides, e hizo un uso más refinado el método exhaustivo, y este método tuvo su origen por el sofista Antifón (480-411 a.C.) y el matemático Eudoxo de Cnido (408-355 a.C.), para demostrar teoremas sobre áreas y volúmenes, el cual se realizó en sus obras de *On the sphere and cylinder*. En donde se encuentra uno de los resultados famosos (Proposición 34 de la obra *On the sphere and cylinder, Book I* (cf. [Arc02])):

“El volumen de cualesquiera esfera es igual a cuatro veces del cono que tiene base como el círculo máximo de la esfera y tiene altura igual al radio de la esfera.”

Y el corolario de esta proposición es el siguiente famoso resultado sobre el área y volumen entre la esfera y el cilindro circunscrito:

“El volumen de cualquier cilindro que tiene base como el círculo máximo de una esfera y tiene altura igual al diámetro de esta esfera, es igual al $\frac{3}{2}$ del volumen de la esfera; y su superficie junto con sus bases es igual al $\frac{3}{2}$ de la superficie de la esfera.”

Este quizá fue el resultado más importante para Arquímedes, y lo apreciaba tanto que a su muerte pidió que este resultado junto con la figura de un cilindro circunscrito a una esfera fuera esculpida en su tumba (Fig. 1.1(a)). De acuerdo a la versión más recurrente de la muerte de Arquímedes, en un ataque sorpresa a Siracusa, el cónsul romano Marcellus

ordenó que no matara a Arquímedes, pero un soldado ignoró las órdenes, quien con su espada quitó la vida al célebre matemático. Al final, Marcellus construyó una tumba de acuerdo con la voluntad de Arquímedes, en donde gravaba la figura simbólica y este famoso teorema, el descubrimiento más orgulloso para él.

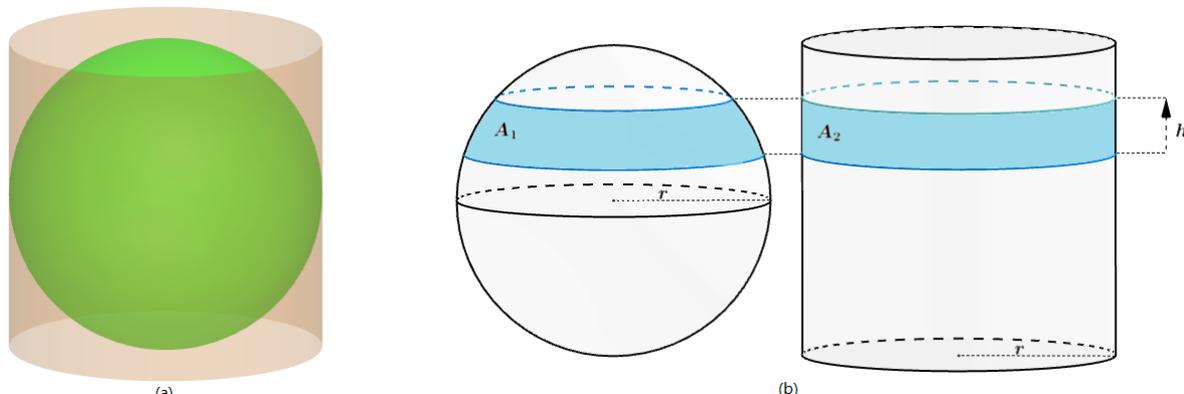


Figura 1.1: (a) Figura simbólica del teorema de Arquímedes; (b) La proyección ortogonal de la superficie lateral del cilindro tiene la misma área como la de la esfera.

De este resultado, Arquímedes también demostró que el área de la superficie de una esfera 2-dimensional es igual al área de la superficie lateral del cilindro circunscrito. Más aún, el área de la superficie de la región de la esfera determinada por cualesquiera dos planos horizontales es igual al área lateral de la superficie determinada por estos dos planos. En otras palabras, el teorema de Arquímedes dice que la proyección ortogonal de cualquier región medible en una esfera en el cilindro circunscrito a la esfera preserva el área, como muestra en la figura 1.1(b), el área de las superficies A_1 y A_2 son iguales: $A_1 = A_2 = 2\pi rh$.

Recientemente, se han hecho investigaciones sobre la generalización del teorema de Arquímedes mencionado anteriormente a las esferas en el espacio euclidiano de dimensión $d \geq 3$ (cf. [AM13] & [CDH17]): si consideramos la esfera unitaria $\mathcal{S}^{d-1} \subset \mathbb{R}^d$, donde

$$\mathcal{S}^{d-1} = \left\{ \mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\| = \sqrt{x_1^2 + \cdots + x_d^2} = 1 \right\};$$

y sea \mathcal{C}^{d-1} el cilindroide de dimensión $d - 1$ circunscrito a esta esfera.

El área de la superficie de la esfera \mathcal{S}^{d-1} es igual al área de la superficie lateral del cilindroide \mathcal{C}^{d-1} circunscrito. También el área de la superficie de la zona de la esfera \mathcal{S}^{d-1} determinada por cualesquiera dos hiperplanos horizontales es igual al área de la superficie de la zona en el cilindroide \mathcal{C}^{d-1} determinada por tales hiperplanos ([AM13], Thm. 3 y 4). En términos de la proyección, si $\mathcal{P} : \mathcal{S}^{d-1} \subset \mathbb{R}^d \rightarrow \mathbb{R}^{d-2}$ es una proyección ortogonal de codimensión 2 en \mathbb{R}^d restringido a \mathcal{S}^{d-1} , se tiene que el rango de \mathcal{P} es la bola unitaria \mathcal{B}^{d-2} de dimensión $d - 2$, donde $\mathcal{B}^{d-2} = \{\mathbf{x} \in \mathbb{R}^{d-2} : \|\mathbf{x}\| \leq 1\}$ (cf. [CDH17]). Más aún, por el principio de momento-volumen ([AM13]), se tiene que para cualquier conjunto medible $U \subset \mathcal{B}^{d-2}$, $\text{Vol}(\mathcal{P}(U)) = 2\pi \text{Vol}(U)$, donde $\text{Vol}(U)$ denota al volumen de U .

Al final de este capítulo, veremos que en la teoría de la probabilidad, la generalización del teorema de Arquímedes nos dice que la proyección ortogonal de \mathcal{S}^{d-1} en la bola \mathcal{B}^{d-2} preserva la medida, esto implica que una distribución uniforme en \mathcal{S}^{d-1} induce una distribución uniforme en la bola \mathcal{B}^{d-2} . Por lo que en la siguiente sección, primero damos la definición de distribuciones uniformes en la esfera \mathcal{S}^{d-1} y en la bola \mathcal{B}^d .

1.2. Distribución uniforme en la esfera \mathcal{S}^{d-1} y en la bola \mathcal{B}^d

Para definir una medida de probabilidad en la esfera unitaria \mathcal{S}^{d-1} , consideremos a $\mathcal{S}^{d-1} \subset \mathbb{R}^d$ como una variedad riemanniana \mathcal{M} junto con la métrica riemanniana g inducida del espacio euclidiano \mathbb{R}^d ([dC92], Chap 1.); y esta métrica induce una medida volumen en la variedad, donde un *elemento de volumen*, $d\sigma$ en \mathcal{M} está dada por:

$$d\sigma(\mathbf{x}) = \sqrt{\det(g(\mathbf{x}))} d\mathbf{x}.$$

Definición 1.1. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Una variable aleatoria en la variedad riemanniana $(\mathcal{M}, \mathcal{B}(\mathcal{M}))^1$ es una función Borel medible $\mathbf{X}: \Omega \rightarrow \mathcal{M}$. La variable aleatoria \mathbf{X} admite una función de densidad $p_X(\mathbf{x})$ si

$$\mathbb{P}(\mathbf{X} \in \mathcal{M}) = \int_{\mathcal{M}} p_X(\mathbf{x}) d\sigma(\mathbf{x}) = 1, \quad \text{y} \quad \mathbb{P}(\mathbf{X} \in \mathcal{A}) = \int_{\mathcal{A}} p_X(\mathbf{x}) d\sigma(\mathbf{x}) \geq 0,$$

para todo $\mathcal{A} \in \mathcal{B}(\mathcal{M})$.

En particular, una distribución uniforme en la esfera unitaria \mathcal{S}^{d-1} tiene función de densidad dada por (cf. [Pen04]):

$$p_X(\mathbf{x}) = \frac{1}{\int_{\mathcal{S}^{d-1}} d\sigma} = \frac{1}{\mathcal{A}(\mathcal{S}^{d-1})}, \quad \forall \mathbf{x} \in \mathcal{S}^{d-1}, \quad (1.1)$$

donde $\mathcal{A}(\mathcal{S}^{d-1})$ es el ‘volumen’ de la esfera \mathcal{S}^{d-1} ; es decir, el área de la superficie de la esfera \mathcal{S}^{d-1} , y la fórmula para ésta es conocida:

$$\mathcal{A}(\mathcal{S}^{d-1}) = \frac{2\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})}, \quad (1.2)$$

donde $\Gamma(\cdot)$ denota la función gamma dada por $\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} dx$, $t > 0$.

Para la existencia de una distribución uniforme en la esfera \mathcal{S}^{d-1} , se tiene el siguiente resultado.

Proposición 1.2. Sean X_1, X_2, \dots, X_d variables aleatorias independientes e idénticamente distribuidas como una normal estándar; es decir, $\mathbf{X} = \{X_1, \dots, X_d\} \sim \mathcal{N}_d(\mathbf{0}, I)$, donde $\mathbf{0}$ el vector de ceros en \mathbb{R}^d e I denota a la matriz de identidad de dimensión $d \times d$. Sea $R = \|\mathbf{X}\|$, entonces el vector aleatorio definido por $\mathbf{U} = \frac{\mathbf{X}}{R}$ tiene una distribución uniforme en la esfera $\mathcal{S}^{d-1} \subset \mathbb{R}^d$.

Demostración. Notemos que $\|\mathbf{U}\| = 1$ c.s.², entonces \mathbf{U} tiene una distribución en la esfera unitaria $\mathcal{S}^{d-1} \subset \mathbb{R}^d$. Ahora verificamos que \mathbf{U} está distribuida uniformemente en \mathcal{S}^{d-1} .

En efecto, sea $f: \mathcal{S}^{d-1} \subset \mathbb{R}^d \rightarrow \mathbb{R}$ una función medible, y considerando $f(\mathbf{U}) = f\left(\frac{\mathbf{X}}{R}\right)$ como función de \mathbf{X} , tenemos que

$$\begin{aligned} \mathbb{E}[f(\mathbf{U})] &= \int_{\mathbb{R}^d} f\left(\frac{x_1}{r}, \dots, \frac{x_d}{r}\right) (2\pi)^{-\frac{d}{2}} e^{-\frac{\|\mathbf{x}\|^2}{2}} d\mathbf{x} \\ &= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} f\left(\frac{x_1}{r}, \dots, \frac{x_d}{r}\right) e^{-\frac{r^2}{2}} d\mathbf{x}. \end{aligned} \quad (1.3)$$

Usando las coordenadas polares y el siguiente resultado:

¹ $\mathcal{B}(E)$ denota a la σ -álgebra de Borel generada por los subconjuntos abiertos de E .

² c.s. := casi seguramente; un evento E en el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ ocurre casi seguramente si $\mathbb{P}(E) = 1$.

Teorema 1.3 ([Fol99], Thm. 2.49). Si $\mathbf{x} \in \mathbb{R}^d \setminus \{0\}$, las coordenadas polares de \mathbf{x} son

$$r = \|\mathbf{x}\| \in (0, \infty), \quad \mathbf{u} = \frac{\mathbf{x}}{\|\mathbf{x}\|} \in \mathcal{S}^{d-1}. \quad (1.4)$$

Existe una única medida de Borel $\sigma = \sigma_{d-1}$ en \mathcal{S}^{d-1} tal que $\mu_* = \rho \times \sigma$ es una medida de Borel en $(0, \infty) \times \mathcal{S}^{d-1}$. Si f es una función medible no negativa o Lebesgue integrable en \mathbb{R}^d , entonces

$$\int_{\mathbb{R}^d} f(\mathbf{x}) d\mathbf{x} = \int_0^\infty \int_{\mathcal{S}^{d-1}} f(r\mathbf{u}) r^{d-1} d\sigma(\mathbf{u}) dr \quad (1.5)$$

Notemos que esta única medida de Borel σ en \mathcal{S}^{d-1} es la llamada *medida de superficie* y coincide con la medida de volumen en la variedad riemanniana \mathcal{S}^{d-1} definida anteriormente. Así, se tiene la esperanza de la ecuación (1.3) en coordenadas polares,

$$\begin{aligned} \mathbb{E}[f(\mathbf{U})] &= (2\pi)^{-\frac{d}{2}} \int_0^\infty \int_{\mathcal{S}^{d-1}} f(\mathbf{u}) e^{-\frac{r^2}{2}} r^{d-1} d\sigma(\mathbf{u}) dr \\ &= K \int_{\mathcal{S}^{d-1}} f(\mathbf{u}) d\sigma(\mathbf{u}), \end{aligned}$$

donde

$$\begin{aligned} K &= (2\pi)^{-\frac{d}{2}} \int_0^\infty r^{d-1} e^{-r^2/2} dr = (2\pi)^{-\frac{d}{2}} \int_0^\infty \left((2t)^{\frac{1}{2}}\right)^{d-1} e^{-t} (2t)^{-\frac{1}{2}} dt \\ &= 2^{-1} \pi^{-\frac{d}{2}} \int_0^\infty t^{\frac{d}{2}-1} e^{-t} dt = \frac{\Gamma\left(\frac{d}{2}\right)}{2\pi^{\frac{d}{2}}}. \end{aligned}$$

Por (1.2), tenemos que $K = \frac{1}{\sigma(\mathcal{S}^{d-1})} = \frac{1}{\sigma(\mathcal{S}^{d-1})}$. Así, para cualquier $\mathcal{A} \subset \mathcal{S}^{d-1}$ conjunto medible y considerando a f como la función característica de \mathcal{A} , i.e., $f \equiv \chi_{\mathcal{A}}$, la esperanza $\mathbb{E}[f(\mathbf{U})]$ corresponde a la probabilidad $\mathbb{P}(\mathbf{U} \in \mathcal{A})$, y se obtiene

$$\begin{aligned} \mathbb{P}(\mathbf{U} \in \mathcal{A}) &= \frac{1}{\sigma(\mathcal{S}^{d-1})} \int_{\mathcal{S}^{d-1}} \chi_{\mathcal{A}}(\mathbf{u}) d\sigma(\mathbf{u}) \\ &= \frac{1}{\sigma(\mathcal{S}^{d-1})} \int_{\mathcal{A}} d\sigma(\mathbf{u}) = \frac{\sigma(\mathcal{A})}{\sigma(\mathcal{S}^{d-1})}. \end{aligned} \quad (1.6)$$

Por lo tanto, \mathbf{U} sigue una distribución uniforme en la esfera \mathcal{S}^{d-1} . □

En la proposición anterior, notemos que \mathbf{X} es un vector aleatorio con una distribución normal multivariada en \mathbb{R}^d con distribución $\mathcal{N}_d(\mathbf{0}, I)$, entonces para cualquier matriz ortogonal B de dimensión $d \times d$, la transformación lineal $B\mathbf{X}$ sigue una distribución multivariada $\mathcal{N}_d(B\mathbf{0}, BIB^\top) = \mathcal{N}_d(\mathbf{0}, I)$. Es decir, $B\mathbf{X}$ es idénticamente distribuida como \mathbf{X} . Por consiguiente, para $\mathbf{U} = \frac{\mathbf{X}}{\|\mathbf{X}\|}$, $B\mathbf{U} = \frac{B\mathbf{X}}{\|\mathbf{X}\|} = \frac{B\mathbf{X}}{\|B\mathbf{X}\|}$ tiene la misma distribución que \mathbf{U} . Esto es, \mathbf{U} es invariante bajo rotaciones en la esfera \mathcal{S}^{d-1} .

Por otro lado, notemos que la esfera \mathcal{S}^{d-1} es invariante bajo el grupo de rotaciones $SO(d)$, y el grupo $SO(d)$ es un grupo de Lie compacto ([CCL99], Chap. 6), entonces existe una única medida de Haar que es el elemento de volumen para el grupo ([Pen04], Sec. 3). Además, como el grupo $SO(d)$ actúa transitivamente sobre \mathcal{S}^{d-1} que preserva el área de la esfera, y también la medida de volumen σ en la esfera \mathcal{S}^{d-1} define una distribución uniforme que es invariante bajo rotaciones, entonces por el teorema para medidas invariantes ([Kal21], Thm. 3.12), σ es la única medida hasta una normalización en \mathcal{S}^{d-1} , e invariante bajo rotaciones. Por lo tanto, la distribución uniforme en la esfera \mathcal{S}^{d-1} es la única distribución en la esfera invariante bajo rotaciones.

Ahora, consideramos a una bola unitaria $\mathcal{B}^d \subset \mathbb{R}^d$, que equivale a una esfera unitaria \mathcal{S}^{d-1} junto con su interior, como conocemos la relación entre el volumen de la bola $\text{Vol}(\mathcal{B}^d)$ y la superficie de su frontera $\mathcal{A}(\mathcal{S}^{d-1})$ dada por (ver [AM13]):

$$\text{Vol}(\mathcal{B}^d) = \frac{\mathcal{A}(\mathcal{S}^{d-1})}{d} \stackrel{(1.2)}{=} \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2} + 1)}. \quad (1.7)$$

Así, podemos definir una distribución uniforme en la bola \mathcal{B}^d en términos de una distribución uniforme en la esfera \mathcal{S}^{d-1} .

Proposición 1.4. *Sea \mathbf{U} una variable aleatoria uniformemente distribuida en la esfera \mathcal{S}^{d-1} definida como en la Proposición 1.2 y sea $U \sim \mathcal{U}(0, 1)$ independiente de \mathbf{U} , entonces el producto $\mathbf{T} = U^{1/d}\mathbf{U}$ tiene distribución uniforme en la bola \mathcal{B}^d .*

Demostración. Primero, notamos que $U \sim \mathcal{U}(0, 1)$ tiene densidad dada por $f_U(u) = I_{[0,1]}(u)$; sea $R = U^{1/d}$, entonces por el teorema de cambio de variable (T.C.V.), la densidad de R está dada por

$$f(r) = f_U(u) \frac{du}{dr} = \frac{f_U(u)}{dr/du} \Big|_{u=r^d} = \frac{1}{(1/d)u^{1/d-1}} = dr^{d-1}, \quad \forall r \in [0, 1].$$

Ahora consideremos a las coordenadas polares como en el Teorema 1.3, y considera la función $\Phi : \mathbb{R}^d \setminus \{\mathbf{0}\} \ni \mathbf{x} \mapsto (r, \mathbf{u}) \in (0, \infty) \times \mathcal{S}^{d-1}$ continua y biyectiva, con la inversa $\Phi^{-1}(r, \mathbf{u}) = r\mathbf{u}$. Si E es un conjunto boreliano en \mathcal{S}^{d-1} , para $0 < a \leq 1$, sea

$$E_a = \Phi^{-1}((0, a) \times E) = \{r\mathbf{u} : 0 < r \leq a, \mathbf{u} \in E\}.$$

Así, como $\mathbf{U} \sim \mathcal{U}(\mathcal{S}^{d-1})$ con densidad dada en (1.6), se obtiene la probabilidad de \mathbf{T} en E_a :

$$\begin{aligned} \mathbb{P}(\mathbf{T} \in E_a) &= \mathbb{P}(\Phi^{-1}(U^{1/d}, \mathbf{U}) \in E_a) = \mathbb{P}(U^{1/d} \in (0, a), \mathbf{U} \in E) \\ &= \mathbb{P}(R \in (0, a))\mathbb{P}(\mathbf{U} \in E) = \left(\int_0^a d \cdot r^{d-1} dr \right) \left(\frac{1}{\sigma(\mathcal{S}^{d-1})} \int_E d\sigma(\mathbf{u}) \right) \\ &= \frac{d}{\sigma(\mathcal{S}^{d-1})} \int_0^a \int_E r^{d-1} d\sigma(\mathbf{u}) dr = \frac{1}{\text{Vol}(\mathcal{B}^d)} \int_0^a \int_E r^{d-1} d\sigma(\mathbf{u}) dr. \end{aligned} \quad (1.8)$$

Notando que si $a = 1$ y $E = \mathcal{S}^{d-1}$, entonces $E_a = \mathcal{B}^d$, y se tiene que

$$\mathbb{P}(\mathbf{T} \in \mathcal{B}^d) = \frac{d}{\sigma(\mathcal{S}^{d-1})} \frac{\sigma(\mathcal{S}^{d-1}) \cdot 1}{d} = 1.$$

Por tanto, \mathbf{T} admite una densidad en la bola \mathcal{B}^d y se distribuye uniformemente por (1.8). \square

Ejemplo 1.5. Consideramos el caso cuando $d = 1$ y denotamos por $\|x\| := |x|$, para $x \in \mathbb{R}$, entonces, la esfera $\mathcal{S}^{d-1} \equiv \mathcal{S}^0$ consiste en un par de puntos $\{-1, 1\}$, que son puntos límites de la bola \mathcal{B}^1 , i.e., el intervalo $[-1, 1]$. Luego, una distribución uniforme \mathbf{U} en la esfera \mathcal{S}^0 tiene función de densidad $\mathbb{P}(\mathbf{U} = \mathbf{u}) = \frac{1}{2} \mathbb{1}_{\{-1, 1\}}(\mathbf{u})$.³

Sea $U \sim \mathcal{U}(0, 1)$, entonces $\mathbf{T} = U^{1/d}\mathbf{U} = UU$ y toma valores en el intervalo $[-1, 1]$. Para verificar que \mathbf{T} tiene distribución uniforme en la bola \mathcal{B}^1 , se tiene que para $t \in [-1, 1]$,

$$\mathbb{P}(\mathbf{T} < t) = \mathbb{P}(UU < t | \mathbf{U} = -1)\mathbb{P}(\mathbf{U} = -1) + \mathbb{P}(UU < t | \mathbf{U} = 1)\mathbb{P}(\mathbf{U} = 1)$$

³ $\mathbb{1}_A : \mathcal{E} \rightarrow \{0, 1\}$ denota a la función característica de un subconjunto $A \subseteq \mathcal{E}$ que está definida por: $\mathbb{1}_A(x) = 1$ si $x \in A$ y $\mathbb{1}_A(x) = 0$ si $x \notin A$.

$$= \frac{1}{2} \left(\mathbb{P}(-U < t) + \mathbb{P}(U < t) \right) = \frac{1}{2} \left(\mathbb{P}(U > -t) + \mathbb{P}(U < t) \right),$$

notemos que si $0 \leq t \leq 1$, $\mathbb{P}(U < t) = t$ y $\mathbb{P}(U > -t) = 1$; si $-1 \leq t < 0$, tenemos que $\mathbb{P}(U < t) = 0$ y $\mathbb{P}(U > -t) = 1 - (-t) = 1 + t$. En ambos casos se tiene que $\mathbb{P}(U > -t) + \mathbb{P}(U < t) = t + 1$. Por lo tanto,

$$\mathbb{P}(\mathbf{T} < t) = \frac{t + 1}{2} = \frac{t - (-1)}{2}, \quad \forall t \in [-1, 1].$$

Esto es, \mathbf{T} se distribuye uniformemente en el intervalo $[-1, 1]$. \square

1.2.1. Simulación de puntos uniformes en \mathcal{S}^{d-1} y en \mathcal{B}^d

Por la Proposición 1.2 y la Proposición 1.4, es posible simular muestras de puntos uniformes (i.e., puntos que se distribuyen uniformemente) en la esfera \mathcal{S}^{d-1} y en la bola \mathcal{B}^d . El procedimiento para la simulación de cada uno de ellos se resume en el Algoritmo 1 y en Algoritmo 2, respectivamente.

Algoritmo 1: Simulación de puntos uniformes en la esfera unitaria \mathcal{S}^{d-1}

Init: Dimensión $d > 0$,

1 Simular: $X_1, \dots, X_d \sim \mathcal{N}(0, 1)$.

2 Calcular: $R = \sqrt{\sum_{i=1}^d X_i^2}$.

3 **return** $\mathbf{U} = (X_1, \dots, X_d)/R$

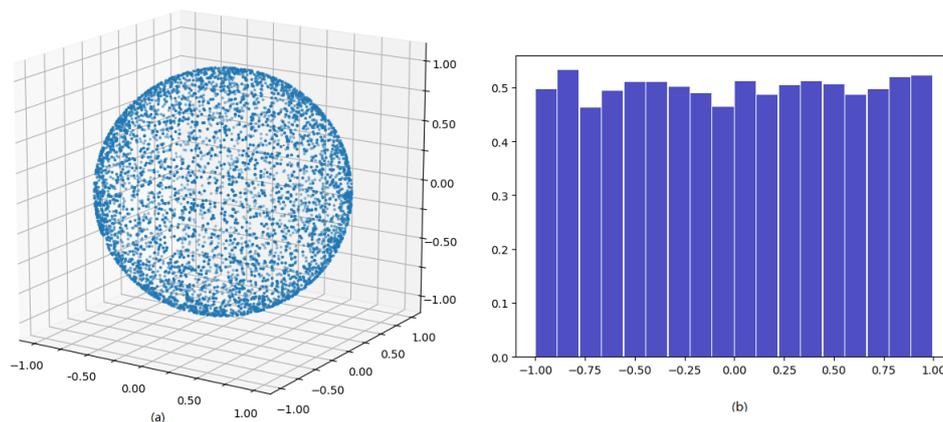


Figura 1.2: (a) Simulación de 5,000 muestras usando el Algoritmo 1; (b) Distribución de la proyección ortogonal sobre la primera coordenada de las muestras generadas en (a).

Al implementar el Algoritmo 1 (ver Apéndice Alg.1), obtenemos una muestra de puntos que se distribuye uniformemente en la esfera \mathcal{S}^2 , como se ilustra en la Figura 1.2(a) para el caso $d = 3$. La distribución de la proyección ortogonal sobre la primera coordenada de esta muestra de puntos está presentada en la Figura 1.2(b), vemos que se distribuye de manera uniforme en el intervalo $[-1, 1]$ c.s.. Esto es, el teorema de Arquímedes en el caso $d = 3$.

Usando el Algoritmo 1, también podemos implementar el Algoritmo 2 (ver Apéndice Alg.2) y generar una muestra de puntos uniformes en la bola \mathcal{B}^d , para todo $d \geq 1$.

Algoritmo 2: Simulación de puntos uniformes en la bola unitaria \mathcal{B}^d

Init: Dimensión $d > 0$,

1 Simular: $X_1, \dots, X_d \sim \mathcal{N}(0, 1)$.

▷ Alg. 1

2 Simular: $U \sim \mathcal{U}(0, 1)$.

3 **return** $\mathbf{T} = U^{1/d}\mathbf{U}$

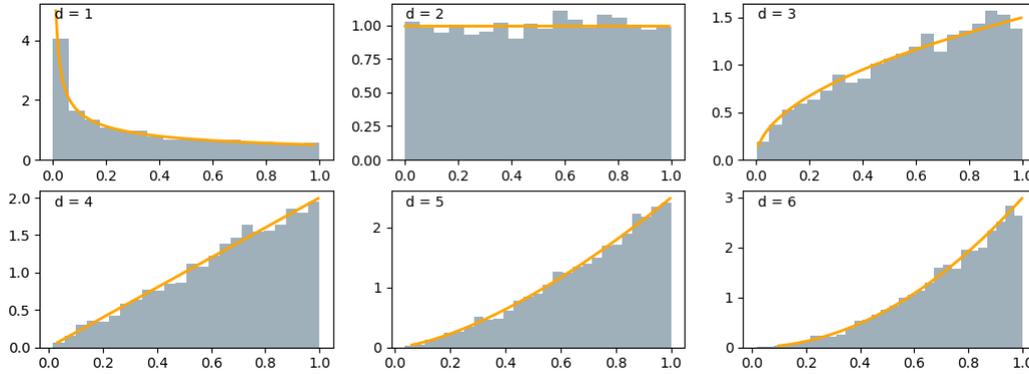


Figura 1.3: Distribución de $\|\mathbf{T}\|^2$, con \mathbf{T} puntos uniformes en \mathcal{B}^d de una muestra generada usando el Algoritmo 2, comparando contra la distribución de beta $\mathfrak{B}\mathfrak{e}(d/2, 1)$ (—), para $d \in \{1, \dots, 6\}$.

En la Figura 1.3 se ilustra la distribución de la norma de \mathbf{T} , puntos uniformes en la bola \mathcal{B}^d de una muestra generada usando el Algoritmos 2 para distintos valores de $d \geq 1$, y comparando contra la distribución beta $\mathfrak{B}\mathfrak{e}(d/2, 1)$ (ver Sec. 1.3).

Por otro lado, también se tiene el siguiente resultado que nos dice la distribución de la norma de las primeras d coordenadas de una uniforme en la esfera \mathcal{S}^{d+1} es exactamente la distribución beta $\mathfrak{B}\mathfrak{e}(d/2, 1)$, esto es:

Lema 1.6 ([HL10], Lem. 1; [CHS81]). *Sea $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2) \sim \mathcal{N}_d(\mathbf{0}, I)$, donde la dimensión de \mathbf{X}_1 es m ($1 \leq m < d$), y sea $\mathbf{U} = (\mathbf{X}_1, \mathbf{X}_2) / \|\mathbf{X}\| = (\mathbf{U}_1, \mathbf{U}_2)$. Entonces \mathbf{U} distribuye uniformemente en la esfera \mathcal{S}^{d-1} por la Proposición 1.2.*

Además, $\mathbf{U}_1 \stackrel{(D)}{=} R_m \mathbf{U}'_1$, donde $R_m = \|\mathbf{X}_1\| / \|\mathbf{X}\|$ es independiente de $\mathbf{U}'_1 = \mathbf{X}_1 / \|\mathbf{X}_1\|$, de manera que $R_m^2 \sim \mathfrak{B}\mathfrak{e}(\frac{m}{2}, \frac{d-m}{2})$ y $\mathbf{U}'_1 \sim \mathcal{U}(\mathcal{S}^{m-1})$.

Por consiguiente, la distribución beta se relaciona con la distribución uniforme en la bola \mathcal{B}^d y la distribución de la proyección ortogonal codimensión 1 de una distribución uniforme en la esfera \mathcal{S}^{d+1} . Esto es el teorema de Arquímedes vía álgebra Beta-Gamma y lo veremos en el siguiente apartado.

1.3. Teorema de Arquímedes vía álgebra Beta-Gamma

Como se mencionó anteriormente, el teorema de Arquímedes expresada en la teoría de la probabilidad es: la proyección ortogonal de \mathcal{S}^{d-1} en la bola \mathcal{B}^{d-2} preserva la medida; por lo tanto, una medida uniforme en \mathcal{S}^{d-1} induce una medida uniforme en \mathcal{B}^{d-2} . Para esto, primero recordamos la definición de la distribución gamma y la de beta:

⁴ $\stackrel{(D)}{=}$ significa ambos lados de la igualdad tiene la misma distribución.

Definición 1.7. Una variable aleatoria continua V tiene una distribución beta, y lo denotamos por $V \sim \mathfrak{B}\mathfrak{e}(\alpha, \beta)$, con parámetros $\alpha > 0$ y $\beta > 0$, si su función de densidad está dada por

$$p_V(v) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} v^{\alpha-1}(1-v)^{\beta-1} = \frac{1}{B(\alpha, \beta)} v^{\alpha-1}(1-v)^{\beta-1}, \quad 0 < v < 1, \quad (1.9)$$

donde $B(\cdot, \cdot)$ denota a la función beta, dada por $B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx$.

Una variable aleatoria continua W tiene una distribución gamma, y lo denotamos por $W \sim \mathfrak{G}\mathfrak{a}(\alpha, \beta)$, con parámetros $\alpha > 0$ y $\beta > 0$, si su función de densidad está dada por

$$p_W(w) = \frac{\beta^\alpha}{\Gamma(\alpha)} w^{\alpha-1} e^{-\beta w}, \quad \forall w > 0. \quad (1.10)$$

Se tiene el siguiente resultado sobre las propiedades de la familia de distribuciones gamma y la de beta, y así, la relación entre la distribución beta y la distribución uniforme en $[0, 1]$.

Proposición 1.8. Sea $\mathbf{X} = (X_1, \dots, X_{d+2}) \sim \mathcal{N}_{d+2}(\mathbf{0}, I)$ y escribe $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$, donde \mathbf{X}_1 corresponde a las primeras d coordenadas de \mathbf{X} , y \mathbf{X}_2 sus últimas 2 coordenadas. Sean $X = \|\mathbf{X}_1\|^2 = \sum_{i=1}^d X_i^2$ y $Y = \|\mathbf{X}_2\|^2 = \sum_{i=d+1}^{d+2} X_i^2$, entonces

$$\frac{X}{X+Y} \sim \mathfrak{B}\mathfrak{e}\left(\frac{d}{2}, 1\right), \quad y \quad \left(\frac{X}{X+Y}\right)^{\frac{d}{2}} \sim \mathcal{U}(0, 1). \quad (1.11)$$

Demostración. Por la relación entre la distribución normal estándar y la chi-cuadrada, y la relación entre la distribución chi-cuadrada y la gamma, tenemos que

$$X \sim \chi^2(d), \quad Y \sim \chi^2(2), \quad \implies \quad X \sim \mathfrak{G}\mathfrak{a}\left(\frac{d}{2}, \frac{1}{2}\right), \quad Y \sim \mathfrak{G}\mathfrak{a}\left(1, \frac{1}{2}\right)$$

además, por la independencia de las variables X_i 's, X es independiente de Y , y tiene densidad conjunta $f(x, y) = p_X(x)p_Y(y)$, donde p_X y p_Y son funciones de densidad de una distribución gamma con parámetros $(\frac{d}{2}, \frac{1}{2})$ y $(1, \frac{1}{2})$ respectivamente.

Ahora, consideramos a las variables $S = \frac{X}{X+Y}$ y $Z = X+Y$, afirmamos que S y Z son independientes con $S \sim \mathfrak{B}\mathfrak{e}\left(\frac{d}{2}, 1\right)$ y $Z \sim \mathfrak{G}\mathfrak{a}\left(\frac{d}{2} + 1, \frac{1}{2}\right)$.

En efecto, sea $g: (0, 1) \times (0, \infty) \rightarrow (\mathbb{R}^+)^2$ una transformación dada por $g(s, z) = (sz, (1-s)z)$, entonces g es la inversa para (X, Y) tal que $g(S, Z) = (X, Y)$. Por lo tanto, para cualquier conjunto medible $B \in \mathcal{B}((\mathbb{R}^+)^2)$, se tiene que

$$\begin{aligned} \mathbb{P}((S, Z) \in B) &= \mathbb{P}(g^{-1}(X, Y) \in B) = \mathbb{P}((X, Y) \in g(B)) = \int_{g(B)} p_X(x)p_Y(y) dx dy \\ &= \int_{g(B)} f(x, y) dx dy = \int_B f \circ g(s, z) \left| \det J_g(s, z) \right| ds dz, \end{aligned}$$

donde $J_g(s, z) = \begin{pmatrix} z & s \\ -z & 1-s \end{pmatrix}$ es el jacobiano de la transformación con $\det J_g(s, z) = z > 0$.

Luego, sustituyendo la densidad $f(x, y)$ en el integrando de la integral anterior, se obtiene

$$\begin{aligned} f \circ g(s, z) \left| \det J_g(s, z) \right| &= \left[\frac{\left(\frac{1}{2}\right)^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)} (sz)^{\frac{d}{2}-1} e^{-\frac{sz}{2}} \right] \left[\frac{1}{\Gamma(1)} \left((1-s)z\right)^{1-1} e^{-\frac{(1-s)z}{2}} \right] z \\ &= \frac{\left(\frac{1}{2}\right)^{\frac{d}{2}+1}}{\Gamma\left(\frac{d}{2}\right) \Gamma(1)} s^{\frac{d}{2}-1} (1-s)^{1-1} z^{\left(\frac{d}{2}+1\right)-1} e^{-\frac{z}{2}} \\ &= \left[\frac{\Gamma\left(\frac{d}{2}+1\right)}{\Gamma\left(\frac{d}{2}\right) \Gamma(1)} s^{\frac{d}{2}-1} (1-s)^{1-1} \right] \left[\frac{\left(\frac{1}{2}\right)^{\frac{d}{2}+1}}{\Gamma\left(\frac{d}{2}+1\right)} z^{\left(\frac{d}{2}+1\right)-1} e^{-\frac{z}{2}} \right]. \end{aligned}$$

Esto es el producto de dos densidades, por lo que integra 1 en el espacio $(0, \infty) \times (0, \infty)$, entonces (S, Z) admite una densidad $f'(s, z) = f'_S(s) f'_Z(z)$, donde f'_S es la densidad de una distribución beta $\mathfrak{B}\mathfrak{e}\left(\frac{d}{2}, 1\right)$ y f'_Z es la densidad de una distribución gamma $\mathfrak{G}\mathfrak{a}\left(\frac{d}{2}+1, \frac{1}{2}\right)$. Por lo tanto, S es independiente de Z con $S = \frac{X}{X+Y} \sim \mathfrak{B}\mathfrak{e}\left(\frac{d}{2}, 1\right)$ y $Z \sim \mathfrak{G}\mathfrak{a}\left(\frac{d}{2}+1, \frac{1}{2}\right)$.

Por otro lado, sea $h: (0, 1) \ni s \rightarrow s^{\frac{d}{2}} \in (0, 1)$ invertible, y para $U = h(S) = S^{\frac{d}{2}}$, por el T.C.V., se obtiene la densidad de U :

$$f_U(u) = f'_S(s) \frac{ds}{du} \Big|_{s=h^{-1}(u)} = \frac{f'_S(s)}{du/ds} \Big|_{u=h(s)} = \frac{\Gamma\left(\frac{d}{2}+1\right)}{\Gamma\left(\frac{d}{2}\right) \Gamma(1)} \frac{s^{\frac{d}{2}-1}}{(d/2)s^{\frac{d}{2}-1}} = 1,$$

⁵para $0 < u < 1$. Por lo tanto, tenemos que $S^{\frac{d}{2}} = \left(\frac{X}{X+Y}\right)^{\frac{d}{2}} \sim \mathcal{U}(0, 1)$. □

Observación 1.1. De la proposición anterior, como $\left(\frac{X}{X+Y}\right)^{\frac{d}{2}} \sim \mathcal{U}(0, 1)$, tenemos que si $U \sim \mathcal{U}(0, 1)$, entonces

$$U^{\frac{1}{d}} \stackrel{(D)}{=} \sqrt{\frac{X}{X+Y}} = \frac{\|\mathbf{X}_1\|}{\|\mathbf{X}\|}. \quad (1.12)$$

Así, podemos demostrar el teorema de Arquímedes generalizado a la esfera \mathcal{S}^{d-1} de dimensión $d > 3$ en términos de la teoría de la probabilidad y teniendo el siguiente resultado.

Teorema 1.9. *Sea \mathbf{U} un vector aleatorio con distribución uniforme en la esfera \mathcal{S}^{d-1} definido como en la Proposición 1.2, y sea $\mathcal{P}: \mathcal{S}^{d-1} \rightarrow \mathcal{B}^{d-2}$ una proyección ortogonal de codimensión 2 en \mathbb{R}^d restringido a \mathcal{S}^{d-1} . Entonces $\mathbf{T} = \mathcal{P}(\mathbf{U})$ tiene distribución uniforme en la bola \mathcal{B}^{d-2} .*

Demostración. Sea $\mathbf{X} = (X_1, \dots, X_d) \sim \mathcal{N}_d(\mathbf{0}, I)$, entonces $\mathbf{U} = (X_1, \dots, X_d) / \|\mathbf{X}\| \sim \mathcal{U}(\mathcal{S}^{d-1})$. Denotamos a $\mathbf{X} \equiv (\mathbf{X}_1, \mathbf{X}_2)$, donde \mathbf{X}_1 corresponde a las primeras $d-2$ coordenadas de \mathbf{X} y \mathbf{X}_2 sus últimas 2 coordenadas. Sin pérdida de la generalidad, supongamos que la proyección ortogonal \mathcal{P} es simplemente quitar las últimas dos coordenadas del vector en \mathbb{R}^d . Entonces, usando el Lema 1.6 para $m = d-2$, se tiene que

$$\begin{aligned} \mathbf{T} = \mathcal{P}(\mathbf{U}) &\stackrel{(D)}{=} \left(\frac{X_1}{\|\mathbf{X}\|}, \dots, \frac{X_{d-2}}{\|\mathbf{X}\|} \right) \stackrel{(D)}{=} \frac{\|\mathbf{X}_1\|}{\|\mathbf{X}\|} \frac{\mathbf{X}_1}{\|\mathbf{X}_1\|} \\ &\stackrel{(D)}{=} U^{\frac{1}{d-2}} \frac{\mathbf{X}_1}{\|\mathbf{X}_1\|}, \end{aligned}$$

donde $U \sim \mathcal{U}(0, 1)$ y la última igualdad se obtiene por (1.12) con $U^{\frac{1}{d-2}} \stackrel{(D)}{=} \|\mathbf{X}_1\| / \|\mathbf{X}\|$. Como $\mathbf{X}_1 / \|\mathbf{X}_1\|$ se distribuye uniformemente en $\mathcal{S}^{d-3} \subset \mathbb{R}^{d-2}$, entonces por la Proposición 1.4, concluimos que \mathbf{T} se distribuye uniformemente en la bola \mathcal{B}^{d-2} . □

⁵ Por propiedades de la función gamma: $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$ y $\Gamma(1) = 1$.

Observación 1.2. Notemos que el Teorema 1.9 establece que la proyección ortogonal de la esfera \mathcal{S}^{d-1} en la bola \mathcal{B}^{d-2} preserva la medida uniforme, por lo que una medida uniforme en \mathcal{S}^{d-1} induce una medida uniforme en \mathcal{B}^{d-2} .

Para finalizar este capítulo, consideramos a la proyección ortogonal de la esfera \mathcal{S}^{d-1} en la bola \mathcal{B}^m con $1 \leq m \leq d-1$, y nos interesa la función de densidad conjunta de las primeras m coordenadas de una distribución uniforme $\mathbf{U} \sim \mathcal{U}(\mathcal{S}^{d-1})$; en efecto, se tiene el siguiente resultado como consecuencia del Lema 1.6.

Proposición 1.10. *Sea \mathbf{U} un vector aleatorio con distribución uniforme en la esfera \mathcal{S}^{d-1} definido como en la Proposición 1.2, sean $1 \leq m < d$ y \mathbf{Y} denota a las primeras m coordenadas de \mathbf{U} , entonces el vector aleatorio \mathbf{Y} admite la densidad*

$$f(\mathbf{y}) = \frac{\Gamma\left(\frac{d}{2}\right)}{\pi^{\frac{m}{2}} \Gamma\left(\frac{d-m}{2}\right)} \left(1 - \|\mathbf{y}\|^2\right)^{\frac{d-m}{2}-1}, \quad \forall \mathbf{y} \in \mathcal{B}^m. \quad (1.13)$$

Demostración. Por la Proposición 1.2, sea $\mathbf{X} = (X_1, \dots, X_d) \sim \mathcal{N}_d(\mathbf{0}, I)$, entonces $\mathbf{U} = \mathbf{X} / \|\mathbf{X}\| \sim \mathcal{U}(\mathcal{S}^{d-1})$, y sean \mathbf{X}_1 y \mathbf{Y} que denotan a las primeras m coordenadas de \mathbf{X} y \mathbf{U} , respectivamente. Por el Lema 1.6, se tiene que $\mathbf{Y} \stackrel{(D)}{=} R_m \mathbf{Y}'$, donde $R_m = \|\mathbf{X}_1\| / \|\mathbf{X}\|$ es independiente de $\mathbf{Y}' = \mathbf{X}_1 / \|\mathbf{X}_1\|$, con $R_m^2 \sim \mathfrak{B}\epsilon\left(\frac{m}{2}, \frac{d-m}{2}\right)$ y $\mathbf{Y}' \sim \mathcal{U}(\mathcal{S}^{m-1})$.

Sea $V = R_m^2$, entonces por el T.C.V., se obtiene la función de densidad para R_m

$$\begin{aligned} f_R(r) &= 2v^{\frac{1}{2}} p_V(v) \Big|_{v=r^2} = \frac{(2r)\Gamma\left(\frac{d}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{d-m}{2}\right)} r^{m-2} (1-r^2)^{\frac{d-m}{2}-1} \\ &= \frac{2\Gamma\left(\frac{d}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{d-m}{2}\right)} r^{m-1} (1-r^2)^{\frac{d-m}{2}-1}, \quad \forall r \in (0, 1). \end{aligned}$$

Sea $f_{\mathbf{Y}'}(\mathbf{y}')$ la densidad de $\mathbf{Y}' \sim \mathcal{U}(\mathcal{S}^{m-1})$, y se tiene que $f_{\mathbf{Y}'}(\mathbf{y}') = \Gamma\left(\frac{m}{2}\right) / (2\pi^{\frac{m}{2}})$ por (1.1). Notemos que $R_m = \|\mathbf{Y}\|$ y por la independencia de R_m y \mathbf{Y}' , bajo las coordenadas polares para la variable \mathbf{Y} , se obtiene la función de densidad para \mathbf{Y} dada por:

$$\begin{aligned} f_{\mathbf{Y}'}(\mathbf{y}') f_R(r) &= \frac{\Gamma\left(\frac{m}{2}\right)}{2\pi^{\frac{m}{2}}} \frac{2\Gamma\left(\frac{d}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{d-m}{2}\right)} r^{m-1} (1-r^2)^{\frac{d-m}{2}-1} \\ &= \frac{\Gamma\left(\frac{d}{2}\right)}{\pi^{\frac{m}{2}} \Gamma\left(\frac{d-m}{2}\right)} (1-r^2)^{\frac{d-m}{2}-1} r^{m-1}, \quad \forall r \in (0, 1), \mathbf{y}' \in \mathcal{S}^{m-1}. \end{aligned}$$

Entonces, por el Teorema 1.3 y el cambio de coordenadas dada en (1.5), se obtiene la densidad de \mathbf{Y} con respecto a la medida de Lebesgue en \mathbb{R}^m , la cual está dada por (1.13). \square

Observación 1.3. En el caso cuando $m = d-2$, tenemos que

$$f(\mathbf{y}) = \frac{\Gamma\left(\frac{d}{2}\right)}{\pi^{\frac{d-2}{2}} \Gamma\left(\frac{d-(d-2)}{2}\right)} \left(1 - \|\mathbf{y}\|^2\right)^{\frac{d-(d-2)}{2}-1} = \frac{\Gamma\left(\frac{d-2}{2} + 1\right)}{\pi^{\frac{d-2}{2}}} = \frac{1}{\text{Vol}(\mathcal{B}^{d-2})}.$$

Es decir, la distribución de \mathbf{Y} que denota las primeras $d-2$ coordenadas de $\mathbf{U} \sim \mathcal{U}(\mathcal{S}^{d-1})$ es una distribución uniforme en la bola \mathcal{B}^{d-2} . Esto es, el teorema de Arquímedes y el contenido del Teorema 1.9.

Capítulo 2

El movimiento browniano en \mathbb{R}^d y en la esfera \mathcal{S}^{d-1}

El movimiento browniano surgió primeramente para explicar el fenómeno del movimiento incesante e irregular de las partículas pequeñas suspendidas en un líquido, que fue observado por el botánico escocés R. Brown en 1827. Sin embargo, el modelo matemático más preciso para el movimiento browniano fue desarrollado por N. Wiener (1918), y así, el movimiento browniano también se denomina *proceso de Wiener*. Es uno de los procesos estocásticos más conocido y el primero que se ha investigado a fondo en la historia (cf. [RY99], [Øk03], [KT75], [Kal21], [IW89], [Fol99], etc). Hasta hoy en día, el movimiento browniano ha resultado ser de suma importancia en la teoría de probabilidad y en diversas áreas tanto de la ciencia pura como de la aplicada: física, biología, economía, estadística matemática y entre otras.

En este capítulo, estudiaremos el movimiento browniano en el espacio euclidiano de dimensión d ; entenderemos el proceso de Bessel como la parte radial de un movimiento browniano; y al final, investigaremos al movimiento browniano en la esfera unitaria \mathcal{S}^{d-1} y su relación con el movimiento browniano en el espacio euclidiano.

2.1. El movimiento browniano en \mathbb{R}^d

El modelo matemático para el fenómeno del movimiento browniano es un proceso estocástico a tiempo continuo, interpretado como la posición en tiempo t de una partícula ω . Un proceso estocástico se define como una familia de variables aleatorias indexada por el tiempo T , aquí consideraremos el tiempo continuo, i.e., $T = [0, \infty)$, y las variables asumen valores en \mathbb{R}^d .

Definición 2.1. Un *proceso estocástico* es una colección de variables aleatorias $(\mathbf{X}_t)_{t \geq 0}$ parametrizada por el conjunto de tiempo $T = [0, \infty)$, definida en un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ y con valores en el espacio medible $(\mathbb{R}^d, \mathcal{B}^d)$, donde \mathcal{B}^d es la σ -álgebra de Borel en el espacio euclidiano \mathbb{R}^d . Más aún, el proceso $\mathbf{X} = (\mathbf{X}_t)_{t \geq 0}$ se dice *adaptado a una filtración*¹ $(\mathcal{F}_t)_{t \geq 0}$ (denotado por (\mathcal{F}_t) -adaptado) si para cada $t \geq 0$, la variable aleatoria $\mathbf{X}_t : \Omega \rightarrow \mathbb{R}^d$ es \mathcal{F}_t -medible.

Observación 2.1. Un proceso estocástico $\mathbf{X} = (\mathbf{X}_t)_{t \geq 0}$ también puede definirse como una variable aleatoria definida en el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ y con valores en el espacio

¹ Una filtración completo $(\mathcal{F}_t)_{t \geq 0}$ es tal que para cada $t \geq 0$, \mathcal{F}_t es una sub- σ -álgebra de \mathcal{F} que contiene a todos los eventos nulos de \mathcal{F} y $\mathcal{F}_t = \bigcap_{\epsilon > 0} \mathcal{F}_{t+\epsilon}$; y $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ se llama un espacio de probabilidad completo filtrado.

de trayectorias $(\mathbb{W}^d, \mathcal{B}(\mathbb{W}^d))$, es decir, $\mathbf{X} : \Omega \rightarrow \mathbb{W}^d$ es $\mathcal{F}/\mathcal{B}(\mathbb{W}^d)$ -medible; donde \mathbb{W}^d es el espacio de todas las funciones continuas $\mathbf{w} : [0, \infty) \ni t \mapsto \mathbf{w}(t) \in \mathbb{R}^d$ y $\mathcal{B}(\mathbb{W}^d)$ es la σ -álgebra generada por los conjuntos cilíndricos de Borel en \mathbb{W}^d ([IW89], Chap. 1.4).

Definición 2.2. Sea μ una medida de probabilidad en el espacio medible $(\mathbb{R}^d, \mathcal{B}^d)$. Un proceso estocástico $\mathbf{B} = (\mathbf{B}_t)_{t \geq 0}$ adaptado a una filtración $(\mathcal{F}_t)_{t \geq 0}$ con valores en $(\mathbb{R}^d, \mathcal{B}^d)$ se denomina un *movimiento browniano de dimensión d con distribución inicial μ* si se satisfacen las siguientes propiedades:

- (i) $\mathbb{P}(\mathbf{B}_0 \in A) = \mu(A)$, $\forall A \in \mathcal{B}^d$;
- (ii) Las trayectorias $t \mapsto \mathbf{B}_t$ son continuas, *c.s.*, $\forall t \geq 0$;
- (iii) Para cualesquiera $s, t \geq 0$, el incremento $\mathbf{B}_{t+s} - \mathbf{B}_s$ tiene distribución normal multivariada de dimensión d con media $\mathbf{0}$ y matriz de covarianza tI_d ;
- (iv) Para cualquier conjunto finito de tiempos $0 < t_1 < t_2 < \dots < t_k$, los incrementos $\mathbf{B}_{t_2} - \mathbf{B}_{t_1}$, $\mathbf{B}_{t_3} - \mathbf{B}_{t_2}, \dots, \mathbf{B}_{t_k} - \mathbf{B}_{t_{k-1}}$ son variables aleatorias independientes con distribuciones dada en (iii).

La distribución de probabilidad de \mathbf{B} en el espacio medible $(\mathbb{W}^d, \mathcal{B}(\mathbb{W}^d))$ se llama la *medida de Wiener con la distribución inicial μ* y denotada por \mathbb{P}_μ .

La condición (ii) de la definición afirma que el movimiento browniano es un proceso continuo, y las condiciones (iii) y (iv) postulan que el desplazamiento del movimiento browniano $\mathbf{B}_{t+s} - \mathbf{B}_s$, $s, t \geq 0$ es independiente del pasado, y tiene distribución $\mathcal{N}_d(\mathbf{0}, tI_d)$, sólo depende de la longitud del intervalo de tiempos. Así, el movimiento browniano es un proceso continuo con incrementos independientes y estacionarios.

Observación 2.2. En la definición del movimiento browniano, si la medida inicial es² $\mu = \delta_x$, decimos que \mathbf{B} es un movimiento browniano de dimensión d que con valor inicial en \mathbf{x} ; cuando $\mathbf{x} = \mathbf{0}$ *c.s.*, le decimos a \mathbf{B} un *movimiento browniano estándar de dimensión d* , además, le denotaremos por $\mathbf{MB}(\mathbb{R}^d)$.

Observación 2.3. Notemos que si $\mathbf{B}_t = (B_t^1, \dots, B_t^d)$, $t \geq 0$ es un movimiento browniano de dimensión d , entonces los procesos unidimensionales $(B_t^k)_{t \geq 0}$, $1 \leq k \leq d$ son movimientos brownianos independientes de dimensión uno. También le diremos *movimiento browniano escalar* al movimiento browniano de dimensión uno.

La existencia y unicidad del movimiento browniano \mathbf{X} , dada la distribución inicial μ , se tiene por el siguiente teorema.

Teorema 2.3. *Para cualquier medida de probabilidad μ en $(\mathbb{R}^d, \mathcal{B}^d)$, existe una única medida de Wiener \mathbb{P}_μ en el espacio de trayectorias $(\mathbb{W}^d, \mathcal{B}(\mathbb{W}^d))$ con la distribución inicial μ .*

Demostración. Para demostrar la existencia, primero consideremos el caso $\mu = \delta_x$ para algún $\mathbf{x} \in \mathbb{R}^d$ y definimos

$$p(t, \mathbf{y}) = (2\pi t)^{-\frac{d}{2}} e^{-\frac{\|\mathbf{y}\|^2}{2t}}, \quad t > 0, \mathbf{y} \in \mathbb{R}^d, \quad (2.1)$$

donde $\|\cdot\|$ denota la norma euclidiana en \mathbb{R}^d , y $\|x\| := |x|$, para $x \in \mathbb{R}$ en el caso $d = 1$. Para cada conjunto finito de tiempos $0 < t_1 < t_2 \dots < t_k$, se define una medida de probabilidad ν_{t_1, \dots, t_k} en el espacio medible $(\mathbb{R}^{kd}, \mathcal{B}^{kd})$ dada por

$$\begin{aligned} \nu_{t_1, \dots, t_k}(F_1 \times \dots \times F_k) &= \int_{\mathbb{R}^d} \delta_x(d\mathbf{x}) \int_{F_1 \times \dots \times F_k} p(t_1, \mathbf{x}_1 - \mathbf{x}) p(t_2 - t_1, \mathbf{x}_2 - \mathbf{x}_1) \\ &\quad \dots p(t_k - t_{k-1}, \mathbf{x}_k - \mathbf{x}_{k-1}) d\mathbf{x}_1 \dots d\mathbf{x}_k. \end{aligned}$$

² δ_x denota una medida de Dirac en $(\mathbb{R}^d, \mathcal{B}^d)$ y definida como $\delta_x(A) = \mathbf{1}_A(\mathbf{x})$, p.a. $\mathbf{x} \in \mathbb{R}^d$, $\forall A \in \mathcal{B}^d$.

donde $F_i \in \mathcal{B}^d$ para todo $i = 1, \dots, k$.

Luego, extendemos esta definición para toda sucesión finita de tiempos de manera que

$$\nu_{t_{\pi(1)}, \dots, t_{\pi(k)}}(F_{\pi(1)} \times \dots \times F_{\pi(k)}) = \nu_{t_1, \dots, t_k}(F_1 \times \dots \times F_k), \quad (2.2)$$

para π cualquier permutación en $\{1, 2, \dots, k\}$. Además, para cada $0 \leq u < s < t$ y $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^d$, notemos que

$$\begin{aligned} & \int_{\mathbb{R}^d} \left(2\pi(t-s)\right)^{-\frac{d}{2}} \exp\left\{-\frac{\|\mathbf{y}-\mathbf{x}\|^2}{2(t-s)}\right\} \left(2\pi(s-u)\right)^{-\frac{d}{2}} \exp\left\{-\frac{\|\mathbf{x}-\mathbf{z}\|^2}{2(s-u)}\right\} d\mathbf{x} \\ &= \left(2\pi(t-u)\right)^{-\frac{d}{2}} \exp\left\{-\frac{\|\mathbf{y}-\mathbf{z}\|^2}{2(t-u)}\right\} = \int_{\mathbb{R}^d} p(t-s, \mathbf{y}-\mathbf{x}) p(s-u, \mathbf{x}-\mathbf{z}) d\mathbf{x}, \end{aligned}$$

donde la primera igualdad se obtiene usando el siguiente resultado:

$$\begin{aligned} (\mathbf{x}-\mathbf{a})^\top A(\mathbf{x}-\mathbf{a}) + (\mathbf{x}-\mathbf{b})^\top B(\mathbf{x}-\mathbf{b}) &= (\mathbf{a}-\mathbf{b})^\top A(A+B)^{-1}B(\mathbf{a}-\mathbf{b}) \\ &+ \left[\mathbf{x} - (A+B)^{-1}(A\mathbf{a} + B\mathbf{b})\right]^\top (A+B) \left[\mathbf{x} - (A+B)^{-1}(A\mathbf{a} + B\mathbf{b})\right], \end{aligned}$$

para A, B matrices simétricas invertibles de dimensión d y $\mathbf{x}, \mathbf{a}, \mathbf{b} \in \mathbb{R}^d$; y por $p(t, \mathbf{y})$ es la densidad de una distribución normal multivariada, entonces $\int_{\mathbb{R}^d} p(t, \mathbf{y}-\mathbf{z}) d\mathbf{y} = 1$, $\forall t > 0$ y $\mathbf{y}, \mathbf{z} \in \mathbb{R}^d$; así, para todo $m \in \mathbb{N}$,

$$\nu_{t_1, \dots, t_k}(F_1 \times \dots \times F_k) = \nu_{t_1, \dots, t_k, t_{k+1}, \dots, t_{k+m}}(F_1 \times \dots \times F_k \times \mathbb{R}^d \times \dots \times \mathbb{R}^d). \quad (2.3)$$

Entonces, (2.2) y (2.3) afirman que la medida ν_{t_1, \dots, t_k} cumple las dos condiciones del teorema de extensión de Kolmogorov ([Ok03], Thm. 2.1.5), y así la existencia de un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P}_x)$ y un proceso estocástico $\mathbf{X} = (\mathbf{X}_t)_{t \geq 0}$ definido en él con $\mathbf{X}_t: \Omega \rightarrow \mathbb{R}^d$ tal que

$$\begin{aligned} \mathbb{P}_x(\mathbf{X}_{t_1} \in F_1, \dots, \mathbf{X}_{t_k} \in F_k) &= \int_{F_1} p(t_1, \mathbf{x}_1 - \mathbf{x}) d\mathbf{x}_1 \int_{F_2} p(t_2 - t_1, \mathbf{x}_2 - \mathbf{x}_1) d\mathbf{x}_2 \\ &\dots \int_{F_k} p(t_k - t_{k-1}, \mathbf{x}_k - \mathbf{x}_{k-1}) d\mathbf{x}_k, \end{aligned} \quad (2.4)$$

para $0 = t_0 < t_1 < \dots < t_k$. Por lo tanto, las variables $\mathbf{X}_{t_1} - \mathbf{X}_{t_0}, \dots, \mathbf{X}_{t_k} - \mathbf{X}_{t_{k-1}}$ son mutuamente independientes y cada uno tiene distribución normal multivariada. Además, notemos que $\mathbb{P}_x(\mathbf{X}_0 = \mathbf{x}) = 1$.

Además, para cualesquiera tiempos $0 \leq s < t$, es fácil demostrar que

$$\mathbb{E} [\|\mathbf{X}_t - \mathbf{X}_s\|^4] = d(d+2)|t-s|^2.$$

Entonces por el teorema de continuidad de Kolmogorov ([Ok03], Thm. 2.6), existe un proceso $\mathbf{B} = (\mathbf{B}_t)_{t \geq 0}$ con trayectorias continuas tal que para cada $t > 0$, $\mathbf{X}_t = \mathbf{B}_t$ c.s..

Por tanto, la distribución de probabilidad \mathbb{P}_x de \mathbf{B} en el espacio $(\mathbb{W}^d, \mathcal{B}(\mathbb{W}^d))$ es una medida de Wiener con la distribución inicial δ_x . En general, para cualquier medida de probabilidad μ en $(\mathbb{R}^d, \mathcal{B}^d)$, y $\forall A \in \mathcal{B}(\mathbb{W}^d)$, $\mathbb{P}_\mu(A) = \int_{\mathbb{R}^d} \mathbb{P}_x(A) \mu(d\mathbf{x})$ define una medida de Wiener con la distribución inicial μ .

Finalmente, la unicidad de la medida es inmediata de la definición y es consecuencia del teorema de unicidad de extensión para medidas ([Fol99], Thm. 1.14). \square

Una propiedad simple de un movimiento browniano $\mathbf{B} = (\mathbf{B}_t)_{t \geq 0}$ en \mathbb{R}^d es la siguiente:

Proposición 2.4. *Sea H una transformación ortogonal en \mathbb{R}^d , entonces $H(\mathbf{B}) = \{H(\mathbf{B}_t) : t \geq 0\}$ también es un movimiento browniano en \mathbb{R}^d .*

Demostración. Como \mathbf{B} es continua en el tiempo t , para todo $t \geq 0$, c.s., entonces $H(\mathbf{B})$ también lo es. Por otro lado, como $H(\mathbf{B}_{s+t}) - H(\mathbf{B}_s) = H(\mathbf{B}_{s+t} - \mathbf{B}_s)$ sólo depende de $\mathbf{B}_{s+t} - \mathbf{B}_s$, entonces por la independencia de incrementos de \mathbf{B} , los incrementos de $H(\mathbf{B})$ también son independientes. Y como $\mathbf{B}_{s+t} - \mathbf{B}_s$ tiene distribución normal multivariada $\mathcal{N}_d(\mathbf{0}, tI_d)$ con la función característica dada por

$$\mathbb{E}[\exp\{i\langle \mathbf{y}, \mathbf{B}_{s+t} - \mathbf{B}_s \rangle\} | \mathcal{F}_s] = \exp\{-\langle \mathbf{y}, \mathbf{y} \rangle t/2\}, \quad \mathbf{y} \in \mathbb{R}^d,$$

donde $\langle \cdot, \cdot \rangle$ denota al producto punto, y sabemos que el producto interior es invariante bajo transformaciones ortogonales, utilizando la ecuación anterior y por H^{-1} también es una transformación ortogonal, obtenemos

$$\begin{aligned} \mathbb{E}[\exp\{i\langle \mathbf{y}, H(\mathbf{B}_{s+t} - \mathbf{B}_s) \rangle\} | \mathcal{F}_s] &= \mathbb{E}[\exp\{i\langle H^{-1}(\mathbf{y}), \mathbf{B}_{s+t} - \mathbf{B}_s \rangle\} | \mathcal{F}_s] \\ &= \exp\{-\langle H^{-1}(\mathbf{y}), H^{-1}(\mathbf{y}) \rangle \frac{t}{2}\} = \exp\{-\langle \mathbf{y}, \mathbf{y} \rangle t/2\}. \end{aligned}$$

Por lo tanto, $H(\mathbf{B}_{s+t}) - H(\mathbf{B}_s)$ también tiene distribución normal multivariada, esto es, $H(\mathbf{B})$ también es un movimiento browniano de dimensión n . \square

Notemos que en la demostración del Teorema 2.3, la densidad conjunta del movimiento browniano $\mathbf{B} = (\mathbf{B}_t)_{t \geq 0}$ para $0 < t_1 < \dots < t_k$ está dada por (2.4). Y sea $f: \mathbb{R}^d \rightarrow \mathbb{R}$ una función acotada y continua (lo denotaremos por $f \in \mathcal{C}_a(\mathbb{R}^d)$), definimos el operador de transición P_t como sigue:

$$P_t f(\mathbf{x}) = \int_{\mathbb{R}^d} f(\mathbf{y}) p_t(\mathbf{x}, \mathbf{y}) d\mathbf{y}, \quad t > 0, \mathbf{x}, \mathbf{y} \in \mathbb{R}^d; \quad (2.5)$$

donde $p_t(\mathbf{x}, \mathbf{y}) := p(t, \mathbf{y} - \mathbf{x})$ es la densidad de transición dada por (2.1) para el movimiento browniano \mathbf{B} . Luego, es inmediato por la definición del movimiento browniano que \mathbf{B} cumple la siguiente *propiedad de Markov*:

$$\mathbb{E}[f(\mathbf{B}_t) | \mathcal{F}_s] = P_{t-s} f(\mathbf{B}_s), \quad \forall s < t, f \in \mathcal{C}_a(\mathbb{R}^d). \quad (2.6)$$

Esto es, el movimiento browniano \mathbf{B} es un proceso de Markov (tiempo homogéneo). Más aún, si definimos $P_0 = I$, el operador identidad y por la definición de (2.5), se puede demostrar fácilmente que la familia $(P_t)_{t \geq 0}$ forma un semigrupo por:

$$P_{t+s} = P_s P_t = P_t P_s, \quad s, t \geq 0, \quad (2.7)$$

llamada la ecuación de Chapman-Kolmogorov. Así, la familia $(P_t)_{t \geq 0}$ se denomina un *semigrupo de transición browniano*, y satisface las siguientes propiedades.

Proposición 2.5. *Sea $(P_t)_{t \geq 0}$ el semigrupo de transición para el movimiento browniano $\mathbf{B} = (\mathbf{B}_t)_{t \geq 0}$, entonces se cumple las siguientes propiedades:*

$$(1). \|P_t f\|_\infty \leq \|f\|_\infty, \|P_t\|_\infty \leq 1, \text{ donde } \|f\|_\infty := \sup_{\mathbf{y} \in \mathbb{R}^d} |f(\mathbf{y})|, \text{ para } f \in \mathcal{C}_a(\mathbb{R}^d).$$

$$(2). P_t \mathcal{C}_0 \subset \mathcal{C}_0 \text{ para cada } t \geq 0, \text{ donde } \mathcal{C}_0 := \left\{ f : f \in \mathcal{C}_a(\mathbb{R}^d) \text{ y } \lim_{\|\mathbf{x}\| \rightarrow \infty} f(\mathbf{x}) = 0 \right\}.$$

$$(3). \text{ Si } f \in \mathcal{C}_0, \text{ entonces } \lim_{t \rightarrow 0} \|P_t f - f\|_\infty = 0.$$

Demostración. (1) Es claro por la definición de $P_t f$ dada en (2.5).

(2) Primero supongamos que f es una función acotada, entonces

$$\left| P_t f(\mathbf{x}) - P_t f(\mathbf{x}_0) \right| \leq \frac{\|f\|_\infty}{(2\pi t)^{d/2}} \int_{\mathbb{R}^d} \left| \exp \left\{ -\frac{\|\mathbf{y}-\mathbf{x}\|^2}{2t} \right\} - \exp \left\{ -\frac{\|\mathbf{y}-\mathbf{x}_0\|^2}{2t} \right\} \right| d\mathbf{y}.$$

Notemos que el integrando tiende a cero cuando $\mathbf{x} \rightarrow \mathbf{x}_0$ y está acotada por $g(\mathbf{y}) = e^{-\frac{\|\mathbf{y}-\mathbf{x}_0\|^2}{2t}}$ con $g(\mathbf{y})$ integrable, entonces el teorema de convergencia acotada afirma que la parte derecha de la desigualdad tiende a cero cuando $\mathbf{x} \rightarrow \mathbf{x}_0$; así, $P_t f(\mathbf{x})$ es continua.

Ahora, para $f \in \mathcal{C}_0$ y $N > 0$, se tiene que

$$\begin{aligned} \left| P_t f(\mathbf{x}) \right| &\leq \int_{\|\mathbf{y}\| \geq N} \frac{1}{(2\pi t)^{d/2}} \exp \left\{ -\frac{\|\mathbf{y}-\mathbf{x}\|^2}{2t} \right\} |f(\mathbf{y})| d\mathbf{y} \\ &\quad + \|f(\mathbf{x})\|_\infty \int_{\|\mathbf{y}\| < N} \frac{1}{(2\pi t)^{d/2}} \exp \left\{ -\frac{\|\mathbf{y}-\mathbf{x}\|^2}{2t} \right\} d\mathbf{y}, \end{aligned}$$

como $f \in \mathcal{C}_0$ es tal que se hace cero en el infinito, entonces para todo $\epsilon > 0$, existe $N = N_\epsilon$ suficientemente grande tal que la primera integral es menor que $\frac{\epsilon}{2}$; entonces fijo N_ϵ , la segunda integral corresponde a la función de distribución de una normal multivariada, y cuando $\|\mathbf{x}\| \rightarrow \infty$, tenemos que ésta integral es menor que $\frac{\epsilon}{2}$, por lo que $\lim_{\|\mathbf{x}\| \rightarrow \infty} P_t f(\mathbf{x}) = 0$, es decir, $P_t \mathcal{C}_0 \subset \mathcal{C}_0$.

(3) Para $\mathbf{x} \in \mathbb{R}^d$, se tiene que

$$\begin{aligned} (P_t f - f)(\mathbf{x}) &= \int_{\mathbb{R}^d} f(\mathbf{y}) \frac{1}{(2\pi t)^{d/2}} \exp \left\{ -\frac{\|\mathbf{y}-\mathbf{x}\|^2}{2t} \right\} d\mathbf{y} - f(\mathbf{x}) \int_{\mathbb{R}^d} \frac{1}{(2\pi t)^{d/2}} e^{-\frac{\|\mathbf{y}\|^2}{2t}} d\mathbf{y} \\ &= \int_{\mathbb{R}^d} (f(\mathbf{x} + \mathbf{y}) - f(\mathbf{x})) \frac{1}{(2\pi t)^{d/2}} \exp \left\{ -\frac{\|\mathbf{y}\|^2}{2t} \right\} d\mathbf{y}. \end{aligned}$$

Como $f \in \mathcal{C}_0$ es continua y acotada, entonces $\forall \epsilon > 0, \exists \delta > 0$ tal que para todo $\mathbf{x} \in \mathbb{R}^d$, se tiene que $|f(\mathbf{x} + \mathbf{y}) - f(\mathbf{x})| < \frac{\epsilon}{2}$ si $\|\mathbf{y}\| < \delta$, por lo tanto,

$$\begin{aligned} \|P_t f - f\|_\infty &\leq \sup_{\mathbf{x}} \left[\int_{\|\mathbf{y}\| < \frac{\delta}{2}} + \int_{\|\mathbf{y}\| \geq \frac{\delta}{2}} \right] \left| f(\mathbf{x} + \mathbf{y}) - f(\mathbf{x}) \right| \frac{1}{(2\pi t)^{d/2}} \exp \left\{ -\frac{\|\mathbf{y}\|^2}{2t} \right\} d\mathbf{y} \\ &\leq \frac{\epsilon}{2} + 2 \|f\|_\infty \int_{\|\mathbf{y}\| \geq \frac{\delta}{2}} \frac{1}{(2\pi t)^{d/2}} \exp \left\{ -\frac{\|\mathbf{y}\|^2}{2t} \right\} d\mathbf{y} \\ &= \frac{\epsilon}{2} + 2 \|f\|_\infty \int_{\|\mathbf{u}\| \geq \frac{\delta}{2\sqrt{t}}} \frac{1}{(2\pi)^{d/2}} \exp \left\{ -\frac{\|\mathbf{u}\|^2}{2} \right\} d\mathbf{u}, \end{aligned}$$

en donde, cuando $t \rightarrow 0$, vemos que la segunda integral de la última igualdad es menor que $\frac{\epsilon}{2}$, entonces se concluye que $\lim_{t \rightarrow 0} \|P_t f - f\|_\infty = 0$. Esto es, el semigrupo $(P_t)_{t \geq 0}$ es fuertemente continuo en el espacio \mathcal{C}_0 . \square

Observación 2.4. Cualquier semigrupo de funciones de transición $P_t, t \geq 0$ en \mathcal{C}_0 que satisface las condiciones 2.5(1)–2.5(3) se conoce como *función de transición de Feller*; un proceso de Markov que tiene una función de transición de Feller se denomina un *proceso de Feller* (ver [RY99]). Por lo que un movimiento browniano también es un proceso de Feller.

Por la proposición anterior, vemos que el semigrupo $(P_t)_{t \geq 0}$ asociado con un movimiento browniano $\mathbf{B} = (\mathbf{B}_t)_{t \geq 0}$ es un semigrupo fuertemente continuo en el espacio \mathcal{C}_0 , entonces se puede definir el generador infinitesimal del semigrupo en el espacio de Banach $(\mathcal{C}_0, \|\cdot\|_\infty)$.

Definición 2.6. Sea $f \in \mathcal{C}_0$, y denotamos por

$$Af := \lim_{t \downarrow 0} \frac{P_t f - f}{t}, \quad (2.8)$$

cuando el límite existe (convergencia uniforme), sea $\mathcal{D}_A = \{f; f \in \mathcal{C}_0 \text{ y } Af \in \mathcal{C}_0\}$, entonces el operador $A: \mathcal{D}_A \rightarrow \mathcal{C}_0$ definido por (2.8) se llama *el generador infinitesimal* del semigrupo $(P_t)_{t \geq 0}$, y denotado por (A, \mathcal{D}_A) .

Mientras que el operador A tiene las siguientes propiedades (ver [RY99], Chap. 7).

Proposición 2.7. Sea $f \in \mathcal{D}_A$, entonces

- (i) $P_t f \in \mathcal{D}_A$ y $\frac{d}{dt} P_t f = A P_t f = P_t A f$, para todo $t > 0$.
- (ii) $P_t f - f = \int_0^t P_s A f ds = \int_0^t A P_s f ds$.

En particular, el generador infinitesimal del movimiento browniano $\mathbf{B} = (\mathbf{B}_t)_{t \geq 0}$ está relacionado con el operador de Laplace.

Teorema 2.8. Sea $f \in \mathcal{C}_0^2(\mathbb{R}^d)$.³ Entonces,

$$A f(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^d \frac{\partial^2 f(\mathbf{x})}{\partial x_i^2} := \frac{1}{2} \Delta f(\mathbf{x}), \quad (2.9)$$

donde $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$, y Δ es conocido como el operador de Laplace.

Demostración. Para $f \in \mathcal{C}_0$, sea $\mathbf{y} = \mathbf{x} + \sqrt{t}\mathbf{z}$, entonces

$$P_t f(\mathbf{x}) = \frac{1}{(2\pi t)^{d/2}} \int_{\mathbb{R}^d} e^{-\frac{\|\mathbf{y}-\mathbf{x}\|^2}{2t}} f(\mathbf{y}) d\mathbf{y} = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-\frac{\|\mathbf{z}\|^2}{2}} f(\mathbf{x} + \sqrt{t}\mathbf{z}) d\mathbf{z}. \quad (2.10)$$

Si $f \in \mathcal{C}_0^2(\mathbb{R}^d)$, por la expansión de Taylor, se tiene que

$$f(\mathbf{x} + \sqrt{t}\mathbf{z}) = f(\mathbf{x}) + \sqrt{t} \sum_{i=1}^d z_i \frac{\partial f(\mathbf{x})}{\partial x_i} + \frac{t}{2} \sum_{i,j=1}^d z_i z_j \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} + \frac{t}{2} \sum_{i,j=1}^d z_i z_j \left[\frac{\partial^2 f(\mathbf{x}')}{\partial x_i \partial x_j} - \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} \right],$$

donde \mathbf{x}' es algún punto en el segmento $[\mathbf{x}, \mathbf{x} + \sqrt{t}\mathbf{z}]$. Entonces sustituyendo en (2.10), y es fácil de mostrar que

$$\int_{\mathbb{R}^d} e^{-\frac{\|\mathbf{z}\|^2}{2}} z_i \frac{\partial f(\mathbf{x})}{\partial x_i} d\mathbf{z} = 0, \quad \int_{\mathbb{R}^d} e^{-\frac{\|\mathbf{z}\|^2}{2}} z_i^2 \frac{\partial^2 f(\mathbf{x})}{\partial x_i^2} d\mathbf{z} = (2\pi)^{\frac{d}{2}} \frac{\partial^2 f(\mathbf{x})}{\partial x_i^2}, \quad \forall i = 1, \dots, d,$$

y

$$\int_{\mathbb{R}^d} e^{-\frac{\|\mathbf{z}\|^2}{2}} z_i z_j \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} d\mathbf{z} = 0, \quad \forall i \neq j,$$

obtenemos que

$$P_t f(\mathbf{x}) = f(\mathbf{x}) + \frac{t}{2} \Delta f(\mathbf{x}) + \frac{t}{2} \frac{1}{(2\pi)^{d/2}} J(t, \mathbf{x}), \quad (2.11)$$

donde $J(t, \mathbf{x}) = \int_{\mathbb{R}^d} e^{-\frac{\|\mathbf{z}\|^2}{2}} \sum_{i,j=1}^d z_i z_j \left[\frac{\partial^2 f(\mathbf{x}')}{\partial x_i \partial x_j} - \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} \right] d\mathbf{z}$. Sea

$$F(\mathbf{x}, \mathbf{z}, t) = \max_{i,j} \left| \frac{\partial^2 f(\mathbf{x}')}{\partial x_i \partial x_j} - \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} \right|,$$

³ $\mathcal{C}_0^2(\mathbb{R}^d) := \{f : f \in \mathcal{C}_0 \text{ es dos veces continuamente diferenciable con sus derivadas parciales de primer y de segundo orden en } \mathcal{C}_0\}$.

entonces, para cualquier $R > 0$, se tiene que

$$\begin{aligned} |J(t, \mathbf{x})| &\leq \int_{\mathbb{R}^d} e^{-\frac{\|\mathbf{z}\|^2}{2}} \sum_{i=1}^d \sum_{j=1}^d \frac{z_i^2 + z_j^2}{2} F(\mathbf{x}, \mathbf{z}, t) d\mathbf{z} = d \int_{\mathbb{R}^d} \left(\sum_{i=1}^d z_i^2 \right) e^{-\frac{\|\mathbf{z}\|^2}{2}} F(\mathbf{x}, \mathbf{z}, t) d\mathbf{z} \\ &\leq d \int_{\|\mathbf{z}\| < R} \|\mathbf{z}\|^2 e^{-\frac{\|\mathbf{z}\|^2}{2}} F(\mathbf{x}, \mathbf{z}, t) d\mathbf{z} + 2d \max_{i,j} \left| \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} \right| \int_{\|\mathbf{z}\| \geq R} \|\mathbf{z}\|^2 e^{-\frac{\|\mathbf{z}\|^2}{2}} d\mathbf{z} \\ &= I_1(R) + I_2(R). \end{aligned}$$

Por la continuidad uniforme de la segunda derivada parcial para $f \in \mathcal{C}_0^2(\mathbb{R}^d)$, cuando $t \rightarrow 0$, $I_1 \rightarrow 0$ uniformemente en \mathbf{x} , entonces para cualquier $R > 0$, se tiene que

$$\limsup_{t \downarrow 0} \sup_{\mathbf{x}} |J(t, \mathbf{x})| \leq I_2(R),$$

de donde se tiene que $I_2(R) \rightarrow 0$ cuando $R \rightarrow \infty$, por lo tanto, $\limsup_{t \downarrow 0} \sup_{\mathbf{x}} |J(t, \mathbf{x})| = 0$.

En consecuencia, por (2.11), obtenemos

$$\lim_{t \downarrow 0} \left\| \frac{P_t f - f}{t} - \frac{1}{2} \Delta f \right\|_{\infty} = 0,$$

para cualquier $f \in \mathcal{C}_0^2(\mathbb{R}^d)$. □

Por la definición de un generador infinitesimal, podemos considerar un movimiento browniano \mathbf{B} como solución a un problema de martingala. Antes de ver este resultado, tenemos la definición de una martingala y martingala local.

Definición 2.9. Sea $\mathbf{X} = (\mathbf{X}_t)_{t \geq 0}$ un proceso estocástico definido en el espacio de probabilidad filtrado $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ que es continuo y (\mathcal{F}_t) -adaptado.

(1) \mathbf{X} se llama una *martingala con respecto a* $(\mathcal{F}_t)_{t \geq 0}$ (denotada por (\mathcal{F}_t) -martingala) si \mathbf{X}_t es integrable para todo $t \geq 0$ y satisface:

$$\mathbb{E}[\mathbf{X}_t | \mathcal{F}_s] = \mathbf{X}_s, \quad c.s. \quad \forall 0 \leq s < t. \quad (2.12)$$

(2) Una variable aleatoria $\tau : \Omega \rightarrow [0, \infty) \cup \{\infty\}$ se denomina un (\mathcal{F}_t) -tiempo de paro si para cada $t \geq 0$, $\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t$.

(3) \mathbf{X} es una (\mathcal{F}_t) -martingala local si existe una sucesión no decreciente de tiempos de paro $\{\tau_k\}_{k=1}^{\infty}$, tal que $\lim_{k \uparrow \infty} \tau_k = \infty$, c.s., y para cada k , el proceso parado definido por

$$\mathbf{X}_t^k = \mathbf{X}_{t \wedge \tau_k} - \mathbf{X}_0, \quad \forall t \geq 0,$$

es una (\mathcal{F}_t) -martingala; donde $t \wedge \tau_k := \min\{t, \tau_k\}$.

Observación 2.5. Para \mathbf{B} un movimiento browniano estándar de dimensión d , por las propiedades 2.2(i) y 2.2(iii), $\mathbf{B}_t \sim \mathcal{N}_d(\mathbf{0}, tI_d)$, se tiene que \mathbf{B}_t es integrable para todo t ; y si definimos⁴ $\mathcal{F}_t^{\mathbf{B}} = \bigcap_{\epsilon > 0} \sigma(\{\mathbf{B}_s : s \leq t + \epsilon\})$, $\forall t \geq 0$, por las propiedades 2.2(iii) y 2.2(iv), $\mathbf{B}_t - \mathbf{B}_s$ es independiente de $\mathcal{F}_s^{\mathbf{B}}$, $\forall 0 \leq s < t$, es decir,

$$\mathbb{E}[\mathbf{B}_t - \mathbf{B}_s | \mathcal{F}_s^{\mathbf{B}}] = \mathbf{0}, \quad c.s. \quad \iff \quad \mathbb{E}[\mathbf{B}_t | \mathcal{F}_s^{\mathbf{B}}] = \mathbf{B}_s, \quad c.s. \quad \forall 0 \leq s < t.$$

Así, \mathbf{B} es una martingala con respecto a $(\mathcal{F}_t^{\mathbf{B}})_{t \geq 0}$.

⁴ $\sigma(\mathbf{B}_t) = \{\mathbf{B}_t^{-1}(A) : A \in \mathcal{B}^d\}$, es la σ -álgebra generada por la variable \mathbf{B}_t , $t \geq 0$.

Notación. Si para una filtración $(\mathcal{F}_t)_{t \geq 0}$ en el espacio de probabilidad completo $(\Omega, \mathcal{F}, \mathbb{P})$ donde está definido el movimiento browniano \mathbf{B} es tal que \mathbf{B} es una martingala con respecto a $(\mathcal{F}_t)_{t \geq 0}$, le decimos a \mathbf{B} un (\mathcal{F}_t) -movimiento browniano.

Proposición 2.10. *Sea $\mathbf{B} = (\mathbf{B}_t)_{t \geq 0}$ un (\mathcal{F}_t) -movimiento browniano en \mathbb{R}^d con distribución inicial $\mu = \delta_x$, para algún $\mathbf{x} \in \mathbb{R}^d$; y con el generador (A, \mathcal{D}_A) dado en (2.9). Entonces el movimiento browniano \mathbf{B} es una solución al siguiente problema de martingala; es decir, para todo $f \in \mathcal{D}_A$,*

$$M_t^f := f(\mathbf{B}_t) - f(\mathbf{B}_0) - \int_0^t \frac{1}{2} \Delta f(\mathbf{B}_s) ds \quad (2.13)$$

es una (\mathcal{F}_t) -martingala, donde $f(\mathbf{B}_0) = f(\mathbf{x})$.

Demostración. Primero notemos que por el Teorema 2.8, f y Af están en \mathcal{C}_0 , son funciones acotadas, entonces M_t^f dada por (2.13) es integrable para cada t . Así, para demostrar M_t^f es una (\mathcal{F}_t) -martingala, debemos demostrar que para cualesquiera $0 \leq s \leq t$,

$$\mathbb{E}[M_t^f | \mathcal{F}_s] = M_s^f, \text{ c.s.} \quad \iff \quad \mathbb{E}[M_t^f - M_s^f | \mathcal{F}_s] = 0, \text{ c.s.}$$

Como se tiene que

$$M_t^f - M_s^f = f(\mathbf{B}_t) - f(\mathbf{B}_s) - \int_s^t \frac{1}{2} \Delta f(\mathbf{B}_u) du.$$

Entonces, por la propiedad de Markov del movimiento browniano (2.6), tenemos

$$\begin{aligned} \mathbb{E}[M_t^f - M_s^f | \mathcal{F}_s] &= \mathbb{E}[f(\mathbf{B}_t) | \mathcal{F}_s] - \mathbb{E}[f(\mathbf{B}_s) | \mathcal{F}_s] - \mathbb{E}\left[\int_s^t \frac{1}{2} \Delta f(\mathbf{B}_u) du \mid \mathcal{F}_s\right] \\ &= P_{t-s} f(\mathbf{B}_s) - f(\mathbf{B}_s) - \int_0^{t-s} \mathbb{E}\left[\frac{1}{2} \Delta f(\mathbf{B}_{h+s}) \mid \mathcal{F}_s\right] dh \quad (\text{sea } h = u - s) \\ &= P_{t-s} f(\mathbf{B}_s) - f(\mathbf{B}_s) - \int_0^{t-s} P_h \left(\frac{1}{2} \Delta f(\mathbf{B}_s)\right) dh, \end{aligned}$$

así, por la propiedad (ii) del operador $A = \frac{1}{2} \Delta$ en la Proposición 2.7, la última igualdad de la ecuación anterior se hace cero. Es decir, $\mathbb{E}[M_t^f - M_s^f | \mathcal{F}_s] = 0$. \square

Observación 2.6. Notemos que por la independencia de incrementos del movimiento browniano, en la proposición anterior, M_t^f es una (\mathcal{F}_t) -martingala es equivalente a que $\mathbb{E}_x[M_t^f] = 0$, para cualquier $\mathbf{x} \in \mathbb{R}^d$ y $t \geq 0$; es decir,

$$\mathbb{E}_x[f(\mathbf{B}_t)] = f(\mathbf{x}) + \mathbb{E}_x\left[\int_0^t \frac{1}{2} \Delta f(\mathbf{B}_s) ds\right], \quad (2.14)$$

donde $\mathbb{E}_x[\cdot]$ corresponde a la esperanza con respecto a la distribución \mathbb{P}_x del movimiento browniano con punto inicial en \mathbf{x} .

Finalmente, concluimos que un movimiento browniano $\mathbf{B} = (\mathbf{B}_t)_{t \geq 0}$ es un proceso de Markov y es a la vez un proceso de difusión por la siguiente definición ([RY99]).

Definición 2.11. Un proceso de Markov $\mathbf{X} = (\mathbf{X}_t)_{t \geq 0}$ definido en el espacio de probabilidad filtrado $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ con valores en \mathbb{R}^d y distribución inicial \mathbb{P}_x se denomina un *proceso de difusión con generador A* si

- i) \mathbf{X} tiene trayectorias continuas, *c.s.*
 ii) para cualquier $\mathbf{x} \in \mathbb{R}^d$ y $f \in \mathcal{C}_a^\infty$,

$$\mathbb{E}_x[f(\mathbf{X}_t)] = f(\mathbf{x}) + \mathbb{E}_x \left[\int_0^t \mathcal{A}f(\mathbf{X}_s) ds \right],$$

donde \mathcal{A} es un operador diferencial de segundo orden dado por

$$\mathcal{A}f(\mathbf{x}) = \frac{1}{2} \sum_{i,j=1}^d a_{ij}(\mathbf{x}) \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} + \sum_{i=1}^d b_i(\mathbf{x}) \frac{\partial f(\mathbf{x})}{\partial x_i},$$

para $a_{ij}: \mathbb{R}^d \rightarrow \mathbb{R}$ y $b_i: \mathbb{R}^d \rightarrow \mathbb{R}$ funciones Borel medibles y localmente acotadas, $i, j = 1, \dots, d$; además, para cada \mathbf{x} , la matriz $(a_{ij}(\mathbf{x}))$ es simétrica y definida positiva de dimensión $d \times d$.

Observación 2.7. La propiedad (ii) anterior es equivalente a decir que el proceso \mathbf{X} es una solución a un *problema de martingala con distribución inicial* \mathbb{P}_x , es decir, $\mathbb{P}_x(\mathbf{X}_0 = \mathbf{x}) = 1$ y para cualquier $f \in \mathcal{C}^2$, el proceso

$$M_t^f := f(\mathbf{X}_t) - f(\mathbf{X}_0) - \int_0^t \mathcal{A}f(\mathbf{X}_s) ds \quad (2.15)$$

es una martingala local con respecto a la filtración incompleta $(\mathcal{F}_t^0) = (\sigma(\mathbf{X}_s, s \leq t))$.

En general, un generador \mathcal{A} de un proceso de difusión \mathbf{X}_t se define como:

Definición 2.12. Sea $(\mathbf{X}_t)_{t \geq 0}$ un proceso de difusión (tiempo homogéneo) en \mathbb{R}^d , el *generador infinitesimal* \mathcal{A} de \mathbf{X}_t está definida como

$$\mathcal{A}f(\mathbf{x}) = \lim_{t \downarrow 0} \frac{\mathbb{E}_x[f(\mathbf{X}_t)] - f(\mathbf{x})}{t}, \quad \mathbf{x} \in \mathbb{R}^d.$$

Se denota por $\mathcal{D}_\mathcal{A}(\mathbf{x})$ al conjunto de funciones $f: \mathbb{R}^d \rightarrow \mathbb{R}$ tales que el límite existe en \mathbf{x} ; si el límite existe para todo $\mathbf{x} \in \mathbb{R}^d$, lo denotamos por $\mathcal{D}_\mathcal{A}$.

Observación 2.8. En consecuencia, por (2.14), un movimiento browniano $\mathbf{B} = (\mathbf{B}_t)_{t \geq 0}$ es un proceso de difusión con generador $\mathcal{A} = A = \frac{1}{2}\Delta$.

Casos más generales de los movimientos brownianos son los que toman valores en una superficie de dimensión d ; en particular, nos interesa el movimiento browniano en la esfera \mathcal{S}^{d-1} . Como hemos mencionado en la Sección 1.2 que podemos considerar a la esfera $\mathcal{S}^{d-1} \subset \mathbb{R}^d$ como una variedad riemanniana junto con la métrica riemanniana inducida del espacio euclidiano \mathbb{R}^d . Y también se puede definir el operador de Laplace de manera similar sobre ella, llamado el operador de Laplace-Beltrami, que lo denotaremos por Δ_S . Luego, un movimiento browniano $\mathbf{B} = (\mathbf{B}_t)_{t \geq 0}$ en la esfera \mathcal{S}^{d-1} es solución de un problema de martingala con generador $\frac{1}{2}\Delta_S$. Se verá la definición precisa en la Sección 2.3.

2.2. La parte radial del movimiento browniano

Un movimiento browniano en \mathbb{R}^d tiene descomposición en coordenadas polares con su parte radial y la parte angular, conocida como la representación *skew-product* (cf. [IM65], [PY81], [RW00]), y lo veremos con detalle en la siguiente sección. Mientras en esta sección, veremos que la parte radial de un movimiento browniano es un proceso de Bessel con la propiedad de Markov, y satisface una ecuación diferencial estocástica.

Definición 2.13. Sea $\mathbf{B} = (\mathbf{B}_t)_{t \geq 0}$ un (\mathcal{F}_t) -movimiento browniano en \mathbb{R}^d , el proceso estocástico $R = (R_t)_{t \geq 0}$ definido por

$$R_t = \|\mathbf{B}_t\| = \left[(B_t^1)^2 + \cdots + (B_t^d)^2 \right]^{1/2}, \quad t \geq 0$$

es no negativo y continuo con valores en $[0, \infty)$, se denomina la *norma del movimiento browniano* o el *proceso de Bessel de dimensión d* y se denota por $\text{BES}(d)$.

Para verificar la propiedad de Markov de un proceso de Bessel $\text{BES}(d)$, se tiene el siguiente resultado ([RW94], Chap. 1).

Lema 2.14. Sean (E, \mathcal{E}) y (E', \mathcal{E}') espacios medibles, y supongamos que $\Phi: (E, \mathcal{E}) \rightarrow (E', \mathcal{E}')$ es una función medible sobreyectiva, $(P_t)_{t \geq 0}$ es un semigrupo de transición de Markov en (E, \mathcal{E}) y $(Q_t)_{t \geq 0}$ es una colección de kernels de probabilidad⁵ en el espacio (E', \mathcal{E}') de manera que, para toda $f \in \mathcal{C}_a(E')$,

$$P_t(f \circ \Phi) = (Q_t f) \circ \Phi, \quad t \geq 0. \quad (2.16)$$

Entonces $(Q_t)_{t \geq 0}$ es un semigrupo de transición de Markov. Además, si \mathbf{X} es un proceso de Markov con semigrupo de transición $(P_t)_{t \geq 0}$, entonces $\Phi(\mathbf{X})$ es un proceso de Markov con semigrupo de transición $(Q_t)_{t \geq 0}$.

Demostración. Por (2.16) y usando la propiedad de semigrupo de $(P_t)_{t \geq 0}$, se tiene que

$$(Q_{t+s} f) \circ \Phi = P_{t+s}(f \circ \Phi) = P_t P_s(f \circ \Phi) = P_t((Q_s f) \circ \Phi) = (Q_t Q_s f) \circ \Phi, \quad s, t > 0$$

como Φ es sobreyectiva, entonces $Q_{t+s} f = Q_t Q_s f$; es decir, $(Q_t)_{t \geq 0}$ es un semigrupo de transición. Para demostrar que $\Phi(\mathbf{X})$ es un proceso de Markov, sea \mathbf{X} un proceso con distribución inicial δ_x , para algún $x \in E$, tomamos $0 = t_0 < t_1 < \cdots < t_k$, y para cada $i = 1, \dots, k$, sean $s_i \equiv t_i - t_{i-1}$, $f_i \in \mathcal{C}_a(E')$, por resultado de la Prop. 1.4, Chap. III, [RY99], se tiene que

$$\begin{aligned} \mathbb{E}_x \left[\prod_{i=1}^k f_i(\Phi(\mathbf{X}_{t_i})) \right] &= \left(P_{s_1}(f_1 \circ \Phi) P_{s_2}(f_2 \circ \Phi) \cdots P_{s_k}(f_k \circ \Phi) \right)(x) \\ &\stackrel{(2.16)}{=} \left((Q_{s_1} f_1)(Q_{s_2} f_2) \cdots (Q_{s_k} f_k) \right)(\Phi(x)). \end{aligned}$$

Entonces, se concluye que $\Phi(\mathbf{X})$ es un proceso de Markov. \square

Así, aplicando el Lema 2.14 para el movimiento browniano y la función $\Phi(\mathbf{x}) = \|\mathbf{x}\|$, se tiene que el proceso de Bessel en la Definición 2.13 es un proceso de Markov.

⁵ Un kernel de probabilidad Q en (E', \mathcal{E}') es tal que $Q: E' \times \mathcal{E}' \rightarrow [0, 1]$, donde $Q(x, \cdot)$ es una medida de probabilidad para cada $x \in E'$, y $Q(\cdot, A)$ es medible para cada $A \in \mathcal{E}'$.

Proposición 2.15. Sea $\mathbf{B} = (\mathbf{B}_t)_{t \geq 0}$ un (\mathcal{F}_t) -movimiento browniano en \mathbb{R}^d con distribución inicial δ_x , para algún $\mathbf{x} \in \mathbb{R}^d$. La parte radial del movimiento browniano $R_t = \|\mathbf{B}_t\|$, $t \geq 0$, es un proceso de Markov; con semigrupo de transición $(Q_t)_{t \geq 0}$ dada por las densidades de transición

$$q_t(x, y) = \frac{x}{t} \left(\frac{y}{x}\right)^{d/2} \exp\left\{-\frac{x^2 + y^2}{2t}\right\} I_{d/2-1}\left(\frac{xy}{t}\right), \quad t > 0, x, y > 0, \quad (2.17)$$

donde $I_\nu(z)$ es la función de Bessel modificada de orden ν dada por

$$I_\nu(z) = \sum_{k=0}^{\infty} \frac{(z/2)^{2k+\nu}}{k! \Gamma(k + \nu + 1)}, \quad \nu \geq -1, z > 0. \quad (2.18)$$

Demostración. Aplicando el Lema 2.14, tenemos que el semigrupo de transición $(P_t)_{t \geq 0}$ es el semigrupo de transición del movimiento browniano \mathbf{B} con densidad de transición dada por (2.1) y (2.5), y la transformación $\Phi: \mathbb{R}^d \ni \mathbf{x} \mapsto \|\mathbf{x}\| \in [0, \infty)$ es medible y sobreyectiva.

Cuando $d = 1$, se tiene que $\Phi(\cdot) = |\cdot|$; si definimos $Q_t f(x) := \int_{\mathbb{R}^d} q_t(x, y) f(y) dy$, para todo $f \in \mathcal{C}_a(\mathbb{R}^+)$, y como $I_{-\frac{1}{2}}(z) = \sqrt{2}(\pi z)^{-\frac{1}{2}} \cosh z$ ([Wat44]), entonces $q_t(x, y)$ tiene forma

$$q_t(x, y) = (2\pi t)^{-\frac{1}{2}} \exp\left\{-\frac{x^2 + y^2}{2t}\right\} 2 \cosh\left(\frac{xy}{t}\right) = p_t(x, y) + p_t(x, -y),$$

donde $p_t(x, y)$ es la densidad de transición para el movimiento browniano escalar. Por lo tanto $P_t(f \circ \Phi) = (Q_t f) \circ \Phi$.

Por otro lado, cuando $d \geq 2$, para $f \in \mathcal{C}_a(\mathbb{R}^+)$, tenemos

$$P_t(f \circ \Phi)(\mathbf{x}) = \int_{\mathbb{R}^d} p_t(\mathbf{x}, \mathbf{y}) f(\Phi(\mathbf{y})) d\mathbf{y} = \int_{\mathbb{R}^d} (2\pi t)^{-d/2} \exp\left\{-\frac{\|\mathbf{y} - \mathbf{x}\|^2}{2t}\right\} f(\|\mathbf{y}\|) d\mathbf{y}.$$

Usando coordenadas polares y por (1.5), sean $\|\mathbf{x}\| = x$ y $\|\mathbf{y}\| = y$, entonces

$$\begin{aligned} P_t(f \circ \Phi)(\mathbf{x}) &\stackrel{(*)}{=} \int_0^\infty (2\pi t)^{-d/2} \int_{\mathcal{S}^{d-1}} \exp\left\{-\frac{\|y\boldsymbol{\xi} - x\mathbf{e}_1\|^2}{2t}\right\} d\sigma(\boldsymbol{\xi}) y^{d-1} f(y) dy \\ &\stackrel{(**)}{=} \int_0^\infty (2\pi t)^{-d/2} \int_{\mathcal{S}^{d-1}} \exp\left\{-\frac{x^2 + y^2 - 2xy \cos(\phi_1)}{2t}\right\} d\sigma(\boldsymbol{\xi}) y^{d-1} f(y) dy \\ &= \int_0^\infty (2\pi t)^{-d/2} e^{-\frac{x^2 + y^2}{2t}} y^{d-1} f(y) \int_{\mathcal{S}^{d-1}} \exp\left\{\frac{xy \cos(\phi_1)}{t}\right\} d\sigma(\boldsymbol{\xi}) dy. \end{aligned} \quad (2.19)$$

Al desarrollar la integral sobre la esfera \mathcal{S}^{d-1} , hacemos la transformación en coordenadas esféricas por $G: (0, \pi)^{d-2} \times (0, 2\pi) \rightarrow \mathcal{S}^{d-1}$, tal que $G(\phi_1, \dots, \phi_{d-2}, \theta) = (\xi_1, \dots, \xi_d) = \boldsymbol{\xi} \in \mathcal{S}^{d-1}$, donde el jacobiano de la transformación está dada por $J = \prod_{k=1}^{d-2} \sin^k \phi_{d-1-k}$ (cf. [Fol99], Sec. 2.7), sea $z = \frac{xy}{t}$, entonces

$$\begin{aligned} \int_{\mathcal{S}^{d-1}} e^{\frac{xy \cos(\phi_1)}{t}} d\sigma(\boldsymbol{\xi}) &= \int_0^{2\pi} \underbrace{\left[\prod_{k=1}^{d-3} \int_0^\pi \sin^k \phi_{d-1-k} d\phi_{d-1-k} \right]}_{\text{integral sobre la esfera } \mathcal{S}^{d-2}} d\theta \int_0^\pi e^{z \cos(\phi_1)} \sin^{d-2} \phi_1 d\phi_1 \\ &= \sigma(\mathcal{S}^{d-2}) \int_0^\pi e^{z \cos(\phi_1)} \sin^{d-2} \phi_1 d\phi_1 \\ &= \frac{2\pi^{(d-1)/2}}{\Gamma\left(\frac{d-1}{2}\right)} \int_0^\pi e^{z \cos(\phi_1)} \sin^{2\nu} \phi_1 d\phi_1. \quad \left(\nu = \frac{d}{2} - 1\right) \end{aligned} \quad (2.20)$$

* $\boldsymbol{\xi} \in \mathcal{S}^{d-1}$, $\mathbf{e}_1 = (1, 0, \dots, 0) \in \mathbb{R}^d$.

** ϕ_1 es el ángulo entre $\boldsymbol{\xi}$ y \mathbf{e}_1 , y para $\boldsymbol{\xi}$ en coordenadas esféricas.

Como se sabe que la representación integral de la función de Bessel modificada I_ν está dada por (ver [Wat44], Sec. 3.71)

$$I_\nu(z) = \frac{(z/2)^\nu}{\Gamma(\nu + \frac{1}{2}) \Gamma(\frac{1}{2})} \int_0^\pi e^{\pm z \cos \theta} \sin^{2\nu} \theta d\theta, \quad (2.21)$$

donde la parte real de ν , $\Re(\nu) > -\frac{1}{2}$. Así, sustituyendo (2.21) en (2.20), se obtiene

$$\begin{aligned} \int_{\mathcal{S}^{d-1}} e^{\frac{xy \cos(\phi_1)}{t}} d\sigma(\boldsymbol{\xi}) &= \frac{2\pi^{(d-1)/2}}{\Gamma(\frac{d-1}{2})} \Gamma\left(\nu + \frac{1}{2}\right) \Gamma\left(\frac{1}{2}\right) \left(\frac{2}{z}\right)^\nu I_\nu(z) \\ &= (2\pi)^{d/2} \left(\frac{t}{xy}\right)^{d/2-1} I_{d/2-1}\left(\frac{xy}{t}\right). \quad (\text{usando } \Gamma(\frac{1}{2}) = \sqrt{\pi}) \end{aligned}$$

Y sustituyendo esta última igualdad en (2.19), obtenemos

$$\begin{aligned} P_t(f \circ \Phi)(\mathbf{x}) &= \int_0^\infty (2\pi t)^{-d/2} e^{-\frac{x^2+y^2}{2t}} y^{d-1} f(y) \left[(2\pi)^{d/2} \left(\frac{t}{xy}\right)^{d/2-1} I_{d/2-1}\left(\frac{xy}{t}\right) \right] dy \\ &= \int_0^\infty \frac{x}{t} \left(\frac{y}{x}\right)^{d/2} \exp\left\{-\frac{x^2+y^2}{2t}\right\} I_{d/2-1}\left(\frac{xy}{t}\right) f(y) dy := Q_t f(x). \end{aligned}$$

Entonces, $P_t(f \circ \Phi) = (Q_t f) \circ \Phi$, y así por el Lema 2.14, concluimos que $R_t = \|\mathbf{B}_t\| = \Phi(\mathbf{B}_t)$, $t \geq 0$ es un proceso de Markov con semigrupo de transición $(Q_t)_{t \geq 0}$, y las densidades de transición dada por (2.17). \square

En lo siguiente, veremos que el proceso de Bessel es solución a una ecuación diferencial estocástica (lo abreviaremos por EDE), primero se tiene las siguientes definiciones.

Definición 2.16. Dado $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ un espacio de probabilidad completo filtrado, definimos a $\mathcal{N}^2(\Omega, \mathbb{R}^k)$ como la clase de funciones $f(t, \omega): [0, \infty) \times \Omega \rightarrow \mathbb{R}^k$ tales que

- (1) $f: (t, \omega) \mapsto f(t, \omega)$ es $\mathcal{P}/\mathcal{B}(\mathbb{R}^k)$ -medible; donde \mathcal{P} es la mínima σ -álgebra generada por los subconjuntos en $[0, \infty) \times \Omega$ de forma: $(s, t] \times F$, $F \in \mathcal{F}_s$ y $\{0\} \times F_0$ con $F_0 \in \mathcal{F}_0$.
- (2) para cada $t \geq 0$, $f(t, \omega)$ es (\mathcal{F}_t) -adaptado.
- (3) para todo $t \geq 0$, $\int_0^t f(t, \omega)^2 dt < \infty$ c.s..

Observación 2.9. Sea $\mathbf{X} = (\mathbf{X}_t)_{t \geq 0}$ un proceso estocástico continuo de dimensión d y definido en el espacio $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ tal que podemos considerarlo como una función $\mathbf{X}_t(\omega) = \mathbf{X}(t, \omega): [0, \infty) \times \Omega \rightarrow \mathbb{R}^d$. Si \mathbf{X} cumple la condición (1) de la definición anterior, se dice que \mathbf{X} es un proceso (\mathcal{F}_t) -predecible. Y retomando la notación \mathbb{W}^d (Sec. 2.1) para el espacio de funciones continuas de $[0, \infty)$ en \mathbb{R}^d , se tiene que $\mathbf{X} \in \mathbb{W}^d$.

En la definición anterior, si identificamos a Ω como el espacio \mathbb{W}^d , sea $\mathcal{P}_t = \sigma(\mathbf{w}(s) : 0 \leq s \leq t)$, donde $\mathbf{w} \in \mathbb{W}^d$. Entonces $f: [0, \infty) \times \mathbb{W}^d \rightarrow \mathbb{R}$ que satisface (1) si es predecible con respecto a la filtración $(\mathcal{P}_t)_{t \geq 0}$ como un proceso definido en \mathbb{W}^d .

Definición 2.17. Sean $\sigma = (\sigma_{ij}): [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times n}$, $b = (b_i): [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ funciones tales que para cada $i = 1, \dots, d$, $j = 1, \dots, n$, σ_{ij} es de la clase $\mathcal{N}^2(\mathbb{W}^d, \mathbb{R})$, b_i es predecible como un proceso con valores en \mathbb{R} tal que $\int_0^t |b_i(s, \mathbf{x})| ds < \infty$ c.s. para todo $t \geq 0$. Una *solución a la ecuación diferencial estocástica* EDE(σ, b) es la pareja (\mathbf{X}, \mathbf{B}) de procesos definidos en un espacio de probabilidad filtrado $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$, donde \mathbf{B} es un (\mathcal{F}_t) -movimiento browniano estándar de dimensión n , y \mathbf{X} es un proceso continuo y (\mathcal{F}_t) -adaptado tal que para cada $i = 1, \dots, d$,

$$X_t^i = X_0^i + \sum_{j=1}^n \int_0^t \sigma_{ij}(s, \mathbf{X}_s) dB_s^j + \int_0^t b_i(s, \mathbf{X}_s) ds, \quad (2.22)$$

donde $\mathbf{X}_0 \in \mathcal{F}_0$, el valor inicial; la primera integral es la integral de Itô y la segunda es la integral de Riemann-Stieltjes (más detalle sobre la integral de Itô, ver e.g. [Ok03], [RY99]). En forma vectorial y diferencial, se tiene que

$$d\mathbf{X}_t = \sigma(t, \mathbf{X}_t) d\mathbf{B}_t + b(t, \mathbf{X}_t) dt. \quad (2.23)$$

El proceso \mathbf{X} que satisface la EDE (2.23) se denomina un *proceso de Itô* de dimensión d .

Definición 2.18. (1) Una solución (\mathbf{X}, \mathbf{B}) a una ecuación EDE(σ, b) se dice *solución fuerte* si dado un espacio de probabilidad filtrado $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ y \mathbf{B} , un (\mathcal{F}_t) -movimiento browniano definido en él de manera que \mathbf{X} es adaptado a $(\overline{\mathcal{F}}_t^{\mathbf{B}})_{t \geq 0}$, la filtración natural y completa generada por \mathbf{B} , que fue completada con todos los conjuntos de medida cero bajo \mathbb{P} . Por el contrario, una solución que no es fuerte se dice una *solución débil*.

(2) La *unicidad por trayectorias o fuerte* de soluciones a la ecuación (2.22) se cumple si para cualesquiera dos soluciones (\mathbf{X}, \mathbf{B}) y $(\mathbf{X}', \mathbf{B}')$ definidas en el mismo espacio de probabilidad filtrado con $\mathbf{X}_0 = \mathbf{X}'_0$ y $\mathbf{B} = \mathbf{B}'$ c.s., entonces implica que \mathbf{X} y \mathbf{X}' son indistinguibles, i.e., para todo $t \geq 0$, para casi toda $\omega \in \Omega$, $\mathbf{X}_t(\omega) = \mathbf{X}'_t(\omega)$.

(3) La *unicidad en ley o débil* para (2.22) es tal que para cualesquiera dos soluciones débiles (\mathbf{X}, \mathbf{B}) y $(\mathbf{X}', \mathbf{B}')$, se tiene que \mathbf{X} y \mathbf{X}' son iguales en distribución.

Observación 2.10. La existencia y unicidad de soluciones a una EDE(σ, b) requiere condiciones particulares sobre los coeficientes σ y b . En particular, los teoremas clásicos para ecuaciones diferenciales estocásticas piden que los coeficientes $\sigma(t, \mathbf{x})$ y $b(t, \mathbf{x})$ definidos en $\mathbb{R}^+ \times \mathbb{R}^d$ son funciones medibles y Lipschitz continuas en la variable $\mathbf{x} \in \mathbb{R}^d$, y σ es localmente acotado para afirmar la existencia y unicidad fuerte de solución a la EDE(σ, b) (ver [RY99], [Ok03]). En particular, tiene la siguiente definición de la *difusión de Itô*:

Definición 2.19. Si las funciones σ y b en la hipótesis de la definición 2.17 son tales que $\sigma(t, \mathbf{X}_t) \equiv \sigma(\mathbf{X}_t)$ y $b(t, \mathbf{X}_t) \equiv b(\mathbf{X}_t)$ definidos en \mathbb{R}^d , son predecibles y *Lipschitz continuas*; es decir, $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^d$, existe algún $K < \infty$ tal que

$$\|b(\mathbf{x}) - b(\mathbf{y})\|^2 + \|\sigma(\mathbf{x}) - \sigma(\mathbf{y})\|^2 \leq K \|\mathbf{x} - \mathbf{y}\|^2, \quad (2.24)$$

donde $\|\cdot\|$ es la norma euclidiana, tal que para $\alpha \in \mathbb{R}^{d \times n}$, se define $\|\alpha\| = \sqrt{\sum_{i=1}^d \sum_{j=1}^n |\alpha_{ij}|^2}$.

Dado \mathbf{B} un (\mathcal{F}_t) -movimiento browniano de dimensión n definido en el espacio de probabilidad filtrado $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$, el proceso estocástico \mathbf{X} continuo y (\mathcal{F}_t) -adaptado que satisface la ecuación

$$d\mathbf{X}_t = \sigma(\mathbf{X}_t) d\mathbf{B}_t + b(\mathbf{X}_t) dt, \quad \forall t \geq 0; \quad \mathbf{X}_0 = \mathbf{x} \in \mathbb{R}^d \quad (2.25)$$

se llama una *difusión de Itô* (en tiempo-homogéneo) con *coeficiente de difusión* σ o $a = \sigma \sigma^\top$ y *coeficiente de deriva* b .

Proposición 2.20. Sea \mathbf{X} una difusión de Itô que satisface la EDE (2.25). Si $f \in \mathcal{C}_0^2(\mathbb{R}^d)$, entonces $f \in \mathcal{D}_A$ en el sentido de la Definición 2.12 y

$$\mathcal{A}f(\mathbf{x}) = \frac{1}{2} \sum_{i,j=1}^d (\sigma \sigma^\top)_{ij}(\mathbf{x}) \frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_{i=1}^d b_i(\mathbf{x}) \frac{\partial f}{\partial x_i}. \quad (2.26)$$

Este resultado es la aplicación directa de la siguiente famosa fórmula de Itô.

Teorema 2.21 (Fórmula de Itô, [Ok03], Thm. 4.2.1). Sean \mathbf{X} una solución a la ecuación (2.23) de dimensión d y $g(t, \mathbf{x}) : [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}^p$ una función de clase $\mathcal{C}^{1 \times 2}([0, \infty) \times \mathbb{R}^d)$, entonces el proceso $\mathbf{Y}_t(\omega) = g(t, \mathbf{X}_t) = (g_1(t, \mathbf{X}_t), \dots, g_p(t, \mathbf{X}_t))$ es solución de una ecuación diferencial estocástica dada por

$$\mathbf{Y}_t^k = \mathbf{Y}_0^k + \int_0^t \frac{\partial g_k}{\partial t}(s, \mathbf{X}_s) ds + \sum_{i=1}^d \int_0^t \frac{\partial g_k}{\partial x_i}(s, \mathbf{X}_s) dX_s^i + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\partial^2 g_k}{\partial x_i \partial x_j}(s, \mathbf{X}_s) dX_s^i dX_s^j, \quad (2.27)$$

para cada $k = 1, \dots, p$; donde $dB_t^i dB_t^j = \delta_{ij} dt$, $dB_t^i dt = dt dB_t^i = 0$.

Demostración de la Proposición 2.20. Sea $Z_t = f(\mathbf{X}_t)$, $t \geq 0$ y aplicando la fórmula de Itô, se obtiene

$$\begin{aligned} dZ_t &= \sum_{i=1}^d \frac{\partial f}{\partial x_i}(\mathbf{X}_t) dX_t^i + \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{X}_t) dX_t^i dX_t^j \\ &= \sum_{i=1}^d \frac{\partial f}{\partial x_i}(\mathbf{X}_t) \left[b_i(\mathbf{X}_t) dt + (\sigma(\mathbf{X}_t) d\mathbf{B}_t)_i \right] + \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{X}_t) \left[(\sigma(\mathbf{X}_t) d\mathbf{B}_t)_i (\sigma(\mathbf{X}_t) d\mathbf{B}_t)_j \right]. \end{aligned}$$

Por simplificación, denotamos por $\sigma_{ik} \equiv \sigma_{ik}(\mathbf{X}_t)$, y $b_i \equiv b_i(\mathbf{X}_t)$. Como

$$\begin{aligned} (\sigma d\mathbf{B}_t)_i \cdot (\sigma d\mathbf{B}_t)_j &= \left(\sum_{k=1}^n \sigma_{ik} dB_t^k \right) \left(\sum_{l=1}^n \sigma_{jl} dB_t^l \right) = \sum_{k,l=1}^n \sigma_{ik} \sigma_{jl} \delta_{kl} dt \\ &= \left(\sum_{k=1}^n \sigma_{ik} \sigma_{jk} \right) dt = (\sigma \sigma^\top)_{ij} dt, \end{aligned} \quad (2.28)$$

por lo que la EDE anterior en forma integral:

$$\begin{aligned} f(\mathbf{X}_t) &= f(\mathbf{X}_0) + \int_0^t \left[\sum_{i=1}^d b_i \frac{\partial f}{\partial x_i}(\mathbf{X}_s) + \frac{1}{2} \sum_{i,j=1}^d (\sigma \sigma^\top)_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{X}_s) \right] ds \\ &\quad + \int_0^t \sum_{i=1}^d \sum_{k=1}^n \sigma_{ik} \frac{\partial f}{\partial x_i}(\mathbf{X}_s) dB_s^k. \end{aligned}$$

Al tomar la esperanza a la ecuación anterior dada la condición inicial $\mathbf{X}_0 = \mathbf{x}$; como $\mathbb{E}_x[f(\mathbf{x})] = f(\mathbf{x})$, entonces

$$\begin{aligned} \mathbb{E}_x[f(\mathbf{X}_t)] &= f(\mathbf{x}) + \mathbb{E}_x \left[\int_0^t \left(\sum_{i=1}^d b_i \frac{\partial f}{\partial x_i}(\mathbf{X}_s) + \frac{1}{2} \sum_{i,j=1}^d (\sigma \sigma^\top)_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{X}_s) \right) ds \right] \\ &\quad + \sum_{k=1}^n \mathbb{E}_x \left[\int_0^t \sum_{i=1}^d \sigma_{ik} \frac{\partial f}{\partial x_i}(\mathbf{X}_s) dB_s^k \right]. \end{aligned} \quad (2.29)$$

Por hipótesis, σ es Lipschitz continua, entonces es acotado en conjuntos compactos; y $f \in \mathcal{C}_0^2(\mathbb{R}^d)$ implica que cada derivada parcial $\frac{\partial f}{\partial x_i}$ es una función continua y acotada que tiene valor 0 en el infinito.

Así, si g es una función Borel medible y acotada tal que $|g| \leq M$ para algún $M > 0$, y B_t es un (\mathcal{F}_t) -movimiento browniano escalar; para cada $N \in \mathbb{N}$, sea $t \wedge N := \min\{t, N\}$, tenemos que

$$\mathbb{E}_x \left[\int_0^{t \wedge N} g(\mathbf{X}_s) dB_s^k \right] = \mathbb{E}_x \left[\int_0^N \mathbf{1}_{\{\omega: s < t\}}(s, \omega) g(\mathbf{X}_s) dB_s^k \right] = 0,$$

ya que $g(\mathbf{X}_s)$ y $\mathbf{1}_{\{\omega:s < t\}}(s, \omega)$ son \mathcal{F}_s -medibles; y B_s es una \mathcal{F}_s -martingala. Más aún,

$$\begin{aligned} \mathbb{E}_x \left[\left(\int_0^t g(\mathbf{X}_s) dB_s - \int_0^{t \wedge N} g(\mathbf{X}_s) dB_s \right)^2 \right] &= \mathbb{E}_x \left[\int_{t \wedge N}^t g^2(\mathbf{X}_s) ds \right] \\ &\leq M^2 \mathbb{E}_x [t - (t \wedge N)] \rightarrow 0, \quad \text{cuando } N \rightarrow \infty. \end{aligned}$$

Así, por el teorema de convergencia dominada,

$$\mathbb{E}_x \left[\int_0^t g(\mathbf{X}_s) dB_s \right] = \lim_{N \rightarrow \infty} \mathbb{E}_x \left[\int_0^{t \wedge N} g(\mathbf{X}_s) dB_s \right] = 0.$$

En consecuencia, podemos cancelar el último término de (2.29), obteniendo

$$\lim_{t \downarrow 0} \frac{\mathbb{E}_x[f(\mathbf{X}_t)] - f(\mathbf{x})}{t} = \lim_{t \downarrow 0} \frac{1}{t} \mathbb{E}_x \left[\int_0^t \mathcal{A}f(\mathbf{X}_s) ds \right],$$

donde $\mathcal{A}f$ está dada por (2.26). Aplicando el teorema de Fubini al lado derecho de la igualdad, y notando que \mathbf{X}_t tiene trayectorias continuas *c.s.*, y $\mathbf{x} \mapsto \mathcal{A}f(\mathbf{x})$ es continua, entonces el teorema de convergencia dominada implica la continuidad de $s \mapsto \mathbb{E}_x[\mathcal{A}f(\mathbf{X}_s)]$, $s \geq 0$, por el teorema fundamental de cálculo, se obtiene

$$\lim_{t \downarrow 0} \frac{\mathbb{E}_x[f(\mathbf{X}_t)] - f(\mathbf{x})}{t} = \lim_{t \downarrow 0} \frac{1}{t} \int_0^t \mathbb{E}_x[\mathcal{A}f(\mathbf{X}_s)] ds = \mathbb{E}_x[\mathcal{A}f(\mathbf{X}_0)] = \mathcal{A}f(\mathbf{x}).$$

Esto es, si $f \in \mathcal{C}_0^2(\mathbb{R}^d)$, entonces $f \in \mathcal{D}_A$ y el generador de \mathbf{X}_t está dado por (2.26). \square

Ejemplo 2.22. Sea \mathbf{B} un movimiento browniano definido en $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ y con valor inicial $\mathbf{B}_0 \in \mathbb{R}^d$, notemos que \mathbf{B} es una solución fuerte a la EDE(σ, b), donde $\sigma \equiv I_d$, la matriz identidad; y $b \equiv \mathbf{0}$, el vector cero. Aplicando la fórmula de Itô al proceso $R_t = \|\mathbf{B}_t\|$, $t \geq 0$, considerando a la función $g(t, \mathbf{x}) = \|\mathbf{x}\|$, la cual es de clase \mathcal{C}^2 en $\mathbb{R}^d \setminus \{\mathbf{0}\}$.

Como el punto origen es *polar* para el movimiento browniano en \mathbf{B} en \mathbb{R}^d ($d \geq 2$); es decir, la probabilidad de que \mathbf{B}_t regrese al origen en tiempos positivos es cero ([RY99], Chap. V, Prop. 2.7). Por lo que, $R_t > 0$ *c.s.* para $t > 0$ si $d \geq 2$; entonces podemos aplicar la fórmula de Itô para la función g en el caso $d \geq 2$, donde $\frac{\partial g}{\partial t}(t, \mathbf{x}) = 0$, $\frac{\partial g}{\partial x_i}(t, \mathbf{x}) = \frac{x_i}{\|\mathbf{x}\|}$ y $\frac{\partial^2 g}{\partial x_i^2}(t, \mathbf{x}) = \frac{1}{\|\mathbf{x}\|} - \frac{x_i^2}{\|\mathbf{x}\|^3}$, para todo $i = 1, \dots, d$, y se obtiene

$$\begin{aligned} R_t = g(t, \mathbf{B}_t) &= \|\mathbf{B}_0\| + \sum_{i=1}^d \int_0^t \frac{B_s^i}{R_s} dB_s^i + \frac{1}{2} \sum_{i=1}^d \int_0^t \left[\frac{1}{R_s} - \frac{(B_s^i)^2}{R_s^3} \right] ds \\ &= R_0 + \sum_{i=1}^d \int_0^t \frac{B_s^i}{R_s} dB_s^i + \frac{1}{2} \int_0^t \frac{d-1}{R_s} ds. \end{aligned}$$

Sea $\tilde{B}_t = \sum_{i=1}^d \int_0^t \frac{B_s^i}{R_s} dB_s^i$, y calculamos su variación cuadrática

$$\langle \tilde{B}, \tilde{B} \rangle_t = \sum_{i,j=1}^d \int_0^t \frac{B_s^i B_s^j \delta_{ij}}{R_s} ds = \sum_{i=1}^d \int_0^t \frac{(B_s^i)^2}{R_s} ds = t,$$

entonces por el teorema de caracterización de Lévy ([RY99], Chap. IV, Thm. 3.6), \tilde{B}_t es un (\mathcal{F}_t) -movimiento browniano escalar. Por lo tanto, el proceso de Bessel de dimensión d ($d \geq 2$) satisface la siguiente EDE,

$$R_t = R_0 + \tilde{B}_t + \int_0^t \frac{d-1}{2R_s} ds, \quad R_t > 0 \text{ c.s.} \quad (2.30)$$

Así también, el proceso $R = (R_t)_{t \geq 0}$ es un proceso de difusión con generador

$$\mathcal{A}f(r) = \frac{1}{2} \frac{d^2}{dr^2} f(r) + \frac{d-1}{2r} \frac{d}{dr} f(r), \quad (2.31)$$

Por otro lado, el otro punto de vista del proceso de Bessel es mediante el cuadrado del proceso de Bessel dado por la siguiente definición.

Definición 2.23. Para cada $\delta \geq 0$ y $Z_0 = z \geq 0$, la única solución fuerte a la ecuación

$$Z_t = Z_0 + 2 \int_0^t \sqrt{Z_s} dB_s + \delta t, \quad Z_t \geq 0 \text{ c.s.}, \quad \forall t \geq 0 \quad (2.32)$$

se le llama el *proceso cuadrado de Bessel de dimensión δ* que comienza en z , y se denota por $Z \sim \text{BES}_z^2(\delta)$.

Observación 2.11. Observemos que el coeficiente de difusión en la ecuación (2.32), $\sigma(x) = 2\sqrt{x}$ no es una función de Lipschitz en \mathbb{R}^+ , aunque el coeficiente $b(x) \equiv \delta$ sí lo es. Por lo que los teoremas clásicos para la existencia y unicidad de soluciones fuertes no se aplican (Observación 2.10). Sin embargo, $b(x)$ y $\sigma(x)$ son funciones continuas, el teorema de existencia de Skorokhod para EDE afirma la existencia de soluciones débiles a la EDE (2.32) ([IW89], Thm. 2.3, p.173 & [Dur96], Sec. 8.4). Más aún, como $|\sqrt{x} - \sqrt{y}| \leq \sqrt{|x-y|}$, $\forall x, y \geq 0$, entonces

$$2 \left| \sqrt{|x|} - \sqrt{|y|} \right| \leq 2\sqrt{||x| - |y||} \leq 2\sqrt{|x-y|} = \rho(|x-y|), \quad \forall x, y \in \mathbb{R};$$

con $\rho(u) = 2\sqrt{u}$, claramente se tiene que $\int_0^\epsilon \rho(u)^{-2} du = \int_0^\epsilon \frac{1}{4u} du = \infty$. Así, los coeficientes $b(x) = \delta$ y $\sigma(x) = 2\sqrt{x}$ satisfacen las condiciones del siguiente criterio clásico de Yamada-Watanabe para la unicidad por trayectorias de una EDE uni-dimensional ([RW00], Chap. V, Thm. 40.1):

Proposición 2.24 (Yamada-Watanabe para la unicidad de trayectorias). *Supongamos que los coeficientes σ y b de una EDE (2.25) en \mathbb{R}^1 son medibles y satisfacen las condiciones:*

- (i) b es Lipschitz, i.e., $|b(x) - b(y)| \leq C|x-y|$, para algún constante $C > 0$, $\forall x, y \in \mathbb{R}$,
- (ii) para todo $x, y \in \mathbb{R}$, $|\sigma(x) - \sigma(y)| \leq \rho(|x-y|)$,

donde $\rho: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ es una función estrictamente creciente con $\rho(0) = 0$ y $\int_0^\epsilon \rho^{-2}(u) du = \infty$ para todo $\epsilon > 0$. Entonces se cumple la unicidad por trayectorias para la EDE (2.25).

Por lo que existe una única solución fuerte a la ecuación (2.32); entonces está bien definido el proceso $\text{BES}^2(\delta)$.

Observación 2.12. Notemos que si $\delta \equiv d \geq 1$ es un entero y $\mathbf{B} = (B^1, \dots, B^d)$ es un movimiento browniano de dimensión \mathbb{R}^d , entonces con la aplicación de la fórmula de Itô, el proceso $R_t^2 = \|\mathbf{B}_t\|^2 = (B_t^1)^2 + \dots + (B_t^d)^2$ satisface la EDE (2.32), es decir, el proceso cuadrado de Bessel de dimensión d (ver [RY99], Chap. XI).

Si denotamos por Q_z^δ a la distribución de $Z \sim \text{BES}_z^2(\delta)$ en el conjunto de funciones continuas en \mathbb{R}^+ , se tiene que las distribuciones de proceso cuadrado de Bessel tiene la siguiente propiedad aditiva ([RY99], Chap.XI, Thm. 1.2).

Teorema 2.25. *Para cualesquiera $\delta, \delta' \geq 0$ y $z, z' \geq 0$,*

$$Q_z^\delta * Q_{z'}^{\delta'} = Q_{z+z'}^{\delta+\delta'}.$$

Esto implica que Z es un proceso de Markov y se puede obtener su función de densidad de transición.

Corolario 2.26. *Sea $\delta > 0$, la densidad de transición para un proceso cuadrado de Bessel de dimensión δ está dada por*

$$q_t^\delta(x, y) = \frac{1}{2t} \left(\frac{y}{x}\right)^{\frac{\nu}{2}} e^{-\frac{x+y}{2t}} I_\nu\left(\frac{\sqrt{xy}}{t}\right), \quad x, y > 0, \quad (2.33)$$

donde $\nu = \frac{\delta}{2} - 1$ y $I_\nu(z)$ es la función de Bessel modificada de orden ν dada por (2.18). En particular, cuando $x = 0$, se tiene que

$$q_t^\delta(0, y) = \frac{1}{2t\Gamma(\frac{\delta}{2})} \left(\frac{y}{2t}\right)^{\frac{\delta}{2}-1} e^{-\frac{y}{2t}}, \quad y > 0. \quad (2.34)$$

Demostración. Si escribimos $\mathbb{E}[f(Z)]$ por $Q_x^\delta(f(X))$ para $Z \sim \text{BES}_x^2(\delta)$, con f una funcional del proceso y X representa el proceso de coordenada de Z . Entonces tenemos la transformada de Laplace para el proceso Z , dada por

$$\mathbb{E}[e^{-\lambda Z_t}] = Q_x^\delta(e^{-\lambda X_t}) := \phi(x, \delta).$$

El teorema anterior implica que para todo $x, x', \delta, \delta' \geq 0$,

$$\phi(x + x', \delta + \delta') = \phi(x, \delta)\phi(x', \delta').$$

En particular, se tiene que $\phi(x + x', 0) = \phi(x, 0)\phi(x', 0)$, entonces dado $\phi(0, 0) = 1$, vemos que $\phi(x, 0) = \alpha^x$, para algún $\alpha > 0$. Análogamente, obtenemos que $\phi(0, \delta) = \beta^\delta$, para algún $\beta > 0$. Por lo tanto,

$$\phi(x, \delta) = \phi(x, 0)\phi(0, \delta) = \alpha^x \beta^\delta.$$

Ahora para encontrar los valores de α y β , consideramos $\delta = 1$, y sea $B = (B_t)_{t \geq 0}$ un movimiento browniano de dimensión uno con punto inicial $B_0 = \sqrt{x}$. Así, se obtiene

$$\phi(x, 1) = \mathbb{E}\left[e^{-\lambda B_t^2}\right] = (2\lambda t + 1)^{-\frac{1}{2}} \exp\left\{-\frac{\lambda x}{2\lambda t + 1}\right\}.$$

Por lo que, en general, tenemos

$$\phi(x, \delta) = (2\lambda t + 1)^{-\frac{\delta}{2}} \exp\left\{-\frac{\lambda x}{2\lambda t + 1}\right\}. \quad (2.35)$$

Notemos que si $x = 0$ y $\delta > 0$, el factor exponencial es igual a 1, y la transformada de Laplace es en realidad la transformada de una distribución Gamma con parámetros $\frac{\delta}{2}$ y $\frac{1}{2t}$, esto es, su densidad de transición está dada por (2.34).

Por otro lado, cuando $x, \delta > 0$, si consideramos a Y_1 una variable aleatoria Poisson con parámetro x , entonces la mezcla de la distribución Gamma $Y \sim \mathfrak{Ga}(\nu + Y_1 + 1, 1)$, con $\nu > -1$ tiene función de densidad dada por

$$f_x^{(\nu)}(y) = \sum_{k=0}^{\infty} \frac{x^k e^{-x}}{k!} \frac{1}{\Gamma(\nu + k + 1)} y^{\nu+k} e^{-y} = e^{-(x+y)} \left(\frac{y}{x}\right)^{\frac{\nu}{2}} \sum_{k=0}^{\infty} \frac{(xy)^{\frac{\nu}{2}+k}}{k! \Gamma(\nu + k + 1)}$$

$$= e^{-(x+y)} \left(\frac{y}{x}\right)^{\frac{\nu}{2}} I_{\nu}(2\sqrt{xy}).$$

Sabemos que la transformada de Laplace para la distribución Gamma $\mathfrak{Ga}(a, 1)$ es igual a $(\lambda + 1)^{-a}$, entonces se obtiene la transformada de Laplace para $f_x^{(\nu)}(y)$:

$$\begin{aligned} \int_0^{\infty} f_x^{(\nu)}(y) e^{-\lambda y} dy &= \int_0^{\infty} \sum_{k=0}^{\infty} \frac{x^k e^{-x}}{k!} \frac{1}{\Gamma(\nu + k + 1)} y^{\nu+k} e^{-y} e^{-\lambda y} dy \\ &= \sum_{k=0}^{\infty} \frac{x^k e^{-x}}{k!} (\lambda + 1)^{-(\nu+k+1)} = (\lambda + 1)^{-(\nu+1)} e^{-x} \sum_{k=0}^{\infty} \frac{(x/(\lambda + 1))^k}{k!} \\ &= (\lambda + 1)^{-(\nu+1)} e^{-\frac{\lambda x}{\lambda+1}}. \end{aligned}$$

Haciendo el cambio de variable, reemplazando x, y por $\frac{x}{2t}$ y $\frac{y}{2t}$ en $f_x^{(\nu)}(y)$ y así en la transformada de Laplace, se obtiene la densidad q_t^{δ} dada por (2.33) y la transformada de Laplace $\phi(x, \delta)$ dada por (2.35), respectivamente; donde $\nu = \frac{\delta}{2} - 1$.

Por lo tanto, para $x, \delta > 0$, la densidad de transición para el proceso cuadrado de Bessel está dada por (2.33). \square

Definición 2.27. Sean $\delta \geq 0$, $x \geq 0$ y $Z \sim \text{BES}_{x^2}^2(\delta)$. El proceso $R_t = \sqrt{Z_t}$ se llama el *proceso de Bessel de dimensión δ* que comienza en x y se denota por $R \sim \text{BES}_x(\delta)$.

Observación 2.13. Observemos que, para $\delta > 0$, por el cambio de variable, reemplazando x e y por \sqrt{x} y \sqrt{y} en la densidad de transición (2.33), vemos que la densidad de transición del proceso de Bessel $\text{BES}_x(\delta)$ está dada por (2.17), con $d = \delta$.

Luego, se puede obtener la existencia y unicidad de solución fuerte a la ecuación (2.30) para el proceso de Bessel por el siguiente resultado, los detalles se pueden encontrar en [Che00] y [RY99], Chap. XI.

Proposición 2.28. Sean $\delta \geq 2$ y $R_0 > 0$, el proceso de Bessel de dimensión δ es la única solución fuerte a la ecuación

$$R_t = R_0 + B_t + \int_0^t \frac{\delta - 1}{2R_s} ds, \quad R_t > 0 \text{ c.s.} \quad (2.36)$$

y se cumple con la unicidad fuerte. Sin embargo, cuando $R_0 = 0$ o $1 < \delta < 2$ se tiene la no unicidad de soluciones a la ecuación (2.36); además, si $\delta = 1$, el proceso $\text{BES}(1)$ no es solución a la ecuación (2.36).

Por consiguiente, en lo sucesivo nos restringiremos a los procesos de Bessel de dimensión mayor o igual a 2 si no se especifica la dimensión del proceso.

2.2.1. Simulación exacta del proceso de Bessel

Sea $Z = (Z_t)_{t \geq 0} \sim \text{BES}_x^2(\delta)$ un proceso cuadrado de Bessel con $x > 0$ y $\delta > 0$, y se tiene su densidad de transición está dada explícitamente en el corolario 2.26. Como vimos en la demostración de este corolario, esta densidad de transición se puede considerar como la densidad de una mezcla de distribución Poisson/Gamma:

$$Z_t \sim \mathfrak{Ga}\left(\nu + Y + 1, \frac{1}{2t}\right), \quad \text{donde } Y \sim \text{Poisson}\left(\frac{x}{2t}\right), \quad \nu = \frac{\delta}{2} - 1 \geq 0, \quad t > 0.$$

Por lo tanto, podemos simular muestras exactas de la trayectoria de un proceso cuadrado de Bessel a través de la distribución de Z_t . Esto es, considera cualquier partición de tiempo $0 = t_0 < t_1 < \dots < t_N = s$, dado $Z_0 = x > 0$, generamos una muestra $\mathbf{Z} \equiv (Z_0, Z_1, \dots, Z_N)$ en cada punto de tiempo de la partición con $Z_n \equiv Z_{t_n}, \forall n = 0, 1, \dots, N$. Así, \mathbf{Z} forma un esqueleto de la trayectoria de un proceso cuadrado de Bessel en el intervalo de tiempo $[0, s]$. En el Algoritmo 3 se presenta este procedimiento con detalle.

Algoritmo 3: Simulación exacta del proceso cuadrado de Bessel $\text{BES}_{x^2}^2(\delta)$

Init: $Z_0 = x^2 > 0, \nu = \frac{\delta}{2} - 1 \geq 0, N > 0$.

- 1 **for** $n \leftarrow 1$ **to** N **do**
- 2 $Y_n \sim \text{Poisson}\left(\frac{Z_{n-1}}{2(t_n - t_{n-1})}\right)$
- 3 $Z_n \sim \mathfrak{Ga}\left(Y_n + \nu + 1, \frac{1}{2(t_n - t_{n-1})}\right)$
- 4 **end for**
- 5 **return** (Z_0, Z_1, \dots, Z_N)

Así mismo, por la Definición 2.27 para el proceso de Bessel y utilizando el Algoritmo 3, también podemos simular muestras exactas de la trayectoria de un proceso de Bessel $R \sim \text{BES}_x(\delta)$ simplemente haciendo $R_t = \sqrt{Z_t}$, para $t \geq 0$. Esto es, el Algoritmo 4.

Algoritmo 4: Simulación exacta del proceso de Bessel $\text{BES}_x(\delta)$

Init: $Z_0 = x^2 > 0, \nu = \frac{\delta}{2} - 1 \geq 0, N > 0$.

- 1 Simular: $(Z_0, Z_1, \dots, Z_N) \sim \text{BES}_{x^2}^2(\delta)$ ▷ Alg.3
- 2 **for** $n \leftarrow 1$ **to** N **do**
- 3 Calcular: $R_n = \sqrt{Z_n}$
- 4 **end for**
- 5 **return** (R_0, R_1, \dots, R_N)

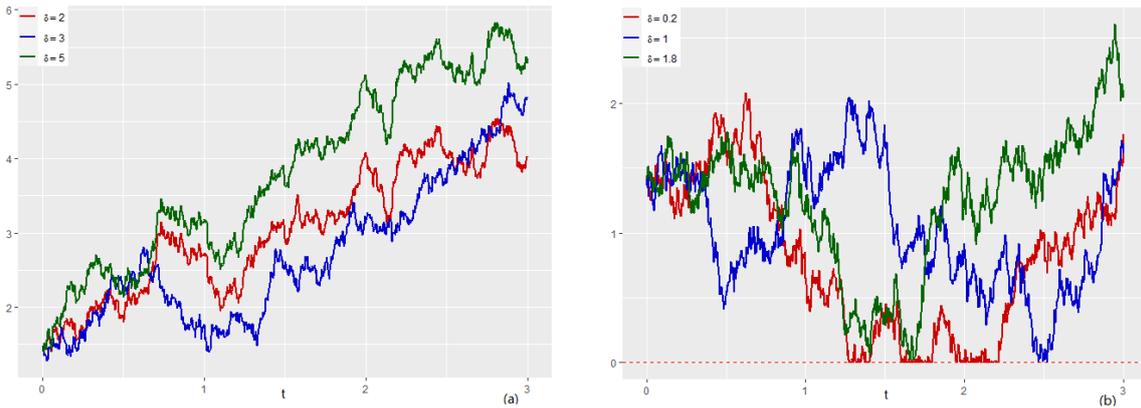


Figura 2.1: Muestras exactas de trayectoria de un proceso de Bessel $\text{BES}_x(\delta)$ simuladas usando el Algoritmo 4 para distintos valores de (a) $\delta \geq 2$ y (b) $\delta \in (0, 2)$, en el intervalo de tiempo $[0, 3]$, con punto inicial $x = \sqrt{2}$.

Al implementar el Algoritmo 4 (ver Apéndice Alg.3 y 4), podemos obtener muestras exactas de la trayectoria de un proceso de Bessel como se muestra en la Figura 2.1, en un intervalo de tiempo $[0, 3]$ para distintos valores de $\delta > 0$. Observando los comportamientos de las trayectorias de un proceso de Bessel simulado en la Figura 2.1 tanto para $\delta \in (0, 2)$ (Fig. 2.1(a)), como para $\delta \geq 2$ (Fig. 2.1(b)), podemos notar que estas muestras de trayectoria siguen los siguientes hechos sobre las trayectorias de procesos de Bessel $\text{BES}(\delta)$ (ver [RY99], Cap.XI, Pag.442):

- (1) para $0 < \delta < 2$, el punto 0 es punto de reflexión, que el proceso se mueve instantáneamente del punto; más aún, para $\delta = 0$, el punto 0 es absorbente (i.e., se queda en el

- punto 0 por siempre una vez el proceso llegue al punto);
- (2) para $\delta \leq 1$, el punto 0 es alcanzable *c.s.*;
 - (3) para $\delta \geq 2$, el punto 0 es inalcanzable;
 - (4) para $\delta \geq 3$, el proceso BES(δ) es transitorio, y es recurrente para $\delta \leq 2$.

2.3. El movimiento browniano en la esfera \mathcal{S}^{d-1}

En esta sección, estudiaremos el movimiento browniano en la esfera unitaria $\mathcal{S}^{d-1} \subset \mathbb{R}^d$, para $d \geq 2$. Veremos que el movimiento browniano en \mathcal{S}^{d-1} es un proceso de difusión que es la única solución fuerte a una ecuación diferencial estocástica llamada la *representación de Stroock para el movimiento browniano esférico* ([Hsu02]).

También, veremos que un movimiento browniano en el espacio euclidiano \mathbb{R}^d tiene una descomposición en coordenadas polares, conocida como la *representación skew-product* (en inglés), donde su parte radial es una difusión, el proceso de Bessel; y su parte angular es un movimiento browniano en la esfera \mathcal{S}^{d-1} con un cambio de escala de tiempo ([Gal63], [IM65], [PY81]). Además, presentamos una demostración en este trabajo distinta a aquellas referencias mencionadas, y usando la técnica que se utiliza en [MMUB18] para la demostración de la descomposición skew-product de un proceso como resultado de la proyección de un movimiento browniano en la esfera (ver Sec. 4.1).

Por otro lado, el movimiento browniano tiene descomposición skew-product en diferentes tipos de variedades ([PR88], [RW00], [Hsu02]). En particular, aquí consideraremos a la esfera \mathcal{S}^{d-1} como una variedad riemanniana \mathcal{M} con la métrica inducida del espacio euclidiano \mathbb{R}^d , el movimiento browniano en la esfera \mathcal{S}^{d-1} es una difusión generada por el generador $\frac{1}{2}\Delta_S$, con Δ_S el operador de Laplace-Beltrami en la esfera \mathcal{S}^{d-1} ; por su representación en coordenadas polares, se puede construir el movimiento browniano en la esfera mediante su representación skew-product ([IM65], [PR88], [Hsu02]). Hacemos su demostración con la misma técnica como se utiliza en la demostración de la representación de skew-product del movimiento browniano en \mathbb{R}^d que se da en la Sección 2.3.2.

Ejemplo 1. El movimiento browniano en el círculo unitario \mathcal{S}^1

Primero consideremos el caso $d = 2$, es decir, el movimiento browniano en el círculo unitario \mathcal{S}^1 . Sea $B = (B_t)_{t \geq 0}$ un movimiento browniano de dimensión uno con valor inicial $B_0 = x$, y considera a la función $g: \mathbb{R} \ni x \mapsto e^{ix} = (\cos(x), \sin(x)) \in \mathbb{R}^2$. Entonces aplicando la fórmula de Itô (Thm. 2.21) a $\mathbf{Y}_t = (Y_t^1, Y_t^2) = (\cos(B_t), \sin(B_t)) = g(B_t)$, se tiene que

$$\begin{cases} dY_t^1 = -\sin(B_t) dB_t - \frac{1}{2} \cos(B_t) dt \\ dY_t^2 = \cos(B_t) dB_t - \frac{1}{2} \sin(B_t) dt \end{cases} \implies \begin{cases} dY_t^1 = -Y_t^2 dB_t - \frac{1}{2} Y_t^1 dt \\ dY_t^2 = Y_t^1 dB_t - \frac{1}{2} Y_t^2 dt \end{cases}$$

Esto es, en notación matricial, \mathbf{Y}_t satisface la EDE:

$$d\mathbf{Y}_t = \sigma(\mathbf{Y}_t) dB_t + b(\mathbf{Y}_t) dt, \quad \mathbf{Y}_0 \in \mathcal{S}^1,$$

donde $\sigma(\mathbf{y}) = (-y_2, y_1)^\top$ y $b(\mathbf{y}) = -\frac{1}{2}\mathbf{y}$, es claro que son funciones acotadas con $\|\mathbf{y}\| = 1$. Entonces por la proposición 2.20, $\mathbf{Y} = (Y_t)_{t \geq 0}$ es una difusión con generador

$$\mathcal{A}f(\mathbf{y}) = \frac{1}{2} \left((y_2)^2 \frac{\partial^2}{\partial y_1^2} - 2y_1y_2 \frac{\partial^2}{\partial y_1 \partial y_2} + (y_1)^2 \frac{\partial^2}{\partial y_2^2} \right) f(\mathbf{y}) - \frac{1}{2} \sum_{i=1}^2 y_i \frac{\partial}{\partial y_i} f(\mathbf{y})$$

$$= \frac{1}{2} \sum_{i,j=1}^2 (\delta_{ij} - y_i y_j) \frac{\partial^2 f(\mathbf{y})}{\partial y_i \partial y_j} - \frac{1}{2} \sum_{i=1}^2 y_i \frac{\partial f(\mathbf{y})}{\partial y_i}, \quad \|\mathbf{y}\| = 1,$$

para $f \in \mathcal{C}_0^\infty(\mathcal{S}^1)$. El proceso \mathbf{Y} así obtenido se dice *el movimiento browniano en el círculo unitario \mathcal{S}^1* .

Ejemplo 2. El movimiento browniano en la esfera unitaria \mathcal{S}^{d-1} , $d > 2$

Sea $\mathbf{B} = (\mathbf{B}_t)_{t \geq 0}$ un movimiento browniano en \mathbb{R}^d está definido en el espacio de probabilidad completo filtrado $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ y con valor inicial $\mathbf{B}_0 = \mathbf{x}$.

Como el movimiento browniano \mathbf{B} en \mathbb{R}^d es transitorio si $d \geq 3$ ([Ok03], Ex.7.4.2), entonces consideramos a la función $g(\mathbf{x}): \mathbb{R}^d \setminus \{\mathbf{0}\} \ni \mathbf{x} \mapsto \frac{\mathbf{x}}{\|\mathbf{x}\|} \in \mathcal{S}^{d-1}$, es de clase \mathcal{C}^∞ y para cada $1 \leq k \leq d$, sea $g_k(\mathbf{x}) = \frac{x_k}{\|\mathbf{x}\|}$ y se tiene que

$$\frac{\partial g_k(\mathbf{x})}{\partial x_i} = \frac{\delta_{ik} \|\mathbf{x}\|^2 - x_i x_k}{\|\mathbf{x}\|^3} \quad \text{y} \quad \frac{\partial^2 g_k(\mathbf{x})}{\partial x_i^2} = \frac{3x_k x_i^2}{\|\mathbf{x}\|^5} - \frac{(1 + 2\delta_{ik})x_k}{\|\mathbf{x}\|^3}.$$

Así, aplicando la fórmula de Itô para $\mathbf{Y}_t = g(\mathbf{B}_t)$, y se obtiene

$$\begin{aligned} dY_t^k &= \sum_{i=1}^d \frac{\partial g_k(\mathbf{B}_t)}{\partial x_i} dB_t^i + \frac{1}{2} \sum_{i=1}^d \frac{\partial^2 g_k(\mathbf{B}_t)}{\partial x_i^2} dt \\ &= \frac{1}{\|\mathbf{B}_t\|} \sum_{i=1}^d (\delta_{ik} - Y_t^i Y_t^k) dB_t^i - \frac{1}{\|\mathbf{B}_t\|^2} \frac{d-1}{2} Y_t^k dt, \quad \forall k = 1, \dots, d \end{aligned} \quad (2.37)$$

Esto es, en notación matricial,

$$d\mathbf{Y}_t = \frac{1}{\|\mathbf{B}_t\|} \sigma(\mathbf{Y}_t) d\mathbf{B}_t + \frac{1}{\|\mathbf{B}_t\|^2} b(\mathbf{Y}_t) dt, \quad \mathbf{Y}_0 = \frac{\mathbf{x}}{\|\mathbf{x}\|} \in \mathcal{S}^{d-1}, \quad (2.38)$$

con $\sigma(\mathbf{y}) = I_d - \mathbf{y}\mathbf{y}^\top$ y $b(\mathbf{y}) = -\frac{d-1}{2}\mathbf{y}$, $\forall \mathbf{y} \in \mathcal{S}^{d-1}$. Si definimos una nueva escala de tiempo:

$$l_t = l(t, \omega) := \int_0^t \frac{ds}{\|\mathbf{B}_s\|^2},$$

notemos que $l(t, \omega) > 0$ c.s. y es (\mathcal{F}_t) -adaptado, entonces $t \mapsto l_t(\omega)$ es estrictamente creciente y su inversa $\tau_t = \inf\{s : l_s > t\}$ es continua con $\tau_0 = 0$ y de manera que $\tau(l_t, \omega) = l(\tau_t, \omega) = t$, $\forall t \geq 0$ y para cada $\omega \in \Omega$.

Entonces para la EDE (2.38) en el tiempo τ_t y en forma integral:

$$\mathbf{Y}_{\tau_t} = \mathbf{Y}_{\tau_0} + \int_0^{\tau_t} \frac{1}{\|\mathbf{B}_s\|} \sigma(\mathbf{Y}_s) d\mathbf{B}_s + \int_0^{\tau_t} \frac{1}{\|\mathbf{B}_s\|^2} b(\mathbf{Y}_s) ds. \quad (2.39)$$

Como por el teorema fundamental de cálculo, se tiene que $l'_t = \frac{1}{\|\mathbf{B}_t\|^2}$, donde l'_t denota la derivada con respecto al tiempo t ; y haciendo el cambio del tiempo $s = \tau_u$, entonces

$$l_s = u \Rightarrow l'_s ds = du \Rightarrow \frac{ds}{\|\mathbf{B}_s\|^2} = du \quad \text{y} \quad l_{\tau_u} = u \Rightarrow l'_{\tau_u} \tau'_u = 1 \Rightarrow \tau'_u = l'_{\tau_u}^{-1} = \|\mathbf{B}_{\tau_u}\|^2.$$

Por lo que:

$$\int_0^{\tau_t} \frac{1}{\|\mathbf{B}_s\|^2} b(\mathbf{Y}_s) ds = \int_0^t b(\mathbf{Y}_{\tau_u}) du,$$

y por la fórmula de cambio de tiempo para integrales de Itô ([Ok03], Thm. 8.5.7),

$$\int_0^{\tau_t} \frac{1}{\|\mathbf{B}_s\|} \sigma(\mathbf{Y}_s) d\mathbf{B}_s = \int_0^t \frac{1}{\|\mathbf{B}_{\tau_u}\|} \sigma(\mathbf{Y}_{\tau_u}) \sqrt{\tau'_u} d\tilde{\mathbf{B}}_u = \int_0^t \sigma(\mathbf{Y}_{\tau_u}) d\tilde{\mathbf{B}}_u,$$

donde $\tilde{\mathbf{B}}_t = \int_0^{\tau_t} \frac{1}{\|\mathbf{B}_s\|} d\mathbf{B}_s$ es un (\mathcal{F}_{τ_t}) -movimiento browniano en \mathbb{R}^d .

Sea $\mathbf{Z}_t = \mathbf{Y}_{\tau_t}$ para todo $t \geq 0$, entonces \mathbf{Z}_t es una (\mathcal{G}_t) -martingala por el teorema de muestreo opcional de Doob ([Kal21], Thm. 9.12), donde $\mathcal{G}_t = \mathcal{F}_{\tau_t}$, $t \geq 0$. Además, \mathbf{Z} satisface la EDE:

$$d\mathbf{Z}_t = (I - \mathbf{Z}_t \mathbf{Z}_t^\top) d\tilde{\mathbf{B}}_t - \frac{d-1}{2} \mathbf{Z}_t dt, \quad \mathbf{Z}_0 = \frac{\mathbf{x}}{\|\mathbf{x}\|} \in \mathcal{S}^{d-1}, \quad (2.40)$$

donde $I \equiv I_d$ es la matriz identidad de dimensión $d \times d$.

Notemos que las funciones $\sigma(\mathbf{z}) = I - \mathbf{z}\mathbf{z}^\top$ y $b(\mathbf{z}) = -\frac{d-1}{2}\mathbf{z}$, $\mathbf{z} \in \mathcal{S}^{d-1}$ son acotadas, y como $\mathbf{z}^\top \mathbf{z} = \|\mathbf{z}\|^2 = 1$, entonces

$$\sigma(\mathbf{z})\sigma(\mathbf{z})^\top = (I - \mathbf{z}\mathbf{z}^\top)(I - \mathbf{z}\mathbf{z}^\top)^\top = I - \mathbf{z}\mathbf{z}^\top - \mathbf{z}\mathbf{z}^\top + \mathbf{z}\mathbf{z}^\top \mathbf{z}\mathbf{z}^\top = I - \mathbf{z}\mathbf{z}^\top.$$

En consecuencia, \mathbf{Z} es un proceso de difusión con generador

$$\mathcal{A}f(\mathbf{z}) = \frac{1}{2} \left[\sum_{i,j=1}^d (\delta_{ij} - z_i z_j) \frac{\partial^2 f(\mathbf{z})}{\partial z_i \partial z_j} \right] - \frac{d-1}{2} \sum_{i=1}^d z_i \frac{\partial f(\mathbf{z})}{\partial z_i}, \quad \|\mathbf{z}\| = 1. \quad (2.41)$$

Así, el proceso $\mathbf{Y}_t = \frac{\mathbf{B}_t}{\|\mathbf{B}_t\|}$ después de un cambio de escala de tiempo conveniente, es igual a un proceso de difusión $\mathbf{Z} = (\mathbf{Z}_t)_{t \geq 0}$ que vive en la esfera unitaria \mathcal{S}^{d-1} . Como el movimiento browniano $\tilde{\mathbf{B}}$ es invariante bajo transformaciones ortogonales en \mathbb{R}^d , y para H una matriz ortogonal de dimensión $d \times d$, se tiene que

$$\sigma(H\mathbf{z}) = I - H\mathbf{z}(H\mathbf{z})^\top = I - H\mathbf{z}\mathbf{z}^\top H^\top = H\sigma(\mathbf{z})H^\top,$$

entonces

$$\begin{aligned} d(H\mathbf{Z}_t) &= Hd\mathbf{Z}_t = H \left[\sigma(\mathbf{Z}_t) d\tilde{\mathbf{B}}_t - \frac{d-1}{2} \mathbf{Z}_t dt \right] \\ &= H\sigma(\mathbf{Z}_t)H^\top H d\tilde{\mathbf{B}}_t - \frac{d-1}{2} H\mathbf{Z}_t dt = \sigma(H\mathbf{Z}_t) d(\tilde{\mathbf{B}}_t) - \frac{d-1}{2} (H\mathbf{Z}_t) dt, \end{aligned}$$

por lo que \mathbf{Z} es una difusión en la esfera \mathcal{S}^{d-1} invariante bajo rotaciones, i.e., un movimiento browniano en la esfera.

Definición 2.29. Sea $\mathbf{Z}: \Omega \rightarrow \mathbb{W}(\mathcal{S}^{d-1})$ un proceso estocástico definido en un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ con $d \geq 2$ y sea $\mu = \mathbb{P} \circ \mathbf{Z}_0^{-1}$ su distribución inicial. \mathbf{Z} se llama un *movimiento browniano en la esfera \mathcal{S}^{d-1}* si \mathbf{Z} es un proceso de difusión con generador $\frac{1}{2}\Delta_S$, donde

$$\Delta_S f(\mathbf{z}) = \left[\sum_{i,j=1}^d (\delta_{ij} - z_i z_j) \frac{\partial^2 f(\mathbf{z})}{\partial z_i \partial z_j} \right] - (d-1) \sum_{i=1}^d z_i \frac{\partial f(\mathbf{z})}{\partial z_i}, \quad (2.42)$$

para todo $f \in \mathcal{C}_0^2(\mathcal{S}^{d-1})$. Denotaremos al movimiento browniano en \mathcal{S}^{d-1} por $\mathbf{MB}(\mathcal{S}^{d-1})$.

Observación 2.14. Veremos en la siguiente sección que el operador Δ_S es en realidad el operador de Laplace en la esfera \mathcal{S}^{d-1} .

Observación 2.15. La ecuación diferencial estocástica (2.40) se conoce como la *representación de Stroock para el movimiento browniano esférico* (ver [Hsu02], p.83 & [Ito75]), y la única solución fuerte $(\mathbf{Z}, \tilde{\mathbf{B}})$ a la EDE (2.40) es un movimiento browniano en la esfera \mathcal{S}^{d-1} .

2.3.1. El operador de Laplace-Beltrami en \mathcal{S}^{d-1}

Sea \mathcal{M} una variedad riemanniana de dimensión d junto con una métrica riemanniana g , y la denotaremos por (\mathcal{M}, g) ; en donde g es una forma bilineal en \mathcal{M} que es simétrica y definida positiva tal que para cada punto $p \in \mathcal{M}$, la forma bilineal $g(p)$ define un producto escalar $\langle \cdot, \cdot \rangle_p$ en el espacio tangente $T_p\mathcal{M}$. Si (x_1, \dots, x_d) es un sistema de coordenadas locales en un conjunto abierto $U \subset \mathcal{M}$, entonces $\left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_d}\right)$ es una base para el espacio tangente en todos los puntos de U . Así, se obtiene una matriz simétrica y definida positiva $g \equiv (g_{ij})$ por

$$g_{ij}(p) = \left\langle \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \right\rangle_p, \quad p \in U. \quad (2.43)$$

El operador de Laplace, también llamado el *operador de Laplace-Beltrami* en la variedad riemanniana (\mathcal{M}, g) , y lo denotamos por $\Delta_{\mathcal{M}}$, está definido como la divergencia del gradiente en funciones $f \in \mathcal{C}^\infty(\mathcal{M})$, donde $\mathcal{C}^\infty(\mathcal{M})$ es el conjunto de funciones $f: \mathcal{M} \rightarrow \mathbb{R}$ infinitamente diferenciables. En un sistema de coordenadas locales $(U, (x_i))$ en \mathcal{M} , el operador de Laplace-Beltrami está expresada en términos de la métrica g por

$$\Delta_{\mathcal{M}} f = \frac{1}{\sqrt{G}} \sum_{i,j=1}^d \frac{\partial}{\partial x_j} \left(\sqrt{G} g^{ij} \frac{\partial f}{\partial x_i} \right), \quad f \in \mathcal{C}^\infty(\mathcal{M}) \quad (2.44)$$

donde (g^{ij}) es la matriz inversa de $g = (g_{ij})$ y $G = \det(g)$.

Para ver más detalles sobre el operador de Laplace-Beltrami en una variedad riemanniana, consultar [dC92], [Jos11], [Shu01].

Por ejemplo, en el espacio euclidiano \mathbb{R}^d con su métrica estándar, i.e., $g \equiv I_d$ es la matriz de identidad; por (2.44), obtenemos el operador de Laplace $\Delta f = \sum_{i=1}^d \frac{\partial^2 f}{\partial x_i^2}$ en \mathbb{R}^d .

Considerando a la esfera $\mathcal{S}^{d-1} \subset \mathbb{R}^d$ como una variedad riemanniana con la métrica inducida por la métrica estándar en \mathbb{R}^d y un operador de Laplace-Beltrami correspondiente, que lo denotamos por $\Delta_{\mathcal{S}}$.

Definición 2.30. Sea $f \in \mathcal{C}^\infty(\mathcal{S}^{d-1})$ y define $\hat{f}: \mathbb{R}^d \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}$ por $\hat{f}(\mathbf{x}) = f\left(\frac{\mathbf{x}}{\|\mathbf{x}\|}\right)$.

El operador de Laplace-Beltrami en \mathcal{S}^{d-1} está definido por

$$\Delta_{\mathcal{S}} f = \Delta \hat{f}|_{\mathcal{S}^{d-1}}, \quad \text{o bien} \quad \Delta \hat{f}(\mathbf{x}) = r^{-2} \Delta_{\mathcal{S}} f\left(\frac{\mathbf{x}}{\|\mathbf{x}\|}\right), \quad (2.45)$$

donde $r = \|\mathbf{x}\|$, $\mathbf{x} \in \mathbb{R}^d$ y Δ es el laplaciano en el espacio de ambiente \mathbb{R}^d .

El operador de Laplace Δ en coordenadas polares

Sea $\mathbf{x} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$, como las coordenadas polares de \mathbf{x} son $\mathbf{x} = (r, \mathbf{u})$ con $r = \|\mathbf{x}\| \in (0, \infty)$ y $\mathbf{u} = \frac{\mathbf{x}}{\|\mathbf{x}\|} \in \mathcal{S}^{d-1}$. Sean $f \in \mathcal{C}^\infty((0, \infty))$, $g \in \mathcal{C}^\infty(\mathcal{S}^{d-1})$ y $h(r, \mathbf{u}) := f(r)g(\mathbf{u})$, entonces por la regla producto del operador laplaciano:

$$\Delta(f(r)g(\mathbf{u})) = (\Delta f)g + f(\Delta g) + 2(\nabla f) \cdot (\nabla g) = (\Delta f)g + f(\Delta g), \quad (2.46)$$

en donde ∇ denota al gradiente y $(\nabla f) \cdot (\nabla g) = 0$, ya que ∇f está dirigida a lo largo del vector radial y ∇g está dirigida a lo largo del vector tangente.

Al calcular $\Delta f(r)$, como $\frac{\partial r}{\partial x_i} = \frac{x_i}{r}$, y $\frac{\partial^2 r}{\partial x_i^2} = \frac{1}{r} - \frac{x_i^2}{r^3}$, y por la regla de la cadena,

$$\frac{\partial}{\partial x_i} = \frac{\partial r}{\partial x_i} \frac{\partial}{\partial r} = \frac{x_i}{r} \frac{\partial}{\partial r},$$

$$\frac{\partial^2}{\partial x_i^2} = \frac{\partial}{\partial x_i} \left(\frac{x_i}{r} \frac{\partial}{\partial r} \right) = \frac{x_i}{r} \frac{\partial}{\partial r} \left(\frac{x_i}{r} \frac{\partial}{\partial r} \right) = \left(\frac{x_i}{r} \right)^2 \frac{\partial^2}{\partial r^2} + \left(\frac{1}{r} - \frac{x_i^2}{r^3} \right) \frac{\partial}{\partial r}.$$

Así que

$$\Delta f(r) = \sum_{i=1}^d \frac{\partial^2 f}{\partial x_i^2}(r) = \sum_{i=1}^d \left[\left(\frac{x_i}{r} \right)^2 \frac{\partial^2}{\partial r^2} + \left(\frac{1}{r} - \frac{x_i^2}{r^3} \right) \frac{\partial}{\partial r} \right] f(r) = f''(r) + \frac{d-1}{r} f'(r).$$

Entonces por la definición del operador de Laplace Δ_S en la esfera \mathcal{S}^{d-1} junto con (2.46), se tiene que

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{d-1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \Delta_S. \quad (2.47)$$

Como las combinaciones lineales de funciones de tipo $f(r)g(\mathbf{u})$ son densos en el espacio de funciones $\mathcal{C}^\infty(\mathbb{R}^d \setminus \{\mathbf{0}\})$, la fórmula (2.47) se cumple para cualquier función $h \in \mathcal{C}^\infty(\mathbb{R}^d \setminus \{\mathbf{0}\})$. Por otro lado, con la parametrización $\mathbf{x} = r\mathbf{u} \in \mathbb{R}^d$, y considerando $f(\mathbf{x}) \equiv f(r, \mathbf{u}) \in \mathcal{C}^\infty(\mathbb{R}^d \setminus \{\mathbf{0}\})$, se puede obtener Δ_S explícitamente, y está dado por (2.42). Por lo que el movimiento browniano en la esfera es una difusión generada por el generador $\frac{1}{2}\Delta_S$.

El operador de Laplace-Beltrami Δ_S en coordenadas polares

En la Definición 2.30 del operador de Laplace-Beltrami Δ_S en la esfera se consideró a la esfera $\mathcal{S}^{d-1} \subset \mathbb{R}^d$ como una variedad riemanniana con la métrica inducida de la métrica estándar de \mathbb{R}^d . Otra manera de expresar el operador Δ_S es en coordenadas polares, considerando a la esfera \mathcal{S}^{d-1} como una variedad rotacionalmente simétrica. Esto es, si $\mathbf{o} \in \mathcal{S}^{d-1}$ es el polo de la esfera; sin pérdida de la generalidad, supongamos que $\mathbf{o} = (0, \dots, 0, 1)^\top$, el polo norte de la esfera. Entonces se puede determinar las coordenadas polares $(r, \boldsymbol{\theta}) \in \mathbb{R}^+ \times \mathcal{S}^{d-2}$ en $\mathcal{S}^{d-1} \setminus \{\mathbf{o}, -\mathbf{o}\}$ alrededor de \mathbf{o} . De manera que para cada $\mathbf{x} \in \mathcal{S}^{d-1} \setminus \{\mathbf{o}, -\mathbf{o}\}$, \mathbf{x} tiene su descomposición en coordenadas polares: $\mathbf{x} = (r(\mathbf{x}), \boldsymbol{\theta}(\mathbf{x})) \in (0, \pi) \times \mathcal{S}^{d-2}$, donde la función radial $r(\mathbf{x}) = \text{dist}(\mathbf{o}, \mathbf{x})$ está definida como la distancia geodésica del punto \mathbf{o} a \mathbf{x} en la esfera \mathcal{S}^{d-1} ; es decir, $r(\mathbf{x}) = \arccos(\langle \mathbf{o}, \mathbf{x} \rangle)$, la colatitud en la esfera \mathcal{S}^{d-1} . En este caso, se tiene que la transformación de coordenadas polares a coordenadas cartesianas en \mathcal{S}^{d-1} es $\mathbf{x} = ((\sin r) \boldsymbol{\theta}^\top, \cos r) \in \mathcal{S}^{d-1}$ (c.f. [PR88]). Además, $\mathcal{S}^{d-1} \setminus \{\mathbf{o}, -\mathbf{o}\} \cong (0, \pi) \times \mathcal{S}^{d-2}$ tiene una métrica riemanniana en coordenadas polares dada por

$$g_{(r,\boldsymbol{\theta})} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{sen}^2 r \end{pmatrix}, \quad \text{y su inversa} \quad g^{-1} = g_{(r,\boldsymbol{\theta})}^{-1} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \frac{1}{\mathbf{sen}^2 r} \end{pmatrix},$$

donde $\mathbf{sen}^2 r$ y $\frac{1}{\mathbf{sen}^2 r}$ en $g_{(r,\boldsymbol{\theta})}$ y g^{-1} representan las submatrices diagonales de dimensión $(d-2)$ con $\mathbf{sen}^2 r$ y $\frac{1}{\mathbf{sen}^2 r}$ en la diagonal principal, respectivamente.

Así, por la representación (2.44) del operador de Laplace-Beltrami en coordenadas locales, sustituyendo $\det(g) = (\mathbf{sen}^2 r)^{d-2}$, y g^{-1} dado por anterior, se obtiene el operador de Laplace-Beltrami en la variedad $\mathcal{M} = \mathcal{S}^{d-1} \setminus \{\mathbf{o}, -\mathbf{o}\}$ bajo coordenadas polares:

$$\begin{aligned} \Delta_S &= \frac{1}{(\mathbf{sen} r)^{d-2}} \frac{\partial}{\partial r} \left((\mathbf{sen} r)^{d-2} \frac{\partial}{\partial r} \right) + \frac{1}{(\mathbf{sen} r)^{d-2}} \sum_{i=1}^{d-2} \frac{\partial}{\partial \theta_i} \left((\mathbf{sen} r)^{d-2} \frac{1}{\mathbf{sen}^2 r} \frac{\partial}{\partial \theta_i} \right) \\ &= \frac{\partial^2}{\partial r^2} + (d-2) \cot r \frac{\partial}{\partial r} + \frac{1}{\mathbf{sen}^2 r} \Delta_{\mathcal{S}^{d-2}}, \end{aligned} \quad (2.48)$$

donde $\Delta_{\mathcal{S}^{d-2}} = \sum_{i=1}^{d-2} \frac{\partial^2}{\partial \theta_i^2}$ es el operador de Laplace-Beltrami en \mathcal{S}^{d-2} .

Para ver más detalle sobre el operador de Laplace-Beltrami en coordenadas polares, consultar [Hsu02], Pag. 83 y [PR88].

2.3.2. Representación skew-product del MB(\mathbb{R}^d)

En esta sección, veremos que el movimiento browniano $(\mathbf{B}_t)_{t \geq 0}$ en \mathbb{R}^d ($d \geq 2$) a través del operador de Laplace en coordenadas polares, tiene una descomposición en coordenadas polares $\mathbf{B}_t = (R_t, \mathbf{U}_t)$, conocida como *la representación skew-product del movimiento browniano*; donde R es su *parte radial* y \mathbf{U} es su *parte angular*. Esto es, el siguiente resultado:

Proposición 2.31. *Sea $\mathbf{B} = (\mathbf{B}_t)_{t \geq 0}$ un movimiento browniano en \mathbb{R}^d , $d \geq 2$, con valor inicial $\mathbf{B}_0 = \mathbf{x} \neq \mathbf{0}$. Sea $R_t = \|\mathbf{B}_t\|$ un proceso de Bessel $\text{BES}_x(d)$. Entonces,*

$$\mathbf{B}_t = R_t \boldsymbol{\theta}_t, \quad t \geq 0, \quad (2.49)$$

donde $l_t = \int_0^t R_s^{-2} ds$ y $(\boldsymbol{\theta}_t)_{t \geq 0}$ es un movimiento browniano en la esfera \mathcal{S}^{d-1} con valor inicial $\boldsymbol{\theta}_0 = \frac{\mathbf{x}}{\|\mathbf{x}\|}$ e independiente del proceso $(R_t)_{t \geq 0}$.

Demostración. Supongamos que el movimiento browniano \mathbf{B}_t está definido en el espacio de probabilidad completo filtrado $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$.

Como $\mathbf{B}_0 = \mathbf{x} \neq \mathbf{0}$ y $d \geq 2$, entonces $\mathbf{B}_t > 0$ c.s., y recordemos que para $\mathbf{x} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$, sus coordenadas polares son: $r = \|\mathbf{x}\| \in (0, \infty)$, $\mathbf{u} = \frac{\mathbf{x}}{\|\mathbf{x}\|} \in \mathcal{S}^{d-1}$; así, la transformación $\Phi: \mathbb{R}^d \setminus \{\mathbf{0}\} \ni \mathbf{x} \mapsto (r, \mathbf{u}) \in (0, \infty) \times \mathcal{S}^{d-1}$ es continua y biyectiva, con la inversa $\Phi^{-1}(r, \mathbf{u}) = r\mathbf{u}$. Entonces por el Lema 2.14, $\Phi(\mathbf{B}_t) = (R_t, \mathbf{U}_t)$ es un proceso de Markov; más aún, es un proceso de difusión, y calculamos su generador.

Sea $h(r, \mathbf{u}) = f(r)g(\mathbf{u})$ tal que $f: (0, \infty) \rightarrow \mathbb{R}$ y $g: \mathcal{S}^{d-1} \rightarrow \mathbb{R}$ son funciones de clase \mathcal{C}_0^2 . Aplicando la fórmula de Itô para $f(R_t)$ y $g(\mathbf{U}_t)$ con $R_t = \|\mathbf{B}_t\|$ y $\mathbf{U}_t = \frac{\mathbf{B}_t}{\|\mathbf{B}_t\|}$, respectivamente, donde sabemos que $R_t = \|\mathbf{B}_t\|$ es un proceso de Bessel que satisface

$$dR_t = d\beta_t + \frac{d-1}{2R_t} dt, \quad (2.50)$$

con $\beta_t = \sum_{i=1}^d \int_0^t \frac{B_s^i}{R_s} dB_s^i = \int_0^t \mathbf{U}_s^\top d\mathbf{B}_s$ un movimiento browniano escalar (ver Ejem. 2.22). También, por (2.37), U_t^i satisface la EDE:

$$dU_t^i = \frac{1}{\|\mathbf{B}_t\|} \sum_{k=1}^d (\delta_{ki} - U_t^k U_t^i) dB_t^k - \frac{1}{\|\mathbf{B}_t\|^2} \frac{d-1}{2} U_t^i dt, \quad \text{para cada } i = 1, \dots, d.$$

Así, con el mismo comentario de la sección 2.3.1, la parte del operador de Laplace en coordenadas polares, se obtiene el generador para (R_t, \mathbf{U}_t) ,

$$\mathcal{A} f(r)g(\mathbf{u}) = \frac{1}{2} \left\{ \frac{\partial^2}{\partial r^2} + \frac{d-1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \Delta_S \right\} f(r)g(\mathbf{u}),$$

esto es, el generador $\frac{1}{2}\Delta$, con Δ el operador de Laplace en \mathbb{R}^d y en coordenadas polares.

Como las combinaciones lineales de funciones de tipo $f(r)g(\mathbf{u})$ son densos en el espacio de funciones $\mathcal{C}_0^2((0, \infty) \times \mathcal{S}^{d-1})$, el generador anterior se cumple para cualquier función $h \in \mathcal{C}_0^2((0, \infty) \times \mathcal{S}^{d-1})$. En consecuencia, está bien definido $\mathbf{B}_t = R_t \mathbf{U}_t$, un movimiento browniano en \mathbb{R}^d .

Por otro lado, vimos en el Ejemplo 2 del movimiento browniano en la esfera \mathcal{S}^{d-1} , si definimos $l_t := \int_0^t \frac{ds}{\|\mathbf{B}_s\|^2} = \int_0^t R_s^{-2} ds$ y $\{\tau_t\}$ su inversa, entonces el proceso $\boldsymbol{\theta}_t = \mathbf{U}_{\tau_t}$, $t \geq 0$ es un movimiento browniano en la esfera \mathcal{S}^{d-1} que satisface la EDE:

$$d\boldsymbol{\theta}_t = (I - \boldsymbol{\theta}_t \boldsymbol{\theta}_t^\top) d\tilde{\mathbf{B}}_t - \frac{d-1}{2} \boldsymbol{\theta}_t dt, \quad (2.51)$$

donde $\tilde{\mathbf{B}}_t = \int_0^{\tau_t} \frac{1}{R_s} d\mathbf{B}_s$ es un (\mathcal{G}_t) -movimiento browniano y $\mathcal{G}_t = \mathcal{F}_{\tau_t}$, $t \geq 0$. Por lo que $\boldsymbol{\theta}_t = \mathbf{U}_t$, y $\mathbf{B}_t = R_t \boldsymbol{\theta}_t$, $t \geq 0$.

Ahora demostramos la independencia de los procesos R y $\boldsymbol{\theta}$.

Primero modificamos el movimiento browniano $\widetilde{\mathbf{B}}$ en la EDE (2.51) de manera que si extendemos el espacio de probabilidad para definir un movimiento browniano escalar $\widehat{\beta}$ independiente de \mathbf{B} ([RY99], Chap.V, Sec.1), y definimos una martingala local continua

$$\mathbf{W}_t = \int_0^{\tau_t} \frac{1}{R_s} (I - \mathbf{U}_s \mathbf{U}_s^\top) d\mathbf{B}_s + \int_0^{\tau_t} \frac{1}{R_s} \mathbf{U}_s d\widehat{\beta}_s.$$

Notemos que para cada $1 \leq i, j \leq d$, se tiene $\langle W^i, W^j \rangle_t = \delta_{ij}t$, entonces por el teorema de caracterización de Lévy ([RY99], Chap. IV, Thm. 3.6), \mathbf{W}_t es un (\mathcal{G}_t) -movimiento browniano. Además, como

$$(I - \mathbf{U}_s \mathbf{U}_s^\top) \left[\frac{1}{R_s} (I - \mathbf{U}_s \mathbf{U}_s^\top), \frac{1}{R_s} \mathbf{U}_s \right] = \left[\frac{1}{R_s} (I - \mathbf{U}_s \mathbf{U}_s^\top), 0 \right],$$

usando la fórmula de cambio de tiempo para integrales estocásticas ([RY99], Chap.V, Prop. 1.4), obtenemos que

$$d\boldsymbol{\theta}_t = (I - \boldsymbol{\theta}_t \boldsymbol{\theta}_t^\top) d\mathbf{W}_t - \frac{d-1}{2} \boldsymbol{\theta}_t dt, \quad (2.52)$$

por lo que $\boldsymbol{\theta}$ es un movimiento browniano en la esfera \mathcal{S}^{d-1} .

Luego, verificamos la independencia de los procesos β y \mathbf{W} . Si definimos $\xi_t = \beta_{\tau_t} = \int_0^{\tau_t} \mathbf{U}_s^\top d\mathbf{B}_s$, así que ξ es una (\mathcal{G}_t) -martingala local y se obtiene que $\langle W^i, \xi \rangle_t = 0$ para cada $1 \leq i \leq d$, y $\langle \xi, \xi \rangle_t = \tau_t$, con su inversa l_t . Entonces, por el teorema de Knight ([RY99], Chap.V, Thm. 1.9) para el cambio de tiempo multidimensional, tenemos que \mathbf{W} y $(\xi_{l_t})_{t \geq 0} = (\beta_{l_t})_{t \geq 0}$ son independientes. Y por lo tanto, se sigue la independencia de R y $\boldsymbol{\theta}$.

En efecto, si $(\mathbf{X}, \widehat{\mathbf{B}})$ es una solución fuerte a una EDE, entonces existe una transformación medible Φ tal que $\mathbf{X} = \Phi(\widehat{\mathbf{B}})$ c.s. (c.f. [Che00], [YW71]). Así, podemos encontrar las transformaciones Φ_1, Φ_2 tales que $R = \Phi_1(\beta)$ y $\boldsymbol{\theta} = \Phi_2(\mathbf{W})$, y en consecuencia, la independencia de R y $\boldsymbol{\theta}$ se sigue de la independencia de β y \mathbf{W} . \square

Observación 2.16. Por la proposición anterior, podemos construir un movimiento browniano en \mathbb{R}^d como el producto de un proceso de Bessel $(R_t)_{t \geq 0}$ y un movimiento browniano independiente $(\boldsymbol{\theta}_t)_{t \geq 0}$ en la esfera \mathcal{S}^{d-1} en el tiempo acelerado $l_t = \int_0^t R_s^{-2} ds$.

2.3.3. Representación skew-product del MB(\mathcal{S}^{d-1})

De manera análoga al movimiento browniano en el espacio euclidiano \mathbb{R}^d , el movimiento browniano en la esfera \mathcal{S}^{d-1} ($d \geq 3$) también tiene una representación skew-product, y tenemos el siguiente resultado.

Proposición 2.32. *Sea $\mathbf{X} = (\mathbf{X}_t)_{t \geq 0}$ un movimiento browniano en la esfera \mathcal{S}^{d-1} ($d \geq 3$) con valor inicial $\mathbf{X}_0 \notin \{\mathbf{o}, -\mathbf{o}\}$, entonces \mathbf{X} tiene una representación skew-product*

$$\mathbf{X}_t = (r_t, \boldsymbol{\varphi}_{l_t}), \quad t \geq 0,$$

donde $r = (r_t)_{t \geq 0}$ es un proceso de difusión en $[0, \pi]$ con generador

$$\mathcal{A}_r = \frac{1}{2} \frac{\partial^2}{\partial r^2} + \frac{d-2}{2} \cot r \frac{\partial}{\partial r}, \quad 0 < r < \pi, \quad (2.53)$$

y con valor inicial $0 < r_0 < \pi$, $l_t = \int_0^t (\sin r_s)^{-2} ds$ y $(\boldsymbol{\varphi}_t)_{t \geq 0}$ es un movimiento browniano en la esfera \mathcal{S}^{d-2} independiente del proceso $(r_t)_{t \geq 0}$. Donde r se denomina la parte radial de \mathbf{X} (el proceso radial) y $\boldsymbol{\varphi}$ su parte angular (el proceso angular).

Demostración. Supongamos que el movimiento browniano \mathbf{X} está definido en el espacio de probabilidad completo filtrado $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ y satisface la EDE

$$d\mathbf{X}_t = (I - \mathbf{X}_t \mathbf{X}_t^\top) d\mathbf{B}_t - \frac{d-1}{2} \mathbf{X}_t dt, \quad (2.54)$$

donde \mathbf{B} es un (\mathcal{F}_t) -movimiento browniano en \mathbb{R}^d .

Como $\mathbf{X}_0 \notin \{\mathbf{o}, -\mathbf{o}\}$, entonces $\mathbf{X}_t \notin \{\mathbf{o}, -\mathbf{o}\}$ c.s. y la función radial $r := r(\mathbf{x}) = \text{dist}(\mathbf{o}, \mathbf{x}) = \arccos(\langle \mathbf{o}, \mathbf{x} \rangle)$ toma valores en $(0, \pi)$; además, como $\mathbf{o} = (0, \dots, 0, 1)^\top$, entonces para el movimiento browniano \mathbf{X} en \mathcal{S}^{d-1} , sea $Z_t = X_t^d \neq \pm 1$, la d -ésima coordenada de \mathbf{X}_t , $t \geq 0$, tenemos que $r(\mathbf{X}_t) = \arccos(X_t^d) = \arccos(Z_t) := g(Z_t)$. Donde

$$\frac{\partial r}{\partial x_d}(\mathbf{x}) = \frac{dg(z)}{dz} = -(1-z^2)^{-\frac{1}{2}}, \quad \frac{\partial^2 r}{\partial x_d^2}(\mathbf{x}) = \frac{d^2g(z)}{dz^2} = -z(1-z^2)^{-\frac{3}{2}},$$

y $\frac{\partial r}{\partial x_i}(\mathbf{x}) = \frac{\partial^2 r}{\partial x_i \partial x_j}(\mathbf{x}) = 0$, para todo $i \neq d$ o $j \neq d$.

Entonces, aplicando la fórmula de Itô para $r_t = r(\mathbf{X}_t)$, se obtiene

$$r_t = r_0 - \int_0^t (1 - Z_s^2)^{-\frac{1}{2}} dZ_s - \frac{1}{2} \int_0^t Z_s (1 - Z_s^2)^{-\frac{3}{2}} (dZ_s)^2, \quad (2.55)$$

donde Z_s satisface la ecuación $dZ_s = \sum_{k=1}^d (\delta_{kd} - X_s^k Z_s) dB_s^k - \frac{d-1}{2} Z_s ds$ por (2.54), y $(dZ_s)^2 = (1 - Z_s^2) ds$. Sustituyendo en (2.55) para obtener

$$\begin{aligned} r_t &= r_0 - \int_0^t (1 - Z_s^2)^{-\frac{1}{2}} \left[\sum_{k=1}^d (\delta_{kd} - X_s^k Z_s) dB_s^k - \frac{d-1}{2} Z_s ds \right] - \frac{1}{2} \int_0^t Z_s (1 - Z_s^2)^{-\frac{3}{2}} (1 - Z_s^2) ds \\ &= r_0 - \int_0^t (1 - Z_s^2)^{-\frac{1}{2}} \sum_{k=1}^d (\delta_{kd} - X_s^k Z_s) dB_s^k + \frac{d-2}{2} \int_0^t Z_s (1 - Z_s^2)^{-\frac{1}{2}} ds. \end{aligned}$$

Sea

$$\beta_t = - \int_0^t (1 - Z_s^2)^{-\frac{1}{2}} \sum_{k=1}^d (\delta_{kd} - X_s^k Z_s) dB_s^k,$$

por simples cálculos y se obtiene que $\langle \beta, \beta \rangle_t = t$, entonces por el teorema de caracterización de Lévy, β_t es un (\mathcal{F}_t) -movimiento browniano escalar.

Como $r_t = \arccos(Z_t)$, entonces $Z_t = \cos r_t$, $(1 - Z_t^2)^{-\frac{1}{2}} = (\sin r_t)^{-1}$ y por lo que

$$r_t = r_0 + \beta_t + \frac{d-2}{2} \int_0^t \cot r_s ds, \quad 0 < r_t < \pi. \quad (2.56)$$

Entonces, la parte radial del movimiento browniano \mathbf{X} es un proceso de difusión generada por el generador \mathcal{A}_r dado en (2.53).

Ahora nos fijamos en el proceso angular. Recordemos que el operador de Laplace-Beltrami en la esfera \mathcal{S}^{d-1} y en coordenadas polares $(r, \boldsymbol{\theta}) \in (0, \pi) \times \mathcal{S}^{d-2}$ tiene la forma

$$\Delta_{\mathcal{S}} = \frac{\partial^2}{\partial r^2} + (d-2) \cot r \frac{\partial}{\partial r} + \frac{1}{\sin^2 r} \Delta_{\mathcal{S}^{d-2}}. \quad (2.57)$$

Como el movimiento browniano \mathbf{X} en la esfera \mathcal{S}^{d-1} es una difusión generada por el generador $\frac{1}{2} \Delta_{\mathcal{S}}$, i.e., \mathbf{X} es solución a un problema de martingala. Sea $f(r, \boldsymbol{\theta}) \equiv f(\boldsymbol{\theta}) \in \mathcal{C}_0^2(\mathcal{S}^{d-2})$, entonces

$$f(\boldsymbol{\theta}_t) = f(\boldsymbol{\theta}_0) + M_t^f + \frac{1}{2} \int_0^t \frac{1}{\sin^2 r_s} \Delta_{\mathcal{S}^{d-2}} f(\boldsymbol{\theta}_s) ds, \quad (2.58)$$

para M_t^f una martingala local. Si definimos una nueva escala de tiempo

$$l_t = l(t, \omega) := \int_0^t \frac{ds}{\text{sen}^2 r_s}, \quad (2.59)$$

entonces la función $t \mapsto l_t$ es estrictamente creciente y sea $\{\tau_t\}$ la inversa de $\{l_t\}$ con $\tau_0 = 0$. Evaluando la ecuación (2.58) en el tiempo τ_t , como $l(\tau_u, \cdot) = u$, entonces $\tau'_u = l'(\tau_u, \cdot)^{-1} = \text{sen}^2 r_{\tau_u}$. Haciendo el cambio de la variable $s = \tau_u$, así $ds = \tau'_u du = \text{sen}^2 r_{\tau_u} du$, y $\int_0^{\tau_t} \frac{1}{\text{sen}^2 r_s} \Delta_{\mathcal{S}^{d-2}} f(\boldsymbol{\theta}_s) ds = \int_0^t \Delta_{\mathcal{S}^{d-2}} f(\boldsymbol{\theta}_{\tau_u}) du$. Sea $\boldsymbol{\varphi}_t = \boldsymbol{\theta}_{\tau_t}$, entonces

$$f(\boldsymbol{\varphi}_t) = f(\boldsymbol{\varphi}_0) + M_{\tau_t}^f + \frac{1}{2} \int_0^t \Delta_{\mathcal{S}^{d-2}} f(\boldsymbol{\varphi}_u) du. \quad (2.60)$$

Además, como para cada t , τ_t es un (\mathcal{F}_s) -tiempo de paro, entonces por el teorema del muestreo opcional de Doob ([Kal21], Thm. 9.12), $M_{\tau_t}^f$, $t \geq 0$ es una (\mathcal{F}_{τ_t}) -martingala local. Por lo tanto, vemos que el proceso $\boldsymbol{\varphi}_t = \boldsymbol{\theta}_{\tau_t}$, $t \geq 0$ es un movimiento browniano en la esfera \mathcal{S}^{d-2} que satisface la EDE

$$d\boldsymbol{\varphi}_t = (I - \boldsymbol{\varphi}_t \boldsymbol{\varphi}_t^\top) d\widetilde{\mathbf{W}}_t - \frac{d-2}{2} \boldsymbol{\varphi}_t dt, \quad (2.61)$$

donde $\widetilde{\mathbf{W}}$ es un (\mathcal{G}_t) -movimiento browniano en \mathbb{R}^{d-1} , con $\mathcal{G}_t = \mathcal{F}_{\tau_t}$, $t \geq 0$.

Entonces, hemos demostrado que el movimiento browniano \mathbf{X} en \mathcal{S}^{d-1} tiene una representación skew-product $\mathbf{X}_t = (r_t, \boldsymbol{\varphi}_t)$.

Ahora demostramos la independencia de los procesos $(r_t)_{t \geq 0}$ y $(\boldsymbol{\varphi}_t)_{t \geq 0}$.

Con la misma idea en la demostración de la proposición anterior, demostraremos la independencia de los movimientos brownianos que derivan de las EDEs (2.56) y (2.61) respectivamente. Con el mismo procedimiento de la demostración anterior, primero extendemos el espacio de probabilidad para definir un movimiento browniano escalar $\widehat{\beta}$ independiente de \mathbf{B} en la EDE (2.54). Sea $\widetilde{\mathbf{B}}_t := (B_t^1, \dots, B_t^{d-1})$, las primeras $(d-1)$ coordenadas de \mathbf{B}_t , entonces $\widetilde{\mathbf{B}}$ es un (\mathcal{F}_t) -movimiento browniano en \mathbb{R}^{d-1} , y definimos una martingala local

$$\mathbf{W}_t = \int_0^{\tau_t} \frac{1}{\text{sen } r_s} (I - \boldsymbol{\theta}_s \boldsymbol{\theta}_s^\top) d\widetilde{\mathbf{B}}_s + \int_0^{\tau_t} \frac{1}{\text{sen } r_s} \boldsymbol{\theta}_s d\widehat{\beta}_s, \quad (2.62)$$

donde para cada $1 \leq i, j \leq d-1$, se obtiene $\langle W^i, W^j \rangle_t = \delta_{ij} t$, y por el teorema de caracterización de Lévy, el proceso \mathbf{W} es un (\mathcal{G}_t) -movimiento browniano en \mathbb{R}^{d-1} . Además, se tiene que

$$(I - \boldsymbol{\theta}_s \boldsymbol{\theta}_s^\top) \left[\frac{1}{\text{sen } r_s} (I - \boldsymbol{\theta}_s \boldsymbol{\theta}_s^\top), \frac{1}{\text{sen } r_s} \boldsymbol{\theta}_s \right] = \left[\frac{1}{\text{sen } r_s} (I - \boldsymbol{\theta}_s \boldsymbol{\theta}_s^\top), 0 \right],$$

entonces por el teorema de cambio de tiempo ([Øk03], Thm.8.5.7), tenemos que

$$d\boldsymbol{\varphi}_t = (I - \boldsymbol{\varphi}_t \boldsymbol{\varphi}_t^\top) d\mathbf{W}_t - \frac{d-2}{2} \boldsymbol{\varphi}_t dt.$$

Por otro lado, definimos $\xi_t = \beta_{\tau_t} = \int_0^{\tau_t} -\frac{1}{\text{sen } r_s} \sum_{k=1}^d (\delta_{kd} - X_s^k Z_s) dB_s^k$. Como se tiene que la transformación de coordenadas polares a coordenadas cartesianas en la esfera \mathcal{S}^{d-1} (Sec. 2.3.1) es tal que $\mathbf{X}_t = ((\text{sen } r_t) \boldsymbol{\theta}_t^\top, \cos r_t)^\top$, es decir,

$$X_t^k = \theta_t^k (\text{sen } r_t), \quad k = 1, \dots, d-1, \quad \text{y} \quad X_t^d = \cos r_t, \quad \forall t \geq 0.$$

Entonces,

$$\xi_t = \int_0^{\tau_t} \cos r_s \sum_{k=1}^{d-1} \theta_s^k d\widetilde{B}_s^k - \int_0^{\tau_t} \text{sen } r_s dB_t^d = \int_0^{\tau_t} (\cos r_s) \boldsymbol{\theta}_s^\top d\widetilde{\mathbf{B}}_s - \int_0^{\tau_t} \text{sen } r_s dB_t^d, \quad t \geq 0.$$

Por lo que ξ es una (\mathcal{G}_t) -martingala local con $\langle W^i, \xi \rangle_t = 0$ para cada $1 \leq i \leq (d-1)$, y $\langle \xi, \xi \rangle_t = \tau_t$, con su inversa l_t . Entonces el teorema de Knight ([RY99], Chap.V, Thm.1.9) para el cambio de tiempo multidimensional afirma que \mathbf{W} y $(\xi_t)_{t \geq 0} = (\beta_t)_{t \geq 0}$ son independientes. En consecuencia, se sigue la independencia de los procesos φ y r . \square

Observación 2.17. Por la proposición anterior, podemos construir un $\mathbf{MB}(\mathcal{S}^{d-1})$ como un producto deformado. Sea $(r_t)_{t \geq 0}$ un proceso de difusión generado por \mathcal{A}_r dado en (2.53), también se le llama el *laplaciano radial* de Δ_S , y sea $(\varphi_t)_{t \geq 0}$ un movimiento browniano en \mathcal{S}^{d-2} independiente de r . Se define el cambio de tiempo l como en (2.59), entonces el proceso $t \mapsto \mathbf{X}_t = (r_t, \varphi_{l_t})$ es un movimiento browniano en la esfera \mathcal{S}^{d-1} .

Observación 2.18. Por otro lado, como la transformación de coordenadas polares a coordenadas cartesianas en la esfera \mathcal{S}^{d-1} es mediante $\mathbf{x} = ((\sin r) \boldsymbol{\theta}^\top, \cos r) \in \mathcal{S}^{d-1}$. Sea \mathbf{X} un movimiento browniano en la esfera \mathcal{S}^{d-1} , y $(r_t)_{t \geq 0}$ el proceso radial. Como $f(r) = \cos r$ es una función de clase \mathcal{C}^∞ , entonces aplicando la fórmula de Itô para $Z_t \equiv X_t^d = \cos r_t$, vemos que Z tiene valor inicial $Z_0 = X_0^d \in (-1, 1)$ y satisface la EDE:

$$\begin{aligned} dZ_t &= -\sin r_t \left(d\beta_t + \frac{d-2}{2} \cot r_t dt \right) - \frac{1}{2} \cos r_t dt \\ &= -\sin r_t d\beta_t - \frac{d-1}{2} \cos r_t dt = -\sqrt{1-Z_t^2} d\beta_t - \frac{d-1}{2} Z_t dt. \end{aligned} \quad (2.63)$$

Más aún, si consideramos a la transformación lineal $g(z) = \frac{1-z}{2}$, y aplicando nuevamente la fórmula de Itô para $Y_t = \frac{1-Z_t}{2}$, se tiene que

$$\begin{aligned} dY_t &= -\frac{1}{2} \left[-\sqrt{1-Z_t^2} d\beta_t - \frac{d-1}{2} Z_t dt \right] \\ &= \sqrt{Y_t(1-Y_t)} d\beta_t + \frac{1}{2} \left[\frac{d-1}{2} (1-Y_t) - \frac{d-1}{2} Y_t \right] dt, \quad Y_0 = \frac{1-X_0^d}{2}. \end{aligned} \quad (2.64)$$

Esto es, un proceso de difusión de Wright-Fisher con parámetros de mutación $\theta_1 = \theta_2 = \frac{d-1}{2}$. Veremos la definición precisa en el siguiente capítulo.

Como veremos en el Capítulo 4, la representación skew-product del movimiento browniano en la esfera \mathbf{X} tiene una relación con la difusión de Wright-Fisher Y , podemos expresar \mathbf{X} en términos de Y (ver [MMUB20]). Además, es posible dar una descomposición skew-product para la proyección de un movimiento browniano $\mathbf{MB}(\mathcal{S}^{d-1})$ en la bola \mathcal{B}^n con $n < d$. Se puede consultar el [MMUB18] para ver los detalles, y también citaremos sus resultados principales en el Capítulo 4.

Capítulo 3

La difusión de Wright-Fisher

Por lo que se mencionó en el final del capítulo anterior, a través de una serie de transformaciones sobre la parte radial del movimiento browniano en la esfera $\mathbf{MB}(\mathcal{S}^d)$ se da a un proceso de difusión de Wright-Fisher. Además, veremos en el siguiente capítulo que la proyección de $\mathbf{MB}(\mathcal{S}^d)$ en sus primeras $k < d$ coordenadas tiene una relación con la difusión de Wright-Fisher (cf. [MMUB18], [HJT17]).

Así, en este capítulo, estudiaremos el proceso de difusión de Wright-Fisher empezando con el modelo de Wright-Fisher como un modelo genético que describe la dinámica estocástica de frecuencias alélicas en una población bajo efectos de mutación y/o selección; específicamente, aquí sólo tratamos con el modelo neutral con mutación (para otros casos, ver [Eth11]). El modelo original de Sewall Wright y Ronald A. Fisher trabajaron con una población finita y en tiempo discreto; en la sección 3.1, veremos que la evolución de la frecuencia alélica en una población finita a lo largo de las generaciones $n \in \mathbb{N}$ (c/s efecto de mutación) se puede representar por una *cadena de Markov* (i.e., un proceso de Markov en tiempo discreto con $t \equiv n \in \mathbb{N}$ que toma valores en un espacio de estado finito o numerable). Específicamente, la siguiente definición:

Definición 3.1. Sean $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y \mathcal{E} un conjunto finito o numerable. Una sucesión de variables aleatorias

$$(X_n)_{n \in \mathbb{N}} \equiv \{X_n : \Omega \rightarrow \mathcal{E}; \quad n = 0, 1, 2, \dots\}$$

se llama una *cadena de Markov* con *espacio de estado* \mathcal{E} si satisface la condición de Markov; esto es, para todo $n \geq 1$ y $x_0, \dots, x_{n-1}, x, y \in \mathcal{E}$, se cumple

$$\mathbb{P}(X_{n+1} = y | X_0 = x_0, \dots, X_{n-1} = x_{n-1}, X_n = x) = \mathbb{P}(X_{n+1} = y | X_n = x).$$

Para todo $x, y \in \mathcal{E}$, $n \in \mathbb{N}$, se le llama a $P_{x,y} = \mathbb{P}(X_{n+1} = y | X_n = x)$ las *probabilidades de transición*; y $P = (P_{x,y})_{x,y \in \mathcal{E}}$ se denomina la *matriz de transición* para la cadena $(X_n)_{n \in \mathbb{N}}$.

En la sección 3.2, veremos que la difusión de Wright-Fisher con mutación es la única solución a una ecuación diferencial estocástica, y este resultado se obtiene mediante la aplicación de la convergencia de cadena de Markov a un proceso de difusión para el modelo de Wright-Fisher con mutación (cf. [Dur96]).

Finalmente, basado en el trabajo de Jenkins y Spanó ([JS17]), presentamos en la sección 3.3 un algoritmo para simular exactamente la densidad de transición de la difusión de Wright-Fisher neutral con mutación, y así también, implementamos este algoritmo al final de esta sección.

3.1. El modelo genético de Wright-Fisher

El modelo de Wright-Fisher es el modelo genético más simple que fue introducido por Sewall Wright y Ronald A. Fisher de manera independiente en la década de 1930. En esta sección, empezamos con la descripción del modelo en su forma más simple, se considera una población *haploide* (i.e., cada individuo de la población con un sólo gen se reproduce asexualmente reproduciendo “hijas” genéticamente idénticas a su “madre”) de tamaño fijo y constante N que evoluciona en generaciones discretas, y los N individuos están divididos en dos tipos de alelos: tipo A y tipo B . Además, supongamos que la población es *neutral*, i.e., el éxito reproductivo de un individuo no depende de su tipo, y esa reproducción es aleatoria.

En el modelo de Wright-Fisher neutral, la nueva generación se modela de la siguiente manera: la generación $(n+1)$ se forma escogiendo N genes uniformemente de la generación n de manera aleatoria y con reemplazo; esto es, cada individuo de la generación $(n+1)$ escoge a su “madre” de manera aleatoria e independiente de aquellos presentes en la generación n , copiando su tipo ([Eth11], Chap.2). Como nos interesa la distribución de la proporción de cada tipo de alelos en el proceso de evolución, definimos

$$X_n^N := \text{el número de alelos tipo } A \text{ en la generación } n, \quad n \in \mathbb{N}.$$

Entonces, $(X_n^N)_{n \in \mathbb{N}}$ define un proceso a tiempo discreto y se tiene el siguiente resultado.

Proposición 3.2. *Para todo $n \in \mathbb{N}$ y para todo $i, j \in \{0, 1, \dots, N\}$,*

$$\mathbb{P}(X_{n+1}^N = j | X_n^N = i) = \binom{N}{j} \left(\frac{i}{N}\right)^j \left(1 - \frac{i}{N}\right)^{N-j}. \quad (3.1)$$

Demostración. Por definición, para todo $n \in \mathbb{N}$, la variable X_n^N toma sus valores en el conjunto $\{0, 1, \dots, N\}$; también por definición, es claro que $\mathbb{P}(X_{n+1}^N = 0 | X_n^N = 0) = 1$, e igualmente $\mathbb{P}(X_{n+1}^N = N | X_n^N = N) = 1$.

De manera más general, sea $i \in \{1, \dots, N-1\}$, y supongamos que en la generación n , la población consiste en i alelos de tipo A y $(N-i)$ alelos de tipo B . Es decir, $X_n^N = i$, y la proporción de alelo A en la generación n es $p_i = \frac{i}{N}$, mientras que la del alelo B es $q_i = \frac{(N-i)}{N} = 1 - \frac{i}{N}$. Como en el modelo de Wright-Fisher, cada gen de la generación $(n+1)$ escoge a su “madre” aleatoriamente e independiente de aquellas presentes en la n -ésima generación; esto es, corresponde a una selección con reemplazo. Entonces, dado $X_n^N = i$, el número de individuos del alelo A en la $(n+1)$ -ésima generación sigue una distribución binomial $\mathfrak{Bin}(N, p_i)$. Por lo tanto, deducimos (3.1). \square

Observación 3.1. De acuerdo con la proposición anterior, como $\mathcal{L}(X_{n+1}^N | X_n^N = i) = \mathfrak{Bin}(N, p_i)$, $i \in \{0, 1, \dots, N\}$, entonces

$$\begin{aligned} \mathbb{E}[X_{n+1}^N | X_n^N = i] &= Np_i = i, & \mathbb{E}[X_{n+1}^N - X_n^N | X_n^N = i] &= 0, \\ \text{var}(X_{n+1}^N | X_n^N = i) &= \text{var}(X_{n+1}^N - X_n^N | X_n^N = i) = Np_i(1 - p_i) = i \left(1 - \frac{i}{N}\right). \end{aligned}$$

Y es inmediato el siguiente resultado.

Teorema 3.3. *En el modelo de Wright-Fisher, el proceso $X^N = (X_n^N)_{n \in \mathbb{N}}$ es una cadena de Markov en tiempo discreto con el espacio de estado finito y es una martingala discreta acotada. La matriz de probabilidades de transición $P = (P_{ij})$ de la cadena está dada por*

$$P_{ij} = \binom{N}{j} \left(\frac{i}{N}\right)^j \left(1 - \frac{i}{N}\right)^{N-j}, \quad \text{para todo } i, j \in \{0, 1, \dots, N\}. \quad (3.2)$$

Los estados 0 y N son dos estados absorbentes para la cadena X^N .

Demostración. Por definición, el proceso $X^N = (X_n^N)_{n \in \mathbb{N}}$ toma sus valores en el conjunto finito $\{0, 1, \dots, N\}$. Además, notemos que la distribución de X_{n+1}^N dado los valores de X_m^N , $0 \leq m \leq n$, sólo depende de X_n^N , entonces el proceso X^N define una cadena de Markov; y por la Proposición 3.2, la matriz de probabilidades de transición para la cadena está dada por (3.2).

Por otro lado, para cada $n \in \mathbb{N}$, denotamos por \mathcal{F}_n a la σ -álgebra generada por las variables $X_0^N, X_1^N, \dots, X_n^N$, y con los cálculos de la observación 3.1, tenemos

$$\begin{aligned} \mathbb{E}\left[X_{n+1}^N \mid \mathcal{F}_n\right] &= \sum_{j=0}^N j \mathbb{P}(X_{n+1}^N = j \mid \mathcal{F}_n) = \sum_{j=0}^N j \mathbb{P}(X_{n+1}^N = j \mid X_n^N) \\ &= \sum_{j=0}^N j \binom{N}{j} \left(\frac{X_n^N}{N}\right)^j \left(1 - \frac{X_n^N}{N}\right)^{N-j} = N \frac{X_n^N}{N} = X_n^N. \end{aligned}$$

Por lo tanto, el proceso X^N es una martingala c.r.a. $(\mathcal{F}_n)_{n \in \mathbb{N}}$ y acotado por N .

Observemos que $P_{0,0} = P_{N,N} = 1$, ya que una vez la población con puro alelos de tipo A ($X_n^N = N$) o de tipo B ($X_n^N = 0$), por definición del modelo, la siguiente generación sigue siendo de alelos puros del mismo tipo. Por lo tanto, 0 y N son estados absorbentes de la cadena X^N (más detalles sobre clasificación de estados de una cadena de Markov, ver [CRUBV04]). \square

Observación 3.2. Por el teorema anterior, el proceso $(X_n^N)_{n \in \mathbb{N}}$ es una (\mathcal{F}_n) -martingala acotada, por el teorema de convergencia de martingalas ([RY99], Chap.2, Thm.2.2), (X_n) converge casi seguramente a una variable aleatoria X_∞^N cuando $n \rightarrow \infty$. Además, como $(X_n^N)_{n \in \mathbb{N}}$ es una cadena de Markov con espacio de estados finito, si $i \in \{1, \dots, N-1\}$ y por la Proposición 3.2, tenemos que

$$P_{i,N} = \mathbb{P}(X_1^N = N \mid X_0^N = i) = \left(\frac{i}{N}\right)^N > 0,$$

esto es, del estado i se puede acceder al estado N , pero el estado i no es accesible desde el estado N , ya que N es un estado absorbente; análogamente, lo mismo ocurre para el estado 0. Entonces los dos estados absorbentes $\{0, N\}$ son estados recurrentes y los demás estados son transitorios (c.f. [CRUBV04]). Por lo tanto, X_∞^N toma sus valores en $\{0, N\}$.

Si definimos al *tiempo de fijación* de uno de los dos tipos de alelos en la población por

$$\tau := \min \{n \geq 0 : X_n^N = N \quad \text{o} \quad X_n^N = 0\},$$

así, τ es un tiempo de paro c.r.a. (\mathcal{F}_n) y acotado *c.s.*, y por el teorema de la parada opcional ([RY99], Chap.2, Prop. 1.4), tenemos

$$i = \mathbb{E}[X_0^N \mid X_0^N = i] = \mathbb{E}[X_\infty^N \mid X_0^N = i] = N \mathbb{P}(X_\infty^N = N \mid X_0^N = i) + 0, \quad i \in \{1, \dots, N\},$$

por lo tanto, $\mathbb{P}(X_\infty^N = N \mid X_0^N = i) = \frac{i}{N}$; es decir, la probabilidad de que se fija el alelo tipo A en la población, dado que se inicia con i individuos de este tipo en la población, es justo su frecuencia inicial $\frac{i}{N}$.

3.1.1. El modelo de Wright-Fisher con mutación

Ahora supongamos la presencia de mutación durante la reproducción; es decir, antes de la formación de la nueva generación, cada gen tiene la posibilidad de sufrir la mutación, esto es, transformar un gen al otro tipo. Específicamente, supongamos que para cada gen, la mutación del alelo A al alelo B ocurre con probabilidad α y del alelo B al alelo A ocurre con probabilidad β . Así, en este caso, si la “madre” es de alelo A , el individuo hereda el alelo A con probabilidad $(1 - \alpha)$ o muta y transforma al alelo B con probabilidad α ; análogamente, si la “madre” es de alelo B , el individuo hereda el alelo B con probabilidad $(1 - \beta)$ o sufre la mutación y porta el alelo A con probabilidad β . Entonces la probabilidad de producir un individuo tipo A en la $(n + 1)$ -ésima generación dado que hay i individuos tipo A en la n -ésima es

$$p_i^* = \frac{i}{N}(1 - \alpha) + \frac{N - i}{N}\beta, \quad i \in \{0, 1, \dots, N\}. \quad (3.3)$$

Al igual que en el caso del modelo neutral sin mutación, supongamos que cada individuo tiene la misma ventaja selectiva y la “madre” de cada individuo en la nueva generación siempre se elige uniformemente de la generación anterior, de manera independiente y aleatoria. Sea X_n^N la variable aleatoria que describe al número de alelos A en la generación n , para todo $n \in \mathbb{N}$, tenemos en este caso que

$$\mathcal{L}(X_{n+1}^N | X_n^N = i) = \mathfrak{Bin}(N, p_i^*), \quad i \in \{0, 1, \dots, N\}.$$

Como dado el presente $\{X_n^N = i\}$, el futuro $\{X_{n+1}^N = j\}$ es independiente del pasado $\{X_0^N, \dots, X_{n-1}^N\}$, vemos que la sucesión $(X_n^N)_{n \in \mathbb{N}}$ es una cadena de Markov con matriz de transición

$$P_{ij} = \mathbb{P}(X_{n+1}^N = j | X_n^N = i) = \binom{N}{j} p_i^{*j} (1 - p_i^*)^{N-j}, \quad \forall i, j \in \{0, 1, \dots, N\}. \quad (3.4)$$

En este caso, notemos que si $\alpha > 0$ y $\beta > 0$, no hay posibilidad de tener fijación en la población para alguno de los dos alelos. Ya que con la matriz de transición $P = (P_{ij})$ para la cadena de Markov $(X_n)_{n \in \mathbb{N}}$ dada por (3.4), y como la cadena tiene soporte finito, entonces es irreducible y recurrente positiva, por lo que existe un único vector de probabilidad invariante no trivial $\pi = (\pi_0, \pi_1, \dots, \pi_N)$ tal que

$$\sum_{i=0}^N \pi_i = 1, \quad \text{y} \quad \sum_{i=0}^N \pi_i P_{ij} = \pi_j, \quad \text{para todo } i, j \in \{0, 1, \dots, N\}.$$

Esta distribución de probabilidad π también se conoce como una distribución de la frecuencia alélica en estado estable.

Observación 3.3. Para ver otros casos del modelo de Wright-Fisher con efecto de selección o con más de dos tipos de alelos, puede consultar [KT81] y [Eth11].

3.2. La difusión de Wright-Fisher con mutación

En esta sección consideramos al modelo de Wright-Fisher de haploide con 2-alelos que involucra el efecto de mutación como se mencionó en la sección anterior. Sabemos que el proceso $(X_n^N)_{n \in \mathbb{N}}$ que describe el número de alelos tipo A a lo largo de generaciones $n \in \mathbb{N}$, es una cadena de Markov con probabilidades de transición dadas por (3.4). Con la aplicación de la convergencia de cadenas de Markov a procesos de difusión, si hacemos el tamaño de la población $N \rightarrow \infty$, se puede aproximar las frecuencias alélicas a un límite llamado proceso de difusión. Para esto, llamamos al siguiente resultado para la convergencia de cadenas de Markov al proceso de difusión (cf. [KT81], Chap.15, p.168 y [Dur96], Sec.8.7).

Lema 3.4. *Supongamos que $X^{(N)} = (X_n^{(N)})_{n \in \mathbb{N}}$ es una sucesión de cadenas de Markov en tiempo discreto con espacio de estado confinado a un intervalo $\mathcal{D} \subset \mathbb{R}$, acotado y cerrado. Sea h_N una sucesión de números positivos que tiende a 0 cuando $N \rightarrow \infty$ tal que los siguientes límites existen uniformemente:*

$$\begin{aligned} & \lim_{N \rightarrow \infty} h_N^{-1} \mathbb{E}[X_{n+1}^{(N)} - X_n^{(N)} | X_n^{(N)} = x] = b(x), \\ & \lim_{N \rightarrow \infty} h_N^{-1} \mathbb{E}[(X_{n+1}^{(N)} - X_n^{(N)})^2 | X_n^{(N)} = x] = a(x), \\ & \text{y} \quad \lim_{N \rightarrow \infty} h_N^{-1} \mathbb{E}[(X_{n+1}^{(N)} - X_n^{(N)})^k | X_n^{(N)} = x] = 0, \quad \text{para algún } k \geq 3. \end{aligned}$$

Entonces los procesos a tiempo continuo definidos por $Y^{(N)} = \{Y_t^{(N)} := X_{\lfloor t/h_N \rfloor}^{(N)} : t \geq 0\}$ ¹ convergen en distribución al proceso $Y = (Y_t)_{t \geq 0}$, donde Y es un proceso de difusión con generador

$$\mathcal{A}f(y) = \frac{1}{2}a(y)f''(y) + b(y)f'(y), \quad f \in \mathcal{C}_0^2(\mathcal{D}).$$

Así, para el modelo de Wright-Fisher neutral con mutación, consideremos al proceso

$$Y_t^N := \frac{1}{N} X_{\lfloor Nt \rfloor}^N, \quad t \geq 0.$$

Esto es, el proceso $Y^N = (Y_t^N)_{t \geq 0}$ representa a la evolución de la proporción de alelos A en unidades de tiempo de tamaño de la población. Supongamos que la frecuencia del alelo A en la generación 0 es $Y_0^N = y_0 \in [0, 1]$ y sean

$$\alpha = \frac{\gamma_1}{N}, \quad \beta = \frac{\gamma_2}{N}, \quad \gamma_1, \gamma_2 > 0, \quad (3.5)$$

es decir, la tasa de mutación es constante por unidad de tiempo para el proceso Y^N ; donde $\gamma_1 = \alpha N$ es la intensidad de mutación del alelo A al alelo B , mientras que $\gamma_2 = \beta N$ es la intensidad de mutación del alelo B al alelo A .

Sea $\Delta t = \frac{1}{N}$ y se tiene que

$$Y_{t+\Delta t}^N - Y_t^N = \frac{1}{N} \left(X_{\lfloor N(t+\Delta t) \rfloor}^N - X_{\lfloor Nt \rfloor}^N \right) = \frac{1}{N} \left(X_{\lfloor Nt \rfloor + 1}^N - X_{\lfloor Nt \rfloor}^N \right).$$

Para calcular los momentos condicionales, notemos que $NY_{t+\Delta t}^N = X_{\lfloor Nt \rfloor + 1}^N \mid NY_t^N = X_{\lfloor Nt \rfloor}^N = i \sim \mathfrak{Bin}(N, p_i^*)$, donde $p_i^* = \frac{1}{N} [(1 - \alpha)i + (N - i)\beta]$, entonces

$$\begin{aligned} \mathbb{E} \left[\left(Y_{t+\Delta t}^N - Y_t^N \right) \mid Y_t^N = \frac{i}{N} \right] &= \mathbb{E} \left[Y_{t+\Delta t}^N \mid Y_t^N = \frac{i}{N} \right] - \mathbb{E} \left[Y_t^N \mid Y_t^N = \frac{i}{N} \right] \\ &= \frac{1}{N} (Np_i^*) - \frac{i}{N} = \frac{1}{N} [-\alpha i + (N - i)\beta] \\ &= \left[-\gamma_1 \frac{i}{N} + \gamma_2 \left(1 - \frac{i}{N} \right) \right] \Delta t, \end{aligned} \quad (3.6)$$

y

$$\begin{aligned} \mathbb{E} \left[\left(Y_{t+\Delta t}^N - Y_t^N \right)^2 \mid Y_t^N = \frac{i}{N} \right] &= \mathbb{E} \left[\left(Y_{t+\Delta t}^N \right)^2 \mid Y_t^N = \frac{i}{N} \right] - 2 \frac{i}{N} \mathbb{E} \left[Y_{t+\Delta t}^N \mid Y_t^N = \frac{i}{N} \right] + \left(\frac{i}{N} \right)^2 \\ &= \frac{1}{N^2} [Np_i^*(1 - p_i^*) + N^2(p_i^*)^2] - 2 \frac{i}{N} p_i^* + \left(\frac{i}{N} \right)^2 = \frac{1}{N} p_i^*(1 - p_i^*) + \left(p_i^* - \frac{i}{N} \right)^2 \\ &= \left\{ \frac{i}{N} \left(1 - \frac{i}{N} \right) + \mathcal{O} \left(\frac{1}{N} \right) \right\} \Delta t, \end{aligned} \quad (3.7)$$

¹ $[x]$ denota a la función parte entera de $x \geq 0$, i.e., $[x] = \max\{k \in \mathbb{N} \mid k \leq x\}$.

donde $\mathcal{O}(\frac{1}{N})$ es de orden a lo más $\frac{1}{N}$. Haciendo $\frac{i}{N} \rightarrow \xi$ cuando $N \rightarrow \infty$, como $0 \leq i \leq N$, entonces $0 \leq \xi \leq 1$ y las ecuaciones (3.6) y (3.7) resultan

$$\lim_{\Delta t \downarrow 0^+} \frac{1}{\Delta t} \mathbb{E} \left[(Y_{t+\Delta t}^N - Y_t^N) \mid Y_t^N = \xi \right] = -\gamma_1 \xi + \gamma_2 (1 - \xi), \quad (3.8)$$

y

$$\lim_{\Delta t \downarrow 0^+} \frac{1}{\Delta t} \mathbb{E} \left[(Y_{t+\Delta t}^N - Y_t^N)^2 \mid Y_t^N = \xi \right] = \xi(1 - \xi),$$

donde la convergencia es uniforme para $0 \leq \xi \leq 1$. Además, por un cálculo directo y usando los momentos de la distribución binomial se obtiene

$$\lim_{\Delta t \downarrow 0^+} \frac{1}{\Delta t} \mathbb{E} \left[(Y_{t+\Delta t}^N - Y_t^N)^4 \mid Y_t^N = \xi \right] = 0.$$

En consecuencia, se obtiene el siguiente resultado:

Proposición 3.5. *Bajo la hipótesis de mutaciones (3.5), si $\frac{1}{N}X_0^N$ converge a $Y_0 \in [0, 1]$ en distribución cuando $N \rightarrow \infty$. Entonces los procesos $(Y_t^N)_{t \geq 0} \equiv (\frac{1}{N}X_{[Nt]}^N)_{t \geq 0}$ convergen en distribución al proceso $Y = (Y_t)_{t \geq 0}$, con (Y, B) la única solución fuerte a la ecuación diferencial estocástica*

$$Y_t = Y_0 + \int_0^t \sqrt{Y_s(1 - Y_s)} dB_s + \int_0^t (-\gamma_1 Y_s + \gamma_2(1 - Y_s)) ds, \quad (3.9)$$

donde $B = (B_t)_{t \geq 0}$ es un movimiento browniano escalar y $Y_t \in [0, 1]$ para todo $t \geq 0$, c.s..

Para más detalles sobre la demostración de la Proposición 3.5, se puede consultar [KT81], p.177-p.179; [Mel16], Chap.6 y [Alf15], Chap.6.

Observación 3.4. Notemos que los coeficientes en la EDE (3.9), $b(x) = -\gamma_1 x + \gamma_2(1 - x)$ y $\sigma(x) = \sqrt{x(1 - x)}$ son tales que, b es Lipschitz continua y σ es Hölder continua con exponente $\frac{1}{2}$. Entonces el criterio de Yamada-Watanabe para la unicidad de trayectorias a las EDEs (Prop. 2.24) afirma la existencia de la única solución fuerte a la ecuación (3.9).

Observación 3.5. En general, se escribe la ecuación (3.9) en la forma diferencial

$$dY_t = \sqrt{Y_t(1 - Y_t)} dB_t + [-\gamma_1 Y_t + \gamma_2(1 - Y_t)] dt, \quad Y_0 = y \in [0, 1], \quad (3.10)$$

donde $\gamma_1, \gamma_2 > 0$; y $Y_t \in [0, 1]$, $\forall t \geq 0$ c.s..

Observación 3.6. La solución $Y = (Y_t)_{t \geq 0}$ a la EDE (3.10) es un proceso de Markov que tiene *distribución estacionaria* $\psi \equiv \mathfrak{Bc}(2\gamma_2, 2\gamma_1)$, de manera que si Y_0 tiene distribución ψ , entonces Y_t tiene distribución ψ para todo $t \geq 0$ ([Eth11], Sec. 3.6, & [KT81], Pag. 221).

Observación 3.7. Cuando $0 < 2\gamma_2 < 1$ (y/o $0 < 2\gamma_1 < 1$), el estado $\{0\}$ (y/o $\{1\}$) es una *frontera regular*, de manera que la difusión Y_t puede entrar y salir del estado $\{0\}$ (y/o $\{1\}$) para todo $t \geq 0$. Cuando $2\gamma_2 \geq 1$ (y/o $2\gamma_1 \geq 1$), el estado $\{0\}$ (y/o $\{1\}$) es una *frontera de entrada*, i.e., no es alcanzable para el proceso Y_t desde el intervalo $(0, 1)$ para $t > 0$ al menos con valor inicial en 0 (o en 1) ([KT81], Chap. 15, Pag. 240).

Definición 3.6. Si $Y = (Y_t)_{t \geq 0}$ es solución a la EDE (3.10), entonces le decimos a Y un *proceso de difusión de Wright-Fisher neutral con mutación*, con parámetros de mutación $\gamma_1, \gamma_2 > 0$, y lo denotaremos por $Y \sim \mathbb{WF}_y(2\gamma_2, 2\gamma_1)$ de acuerdo a su distribución estacionaria $\psi \equiv \mathfrak{Bc}(2\gamma_2, 2\gamma_1)$.

3.3. Simulación exacta de la difusión de Wright-Fisher con mutación

Vimos en la sección 2.2.1 una simulación exacta del proceso de Bessel mediante su densidad de transición, de la misma manera, como un proceso de difusión de Wright-Fisher neutral con mutación que satisface el modelo (3.10) tiene una representación probabilística de la función de densidad de transición ([GS10], [JS17]), la que nos permite simular el proceso de Wright-Fisher neutral con mutación a partir de su densidad de transición. El procedimiento de la simulación exacta se presentó con detalle por Jenkins y Spanó ([JS17]) y en esta sección, revisamos brevemente la idea principal sobre la simulación exacta de la difusión de Wright-Fisher neutral con mutación.

Consideremos al proceso $Y = (Y_t)_{t \geq 0}$ que satisface la EDE

$$dY_t = \sqrt{Y_t(1 - Y_t)} dB_t + \frac{1}{2} [\theta_1(1 - Y_t) - \theta_2 Y_t] dt, \quad Y_0 = x \in [0, 1], \quad (3.11)$$

donde $\theta_1, \theta_2 > 0$; es decir, $Y \sim \mathbb{WF}_x(\theta_1, \theta_2)$ ². En donde Y tiene una representación probabilística de la densidad de transición dada por ([GS10], [JS17]):

$$f_t(x, y) = \sum_{m=0}^{\infty} q_m^\theta(t) \sum_{l=0}^m \mathfrak{Bin}(l; m, x) \mathfrak{Be}(y; \theta_1 + l, \theta_2 + m - l), \quad (3.12)$$

donde $\theta = \theta_1 + \theta_2$, $\mathfrak{Bin}(\cdot)$ es la función de probabilidad de una variable binomial y $\mathfrak{Be}(\cdot)$ es la función de densidad de una variable beta que están dadas por

$$\mathfrak{Bin}(l; m, x) = \binom{m}{l} x^l (1 - x)^{m-l}, \quad l = 0, 1, \dots, m,$$

$$y \quad \mathfrak{Be}(y; \theta_1, \theta_2) = \frac{\Gamma(\theta_1 + \theta_2)}{\Gamma(\theta_1)\Gamma(\theta_2)} y^{\theta_1-1} (1 - y)^{\theta_2-1}, \quad 0 < y < 1;$$

además, $q_m^\theta(t)$, $m = 0, 1, \dots$ son funciones de transición de un proceso de muerte puro $A_\infty^\theta(t)$ con frontera de entrada en el infinito, y lo veremos más adelante.

La estrategia de simulación de muestras de la densidad de transición $f_t(x, y)$ para la difusión de Wright-Fisher con mutación es inmediata de la representación (3.12), se resume el procedimiento en el Algoritmo 5 ([JS17], Algorithm 1), y es el método que apareció primero en [GL83] para una población con K alelos.

Algoritmo 5: Simulación de muestras exactas de la densidad de transición $f_t(x, \cdot)$ para la difusión de Wright-Fisher neutral con mutación $\mathbb{WF}_x(\theta_1, \theta_2)$

Init: $\theta_1, \theta_2 > 0$, $\theta = \theta_1 + \theta_2$; $Y_0 = x \in [0, 1]$, $t > 0$

- 1 Simular $M \sim A_\infty^\theta(t)$ de la distribución de $\{q_m^\theta(t) : m = 0, 1, \dots\}$. ▷ Alg.6 o 6'
 - 2 Dado $M = m$, simular $L \sim \mathfrak{Bin}(m, x)$.
 - 3 Dado $L = l$, simular $Y \sim \mathfrak{Be}(\theta_1 + l, \theta_2 + M - l)$.
 - 4 **return** Y
-

Cabe mencionar que el modelo Beta-Binomial que describe en [MW09] (Sec. 5) como la representación de la transición del modelo de Wright-Fisher, es uno distinto de la representación (3.12), con la diferencia en la función de probabilidad $q_m^\theta(t)$ del proceso ancestral $A_\infty^\theta(t)$.

² Usamos la notación de [JS17] para la difusión de Wright-Fisher, en donde se difiere de (3.10) por un factor de 2 y con los parámetros de mutación $\theta_1 = 2\gamma_2$, $\theta_2 = 2\gamma_1$.

Así mismo, la carga computacional en el Algoritmo 5 es en el paso 1, en donde se requiere las funciones de probabilidad $\{q_m^\theta(t) : m = 0, 1, \dots\}$, y los cuales no tienen una forma cerrada. Pero se demostró por Jenkins y Spanó ([JS17]) que es posible de simular muestras exactas de esta distribución del proceso ancestral $A_\infty^\theta(t)$, que es el proceso dual a la difusión de Wright-Fisher ([GS10]), y lo describimos brevemente en el siguiente apartado.

Simulación exacta del proceso puro de muerte A_∞^θ

Bajo el modelo de Wright-Fisher neutral con mutación para la evolución genética de una población descrito en la Sección 3.1.1, el proceso ancestral empieza con una muestra de n genes en el tiempo presente (o bien, con n linajes en el tiempo $t = 0$), y hacia el tiempo atrás, se trazan los linajes ancestrales de la muestra hasta el ancestro común más reciente. Donde un linaje en una generación particular en el tiempo pasado está representado por un individuo cuyo genoma contiene material directamente ancestral a uno o más de la muestra. Tal proceso se puede representar por un árbol coalescente ([Gri06]).

Si $A_n^\theta(t)$ denota al número de linajes ancestrales no mutantes en el tiempo t hacia atrás en el proceso de coalescencia con mutación y $A_n^\theta(0) = n$, el número de linajes que pasa de m a $m - 1$ por coalescencia con tasa $\binom{m}{2} = m(m - 1)/2$ o por mutación con tasa $m\theta/2$; donde $\theta = \theta_1 + \theta_2$ con $\theta_1 \geq 0$ y $\theta_2 \geq 0$ son parámetros de mutación definidos anteriormente. Luego, para cada $n \in \mathbb{N}$, $\{A_n^\theta(t) : t \geq 0\}$ es un proceso de Markov puro de muerte que comienza en el estado n y pasa al estado $n - 1$, y así hasta el estado 1 o el 0 (0 es un estado absorbente, ya que eventualmente, se desaparece el linaje debido al efecto de muestreo y mutación). Para cada $m = 1, 2, \dots, n$, las únicas transiciones del proceso es pasar el estado m al estado $m - 1$ con tasa $m(m + \theta - 1)/2$ (esto es, se pierde la arista en el árbol coalescente por coalescencia o por mutación) (cf. [Gri06], [Tav84]).

Cuando $n = \infty$, el proceso coalescente también se conoce como el coalescente de Kingman con mutación ([Eth11]), y $A_\infty^\theta(t)$ es un proceso puro de muerte con frontera de entrada en ∞ c.s.; sea $q_m^\theta(t) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n^\theta(t) = m)$, entonces $q_m^\theta(t)$ es la función de probabilidad de $A_\infty^\theta(t)$ y tiene una expansión en serie alternante infinita ([Gri80], [Tav84]):

$$q_m^\theta(t) = \mathbb{P}(A_\infty^\theta(t) = m) = \sum_{k=m}^{\infty} (-1)^{k-m} b_k^{(t,\theta)}(m), \quad m = 0, 1, 2, \dots, \quad (3.13)$$

$$\text{donde} \quad b_k^{(t,\theta)}(m) = \begin{cases} 1 & \text{si } m = k = 0, \\ \binom{k}{m} \frac{(\theta+2k-1)}{k!} \frac{\Gamma(\theta+m+k-1)}{\Gamma(\theta+m)} e^{-k(k+\theta-1)t/2} & \text{en otros casos con } m \leq k. \end{cases}$$

Cuando $\theta = 0$, $A_\infty(t)$ es el número de aristas en tiempo t hacia atrás en el árbol coalescente de Kingman, y la fórmula anterior sigue cumpliendo para $q_m^0(t) = \mathbb{P}(A_\infty(t) = m)$, mientras que $q_0^0(t) = 0$, ya que en este caso, siempre existe al menos un linaje en la población.

Notemos que la probabilidad $q_m^\theta(t)$ no tiene una expresión de forma cerrada, sin embargo, podemos simular muestras exactas de la distribución discreta de la variable $A_\infty^\theta(t)$ mediante un método de series alternantes como se describe en [JS17]. En donde se requiere la sucesión de los coeficientes de $q_m^\theta(t)$ dada por (3.13) tal que $b_k^{(t,\theta)}(m) \downarrow 0$ cuando $k \rightarrow \infty$ para cada m ; aunque tal condición no siempre se cumple, pero la Proposición 1 de [JS17] afirma que si se define

$$C_m^{(t,\theta)} := \inf \left\{ i \geq 0 : b_{i+m+1}^{(t,\theta)}(m) < b_{i+m}^{(t,\theta)}(m) \right\}, \quad (3.14)$$

entonces $C_m^{(t,\theta)} < \infty$ para todo m , y $C_m^{(t,\theta)} = 0$ para todo $m > D_0^{(t,\theta)}$, donde

$$D_0^{(t,\theta)} := \inf \left\{ k \geq \left(\frac{1}{t} - \frac{\theta+1}{2} \right) \vee 0 : (\theta+2k+1)e^{-\frac{(2k+\theta)t}{2}} < 1 \right\}; \quad (3.15)$$

además, la sucesión $b_k^{(t,\theta)}(m) \downarrow 0$ cuando $k \rightarrow \infty$, para todo $k \geq m + C_m^{(t,\theta)}$.

Luego, para cada $M \in \mathbb{N}$, existe $k_m \in \mathbb{N}$ para cada $m \in \{0, 1, \dots, M\}$, y denotamos al vector $\mathbf{k} = (k_0, k_1, \dots, k_M) \in \mathbb{N}^{M+1}$ tal que

$$S_{\mathbf{k}}^-(M) := \sum_{m=0}^M \sum_{i=0}^{2k_{m+1}} (-1)^i b_{m+i}^{(t,\theta)}(m) \leq \sum_{m=0}^M q_m^\theta(t) \leq \sum_{m=0}^M \sum_{i=0}^{2k_m} (-1)^i b_{m+i}^{(t,\theta)}(m) =: S_{\mathbf{k}}^+(M), \quad (3.16)$$

Más aún, se tiene que

$$\limsup_{\mathbf{k} \rightarrow (\infty, \dots, \infty)} S_{\mathbf{k}}^-(M) = \mathbb{P}(A_\infty^\theta(t) \leq M) \quad \text{y} \quad \liminf_{\mathbf{k} \rightarrow (\infty, \dots, \infty)} S_{\mathbf{k}}^+(M) = \mathbb{P}(A_\infty^\theta(t) \leq M),$$

y tanto $S_{\mathbf{k}}^-(M)$ como $S_{\mathbf{k}}^+(M)$ se pueden calcular a partir de un número finito de muchos términos. Entonces, dado $U \sim \mathcal{U}(0, 1)$, podemos encontrar un vector $\mathbf{k}^0 \in \mathbb{N}^{M+1}$ con elementos k_m^0 tal que para cada $m \in \{0, \dots, M\}$,

$$k_m^0 := \inf \left\{ k_m \in \mathbb{N} : S_{\mathbf{k}}^-(M) > U \text{ o } S_{\mathbf{k}}^+(M) < U \right\},$$

Ahora, si \mathbf{k}^0 es tal que $S_{\mathbf{k}^0}^-(M) > U$, la cantidad

$$M := \inf \left\{ M \in \mathbb{N} : \sum_{m=0}^M q_m^\theta(t) \geq S_{\mathbf{k}^0}^-(M) > U \right\}$$

está distribuida exactamente de acuerdo a $\{q_m^\theta : m = 0, 1, \dots\}$ por el método de la transformada inversa; esto es, $M \sim A_\infty^\theta(t)$.

Así, resumimos el procedimiento de la simulación exacta del proceso puro de muerte $A_\infty^\theta(t)$ en el Algoritmo 6 ([JS17], Algorithm 2). En la sección 3.2 de [GPHK19] se incorporaron algunas consideraciones que mejoran este algoritmo en reducir el tiempo de cómputo. Sin embargo, en el momento de la implementación del algoritmo ([GPHK19], Algorithm 3), ocasionalmente queda indefinido un ciclo por el signo negativo (ver Apéndice B).

Algoritmo 6: Simulación exacta de muestras de $q_m^\theta(t)$, función de transición del proceso puro de muerte $A_\infty^\theta(t)$

Init: $M \leftarrow 0, k_0 \leftarrow 0, \mathbf{k} \leftarrow (k_0)$

1 Simular $U \sim \mathcal{U}(0, 1)$.

2 **repeat**

3 **for all** $m \in \{0, \dots, M\}$ **do**

4 $k_m \leftarrow \lceil C_m^{(t,\theta)} / 2 \rceil$

▷ Ec. (3.14)^a

5 **end for**

6 **while** $S_{\mathbf{k}}^-(M) < U < S_{\mathbf{k}}^+(M)$ **do**

7 $\mathbf{k} \leftarrow \mathbf{k} + (1, 1, \dots, 1)$

▷ Ec. (3.16)

8 **end while**

9 **if** $S_{\mathbf{k}}^-(M) > U$ **then**

10 **return** M

11 **else if** $S_{\mathbf{k}}^+(M) < U$ **then**

12 $\mathbf{k} \leftarrow (k_0, k_1, \dots, k_M, 0)$

13 $M \leftarrow M + 1$

14 **end if**

15 **until** *false*

^a $\lceil x \rceil$ denota a la función techo para $x \in \mathbb{R}$, i.e., $\lceil x \rceil = \min\{k \in \mathbb{Z} \mid x \leq k\}$.

Una observación sobre la práctica de la implementación del Algoritmo 6, es que el Algoritmo 6 falla en los casos cuando $t < 0.05$, y por lo tanto, no es conveniente aplicar el Algoritmo 6 para el paso 1 del Algoritmo 5 cuando el tiempo t sea muy pequeño. Sin embargo, $A_\infty^\theta(t)$ tiene una distribución normal asintótica ([Gri84], [JS17]), de manera que cuando $t \rightarrow 0$, el proceso $A_\infty^\theta(t)$ converge en distribución a una distribución normal $\mathcal{N}(\mu^{(t,\theta)}, \sigma^{(t,\theta)})$, siempre y cuando $\vartheta = \frac{1}{2}(\theta - 1)t$ está acotada por arriba. En donde

$$\mu^{(t,\theta)} = \frac{2\eta}{t}, \quad \text{y} \quad (\sigma^{(t,\theta)})^2 = \begin{cases} \frac{2\eta}{t\vartheta^2}(\eta + \vartheta)^2 \left(1 + \frac{\eta}{\eta + \vartheta} - 2\eta\right), & \vartheta \neq 0, \\ \frac{2}{3t}, & \vartheta = 0, \end{cases} \quad (3.17)$$

donde $\eta = \frac{\vartheta}{e^{\vartheta}-1}$ si $\vartheta \neq 0$ y $\eta = 1$ en otros casos.

Por lo tanto, podemos aplicar esta aproximación en la implementación del Algoritmo 5 cuando t sea muy pequeño (en particular, $t < 0.05$), esto es, reemplazar el paso 1 del Algoritmo 5 cuando $t < 0.05$ por ([JS17], Sec.4):

Alg. 6'. Simular $A_\infty^\theta(t) \sim \mathcal{N}(\mu^{(t,\theta)}, \sigma^{(t,\theta)})$ y lo redondea al entero no negativo más cercano.

Al implementar el Algoritmo 5 usando las consideraciones anteriores (ver Alg.5), podemos obtener muestras de la trayectoria de un proceso de difusión de Wright-Fisher con mutación mediante su densidad de transición. Siguiendo el mismo método para la simulación de trayectoria del proceso de Bessel (Sec. 2.2.1), y aplicando el Algoritmo 5 para cada punto de la partición de un intervalo de tiempo $[0, t]$, se generan muestras de la trayectoria de una difusión de Wright-Fisher y se obtiene un esqueleto para la trayectoria en $[0, t]$.

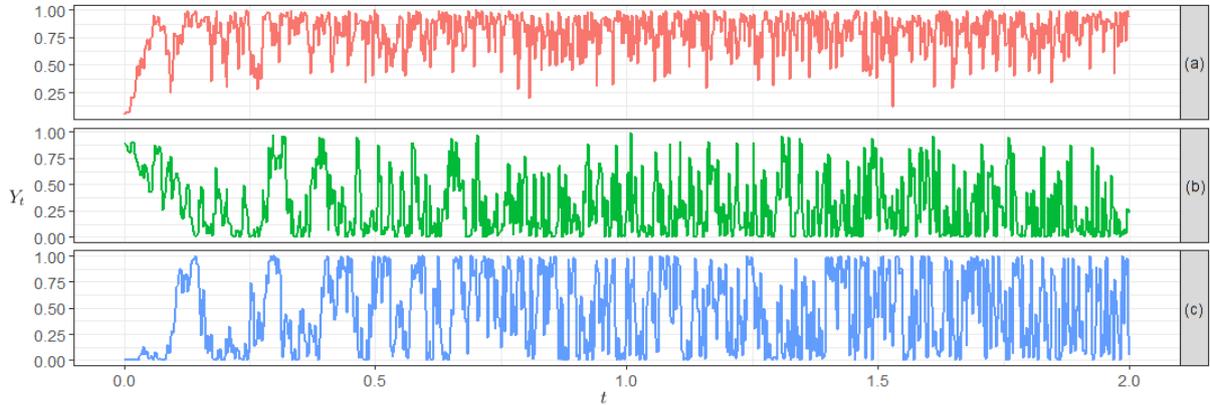


Figura 3.1: Muestras de trayectoria de una difusión de Wright-Fisher Y_t , $t \in [0, 2]$ ($\Delta t = 0.002$) simuladas usando el Algoritmo 5 con parámetros y valor inicial: (a) $\theta_1 = 4$, $\theta_2 = 1$ y $x = 0.05$, (b) $\theta_1 = 0.5$, $\theta_2 = 1.3$ y $x = 0.9$ y (c) $\theta_1 = \theta_2 = 0.5$ y $x = 0$.

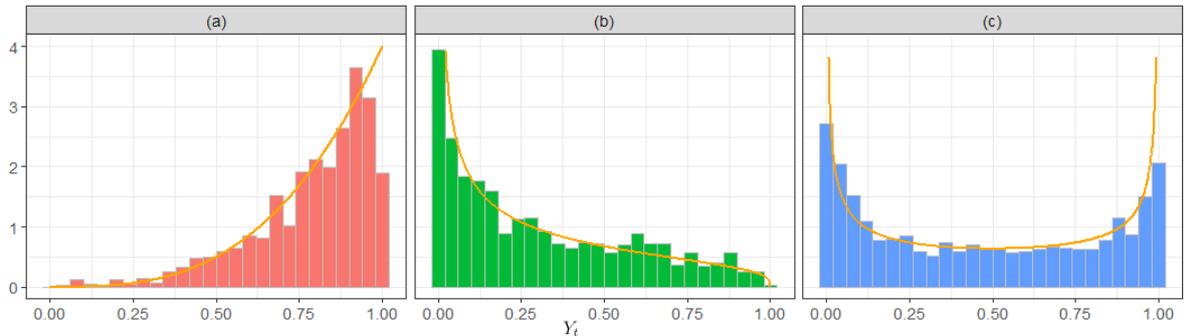


Figura 3.2: Distribución de la muestra de trayectoria Y_t que corresponde a cada uno de los casos de la Figura 3.1, comparando contra la distribución estacionaria $\mathfrak{B}_e(\theta_1, \theta_2)$ (—).

En la Figura 3.1 se presentan muestras de trayectoria de $Y \sim \mathbb{WF}_x(\theta_1, \theta_2)$, simulada usando el Algoritmo 5, en donde representan la frecuencia del alelo A durante el tiempo

$t \in [0, 2]$, con distintos parámetros de mutación $\theta_1, \theta_2 > 0$ y distintos valores iniciales $x \in [0, 1]$. Mientras que en la Figura 3.2, se ilustra la distribución de muestra de trayectoria simulada en cada caso de la Figura 3.1 y contra la distribución estacionaria $\mathfrak{Be}(\theta_1, \theta_2)$.

Observación 3.8. Observamos cada caso de la Figura 3.1, la trayectoria toma valores en $[0, 1]$ y durante la mayor parte del tiempo, contiene al valor medio $\mu = \frac{\theta_1}{\theta_1 + \theta_2}$ de la distribución estacionaria $\mathfrak{Be}(\theta_1, \theta_2)$.

Observación 3.9. Por la Observación 3.7, para una difusión de Wright-Fisher $Y \sim \mathbb{WF}_x(\theta_1, \theta_2)$, tenemos que cuando $\theta_1 \geq 1$, una muestra de trayectoria de Y_t nunca toca a 0 para $t > 0$ y nunca toca a 1 cuando $\theta_2 \geq 1$. La propiedad que podemos notar en la Figura 3.1.

Capítulo 4

Proyección del movimiento browniano $\mathbf{MB}(\mathcal{S}^{d-1})$

Por el teorema de Arquímedes y su generalización en las esferas unitarias en el espacio euclidiano de dimensión $d \geq 3$, vimos en el Capítulo 1 que la proyección ortogonal \mathcal{P} de la esfera unitaria \mathcal{S}^{d-1} en la bola \mathcal{B}^{d-2} preserva la medida uniforme (Thm. 1.9). Además, como la distribución uniforme en la esfera \mathcal{S}^{d-1} es la única distribución en la esfera invariante bajo rotaciones, y así, es una medida invariante para movimientos brownianos en la esfera. El trabajo de [MMUB18] investigó el proceso \mathbf{X} obtenido de la proyección del movimiento browniano $\mathbf{MB}(\mathcal{S}^{n+k-1})$ en la bola \mathcal{B}^n , y tal proceso tiene una descomposición skew-product análoga a la del movimiento browniano en \mathbb{R}^d que vimos en la Sección 2.3.2. Además, con tal descomposición skew-product, es posible dar la medida invariante para el proceso \mathbf{X} , y veremos en la Proposición 4.1 que esta medida invariante admite la densidad igual a la de la distribución uniforme $\mathcal{U}(\mathcal{S}^{n+k-1})$ proyectada en la bola \mathcal{B}^n que vimos en el Capítulo 1, Proposición 1.10. Citaremos estos resultados principales de [MMUB18] en la Sección 4.1.

También, el movimiento browniano \mathbf{Z} en la esfera \mathcal{S}^{d-1} tiene una descomposición skew-product que vimos en la Sección 2.3.3 está expresada en coordenadas polares $(r, \boldsymbol{\theta})$. Con la transformación a coordenadas cartesianas, se tiene que $\mathbf{Z} = ((\text{sen } r)\boldsymbol{\theta}^\top, \cos r)$, y vimos que con una transformación lineal, la última coordenada de \mathbf{Z} se relaciona con la difusión de Wright-Fisher. Veremos en la Sección 4.2 que se puede dar la descomposición skew-product para \mathbf{Z} en términos de una difusión de Wright-Fisher y un movimiento browniano en \mathcal{S}^{d-2} . Y esta descomposición es igual a la dada en [MMUB20], en donde se utilizaron los resultados de la proyección del movimiento browniano esférico en [MMUB18].

Finalmente, con esta descomposición skew-product del movimiento browniano \mathbf{Z} en la esfera \mathcal{S}^{d-1} expresada en términos de la difusión de Wright-Fisher y del movimiento browniano en \mathcal{S}^{d-2} , y usando la simulación exacta de muestra de la difusión Wright-Fisher dada en el Algoritmo 5, es posible simular muestras exactas de un movimiento browniano \mathbf{Z} . En la Sección 4.3 revisamos el procedimiento de simulación exacta del movimiento browniano \mathbf{Z} presentado en [MMUB20] y finalmente, implementamos el algoritmo para la simulación de las trayectorias.

4.1. Proyección del movimiento browniano esférico

Supongamos que $\mathbf{Z} = (\mathbf{Z}_t)_{t \geq 0}$ es un movimiento browniano en la esfera \mathcal{S}^{n+k-1} que satisface la EDE (ver Sección 2.3 y (2.40)):

$$d\mathbf{Z}_t = (I - \mathbf{Z}_t \mathbf{Z}_t^\top) d\tilde{\mathbf{B}}_t - \frac{n+k-1}{2} \mathbf{Z}_t dt, \quad \mathbf{Z}_0 \in \mathcal{S}^{n+k-1}, \quad (4.1)$$

donde $n, k \in \mathbb{Z}^+$, I denota la matriz identidad de dimensión adecuada y $\tilde{\mathbf{B}}$ es un movimiento browniano en \mathbb{R}^{n+k} . Si $\mathbf{X} = (\mathbf{X}_t)_{t \geq 0}$ es el proceso obtenido aplicando la proyección ortogonal $\mathcal{P}_k: \mathcal{S}^{n+k-1} \rightarrow \mathcal{B}^n$ de codimensión k sobre el movimiento browniano \mathbf{Z} , entonces \mathbf{X} satisface una ecuación diferencial estocástica en la bola unitaria \mathcal{B}^n . Es decir, el siguiente resultado:

Proposición 4.1 ([MMUB18], Prop. 1.1). *Si denotamos por \mathbf{X} a las primeras n coordenadas de \mathbf{Z} . Entonces, existe un movimiento browniano \mathbf{B} en \mathbb{R}^n tal que la pareja (\mathbf{X}, \mathbf{B}) satisface la EDE*

$$d\mathbf{X}_t = \sigma(\mathbf{X}_t) d\mathbf{B}_t - \frac{n+k-1}{2} \mathbf{X}_t dt, \quad \mathbf{X}_0 = (Z_0^1, \dots, Z_0^n) \in \mathcal{B}^n, \quad (4.2)$$

donde la función matricial $\sigma(\mathbf{x})$ con valores en $\mathbb{R}^{n \times n}$ está dada por

$$\sigma(\mathbf{x}) = I - \left(1 - \sqrt{1 - \|\mathbf{x}\|^2}\right) \frac{\mathbf{x}\mathbf{x}^\top}{\|\mathbf{x}\|^2} \mathbf{1}_{\{\|\mathbf{x}\| > 0\}}, \quad \mathbf{x} \in \mathcal{B}^n. \quad (4.3)$$

La solución (\mathbf{X}, \mathbf{B}) a la ecuación (4.2) es única por trayectorias, y \mathbf{X} es un proceso de Markov fuerte con una única medida invariante que admite la densidad

$$f(\mathbf{x}) = \frac{\Gamma\left(\frac{n+k}{2}\right)}{\pi^{\frac{n}{2}} \Gamma\left(\frac{k}{2}\right)} \left(1 - \|\mathbf{x}\|^2\right)^{\frac{k}{2}-1}, \quad \forall \mathbf{x} \in \mathcal{B}^n. \quad (4.4)$$

Más aún, $Y = \|\mathbf{X}\|^2$ satisface la EDE

$$dY_t = 2\sqrt{Y_t(1-Y_t)} d\beta_t + [n(1-Y_t) - kY_t] dt, \quad Y_0 = \|\mathbf{X}_0\|^2 \in [0, 1], \quad (4.5)$$

donde β es un movimiento browniano escalar. Esto es, Y es una difusión de Wright-Fisher neutral con mutación.¹

Observación 4.1. La función matricial $\sigma(\mathbf{x})$ definida por (4.3) es acotada y continua en la bola \mathcal{B}^n , además, por simples cálculos, se obtiene $\sigma(\mathbf{x})\sigma(\mathbf{x})^\top = I - \mathbf{x}\mathbf{x}^\top$ para todo $\mathbf{x} \in \mathcal{B}^n$.

Demostración. Se puede encontrar la demostración rigurosa de la proposición en el trabajo de [MMUB18], a continuación se da la idea general.

Primero, para llegar a la EDE (4.2) a partir de la EDE (4.1) para \mathbf{Z} , denotemos por \mathbf{X}' a las últimas k coordenadas de \mathbf{Z} , y análogamente escribimos a $\tilde{\mathbf{B}} = (\tilde{\mathbf{B}}^1, \tilde{\mathbf{B}}^2)$ que separa en sus primeras n , y sus últimas k coordenadas. Entonces el proceso \mathbf{B} dado por

$$\mathbf{B}_t = \int_0^t \sigma(\mathbf{X}_s) d\tilde{\mathbf{B}}_s^1 + \int_0^t \left(-\mathbf{X}_s \mathbf{X}'_s{}^\top (1 - \|\mathbf{X}_s\|^2)^{-\frac{1}{2}} \mathbf{1}_{\{\|\mathbf{X}_s\| < 1\}} + \mathbf{X}_s \mathbf{z}^\top \mathbf{1}_{\{\|\mathbf{X}_s\| = 1\}} \right) d\tilde{\mathbf{B}}_s^2,$$

es un movimiento browniano en \mathbb{R}^n , donde $\mathbf{z} \in \mathcal{S}^{k-1}$ es un vector unitario arbitrario.

En efecto, notemos que $\|\mathbf{X}'_t\|^2 = \|\mathbf{Z}_t\|^2 - \|\mathbf{X}_t\|^2 = 1 - \|\mathbf{X}_t\|^2$, y consideremos a la matriz de dimensión $n \times (n+k)$:

$$A_t := \left[\sigma(\mathbf{X}_t), -\mathbf{X}_t \mathbf{X}'_t{}^\top (1 - \|\mathbf{X}_t\|^2)^{-\frac{1}{2}} \mathbf{1}_{\{\|\mathbf{X}_t\| < 1\}} + \mathbf{X}_t \mathbf{z}^\top \mathbf{1}_{\{\|\mathbf{X}_t\| = 1\}} \right] =: [\sigma(\mathbf{X}_t), D_t].$$

La elección del vector constante \mathbf{z} no es relevante en la definición de A_t , como veremos que el tiempo en que \mathbf{X} toca a la frontera de la bola tiene una medida de Lebesgue cero.

¹ Es diferente de la EDE para la difusión de Wright-Fisher con mutación dada en (3.11) por un factor 2, pero puede llegar a la misma forma por un cambio de tiempo acelerado $t \mapsto 4t$.

Como $\sigma(\mathbf{X}_t)^2 = \sigma(\mathbf{X}_t)\sigma(\mathbf{X}_t)^\top = 1 - \mathbf{X}_t\mathbf{X}_t^\top$, y también se obtiene $D_t D_t^\top = \mathbf{X}_t\mathbf{X}_t^\top$ por simples cálculos matriciales, entonces $A_t A_t^\top = \sigma(\mathbf{X}_t)\sigma(\mathbf{X}_t)^\top + D_t D_t^\top = I$. Usando la notación de A_t y se tiene que \mathbf{B} está definido por $\mathbf{B}_t = \int_0^t A_s d\tilde{\mathbf{B}}_s$, lo cual es una martingala local con variación cuadrática $\langle B^i, B^j \rangle_t = \int_0^t (A_s A_s^\top)_{ij} ds = \delta_{ij}t$, entonces por el teorema de caracterización de Lévy, concluimos que \mathbf{B} es un movimiento browniano de dimensión n .

Como otros cálculos muestra que $\sigma(\mathbf{X}_t)D_t = -\mathbf{X}_t\mathbf{X}_t^\top \mathbb{1}_{\{0 < \|\mathbf{X}_t\| < 1\}}$, y por la definición de \mathbf{Z} dada en (4.1), se obtiene la EDE que satisface por \mathbf{X} :

$$\begin{aligned} d\mathbf{X}_t &= (I - \mathbf{X}_t\mathbf{X}_t^\top) d\tilde{\mathbf{B}}_t^1 - \mathbf{X}_t\mathbf{X}_t^\top d\tilde{\mathbf{B}}_t^2 - \frac{n+k-1}{2}\mathbf{X}_t dt \\ &= \sigma(\mathbf{X}_t)^2 d\tilde{\mathbf{B}}_t^1 + \sigma(\mathbf{X}_t)D_t d\tilde{\mathbf{B}}_t^2 - \frac{n+k-1}{2}\mathbf{X}_t dt \\ &= \sigma(\mathbf{X}_t)A_t d\tilde{\mathbf{B}}_t - \frac{n+k-1}{2}\mathbf{X}_t dt = \sigma(\mathbf{X}_t) d\mathbf{B}_t - \frac{n+k-1}{2}\mathbf{X}_t dt. \end{aligned}$$

Luego, para la unicidad de solución fuerte a la ecuación anterior y las últimas afirmaciones de la proposición, se requieren los siguientes resultados sobre la caracterización de un proceso \mathbf{X} con valores en la bola \mathcal{B}^n que satisface una EDE de tipo (4.6):

Teorema 4.2 ([MMUB18], Thm. 1.5). Sean $n \geq 2$ y \mathbf{X} una solución a la EDE

$$d\mathbf{X}_t = \gamma(\|\mathbf{X}_t\|)\sigma(\mathbf{X}_t) d\mathbf{B}_t - g(\|\mathbf{X}_t\|)\mathbf{X}_t dt, \quad \mathbf{X}_0 = \mathbf{x} \in \mathcal{B}^n, \quad (4.6)$$

donde $\gamma: [0, 1] \rightarrow (0, \infty)$ y $g: [0, 1] \rightarrow \mathbb{R}$ son Lipschitz continuas que satisface $\frac{g(1)}{\gamma^2(1)} \geq \frac{n-1}{2}$. Toma $s \in [0, \infty)$, y suponga que $s > 0$ o $s = 0$ y $\mathbf{X}_0 \neq \mathbf{0}$. Sea $R := \|\mathbf{X}_0\|$ el componente radial de \mathbf{X} , entonces el proceso $\hat{\mathbf{V}} = (\hat{\mathbf{V}}_t)_{t \geq 0}$ dada por $\hat{\mathbf{V}}_t := \frac{\mathbf{X}_{T_s(t)}}{R_{T_s(t)}}$, es un movimiento browniano en la esfera \mathcal{S}^{n-1} con valor inicial $\hat{\mathbf{V}}_0 = \mathbf{X}_s/R_s$ e independiente de R , donde $S_s(t) := \int_s^t \frac{\gamma^2(R_u)}{R_u^2} du$ es un cambio de escala de tiempo que satisface $\lim_{t \rightarrow \infty} S_s(t) = \infty$, y con la inversa $T_s: [0, \infty) \rightarrow [s, \infty)$.

Por lo tanto, el proceso \mathbf{X} tiene una representación skew-product dada por

$$\mathbf{X}_t = R_t \hat{\mathbf{V}}_{S_s(t)}, \quad t \geq s.$$

Más aún, si $\mathbf{X}_0 = \mathbf{0}$, entonces $\hat{\mathbf{V}}_t$ está distribuida uniformemente en \mathcal{S}^{n-1} para cualquier $t > 0$ y en consecuencia, es un movimiento browniano estacionario en la esfera.

Corolario 4.3 ([MMUB18], Corol. 1.6). La unicidad en ley se cumple para la EDE (4.6).

Teorema 4.4 ([MMUB18], Thm. 1.7). Si $\frac{g(1)}{\gamma^2(1)} - \frac{n-1}{2} > \sqrt{2} - 1 \approx 0.4142$, entonces la unicidad por trayectorias se cumple para la EDE (4.6).

Lema 4.5 ([MMUB18], Lem. 2.1). Sea \mathbf{X} una solución a la EDE (4.6) con $\frac{g(1)}{\gamma^2(1)} \geq \frac{n-1}{2}$. Entonces el proceso $Y = \|\mathbf{X}\|^2$ satisface la EDE

$$dY_t = 2\tilde{\gamma}(Y_t)\sqrt{Y_t(1-Y_t)} d\tilde{\beta}_t + \tilde{\gamma}^2(Y_t) \left(n(1-Y_t) - \left(\frac{2\tilde{g}(Y_t)}{\tilde{\gamma}^2(Y_t)} - (n-1) \right) Y_t \right) dt, \quad (4.7)$$

donde $\tilde{g}(y) := g(\sqrt{y})$, $\tilde{\gamma} := \gamma(\sqrt{y})$, y $\tilde{\beta}$ es un \mathcal{F}_t -movimiento browniano escalar definido por

$$\tilde{\beta}_t = \sum_{i=1}^n \int_0^t \frac{X_u^i}{\sqrt{Y_u}} \mathbb{1}_{\{\omega: Y_u > 0\}}(u, \omega) dB_u^i + \int_0^t \mathbb{1}_{\{\omega: Y_u = 0\}}(u, \omega) d\xi_u,$$

donde ξ es un movimiento browniano escalar independiente de \mathbf{B} .

Para $n \geq 2$, el proceso Y nunca toca al 0 y si $\frac{g(r)}{\gamma^2(r)} \geq \frac{n-1}{2} + 1$ se cumple para r alrededor de 1, entonces Y nunca toca al 1.

Observación 4.2. Como la norma euclidiana $\|\cdot\|$ es Lipschitz continua, y también lo son las funciones $\gamma(\cdot)$ y $g(\cdot)$ por definición, entonces $\gamma(\|\cdot\|)$ y $g(\|\cdot\|)$ en la ecuación (4.6) son Lipschitz continuas. Además, como $\sigma(\mathbf{x})$ es acotada y continua en la bola \mathcal{B}^n , se tiene que los coeficientes de la EDE (4.6) son continuas y por lo tanto, se cumple la existencia débil de solución a la EDE ([IW89], Chap. IV, Thm. 2.3).

Observación 4.3. La condición $\frac{g(1)}{\gamma^2(1)} \geq \frac{n-1}{2}$ es necesaria para que una solución a la ecuación (4.6) quede en la bola unitaria \mathcal{B}^n .

Observación 4.4. El Teorema 4.2 dice que el proceso \mathbf{X} disfruta una descomposición skew-product análoga a la del movimiento browniano en \mathbb{R}^n (ver Sec. 2.3.2). La demostración está dada detalladamente en [MMUB18], y hemos usado las mismas técnicas para la demostración de la descomposición skew-product del movimiento browniano en \mathbb{R}^d (Prop. 2.31) y en la esfera \mathcal{S}^{d-1} (Prop. 2.32).

En el caso cuando $\mathbf{X}_0 = \mathbf{0}$, i.e., $R_0 = 0$, se tiene que $\lim_{s \rightarrow 0^+} S_s(t) = \infty$ para cualquier $t > 0$ ([MMUB18], Lem. 2.3), esto significa que la muestra de trayectoria del proceso \mathbf{X}_t que entra desde $\mathbf{0}$ viene en *spinning*, y la parte angular $\widehat{\mathbf{V}}_t$ de \mathbf{X}_t tiene que ser una distribución uniforme en la esfera para cualquier $t > 0$ ([IM65], Sec. 7.16, Pag. 275).

Observación 4.5. Por el Teorema 4.2, el proceso \mathbf{X} puede expresarse en su descomposición skew-product como una funcional medible de una pareja de procesos independientes $(R, \widehat{\mathbf{V}})$ con distribuciones dadas, por lo que es inmediato el Corolario 4.3 de la unicidad débil de solución a la ecuación (4.6).

Observación 4.6. Como la función σ es localmente Lipschitz solamente en el interior de la bola \mathcal{B}^n , entonces por el teorema clásico para la unicidad de solución fuerte a las EDEs (ver [IW89], Chap. IV, Thm. 3.1), la unicidad por trayectorias de solución a la EDE (4.6) se cumple para \mathbf{X} hasta su primer tiempo de golpe a la frontera. Sin embargo, por el Lema 4.5, si se cumple $\frac{g(r)}{\gamma^2(r)} \geq \frac{n-1}{2} + 1$ para r suficientemente cercano a 1 y $\|\mathbf{X}_0\| < 1$, \mathbf{X} nunca visitará a la frontera de la bola \mathcal{B}^n , entonces se cumple la unicidad por la trayectoria. Si $\frac{g(1)}{\gamma^2(1)} = \frac{n-1}{2}$, entonces después del primer tiempo de golpe a la frontera, \mathbf{X} se comporta como un movimiento browniano en la esfera \mathcal{S}^{n-1} en una escala de tiempo-cambiado, por lo tanto, también se cumple la unicidad por trayectorias. Si $n = 1$, la EDE (4.6) se simplifica a $dX_t^1 = \gamma(|X_t^1|)\sqrt{1 - (X_t^1)^2}dB_t^1 - g(|X_t^1|)X_t^1 dt$, entonces se sigue la unicidad por trayectorias por el criterio de Yamada-Watanabe (Prop. 2.24).

Para otros casos, como se establece en el Teorema 4.4, la unicidad de solución fuerte a la ecuación (4.6) es más complicada, y se puede consultar la Sección 2.2 de [MMUB18] para su demostración completa.

Observación 4.7. La ecuación (4.7) es la aplicación directa de la fórmula de Itô (Thm. 2.21) al proceso $Y = \|\mathbf{X}\|^2$ para la función diferenciable $f(\mathbf{x}) = \|\mathbf{x}\|^2$. Considerando un cambio de escala de tiempo que no afecta en las propiedades de entrada en la frontera, definiendo

$$l_t := \int_0^t 4\gamma^2(\|\mathbf{X}_u\|) du, \quad \text{y su inversa} \quad \tau_t := \inf\{u \geq 0 : l_u > t\}.$$

Entonces $\widehat{Y}_t := Y_{\tau_t}$ satisface la EDE

$$d\widehat{Y}_t = \sqrt{\widehat{Y}_t(1 - \widehat{Y}_t)} d\widehat{\beta}_t + \frac{1}{4} \left(n(1 - \widehat{Y}_t) - \left(\frac{2\widetilde{g}(\widehat{Y}_t)}{\widetilde{\gamma}^2(\widehat{Y}_t)} - (n-1) \right) \widehat{Y}_t \right) dt, \quad (4.8)$$

donde $\widehat{\beta}_t = \int_0^{\tau_t} 2\gamma(\|\mathbf{X}_u\|) d\widetilde{\beta}_u$ también es un movimiento browniano escalar. Como el coeficiente de deriva es Lipschitz continua, y el coeficiente de difusión es Hölder continua con

exponente $\frac{1}{2}$, por el criterio de Yamada-Watanabe (Prop. 2.24), se obtiene la unicidad de solución fuerte a la EDE (4.8). Por otro lado, notemos que la EDE (4.8) se parece a la de la difusión de Wright-Fisher (3.11) con la única diferencia en el término de la deriva. Por lo que, si denotamos

$$M := \max_{y \in [0,1]} \left(\frac{2\tilde{g}(y)}{\tilde{\gamma}^2(y)} - (n-1) \right) \geq 0,$$

y aplicando el teorema de comparación para EDEs ([RY99], Chap.IX, Thm. 3.7), se tiene que la solución de la EDE (4.8) siempre es mayor que $\mathbb{WF}_x\left(\frac{n}{2}, \frac{M}{2}\right)$ con el mismo valor inicial x . Por las propiedades de una trayectoria de la difusión de Wright-Fisher (Observación 3.7), notemos que si $n \geq 2$, los procesos \hat{Y} y Y nunca toca a 0 para $t > 0$. Similarmente, si se cumple $\frac{g(r)}{\gamma^2(r)} \geq \frac{n-1}{2} + 1$ cerca de 1, entonces podemos usar el teorema de comparación por lo menos localmente (cerca de 1) para demostrar que \hat{Y} es menor que $\mathbb{WF}_x\left(\frac{n}{2}, 1\right)$ con el mismo valor inicial x , y por lo tanto, \hat{Y} y Y nunca toca a 1 para $t > 0$.

Ahora, regresamos a la Proposición 4.1, notemos que la ecuación (4.2) es un caso especial de la EDE (4.6) con $\gamma \equiv 1$ y $g \equiv \frac{n+k-1}{2}$, y se tiene que $\frac{g(r)}{\gamma^2(r)} \equiv \frac{n-1+k}{2}$ para todo $r \in [0, 1]$, entonces $\frac{g(1)}{\gamma^2(1)} - \frac{n-1}{2} = \frac{k}{2} > \sqrt{2} - 1$ para todo $k \geq 1$. Así, por el Teorema 4.4, se cumple la unicidad de solución por trayectorias a la EDE (4.2). Por otro lado, el Lema 4.5 afirma que $Y = \|\mathbf{X}\|^2$ satisface la EDE (4.5), además, es un proceso de difusión de Wright-Fisher $\mathbb{WF}\left(\frac{n}{2}, \frac{k}{2}\right)$ en el tiempo cambiado por $t \mapsto 4t$. Y sabemos del Capítulo 3 que la medida invariante para una difusión de Wright-Fisher $\mathbb{WF}\left(\frac{n}{2}, \frac{k}{2}\right)$ está dada por la distribución $\mathfrak{B}\mathfrak{e}\left(\frac{n}{2}, \frac{k}{2}\right)$. Así, cuando $n \geq 2$, el componente radial $R = \|\mathbf{X}\|$ de la solución \mathbf{X} a la EDE (4.2) tiene una medida invariante con densidad

$$h(r) = \frac{2\Gamma\left(\frac{n+k}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{k}{2}\right)} r^{n-1} (1-r^2)^{\frac{k}{2}-1}, \quad \forall r \in [0, 1].$$

Entonces, si suponemos que la distribución inicial del proceso \mathbf{X} tiene la densidad f dada por (4.4), usando las coordenadas polares (ver Thm. 1.3) y la descomposición skew-product de \mathbf{X} del Teorema 4.2, vemos que la densidad de R_0 es h , y por ser la densidad invariante para R , R_t tiene la misma densidad h para todo $t \geq 0$. Y como el movimiento browniano en la esfera en el tiempo cambiado sigue teniendo la medida invariante uniforme, ya que el cambio de tiempo es independiente del movimiento browniano. Luego, del sistema de coordenadas polares a cartesianas, se demuestra que \mathbf{X}_t tiene densidad f para todo $t \geq 0$. Es decir, f es la densidad de la medida invariante para el proceso \mathbf{X} .

Para el caso cuando $n = 1$, tenemos que X^1 es un proceso con espacio de estados $[-1, 1]$ que satisface la EDE $dX_t^1 = \sqrt{1 - (X_t^1)^2} dB_t^1 - \frac{k}{2} X_t^1 dt$, i.e., X^1 es un proceso de difusión con coeficiente de difusión $\sigma^2(x) = (1-x^2)$ y de deriva $b(x) = -\frac{k}{2}x$, luego es posible calcular su distribución estacionaria (ver [KT81], Chap. 15, Sec. 5), obteniendo la densidad estacionaria

$$f(x) = \frac{\Gamma\left(\frac{1+k}{2}\right)}{\pi^{\frac{1}{2}}\Gamma\left(\frac{k}{2}\right)} \left(1-x^2\right)^{\frac{k}{2}-1}, \quad \forall x \in [-1, 1],$$

la cual es igual a (4.4) para el caso $n = 1$. □

4.2. Representación skew-product del $\text{MB}(\mathcal{S}^{d-1})$ y su relación con la difusión de Wright-Fisher

En esta sección, consideramos a la representación skew-product del movimiento browniano en la esfera \mathcal{S}^{d-1} que ya se mencionó anteriormente en la Sección 2.3.3. El resultado principal de [MMUB20] es una nueva versión de descomposición skew-product del movimiento browniano en la esfera que se relaciona con la difusión de Wright-Fisher $\text{WF}(\frac{d-1}{2}, \frac{d-1}{2})$, y el cual se obtiene a partir de los resultados de [MMUB18] sobre la proyección del movimiento browniano \mathcal{S}^{n+k-1} en la bola \mathcal{B}^n , es decir, los resultados que se mencionaron en la sección anterior. Sin embargo, la representación skew-product del movimiento browniano $\mathbf{Z} = (r, \boldsymbol{\theta})$ que estudiamos en la Sección 2.3.3 está dada en coordenadas polares, por la transformación a coordenadas cartesianas $\mathbf{Z} = ((\sin r)\boldsymbol{\theta}^\top, \cos r)^\top$. Haciendo la transformación lineal para el componente radial $\frac{1-\cos r}{2}$, usando la fórmula de Itô, podemos obtener la difusión de Wright-Fisher $\text{WF}(\frac{d-1}{2}, \frac{d-1}{2})$ (Observación 2.18).

En consecuencia, primero retomamos el resultado de la representación skew-product del movimiento browniano en la esfera \mathcal{S}^{d-1} de la Sección 2.3.3:

Sea $\mathbf{Z} = (\mathbf{Z}_t)_{t \geq 0}$ un movimiento browniano en la esfera \mathcal{S}^{d-1} y definido en el espacio de probabilidad completo filtrado $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ que satisface la EDE

$$d\mathbf{Z}_t = (I - \mathbf{Z}_t \mathbf{Z}_t^\top) d\tilde{\mathbf{B}}_t - \frac{d-1}{2} \mathbf{Z}_t dt, \quad \mathbf{Z}_0 \in \mathcal{S}^{d-1} \setminus \{\mathbf{o}, -\mathbf{o}\}, \quad (4.9)$$

donde $\tilde{\mathbf{B}}$ es un (\mathcal{F}_t) -movimiento browniano en \mathbb{R}^d , y $\mathbf{o} = (0, \dots, 0, 1)^\top$, $-\mathbf{o} = (0, \dots, 0, -1)^\top$ son el polo de la esfera \mathcal{S}^{d-1} y su antipodal, respectivamente.

Por la Proposición 2.32, \mathbf{Z} tiene una representación skew-product $\mathbf{Z}_t = (r_t, \boldsymbol{\theta}_t)$, donde $r = (r_t)_{t \geq 0}$ es un proceso de difusión en $[0, \pi]$ y $\boldsymbol{\theta}_t \equiv \boldsymbol{\varphi}_{l_t}$, con $\boldsymbol{\varphi} = (\boldsymbol{\varphi}_t)_{t \geq 0}$ un movimiento browniano en la esfera \mathcal{S}^{d-2} e independiente de r , en el tiempo cambiado $l_t = \int_0^t (\sin r_s)^{-2} ds$. Como se demostró en la Proposición 2.32, $r_0 = \arccos(Z_0^d)$ y r es un proceso de difusión que satisface la EDE

$$r_t = r_0 + \beta_t + \frac{d-2}{2} \int_0^t \cot r_s ds, \quad 0 < r_t < \pi, \quad (4.10)$$

donde $\beta_t = -\int_0^t (1 - (Z_s^d)^2)^{-\frac{1}{2}} \sum_{k=1}^d (\delta_{kd} - Z_s^k Z_s^d) d\tilde{B}_s^k$ es un (\mathcal{F}_t) -movimiento browniano. Además, existe un (\mathcal{G}_t) -movimiento browniano \mathbf{W} en \mathbb{R}^{d-1} e independiente de β tal que la pareja $(\boldsymbol{\varphi}, \mathbf{W})$ satisface la EDE

$$d\boldsymbol{\varphi}_t = (I - \boldsymbol{\varphi}_t \boldsymbol{\varphi}_t^\top) d\mathbf{W}_t - \frac{d-2}{2} \boldsymbol{\varphi}_t dt, \quad (4.11)$$

donde $\mathcal{G}_t := \mathcal{F}_{\tau_t}$ para todo $t \geq 0$ y $\tau_t := \inf\{s \geq 0 : l_s > t\}$ es la inversa de l_t . Es decir, $\boldsymbol{\varphi}$ es un (\mathcal{G}_t) -movimiento browniano en la esfera \mathcal{S}^{d-2} .

Como se mencionó en la Observación 2.18, la transformación en coordenadas cartesianas está dada por $\mathbf{z} = ((\sin r)\boldsymbol{\theta}^\top, \cos r)^\top \in \mathcal{S}^{d-1}$, por lo que $Z_t^d = \cos r_t$, donde $Z^d = (Z_t^d)_{t \geq 0}$ denota a la última coordenada del proceso \mathbf{Z} y $\mathbf{X}_t := (\sin r_t) \boldsymbol{\varphi}_{l_t}$ representa a sus primeras $(d-1)$ coordenadas. Luego, si consideramos al proceso $\mathcal{Y} := \frac{1-Z^d}{2}$, por la fórmula de Itô, vemos que $\mathcal{Y} \sim \text{WF}(\frac{d-1}{2}, \frac{d-1}{2})$ y satisface la EDE

$$d\mathcal{Y}_t = \sqrt{\mathcal{Y}_t(1-\mathcal{Y}_t)} d\beta_t + \frac{1}{2} \left[\frac{d-1}{2}(1-\mathcal{Y}_t) - \frac{d-1}{2}\mathcal{Y}_t \right] dt, \quad \mathcal{Y}_0 = \frac{1-Z_0^d}{2}, \quad (4.12)$$

donde β es el mismo movimiento browniano escalar en la EDE (4.10) para el proceso r .

Por lo tanto, $Z^d = 1 - 2\mathcal{Y} = \cos r$, y se obtiene

$$\sin r = \sqrt{1 - \cos^2 r} = \sqrt{1 - (Z^d)^2} = 2\sqrt{\mathcal{Y}(1 - \mathcal{Y})}. \quad (4.13)$$

Así, para el cambio de tiempo $l_t = \int_0^t (\sin r_s)^{-2} ds = \int_0^t \frac{1}{4\mathcal{Y}_s(1-\mathcal{Y}_s)} ds$, como $d \geq 3$, y por las propiedades de trayectorias de una difusión de Wright-Fisher, para $\mathcal{Y} \sim \mathbb{WF}(\theta_1, \theta_2)$ con parámetros de mutación $\theta_1 = \theta_2 = \frac{d-1}{2} \geq 1$, el proceso \mathcal{Y}_t nunca alcanza a 0 ni a 1 para todo $t > 0$. Más aún, como $\mathbf{Z}_0 \notin \{\mathbf{o}, -\mathbf{o}\}$, entonces $Z_0^d \notin \{1, -1\}$, y $\mathcal{Y}_0 = \frac{1-Z_0^d}{2} \notin \{0, 1\}$. Por lo tanto, l_t está bien definido y es estrictamente creciente con $\lim_{t \rightarrow \infty} l_t = \infty$, que tiene inversa $\tau_t: [0, \infty) \rightarrow [0, \infty)$.

Así, expresando el proceso $\mathbf{X}_t = (\sin r_t) \boldsymbol{\varphi}_{l_t}$ en términos de la difusión de Wright-Fisher \mathcal{Y} , tenemos

$$\mathbf{X}_t = 2\sqrt{\mathcal{Y}_t(1 - \mathcal{Y}_t)} \boldsymbol{\varphi}_{l_t}, \quad t \geq 0. \quad (4.14)$$

Notemos que el movimiento browniano $\boldsymbol{\varphi}$ en escala de tiempo $l_t = \int_0^t \frac{1}{4\mathcal{Y}_s(1-\mathcal{Y}_s)} ds$ es independiente de \mathcal{Y} , ya que los movimientos brownianos \mathbf{W} y β que derivan de las EDEs (4.11) y (4.12) para los procesos $\boldsymbol{\varphi}$ y \mathcal{Y} respectivamente son independientes. Consecuentemente, la representación skew-product en coordenadas cartesianas del movimiento browniano \mathbf{Z} expresada en términos de la difusión de Wright-Fisher \mathcal{Y} está dada por

$$\mathbf{Z}_t = ((\sin r_t) \boldsymbol{\varphi}_{l_t}^\top, \cos r_t)^\top \equiv (2\sqrt{\mathcal{Y}_t(1 - \mathcal{Y}_t)} \boldsymbol{\varphi}_{l_t}^\top, 1 - 2\mathcal{Y}_t)^\top, \quad t \geq 0.$$

Esto es, el siguiente resultado sobre la representación skew-product del movimiento browniano en la esfera \mathcal{S}^{d-1} y está dado en [MMUB20]:

Proposición 4.6 ([MMUB20], Prop. 2.1). *Sean $d \geq 3$ y \mathbf{Z} una solución de la EDE (4.1) con $n = d - 1$ y $k = 1$. Tome $s \in [0, \infty)$ y suponga que $s > 0$ o $s = 0$ y $\mathbf{Z}_0 \notin \{\mathbf{o}, -\mathbf{o}\}$. Sea $\mathbf{X} = (Z^1, \dots, Z^{d-1})^\top$ que denota a las primeras $(d - 1)$ coordenadas de \mathbf{Z} . Entonces $R := \|\mathbf{X}\| = 2\sqrt{\mathcal{Y}(1 - \mathcal{Y})}$ y el proceso $\boldsymbol{\varphi} = (\boldsymbol{\varphi}_t)_{t \geq 0}$ dado por $\boldsymbol{\varphi}_t := \frac{\mathbf{X}_{T_s(t)}}{R_{T_s(t)}}$ es un movimiento browniano en la esfera \mathcal{S}^{d-2} con valor inicial $\boldsymbol{\varphi}_0 = \mathbf{X}_s/R_s$, independiente de \mathcal{Y} . Donde $S_s(t) := \int_s^t \frac{1}{4\mathcal{Y}_u(1-\mathcal{Y}_u)} du$ es un cambio de tiempo que satisface $\lim_{t \rightarrow \infty} S_s(t) = \infty$, y tiene inversa $T_s(t): [0, \infty) \rightarrow [s, \infty)$, para todo $0 \leq s \leq t$.*

Por lo tanto, obtenemos la descomposición skew-product para el proceso \mathbf{Z} :

$$\mathbf{Z}_t = (2\sqrt{\mathcal{Y}_t(1 - \mathcal{Y}_t)} \boldsymbol{\varphi}_{S_s(t)}^\top, 1 - 2\mathcal{Y}_t)^\top \quad t \geq s. \quad (4.15)$$

Más aún, si $\mathbf{Z}_0 \in \{\mathbf{o}, -\mathbf{o}\}$ (i.e., $R_0 = 0$), entonces $\boldsymbol{\varphi}_t$ está distribuida uniformemente en \mathcal{S}^{d-2} para cualquier $t > 0$ y por lo tanto es un movimiento browniano estacionario en la esfera.

Observación 4.8. Hemos demostrado este resultado anteriormente a partir de la representación skew-product de un movimiento browniano en la esfera \mathcal{S}^{d-1} dada en la Sección 2.3.3, y para el caso cuando $\mathbf{Z}_0 \notin \{\mathbf{o}, -\mathbf{o}\}$.

Demostración. El método que se usa en [MMUB20] para demostrar esta proposición es utilizar los resultados de la proyección del movimiento browniano esférico de [MMUB18], i.e., los resultados presentados en la sección anterior, en donde se hace uso del Teorema 4.2 para obtener la descomposición skew-product del proceso \mathbf{X} obtenido de la proyección \mathcal{P}_1 al proceso $\mathbf{Z} \sim \mathbf{MB}(\mathcal{S}^{d-1})$ en la bola \mathcal{B}^{d-1} , y utiliza la Proposición 4.1 para obtener una EDE para la última coordenada de \mathbf{Z} .

Esto es, sea \mathbf{Z} un movimiento browniano en la esfera \mathcal{S}^{d-1} dado por la EDE (4.9) con valor inicial $\mathbf{Z}_0 \in \mathcal{S}^{d-1}$ incluyendo los puntos $\{\mathbf{o}, -\mathbf{o}\}$. Si $\mathbf{X} = (\mathbf{X}_t)_{t \geq 0}$ es el proceso que denota a las primeras $(d-1)$ coordenadas de \mathbf{Z} , usando la Proposición 4.1 con $n = d-1$, $k = 1$, tenemos que \mathbf{X} es la única solución fuerte a la EDE

$$d\mathbf{X}_t = \sigma(\mathbf{X}_t) d\mathbf{B}_t - \frac{d-1}{2} \mathbf{X}_t dt, \quad \mathbf{X}_0 = (Z_0^1, \dots, Z_0^{d-1}) \in \mathcal{B}^{d-1}, \quad (4.16)$$

donde \mathbf{B} es un movimiento browniano en \mathbb{R}^{d-1} y $\sigma(\mathbf{x})$ es una función matricial con valores en $\mathbb{R}^{(d-1) \times (d-1)}$ definida por (4.3). Además, $Y = \|\mathbf{X}\|^2$ satisface la EDE

$$dY_t = 2\sqrt{Y_t(1-Y_t)} d\tilde{\beta}_t + [(d-1)(1-Y_t) - Y_t] dt, \quad Y_0 = \|\mathbf{X}_0\|^2 \in [0, 1], \quad (4.17)$$

donde $\tilde{\beta}_t = \sum_{k=1}^{d-1} \int_0^t \frac{X_s^k}{\sqrt{Y_s}} \mathbf{1}_{\{Y_s > 0\}} d\mathbf{B}_s^k + \int_0^t \mathbf{1}_{\{Y_s = 0\}} d\xi_s$ es un (\mathcal{F}_t) -movimiento browniano con ξ otro movimiento browniano escalar independiente de \mathbf{B} . Sea $R := \|\mathbf{X}\|$ el componente radial de \mathbf{X} , y notemos que $\|\mathbf{X}\| = \sqrt{1 - (Z^d)^2}$, entonces $R_0 = \sqrt{1 - (Z_0^d)^2}$, aplicando la fórmula de Itô para $R = \sqrt{Y}$, obtenemos

$$\begin{aligned} dR_t &= \frac{1}{2\sqrt{Y_t}} \left\{ 2\sqrt{Y_t(1-Y_t)} d\tilde{\beta}_t + [(d-1)(1-Y_t) - Y_t] dt \right\} + \frac{1}{2} \left(-\frac{4Y_t(1-Y_t)}{4Y_t^{3/2}} dt \right) \\ &= \frac{1}{R_t} \sqrt{R_t^2(1-R_t^2)} d\tilde{\beta}_t + \frac{1}{2R_t} \left[(d-1)(1-R_t^2) - R_t^2 - (1-R_t^2) \right] dt \\ &= \sqrt{1-R_t^2} d\tilde{\beta}_t + \left[\frac{(d-2)}{2R_t} - \frac{(d-1)}{2} R_t \right] dt. \end{aligned} \quad (4.18)$$

Por otro lado, consideramos a la última coordenada Z^d del proceso \mathbf{Z} , por la simetría de la esfera unitaria \mathcal{S}^{d-1} , supongamos que la proyección del proceso \mathbf{Z} en su ‘primera’ coordenada es Z^d , y por la Proposición 4.1 con $n = 1$, $k = d-1$, vemos que el proceso $Z^d = (Z_t^d)_{t \geq 0}$ satisface la EDE

$$dZ_t^d = \sqrt{1 - (Z_t^d)^2} d\chi_t - \frac{d-1}{2} Z_t^d dt, \quad (4.19)$$

donde $\chi_t = \int_0^t R_s d\tilde{B}_s^d + \int_0^t \left[-\frac{Z_s^d}{R_s} \mathbf{X}_s^\top \mathbf{1}_{\{|Z_s^d| < 1\}} + Z_s^d \mathbf{z}^\top \mathbf{1}_{\{|Z_s^d| = 1\}} \right] d\hat{\mathbf{B}}_s$ es un (\mathcal{F}_t) -movimiento browniano escalar, en donde $\hat{\mathbf{B}} = (\tilde{B}^1, \dots, \tilde{B}^{d-1})$ denota a las primeras $(d-1)$ coordenadas de $\tilde{\mathbf{B}}$ de la EDE (4.9), y $\mathbf{z} \in \mathcal{S}^{d-2}$ es un vector unitario arbitrario. Luego, aplicamos la fórmula de Itô para $\mathcal{Y}_t = \frac{1-Z_t^d}{2}$, se muestra que \mathcal{Y} satisface la EDE (4.12) con el mismo valor inicial pero con el movimiento browniano escalar $\beta = -\chi$. Más aún, usando nuevamente la fórmula de Itô para la transformación $\tilde{\mathcal{Y}} = 2\sqrt{\mathcal{Y}(1-\mathcal{Y})}$, vemos que $\tilde{\mathcal{Y}}$ cumple la misma EDE (4.18) y con el mismo valor inicial, esto es, los procesos R y $\tilde{\mathcal{Y}}$ son iguales, y se tiene que $R = 2\sqrt{\mathcal{Y}(1-\mathcal{Y})}$. Como $d \geq 3$, cualesquiera muestra de trayectoria de \mathcal{Y}_t no toca a 0 ni a 1 para todo $t > 0$, por lo que $0 < R_t \leq 1$ (i.e., $|Z_t^d| < 1$) para todo $t > 0$, entonces $\{(t, \omega) : |Z_t^d| = 1\}$ tiene medida de Lebesgue cero, por lo que, podemos omitir el último sumando de la definición de χ y por simples cálculos, se demuestra que el movimiento browniano β que deriva de la EDE para $\mathcal{Y} = \frac{1-Z^d}{2}$,

$$\beta_t = -\chi_t = \int_0^t (1-2\mathcal{Y}_s) \frac{X_s^\top}{R_s} d\hat{\mathbf{B}}_s - \int_0^t R_s d\tilde{B}_s^d, \quad t \geq 0 \quad (4.20)$$

es exactamente el mismo β de la EDE (4.12), i.e., es igual a la EDE (4.10).

Ahora, consideramos a la parte angular: sea $\boldsymbol{\theta}_t := \frac{\mathbf{X}_t}{\|\mathbf{X}_t\|}$ para todo $t \geq s$, y definimos la función $g(\mathbf{x}) : \mathcal{B}^{d-1} \setminus \{\mathbf{0}\} \ni \mathbf{x} \mapsto \frac{\mathbf{x}}{\|\mathbf{x}\|} \in \mathcal{S}^{d-2}$, que es de clase \mathcal{C}^∞ y para cada $1 \leq k \leq (d-1)$, sea $g_k(\mathbf{x}) = \frac{x_k}{\|\mathbf{x}\|}$, entonces

$$\frac{\partial g_k(\mathbf{x})}{\partial x_i} = \frac{\delta_{ik} \|\mathbf{x}\|^2 - x_i x_k}{\|\mathbf{x}\|^3} \quad \text{y} \quad \frac{\partial^2 g_k(\mathbf{x})}{\partial x_i \partial x_j} = \frac{3x_k x_i x_j}{\|\mathbf{x}\|^5} - \frac{\delta_{jk} x_i + \delta_{ik} x_j + \delta_{ij} x_k}{\|\mathbf{x}\|^3}.$$

Como $\sigma(\mathbf{x})\sigma(\mathbf{x})^\top = I - \mathbf{x}\mathbf{x}^\top$, y la covariación cuadrática $d\langle X^i, X^j \rangle_t = (\sigma\sigma^\top)_{ij}(\mathbf{X}_t) dt$, aplicando la fórmula de Itô para $\boldsymbol{\theta}_t = g(\mathbf{X}_t)$, obtenemos

$$\begin{aligned} d\boldsymbol{\theta}_t^k &= \sum_{i=1}^{d-1} \frac{\partial g_k(\mathbf{X}_t)}{\partial x_i} dX_t^i + \frac{1}{2} \sum_{i,j=1}^{d-1} \frac{\partial^2 g_k(\mathbf{X}_t)}{\partial x_i \partial x_j} dX_t^i dX_t^j \\ &= \frac{1}{\|\mathbf{X}_t\|} \sum_{i=1}^{d-1} (\delta_{ik} - \theta_t^i \theta_t^k) \left[\sum_{j=1}^{d-1} \sigma_{ij}(\mathbf{X}_t) dB_t^j - \frac{d-1}{2} X_t^i dt \right] + \\ &\quad + \frac{1}{2} \sum_{i,j=1}^{d-1} (\sigma\sigma^\top)_{ij}(\mathbf{X}_t) \left(\frac{3X_t^i X_t^j X_t^k}{\|\mathbf{X}_t\|^5} - \frac{\delta_{jk} X_t^i + \delta_{ik} X_t^j + \delta_{ij} X_t^k}{\|\mathbf{X}_t\|^3} \right) dt \\ &= \frac{1}{\|\mathbf{X}_t\|} \sum_{i,j=1}^{d-1} (\delta_{ik} - \theta_t^i \theta_t^k) \sigma_{ij}(\mathbf{X}_t) dB_t^j - \frac{d-1}{2} \sum_{i=1}^{d-1} (\delta_{ik} - \theta_t^i \theta_t^k) \theta_t^i dt + \\ &\quad + \frac{1}{2\|\mathbf{X}_t\|^2} \sum_{i,j=1}^{d-1} (\delta_{ij} - X_t^i X_t^j) \left(3\theta_t^i \theta_t^j \theta_t^k - (\delta_{jk} \theta_t^i + \delta_{ik} \theta_t^j + \delta_{ij} \theta_t^k) \right) dt, \end{aligned}$$

sustituyendo $\sigma_{ij}(\mathbf{x})$ en la última igualdad, podemos mostrar que el primer sumando es igual a $\frac{1}{\|\mathbf{X}_t\|} \sum_{j=1}^{d-1} (\delta_{jk} - \theta_t^j \theta_t^k) dB_t^j$, y por la definición de $\boldsymbol{\theta}$, se tiene que $X_t^i = \theta_t^i \|\mathbf{X}_t\|$, y el último sumando es igual a $\frac{-(d-2)}{2\|\mathbf{X}_t\|^2} \theta_t^k dt$. Es decir, $\boldsymbol{\theta} = \frac{\mathbf{X}}{\|\mathbf{X}\|} \equiv \frac{\mathbf{X}}{R}$ satisface la EDE

$$d\boldsymbol{\theta}_t = \frac{1}{R_t} (I - \boldsymbol{\theta}_t \boldsymbol{\theta}_t^\top) d\mathbf{B}_t - \frac{1}{R_t^2} \frac{d-2}{2} \boldsymbol{\theta}_t dt, \quad t \geq s.$$

Considerando el cambio de escala de tiempo $S_s(t) := \int_s^t \frac{1}{R_u^2} du \equiv \int_s^t \frac{1}{4\mathcal{Y}_u(1-\mathcal{Y}_u)} du$ y su inversa $T_s(t) : [0, \infty) \rightarrow [s, \infty)$, con el mismo método como se hizo en el Ejemplo 2 de la Sección 2.3, podemos demostrar que el proceso $\boldsymbol{\varphi}_t = \boldsymbol{\theta}_{T_s(t)}$, $t \geq 0$ satisface la EDE

$$d\boldsymbol{\varphi}_t = (I - \boldsymbol{\varphi}_t \boldsymbol{\varphi}_t^\top) d\widetilde{\mathbf{W}}_t - \frac{d-2}{2} \boldsymbol{\varphi}_t dt,$$

donde $\widetilde{\mathbf{W}}_t = \int_s^{T_s(t)} R_s^{-1} d\mathbf{B}_s$ es un (\mathcal{G}'_t) -movimiento browniano en \mathbb{R}^{d-1} y $\mathcal{G}'_t := \mathcal{F}_{T_s(t)}$ para todo $t \geq 0$. Esto es, $\boldsymbol{\varphi}$ es un (\mathcal{G}'_t) -movimiento browniano en la esfera \mathcal{S}^{d-2} .

Finalmente, verificamos la independencia entre $\boldsymbol{\varphi}$ y \mathcal{Y} , y esto se sigue de la independencia de los movimientos brownianos que derivan de las EDEs para los procesos $\boldsymbol{\varphi}$ y \mathcal{Y} respectivamente (c.f. [Che00], [YW71]). En particular, como el movimiento browniano β que se deriva de la EDE del proceso \mathcal{Y} está dado por (4.20), es el mismo de la EDE (4.10) para el proceso r . Entonces, si seguimos el procedimiento en la demostración de la Proposición 2.31 y de la 2.32, definimos la martingala local continua $\widehat{\mathbf{W}} = (\widehat{\mathbf{W}}_t)_{t \geq 0}$ en \mathbb{R}^{d-1} dada por

$$\widehat{\mathbf{W}}_t = \int_s^{T_s(t)} \frac{1}{R_u} (I - \boldsymbol{\theta}_u \boldsymbol{\theta}_u^\top) d\widehat{\mathbf{B}}_u + \int_s^{T_s(t)} \frac{\boldsymbol{\theta}_u}{R_u} d\widehat{\beta}_u, \quad (4.21)$$

donde $\widehat{\beta}$ es un (\mathcal{F}_t) -movimiento browniano escalar independiente de $\widehat{\mathbf{B}}$.

Notemos que $(I - \boldsymbol{\theta}\boldsymbol{\theta}^\top)(I - \boldsymbol{\theta}\boldsymbol{\theta}^\top)^\top = I - \boldsymbol{\theta}\boldsymbol{\theta}^\top$, y para cada $1 \leq i, j \leq d-1$, calculamos la covariación cuadrática de $\hat{\mathbf{W}}$ dada por $\langle \hat{W}^i, \hat{W}^j \rangle_t = \int_s^{T_s(t)} \frac{\delta_{ij}}{R_u^2} du = \delta_{ij}t$, entonces por el teorema de caracterización de Lévy, afirmamos que $\hat{\mathbf{W}}$ es un (\mathcal{G}_t) -movimiento browniano en \mathbb{R}^{d-1} . Además, se tiene que

$$(I - \boldsymbol{\theta}_u \boldsymbol{\theta}_u^\top) \left[\frac{1}{R_u} (I - \boldsymbol{\theta}_u \boldsymbol{\theta}_u^\top), \frac{1}{R_u} \boldsymbol{\theta}_u \right] = \left[\frac{1}{R_u} (I - \boldsymbol{\theta}_u \boldsymbol{\theta}_u^\top), 0 \right],$$

por el teorema de cambio de tiempo ([Øk03], Thm. 8.5.7), obtenemos

$$d\boldsymbol{\varphi}_t = (I - \boldsymbol{\varphi}_t \boldsymbol{\varphi}_t^\top) d\hat{\mathbf{W}}_t - \frac{d-2}{2} \boldsymbol{\varphi}_t dt.$$

Como la independencia de $\boldsymbol{\varphi}$ y \mathcal{Y} se sigue de la independencia de los movimientos brownianos $\hat{\mathbf{W}}$ y β , en donde ambos movimientos brownianos corren en diferente escala de tiempo, y notemos que la propiedad de Markov implica que $\hat{\mathbf{W}}$ es independiente de \mathcal{F}_s , ya que el proceso depende de $\mathcal{G}'_0 = \mathcal{F}_s$ solamente para $\hat{\mathbf{W}}$ en el tiempo 0, i.e., $\hat{\mathbf{W}}_0 = \mathbf{0}$. Por lo que $\hat{\mathbf{W}}$ es independiente de $(\beta_t)_{t \in [0, s]}$. Entonces, es suficiente probar que $\hat{\mathbf{W}}$ es independiente de $(\beta_{t+s} - \beta_s)_{t \geq 0}$.

En efecto, si definimos $\eta_t := \beta_{T_s(t)} - \beta_s = \int_s^{T_s(t)} (1 - 2\mathcal{Y}_u) \frac{X_u^\top}{R_u} d\hat{\mathbf{B}}_u - \int_s^{T_s(t)} R_u d\tilde{B}_u^d$ para todo $t \geq 0$, entonces η es una (\mathcal{G}'_t) -martingala local, y es fácil de calcular la covariación cuadrática $\langle \hat{W}^i, \xi \rangle_t = 0$ para cada $1 \leq i \leq (d-1)$, y la variación cuadrática $\langle \eta, \eta \rangle_t = T_s(t) - s$, como $S_s(t)$ es la inversa de $T_s(t)$, entonces $\langle \eta, \eta \rangle_{S_s(t+s)} = t + s - s = t$ para todo $t \geq 0$. Entonces el teorema de Knight ([RY99], Chap.V, Thm. 1.9) para el cambio de tiempo multidimensional afirma que $\hat{\mathbf{W}}$ y $(\eta_{S_s(t+s)})_{t \geq 0} = (\beta_{t+s} - \beta_s)_{t \geq 0}$ son independientes. En consecuencia, se concluye la independencia de los procesos $\boldsymbol{\varphi}$ y \mathcal{Y} .

Más aún, cuando el proceso \mathbf{Z} tiene valor inicial $\mathbf{Z}_0 \in \{\mathbf{o}, -\mathbf{o}\}$, es decir, $\mathbf{X}_0 = \mathbf{0}$ y la evolución del proceso \mathbf{X} está dada por $(R_t \boldsymbol{\phi}_{S_s(t)}, t \geq s)$, donde R es una solución a la EDE (4.18) y $\boldsymbol{\phi}$ es un movimiento browniano en la esfera \mathcal{S}^{d-2} e independiente de R con valor inicial $\boldsymbol{\phi}_0 = \mathbf{X}_s/R_s$. Por la propiedad *spinning* (i.e., $\lim_{s \rightarrow 0^+} S_s(t) = \infty$) del proceso \mathbf{X} , la parte angular $\boldsymbol{\phi}_t$ de \mathbf{X}_t tiene que ser un vector distribuido uniformemente en la esfera para cualquier $t > 0$ (cf. [IM65], [MMUB18]). \square

4.3. Simulación exacta del $\mathbf{MB}(\mathcal{S}^{d-1})$

La Proposición 4.6 establece que el movimiento browniano \mathbf{Z} en la esfera unitaria \mathcal{S}^{d-1} tiene una representación skew-product dada por (4.15), donde su parte angular es un movimiento browniano en la esfera \mathcal{S}^{d-2} y es el mismo de la parte angular de \mathbf{X} , las primeras $(d-1)$ coordenadas de \mathbf{Z} , mientras que su parte radial es la transformación lineal de la difusión de Wright-Fisher. En particular, cuando el valor inicial de \mathbf{Z} es el polo $\mathbf{o} = (0, \dots, 0, 1)^\top \in \mathcal{S}^{d-1}$, el componente angular de \mathbf{Z} en cualquier instante $t > 0$ es una distribución uniforme en \mathcal{S}^{d-2} . En [MMUB20] se presentó un algoritmo de simulación exacta de incrementos del movimiento browniano \mathbf{Z} en la esfera basado en su descomposición skew-product usando la Proposición 4.6, y el procedimiento se resume en el Algoritmo 7 ([MMUB20], Algorithm 1).

Sea \mathbf{Z} un movimiento browniano en la esfera \mathcal{S}^{d-1} con valor inicial $\mathbf{z} \in \mathcal{S}^{d-1} \setminus \{\mathbf{o}\}$, la idea del Algoritmo 7 es usando la descomposición skew-product (4.15) suponiendo que el valor inicial de \mathbf{Z} es el punto del polo norte \mathbf{o} , y después por una transformación ortogonal $O(\mathbf{z})$ que lo transforma al punto \mathbf{z} .

Algoritmo 7: Simulación de muestras exactas del incremento de un movimiento browniano en la esfera \mathcal{S}^{d-1}

Init: Dimensión $d \geq 3$, valor inicial $\mathbf{z} \in \mathcal{S}^{d-1} \setminus \{\mathbf{o}\}$, $t > 0$

- 1 Simular el componente radial: $Y \sim \mathbb{W}\mathbb{F}_0\left(\frac{d-1}{2}, \frac{d-1}{2}\right)$. ▷ Alg.5
- 2 Simular el componente angular: $\mathbf{X} \sim \mathcal{U}(\mathcal{S}^{d-2})$. ▷ Alg.1
- 3 Sea $\mathbf{u} := (\mathbf{o} - \mathbf{z}) / \|\mathbf{o} - \mathbf{z}\|$ y calcula $O(\mathbf{z}) := I - 2\mathbf{u}\mathbf{u}^\top$.
- 4 **return** $\mathbf{Z} = O(\mathbf{z})(2\sqrt{Y(1-Y)}\mathbf{X}^\top, 1 - 2Y)^\top$

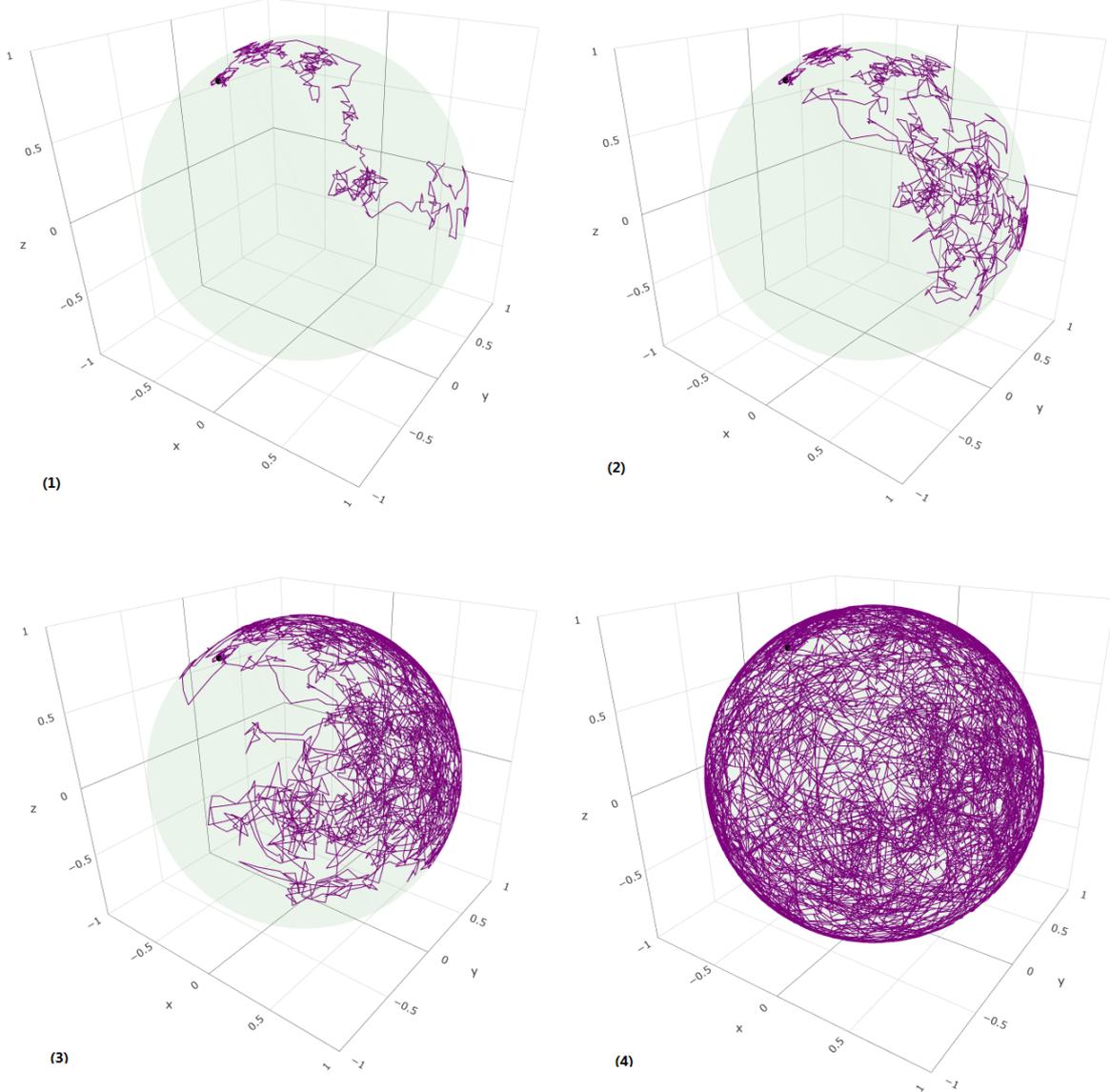


Figura 4.1: Muestra de trayectoria de un movimiento browniano \mathbf{Z} en la esfera $\mathcal{S}^2 \subset \mathbb{R}^3$ de N incrementos con tamaño de tiempo 0.00001 en cada paso usando el Algoritmo 7, donde (1) es resultado con $N = 500$, (2) con $N = 1000$, (3) con $N = 2500$ y (4) con $N = 5000$.

La transformación $O(\mathbf{z})$ en el paso 3 del Algoritmo 7 es una matriz ortogonal de $\mathbb{R}^{d \times d}$, la cual representa la reflexión de $\mathbf{x} \in \mathcal{S}^{d-1}$ con respecto a un hiperplano que pasa por el origen y con el vector normal $\mathbf{u} = \frac{\mathbf{o} - \mathbf{z}}{\|\mathbf{o} - \mathbf{z}\|} \in \mathcal{S}^{d-1}$, y se obtiene

$$O(\mathbf{z})O(\mathbf{z})^\top = (I - 2\mathbf{u}\mathbf{u}^\top)(I - 2\mathbf{u}\mathbf{u}^\top)^\top = I, \quad \text{y} \quad O(\mathbf{z})\mathbf{o} = (I - 2\mathbf{u}\mathbf{u}^\top)\mathbf{o} = \mathbf{z},$$

donde esta última propiedad es la clave para conducir a una muestra exacta del incremento del movimiento browniano en la esfera \mathcal{S}^{d-1} .

En efecto, si suponemos que $O_1(\mathbf{z})$ y $O_2(\mathbf{z})$ son matrices ortogonales de $\mathbb{R}^{d \times d}$ que cumplen las propiedades anteriores, entonces $O_1(\mathbf{z})^\top O_2(\mathbf{z})\mathbf{o} = O_1(\mathbf{z})^\top \mathbf{z} = \mathbf{o}$, i.e., una reflexión en \mathbb{R}^d que fija al punto \mathbf{o} , por lo que el producto $O_1(\mathbf{z})^\top O_2(\mathbf{z})$ es idéntico a una transformación ortogonal $\tilde{O}(\mathbf{z})$ de $\mathbb{R}^{(d-1) \times (d-1)}$ en el complemento ortogonal $\{\mathbf{o}\}^\perp$ en \mathbb{R}^d . En consecuencia, $O_2(\mathbf{z})(2\sqrt{Y(1-Y)}\mathbf{X}^\top, 1-2Y)^\top = O_1(\mathbf{z})(2\sqrt{Y(1-Y)}(\tilde{O}(\mathbf{z})\mathbf{X})^\top, 1-2Y)^\top$, como \mathbf{X} es uniforme en la esfera \mathcal{S}^{d-2} e invariante bajo transformaciones ortogonales en la esfera, por lo que

$$O_2(\mathbf{z})(2\sqrt{Y(1-Y)}\mathbf{X}^\top, 1-2Y)^\top \stackrel{(D)}{=} O_1(\mathbf{z})(2\sqrt{Y(1-Y)}\mathbf{X}^\top, 1-2Y)^\top.$$

En particular, la fórmula de $O(\mathbf{z})$ en el paso 3 en el Algoritmo 7 se escogió por su simplicidad ([MMUB20]).

La implementación del Algoritmo 7 está dada en el Apéndice Alg. 7, y con este algoritmo, generamos una secuencia de muestras exactas de incrementos $(\mathbf{Z}_0 = \mathbf{Z}_{t_0}, \mathbf{Z}_{t_1}, \dots, \mathbf{Z}_{t_N})$ de un movimiento browniano \mathbf{Z} en la esfera \mathcal{S}^{d-1} , donde para generar el valor \mathbf{Z}_{t_n} , $1 \leq n \leq N$, se toma $\mathbf{Z}_{t_{n-1}}$ como el valor inicial en el algoritmo, y así, obteniendo la trayectoria de \mathbf{Z} en la esfera de N desplazamientos. En la Figura 4.1 se presenta una muestra de trayectoria del movimiento browniano en la esfera para el caso $d = 3$ como resultado de implementación del Algoritmo 7 con el tamaño de tiempo 0.00001 en cada incremento y después de (1) $N = 500$, (2) $N = 1000$, (3) $N = 2500$ y (4) $N = 5000$ pasos. Como podemos observar, después de un lapso de tiempo, el movimiento browniano visita a cualquier lado de la esfera uniformemente. En efecto, en la figura 4.2(a) se muestra la distribución de la primera coordenada de estas muestras de incrementos de \mathbf{Z} después de $N = 50000$ pasos, en donde vemos que tiene una distribución uniforme en el intervalo $[-1, 1]$. Esto es, por el hecho de que la distribución uniforme en \mathcal{S}^{d-1} es la medida invariante para el movimiento browniano en la esfera y el teorema de Arquímedes aplicado para la distribución uniforme en la esfera \mathcal{S}^2 (ver Figura 1.2).

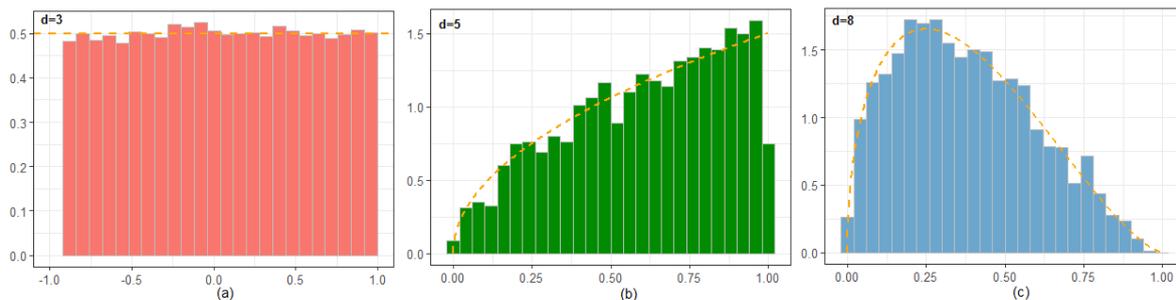


Figura 4.2: (a) Distribución de la primera coordenada de las muestras del movimiento browniano \mathbf{Z} en la esfera \mathcal{S}^{d-1} simulados usando el Algoritmo 7 con $d = 3$; (b) y (c) son distribución de $\|\mathbf{X}\|^2$, con \mathbf{X} las primeras $n = 3$ coordenadas de \mathbf{Z} en \mathcal{S}^{d-1} para $d = 5$ y $d = 8$, respectivamente, comparando contra la distribución beta $\mathfrak{B}\mathfrak{e}(n/2, (d-n)/2)$ (---).

Por otro lado, en las figuras 4.2(b) y 4.2(c) se ilustran la distribución de $\|\mathbf{X}\|^2$, donde \mathbf{X} denota a las primeras $n = 3$ coordenadas de una muestra de trayectoria del movimiento browniano \mathbf{Z} en la esfera \mathcal{S}^{d-1} de $N = 2000$ pasos con tamaño de tiempo 0.0005 en cada incremento, para $d = 5$ y $d = 8$, respectivamente; y comparando con respecto a la distribución $\mathfrak{B}\mathfrak{e}(\frac{n}{2}, \frac{d-n}{2})$. Más aún, para verificar que la muestra resultante de $\|\mathbf{X}\|^2$ sea consistente con la distribución $\mathfrak{B}\mathfrak{e}(\frac{n}{2}, \frac{d-n}{2})$, realizamos la prueba de Kolmogorov-Smirnov (K-S) para una muestra aplicada a $\|\mathbf{X}\|^2$. En la Tabla 4.1 se muestra el p -value de la prueba K-S para la muestra de la primera coordenada de \mathbf{Z} generados usando el Algoritmo 7 en el caso $d = 3$ y $n = 1$ contra la distribución uniforme en $[-1, 1]$. Y para los otros casos, son p -values de la

Tabla 4.1: Tabla del tiempo de cómputo* necesario (en segundos) para la simulación de N muestras de la trayectoria del movimiento browniano \mathbf{Z} en la esfera \mathcal{S}^{d-1} en el intervalo de tiempo $[0, 1]$ usando el Algoritmo 7 y el p -value de la prueba de K-S aplicada a $\|\mathbf{X}\|^2$ comparando contra la distribución $\mathfrak{B}\mathfrak{e}(\frac{n}{2}, \frac{d-n}{2})$, donde \mathbf{X} denota a las primeras n coordenadas de \mathbf{Z} .

		N = 1000	N = 2000	N = 5000	N = 10000
$d = 3$	tiempo	51.61	100.88	246.56	641.89
	$n = 1^{**}$	0.0643	0.0631	0.4578	0.1175
	$n = 2$	0.2859	0.8765	0.0016	0.1134
$d = 5$	tiempo	48.49	90.71	233.95	457.31
	$n = 1$	0.044	0.6548	0.6397	0.2986
	$n = 2$	0.1947	0.4883	0.0651	0.4947
	$n = 3$	0.3181	0.011	0.3553	0.0586
	$n = 4$	0.882	0.6694	0.5165	0.5445
$d = 8$	tiempo	43.01	86.16	197.64	567.64
	$n = 2$	0.9833	0.868	0.0029	0.9728
	$n = 4$	0.6408	0.398	0.048	0.4838
	$n = 6$	0.2583	0.943	0.3645	0.4004
	$n = 7$	0.477	0.1752	0.5996	0.2934

prueba K-S para la muestra de $\|\mathbf{X}\|^2$ contra la distribución $\mathfrak{B}\mathfrak{e}(\frac{n}{2}, \frac{d-n}{2})$, donde \mathbf{X} denota a las primeras n coordenadas del movimiento browniano \mathbf{Z} en la esfera \mathcal{S}^{d-1} .

Observamos de la tabla que con un número suficientemente grande de muestras de la trayectoria de \mathbf{Z} en el intervalo de tiempo $[0, 1]$, i.e. N suficientemente grande, el p -value de la prueba K-S es mayor que 0.05 en todos los casos, es decir, se puede considerar que la muestra de $\|\mathbf{X}\|^2$ pertenece a la distribución $\mathfrak{B}\mathfrak{e}(\frac{n}{2}, \frac{d-n}{2})$ para todos los casos enlistados.

En efecto, la distribución $\mathfrak{B}\mathfrak{e}(\frac{n}{2}, \frac{d-n}{2})$ es la distribución estacionaria de la difusión de Wright-Fisher $\mathbb{W}\mathbb{F}(\frac{n}{2}, \frac{d-n}{2})$, es decir, la distribución estacionaria del proceso $\|\mathbf{X}\|^2$ (ver Prop. 4.1). Sin embargo, también sabemos del Capítulo 1 que la distribución de la norma al cuadrado de las primeras n coordenadas de una distribución uniforme en la esfera \mathcal{S}^{d-1} es una distribución beta $\mathfrak{B}\mathfrak{e}(\frac{n}{2}, \frac{d-n}{2})$ (ver Lema 1.6). Por lo que la caracterización de la distribución uniforme en la esfera como medida invariante para el movimiento browniano en la esfera se preserva bajo proyección ortogonal.

* Característica del equipo de cómputo: sistema operativo Windows 8 64-bit, procesador Core i5, memoria RAM de 6GB, velocidad de procesador 1.70GHz de 2 núcleos.

** p -value de la prueba K-S para las N muestras de Z_1 contra la distribución uniforme en $[-1, 1]$, donde Z_1 es la primera coordenada de \mathbf{Z} .

Apéndice A

Implementación de algoritmos[†]

Alg. 1: Simulación de puntos uniformes en la esfera \mathcal{S}^{d-1}

```
runif_sphere <- function(d, nsim){  
  # d = dimensión del espacio  
  # nsim = número de muestras que se genera  
  
  # 1. Generar nsim normales independientes de dimensión d  
  X = matrix(rnorm(d*nsim, 0,1), ncol = d)  
  
  # 2. Calcular la norma de cada muestra de normal  
  library(wordspace)  
  R = rowNorms(X, method = "euclidean", p = 2)  
  
  # 3.  $U = X/R$   
  return(X/R)  
}
```

Alg. 2: Simulación de puntos uniformes en la bola \mathcal{B}^d

```
runif_bola <- function(d, nsim){  
  # d = dimensión del espacio  
  # nsim = número de muestras que se genera  
  
  # 1. Generar U uniforme en la esfera  
  U = runif_sphere(d, nsim)  
  
  # 2. Generar  $u \sim \text{unif}(0,1)$   
  u = runif(nsim, 0,1)  
  
  # 3.  $T = u^{1/d}U$   
  return((u^(1/d))*U)  
}
```

[†] La implementación de algoritmos del presente trabajo se realizó en el lenguaje de programación R.

Alg. 3 y 4: Simulación exacta del proceso cuadrado de Bessel y del proceso de Bessel

```

BESQ <- function(x, N, nu, max.t) {
  # Z_0 = x > 0: el valor inicial
  # N: número de muestras que va a generar
  # nu>=0: la dimensión del proceso cuadrado de Bessel
  # max.t = s
  Y = Z = rep(NA, N+1)
  Z[1] = x
  t = seq(0, max.t, length.out = N+1)

  for (n in 2:(N+1)) {
    Y[n] = rpois(1, Z[n-1]/(2*(t[n] - t[n-1])))
    Z[n] = rgamma(1, Y[n] + nu + 1, 1/(2*(t[n] - t[n-1])))
  }

  R = sqrt(Z)      # El proceso de Bessel

  return(list(BESQ = Z, BES = bes))
}

```

Alg. 5: Simulación exacta de muestras de la distribución de la difusión de Wright-Fisher

```

SamWF <- function(t, theta1, theta2, x=0) {
  ## t>0 : el tiempo del proceso
  ## theta1, theta2 : parámetros de mutación
  ## x : el valor inicial del proceso
  ## nsim : número de muestras que va a generar

  theta = theta1 + theta2

  if(t >=0.05) {
    # Cuando el tiempo t >= 0.05, se usa el método de Jenkins & Spanó de Alg.6
    An = Alg.JS17(theta, t)
  }else if(t < 0.05) {
    # Cuando el tiempo t<0.05, se utiliza la distribución aproximada de Alg.6
    An = Alg.normal(theta, t, 1)
  }

  L = rbinom(1, An, x)
  Y = rbeta(1, theta1 + L, theta2 + An - L)

  return(list(An = An, Bin = L, WF = Y))
}

```

Alg. 6: Simulación exacta de muestras de la distribución del proceso ancestral $A_{\infty}^{\theta}(t)$

```

#### 1. Definir la función para generar términos  $b_k(m)$  ####
## Los  $b_k(m)$  son los términos de la serie alternante para las funciones
## de probabilidad del proceso ancestral
b_k <- function(t, theta, k, m) {
  ## t > 0 : tiempo
  ## theta >= 0 : parámetro de mutacion, con theta = sum(theta_i), con
  ##           los theta_i son parametros de mutaciones entre alelos
  ## k >= m = 0 son términos en la suma

  if (m == 0 && k == m && theta > 0){

    # Caso 1. m = k = 0, theta != 0
    result = 1

  }else if (m >= 0 && k > m && theta > 0){

    # Caso 2. k > m >= 0, theta >= 0
    A = choose(k, m)
    B = (theta + 2*k-1)/(factorial(k))
    C = gamma(theta+m+k-1)/gamma(theta+m)
    D = exp(-k*(k+theta-1)*t/2)

    result = A*B*C*D

  }else if (m > 0 && k >= m && theta > 0){

    # Caso 3. k >= m > 0, theta >= 0
    A = choose(k, m)
    B = (theta + 2*k-1)/(factorial(k))
    C = gamma(theta+m+k-1)/gamma(theta+m)
    D = exp(-k*(k+theta-1)*t/2)

    result = A*B*C*D

  }else if (m == 0 && k >= m && theta == 0){

    # Caso 4. k >= m = 0, theta = 0
    result = 0

  }else if (m > 0 && k >= m && theta == 0){

    # Caso 5. k >= m > 0, theta = 0
    A = choose(k, m)
    B = (theta + 2*k-1)/(factorial(k))
    C = gamma(theta+m+k-1)/gamma(theta+m)
    D = exp(-k*(k+theta-1)*t/2)

    result = A*B*C*D
  }
}

```

```

    }else {
      result ="Error: No se cumplen las condiciones."
    }
  return(result)
}

#### 2. Definir la función para generar valores de C_m ####
## Los valores C_m están definidas para probar la descendencia de
## los términos b_k(m). Es el valor mínimo de los i en donde
## se valida b_{i+m+1} < b_{i+m}, y si m > D_0, entonces C_m =0
C_m <- function(t, theta, m) {

  # t > 0: el tiempo
  # theta >= 0: parámetro de mutación total
  # m >= 0: estados del proceso ancestral

  ## 2.1. Considerar a la función D_e
  D = 0
  epsilon = 0
  A = 1/t -(theta +1)/2
  B = (theta+2*D+1)*exp(-(2*D+theta)*t/2)

  while(!(D >= A && B < (1-epsilon)) && !(D >=0 && B < (1-epsilon))){

    if(D > m){
      break
    }

    D = D+1
    B = (theta+2*D+1)*exp(-(2*D+theta)*t/2)
  }

  if ((theta != 0) || (m !=0)) {

    if(m <= D){

      ## Siempre y cuando m <= D = D_0, se obtiene que los coeficientes
      ## b_k(m) va decreciendo a partir de algún momento

      i = 0
      A = b_k(t, theta, k=i+m+1, m)
      B = b_k(t, theta, k=i+m, m)

      while(!(i >= 0 && A < B)){

        if(is.infinite(A) || is.infinite(B)){
          i = Inf
          break
        }

        i = i+1
      }
    }
  }
}

```

```

        A = b_k(t, theta, i+m+1, m)
        B = b_k(t, theta, i+m, m)
    }
} else {
    ## Cuando m > D = D_0, se asigna a C_m = 0
    i = 0
}
} else {
    ## Cuando theta = m = 0, asigna a C_m = 0
    i = 0
}
return(i)
}

#### 3. Definir la función para generar sumas sup y sumas inf
## Sumas superiores e inferiores a la distribución del proceso ancestral
sumas <- function(k, M, t, theta) {
    ## M: el valor que esperamos generar de la distribución del proceso ancestral
    ## k=(k[m]): el vector con valores naturales de dimensión M+1,
    ## con m = 0, 1, ..., M
    ## t > 0 : el tiempo del proceso ancestral
    ## theta >= 0: parámetro total de mutación

    if(length(k) == (M+1)) {

        M1 = length(k)
        Bm = list()

        for (m in 1:M1){

            Bm[[m]] = rep(NA, (2*k[m]+2))

            for (i in 1:(2*k[m]+2)){
                Bm[[m]][i] <- ((-1)^(i-1))*b_k(t,theta, k=(m+i-2), m=(m-1))
            }
            sumas_inf = lapply(Bm, sum)
            sumas_sup = lapply(Bm, function(x) sum(x[-length(x)]))
        }
        sum_inf = sum(unlist(sumas_inf))
        sum_sup = sum(unlist(sumas_sup))

        if (is.nan(sum_inf) || is.nan(sum_sup) ||
            is.infinite(sum_inf) || is.infinite(sum_sup)) {
            # Esto pasa cuando t sea muy pequeño
            return(c("Error:la suma es infinito!"))
        } else {
            return(list(sum_inf = sum_inf, sum_sup = sum_sup))
        }
    } else {
        print("Error: dimensión de k es diferente de (M+1)")
    }
}
}

```

```

#### 4. Implementación del algoritmo 6 ####
#### 4.1. El algoritmo de JS17 ####
Alg.JS17 <- function(theta, t){
  ## Paso 1. Simular el numero  $U \sim \text{unif}(0,1)$ 
  U <- runif(1, 0,1)

  k0 = 0      ## Asignar el inicial con el vector  $k = 0$  en la suma
  M = 0      # Inicializa con  $M = 0$ 
  k <- c(k0) #  $k_0 = 0$ 

  repeat{

    M1 = length(k)

    for(m in 1:M1){

      ## Asignar el valor a cada elemento del vector  $k$ 
      k[m] = ceiling(C_m(t, theta, (m-1))/2)
    }

    ## Calcular la suma inferior y la superior
    suma = sumas(k = k, M = (M1-1), t, theta)
    if (class(suma) != "list"){
      return(c("Error: el tiempo t es muy pequeño!"))
      break
    }
    sum_inf = suma$sum_inf
    sum_sup = suma$sum_sup

    while(sum_inf < U && U < sum_sup){

      k <- k + rep(1, length(k))

      # Actualiza la suma cada vez actualiza el vector  $k$ 
      suma2 = sumas(k, (M1-1), t, theta)

      if (class(suma2) != "list"){
        return(c("Error: el tiempo t es muy pequeño!"))
        break
      }
      sum_inf = suma2$sum_inf
      sum_sup = suma2$sum_sup
    }

    if(sum_inf > U){

      return(M)

    }else if(sum_sup < U){

      k <- c(k, 0)
    }
  }
}

```

```
M <- M+1

}else{
  print("Falso")
  break
}
}
}

### 4.2. Simulación exacta' mediante la distribución normal asintótica ###
Alg.normal <- function(theta, t, nsim = 1) {

  beta = (theta-1)*t/2
  if(beta != 0){
    eta = beta/(exp(beta) - 1)
  }else{ eta = 1 }

  mu = 2*eta/t  ## La media de la distribución normal

  ## La varianza de la distribución normal
  if(beta != 0){
    sigma2 = (2*eta/t)*(eta + beta)^2*(1+eta/(eta+beta)
              - 2*eta)*(beta^(-2))
  }else{
    sigma2 = 2/(3*t)
  }

  ## Simular  $A \sim \text{Normal}(\text{mean} = \mu, \text{sd} = \text{sqrt}(\text{sigma2}))$ 
  Anc = rnorm(nsim, mean = mu, sd = sqrt(sigma2))
  Anc = round(Anc)  ## Tomar el entero más cercano
  for(i in 1:nsim)
    if(Anc[i] < 0){
      Anc[i] = 0
    }

  return(Anc)
}
```

Alg. 7: Simulación exacta de muestras del incremento de un movimiento browniano en la esfera \mathcal{S}^{d-1}

```

SamMB <- function(d, z, t) {
  ## d >= 3: dimensión de  $\mathcal{S}^{d-1}$  \subset  $\mathbb{R}^d$ 
  ## z \neq 0: valor inicial de Z
  ## t > 0: el tiempo

  theta1 = theta2 = (d-1)/2

  o = matrix(c(rep(0, d-1), 1)) # El polo de la esfera

  u = (o-as.matrix(z))/norm((o-as.matrix(z)), type = "2")
  O = diag(d) - 2*(u %*% t(u))

  ## Paso 1. Simular el componente radial -----
  Y = SamWF(t, theta1, theta2, x=0)$WF

  ## Paso 2. Simular el componente angular -----
  X = runif_sphere((d-1), nsim = 1)

  ## Paso 3. Salida del MB -----
  Z = as.vector(O %*% matrix(c(2*sqrt(Y*(1-Y))*X, 1-2*Y)))

  return(Z)
}

```

Apéndice B

Sobre la simulación exacta del proceso ancestral[‡]

El algoritmo para la simulación exacta del proceso ancestral $A_\infty^\theta(t)$ descrito en la Sección 3.3 y dado en [GPHK19] es el siguiente Algoritmo 8, en donde se incorporaron algunas modificaciones que mejoran el Algoritmo 6 para la reducción de tiempos de cómputo.

Algoritmo 8: Simulación exacta de muestras de $q_m^\theta(t)$, funciones de transición del proceso ancestral $A_\infty^\theta(t)$ ([GPHK19], Algorithm 3)

Init: $M \leftarrow \hat{q}_{mod}$, $\mathbf{k} \leftarrow (0, \dots, 0)$, $j \leftarrow 1$

```
1 Simular  $U \sim \mathcal{U}(0, 1)$ .
2 repeat
3   for all  $m \in \{0, \dots, M\}$  do
4      $k_m \leftarrow \lceil C_m^{(t, \theta)} / 2 \rceil$  ▷ Ec. (3.14)
5   end for
6   while  $S_{\mathbf{k}}^-(M) < U < S_{\mathbf{k}}^+(M)$  do
7      $\mathbf{k} \leftarrow \mathbf{k} + (1, 1, \dots, 1)$  ▷ Ec. (3.16)
8   end while
9   if  $S_{\mathbf{k}}^-(M) > U$  then
10    return  $M$ 
11  else if  $S_{\mathbf{k}}^+(M) < U$  then
12     $M \leftarrow \hat{q}_{mod} + (-1)^j \lceil \frac{j}{2} \rceil$ 
13    if  $j$  impar then
14       $\mathbf{k} \leftarrow (k_0, \dots, k_M)$ 
15    else if  $j$  par then
16       $\mathbf{k} \leftarrow (\mathbf{k}, 0, \dots, 0)$ 
17    end if
18  end if
19   $j \leftarrow j + 1$ 
20 until false
```

[‡] El algoritmo 3 de [GPHK19].

Una modificación más destacada entre el Algoritmo 6 y el Algoritmo 8 es: en lugar de inicializar la variable M en 0, se inicializa con el número entero más cercano alrededor de la media de una cierta distribución que sirve como una estimación de la moda \hat{q}_{mod} de $q_m^\theta(t)$. En particular, se puede tomar el valor $\hat{q}_{mod} = \lceil \mu^{(t,\theta)} \rceil$, donde $\mu^{(t,\theta)}$ es la media de la distribución normal asintótica $\mathcal{N}(\mu^{(t,\theta)}, \sigma^{(t,\theta)})$ de $A_\infty^\theta(t)$ dada por (3.17).

Además, cuando M se inicia en 0 en el Algoritmo 6, el vector \mathbf{k} se actualiza cada vez más, i.e., se agrega un elemento al vector en cada iteración nueva donde M ha crecido por una unidad. Pero en el Algoritmo 8, M se actualiza de manera telescópica, esto es, para cada nueva iteración j , M mueve del valor \hat{q}_{mod} por una unidad alternativamente para arriba o para abajo, y esto implica la actualización de los correspondientes $M + 1$ elementos del vector \mathbf{k} , haciendo el número de elementos de \mathbf{k} crecientes o decrecientes en cada iteración.

Sin embargo, al implementar el Algoritmo 8, ocasionalmente arroja error por la actualización telescópica. Específicamente, cuando el número de iteraciones j es impar y excede de $2\hat{q}_{mod}$, esto implica que M es negativo en el paso 12, y resulta que no puede tomar M elementos negativos del vector \mathbf{k} . Por lo que implica el error del algoritmo.

El siguiente programa es la implementación del Algoritmo 8 en el lenguaje de programación R, en donde se utilizan las funciones `b_k`, `C_m` y `sumas` definidas en Alg. 6 para la implementación del Algoritmo 6.

```
#### Implementacion del algoritmo de GHK19 ####
Alg.GHK19 <- function(theta, t){
  ## Paso 1. Simular el numero U ~ Unif(0,1)
  U <- runif(1, 0,1)
  k0 = 0    ## Asignar el valor inicial con k = 0

  ## Usar la media de la distribución normal asintótica
  ## para aproximar la moda de q
  vtheta = (theta-1)*t/2

  if(vtheta != 0) {
    eta = vtheta/(exp(vtheta) - 1)
  }else{
    eta = 1
  }
  mu = 2*eta/t    ## La media de la distribución

  ## Toma q_mod como la moda para la distribución q,
  ## es el entero mas grande y menor que mu
  q_mod = abs(round(mu))

  ## M inicializa con el valor de q_mod
  M = q_mod
  k = rep(k0, (M+1))
  j = 1

  repeat{
    M1 = length(k)

    for(m in 1:M1){
      k[m] = ceiling(C_m(t, theta, (m-1))/2)
    }
  }
}
```

```

suma = sumas(k, (M1-1), t, theta)

if (class(suma) != "list"){
  return(c("Error: t es muy pequeño!"))
  break
}

sum_inf = suma$sum_inf
sum_sup = suma$sum_sup

while(sum_inf < U && U < sum_sup){

  k <- k + rep(1, length(k))
  suma2 = sumas(k, (M1-1), t, theta)

  if (class(suma2) != "list"){
    return(c("Error: t es muy pequeño!"))
    break
  }
  sum_inf = suma2$sum_inf
  sum_sup = suma2$sum_sup
}

if(sum_inf > U){

  return(M)

}else if(sum_sup < U){

  M <- q_mod + ((-1)^j)*ceiling(j/2)

  if((j %% 2) != 0) { ## cuando j es impar

    k <- k[1:(M+1)]

  }else if((j %% 2) == 0){ ## cuando j es par

    k <- c(k, rep(0, (M - length(k) + 1)))

  }
}else{
  print("Falso")
  break
}
j = j+1
}
}

```

Bibliografía

- [Alf15] Aurélien Alfonsi, *Affine diffusions and related processes: simulation, theory and applications*, Bocconi & Springer Series, vol. 6, Springer, Cham; Bocconi University Press, Milan, 2015. MR 3363174 [↑45](#)
- [AM13] Tom M. Apostol and Mamikon A. Mnatsakanian, *New balancing principles applied to circum-solids of revolution, and to n -dimensional spheres, cylindroids, and cylindrical wedges*, Amer. Math. Monthly **120** (2013), no. 4, 298–321. MR 3035124 [↑viii](#), [↑2](#), [↑5](#)
- [Arc02] Archimedes, *The works of Archimedes*, Dover Publications, Inc., Mineola, NY, 2002, Reprint of the 1897 edition and the 1912 supplement, Edited by T. L. Heath. MR 2000800 [↑1](#)
- [Bel37] E. T. Bell, *Men of mathematics*, Simon & Schuster, New York, 1937. MR 3728303 [↑1](#)
- [BR05] Alexandros Beskos and Gareth O. Roberts, *Exact simulation of diffusions*, Ann. Appl. Probab. **15** (2005), no. 4, 2422–2444. MR 2187299
- [Bri97] David R. Brillinger, *A particle migrating randomly on a sphere*, vol. 10, 1997, Dedicated to Murray Rosenblatt, pp. 429–443. MR 1455152 [↑viii](#)
- [CCL99] S. S. Chern, W. H. Chen, and K. S. Lam, *Lectures on differential geometry*, Series on University Mathematics, vol. 1, World Scientific Publishing Co., Inc., River Edge, NJ, 1999. MR 1735502 [↑4](#)
- [CDH17] Vincent Coll, Jeff Dodd, and Michael Harrison, *The Archimedean projection property*, Adv. Geom. **17** (2017), no. 3, 363–372. MR 3680359 [↑viii](#), [↑2](#)
- [Che00] A. S. Cherny, *On the strong and weak solutions of stochastic differential equations governing Bessel processes*, Stochastics Stochastics Rep. **70** (2000), no. 3-4, 213–219. MR 1800956 [↑28](#), [↑36](#), [↑59](#)
- [CHS81] Stamatis Cambanis, Steel Huang, and Gordon Simons, *On the theory of elliptically contoured distributions*, J. Multivariate Anal. **11** (1981), no. 3, 368–385. MR 629795 [↑7](#)
- [CRUBV04] María E. Caballero, Victor Rivero, Gerónimo Uribe Bravo, and Carlos Velarde, *Cadenas de markov: un enfoque elemental*, Aportaciones matemáticas: Textos nivel medio, Sociedad Matemática Mexicana, 01 2004. [↑42](#)
- [dC92] Manfredo Perdigão do Carmo, *Riemannian geometry*, Mathematics: Theory & Applications, Birkhäuser Boston, Inc., Boston, MA, 1992, Translated from the second Portuguese edition by Francis Flaherty. MR 1138207 [↑3](#), [↑33](#)
- [Dur96] Richard Durrett, *Stochastic calculus*, Probability and Stochastics Series, CRC Press, Boca Raton, FL, 1996, A practical introduction. MR 1398879 [↑26](#), [↑40](#), [↑43](#)
- [Eth11] Alison Etheridge, *Some mathematical models from population genetics*, Lecture Notes in Mathematics, vol. 2012, Springer, Heidelberg, 2011, Lectures from the 39th Probability Summer School held in Saint-Flour, 2009, École d’Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School]. MR 2759587 [↑40](#), [↑41](#), [↑43](#), [↑45](#), [↑47](#)
- [Fol99] Gerald B. Folland, *Real analysis*, second ed., Pure and Applied Mathematics (New York), John Wiley & Sons, Inc., New York, 1999, Modern techniques and their applications, A Wiley-Interscience Publication. MR 1681462 [↑4](#), [↑11](#), [↑13](#), [↑21](#)
- [Gal63] Alberto Raúl Galmarino, *Representation of an isotropic diffusion as a skew product*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **1** (1962/63), 359–378. MR 148118 [↑30](#)

- [GL83] Robert C Griffiths and Wen-Hsiung Li, *Simulating allele frequencies in a population and the genetic differentiation of populations under mutation pressure*, Theoretical population biology **23** (1983), no. 1, 19–33. [↑46](#)
- [GPHK19] Celia García-Pareja, Henrik Hult, and Timo Koski, *Exact simulation of coupled wright-fisher diffusions*, arXiv:1909.11626, 2019. [↑48](#), [↑73](#)
- [Gri80] Robert C. Griffiths, *Lines of descent in the diffusion approximation of neutral Wright-Fisher models*, Theoret. Population Biol. **17** (1980), no. 1, 37–50. MR 568666 [↑47](#)
- [Gri84] ———, *Asymptotic line-of-descent distributions*, J. Math. Biol. **21** (1984), no. 1, 67–75. MR 770713 [↑49](#)
- [Gri06] ———, *Coalescent lineage distributions*, Adv. in Appl. Probab. **38** (2006), no. 2, 405–429. MR 2264950 [↑47](#)
- [GS10] Robert C. Griffiths and Dario Spanó, *Diffusion processes and coalescent trees*, Probability and mathematical genetics, London Math. Soc. Lecture Note Ser., vol. 378, Cambridge Univ. Press, Cambridge, 2010, pp. 358–379. MR 2744247 [↑ix](#), [↑46](#), [↑47](#)
- [HJT17] Julian Hofrichter, Jürgen Jost, and Tat Dat Tran, *Information geometry and population genetics*, Understanding Complex Systems, Springer, Cham, 2017, The mathematical structure of the Wright-Fisher model. MR 3444602 [↑40](#)
- [HL10] Radoslav Harman and Vladimír Lacko, *On decompositional algorithms for uniform sampling from n -spheres and n -balls*, J. Multivariate Anal. **101** (2010), no. 10, 2297–2304. MR 2719863 [↑7](#)
- [Hsu02] Elton P. Hsu, *Stochastic analysis on manifolds*, Graduate Studies in Mathematics, vol. 38, American Mathematical Society, Providence, RI, 2002. MR 1882015 [↑viii](#), [↑30](#), [↑32](#), [↑34](#)
- [IM65] Kiyoshi Itô and Henry P. McKean, Jr., *Diffusion processes and their sample paths*, Die Grundlehren der Mathematischen Wissenschaften, Band 125, Academic Press, Inc., Publishers, New York; Springer-Verlag, Berlin-New York, 1965. MR 0199891 [↑viii](#), [↑20](#), [↑30](#), [↑54](#), [↑60](#)
- [Ito75] Kiyoshi Itô, *Stochastic differentials*, Appl. Math. Optim. **1** (1975), no. 4, 374–381. MR 388538 [↑32](#)
- [IW89] Nobuyuki Ikeda and Shinzo Watanabe, *Stochastic differential equations and diffusion processes*, second ed., North-Holland Mathematical Library, vol. 24, North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, 1989. MR 1011252 [↑11](#), [↑12](#), [↑26](#), [↑54](#)
- [Jos11] Jürgen Jost, *Riemannian geometry and geometric analysis*, sixth ed., Universitext, Springer, Heidelberg, 2011. MR 2829653 [↑33](#)
- [JS17] Paul A. Jenkins and Dario Spanò, *Exact simulation of the Wright-Fisher diffusion*, Ann. Appl. Probab. **27** (2017), no. 3, 1478–1509. MR 3678477 [↑ix](#), [↑40](#), [↑46](#), [↑47](#), [↑48](#), [↑49](#)
- [Kal21] Olav Kallenberg, *Foundations of modern probability*, third ed., Springer International Publishing, 2021. [↑4](#), [↑11](#), [↑32](#), [↑38](#)
- [KT75] Samuel Karlin and Howard M. Taylor, *A first course in stochastic processes*, second ed., Academic Press, New York, 1975. MR 0356197 [↑11](#)
- [KT81] ———, *A second course in stochastic processes*, Academic Press, New York, 1981. MR 611513 [↑43](#), [↑45](#), [↑55](#)
- [Lin02] Yuanlie Lin, *Applied stochastic process*, Beijing: Tsinghua University Press, 2002, Chino.
- [Mel16] Sylvie Méléard, *Modèles aléatoires en écologie et évolution*, Mathématiques & Applications (Berlin) [Mathematics & Applications], vol. 77, Springer-Verlag, Berlin, 2016. MR 3599924 [↑45](#)
- [MMUB18] Aleksandar Mijatović, Veno Mramor, and Gerónimo Uribe Bravo, *Projections of spherical Brownian motion*, Electron. Commun. Probab. **23** (2018), Paper No. 52, 12. MR 3852266 [↑viii](#), [↑ix](#), [↑30](#), [↑39](#), [↑40](#), [↑51](#), [↑52](#), [↑53](#), [↑54](#), [↑56](#), [↑57](#), [↑60](#)
- [MMUB20] ———, *A note on the exact simulation of spherical Brownian motion*, Statist. Probab. Lett. **165** (2020), 108836, 6. MR 4113846 [↑viii](#), [↑ix](#), [↑39](#), [↑51](#), [↑56](#), [↑57](#), [↑60](#), [↑62](#)

- [MW09] R.H. Mena and S.G. Walker, *On a construction of markov models in continuous time*, *Metron* **67** (2009), no. 3, 303–323, cited By 13. [↑46](#)
- [Øk03] Bernt Øksendal, *Stochastic differential equations*, sixth ed., Universitext, Springer-Verlag, Berlin, 2003, An introduction with applications. MR 2001996 [↑11](#), [↑13](#), [↑23](#), [↑24](#), [↑31](#), [↑32](#), [↑38](#), [↑60](#)
- [PA98] Xavier Pennec and Nicholas Ayache, *Uniform distribution, distance and expectation problems for geometric features processing*, *J. Math. Imaging Vision* **9** (1998), no. 1, 49–67. MR 1631603
- [Pen04] Xavier Pennec, *Probabilities and statistics on riemannian manifolds: A geometric approach*, Research Report 5093, INRIA, January 2004. [↑3](#), [↑4](#)
- [PR88] E. J. Pauwels and L. C. G. Rogers, *Skew-product decompositions of Brownian motions*, *Geometry of random motion* (Ithaca, N.Y., 1987), *Contemp. Math.*, vol. 73, Amer. Math. Soc., Providence, RI, 1988, pp. 237–262. MR 954643 [↑30](#), [↑34](#)
- [PY81] Jim Pitman and Marc Yor, *Bessel processes and infinitely divisible laws*, *Stochastic integrals* (Proc. Sympos., Univ. Durham, Durham, 1980), *Lecture Notes in Math.*, vol. 851, Springer, Berlin, 1981, pp. 285–370. MR 620995 [↑20](#), [↑30](#)
- [RW94] L. C. G. Rogers and David Williams, *Diffusions, Markov processes, and martingales.*, second ed., *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*, vol. 1, John Wiley & Sons, Ltd., Chichester, 1994, Foundations. MR 1331599 [↑20](#)
- [RW00] ———, *Diffusions, markov processes, and martingales*, *Cambridge Mathematical Library*, vol. 2, Cambridge University Press, 2000, Itô calculus, Reprint of the second (1994) edition. MR 1780932 [↑viii](#), [↑20](#), [↑26](#), [↑30](#)
- [RY99] Daniel Revuz and Marc Yor, *Continuous Martingales and Brownian Motion*, third ed., *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, Springer-Verlag, Berlin, 1999. MR 1725357 [↑viii](#), [↑11](#), [↑15](#), [↑16](#), [↑18](#), [↑20](#), [↑23](#), [↑25](#), [↑26](#), [↑27](#), [↑28](#), [↑29](#), [↑36](#), [↑39](#), [↑42](#), [↑55](#), [↑60](#)
- [Shu01] M. A. Shubin, *Pseudodifferential operators and spectral theory*, second ed., Springer-Verlag, Berlin, 2001, Translated from the 1978 Russian original by Stig I. Andersson. MR 1852334 [↑33](#)
- [Tav84] Simon Tavaré, *Line-of-descent and genealogical processes, and their applications in population genetics models*, *Theoret. Population Biol.* **26** (1984), no. 2, 119–164. MR 770050 [↑47](#)
- [VGS17] Aaron R. Voelker, Jan Gosmann, and Terrence C. Stewart, *Efficiently sampling vectors and coordinates from the n -sphere and n -ball*, Technical report, Centre for Theoretical Neuroscience, 2017.
- [Wat44] G. N. Watson, *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, Cambridge, England; The Macmillan Company, New York, 1944. MR 0010746 [↑21](#), [↑22](#)
- [YW71] Toshio Yamada and Shinzo Watanabe, *On the uniqueness of solutions of stochastic differential equations*, *J. Math. Kyoto Univ.* **11** (1971), 155–167. MR 278420 [↑36](#), [↑59](#)