Universidad Nacional Autónoma de México

Facultad de Ciencias

Grandes Desviaciones y Cambios de Medida

T E S I S

QUE PARA OBTENER EL TÍTULO DE:
MATEMÁTICO

PRESENTA:
CÉSAR RAMÍREZ IBAÑEZ

DIRECTOR DE TESIS:
DRA. ANA MEDA GUARDIOLA

2014

1. Datos del alumno.
   Ramírez
   Ibañez
   César
   55 50 41 68
   Universidad Nacional Autónoma de México
   Facultad de Ciencias
   Matemáticas
   409005389
2. Datos del tutor
   Dra.
   Ana
   Meda
   Guardiola
3. Datos del sinodal 1
   Dr.
   Luis Antonio
   Rincón
   Solís
4. Datos del sinodal 2
   M en C
   César
   Almenara
   Martínez
5. Datos del sinodal 3
   Dr.
   Fernando
   Baltazar
   Larios
6. Datos del sinodal 4
   Dr.
   Ramón Gabriel
   Plaza
   Villegas
7. Datos del trabajo escrito
   Grandes Desviaciones y Cambios de Medida
   92 p
   2014

Agradecimientos

Quisiera agradecer a la Dra. Ana Meda Guardiola por su manera amigable y alivianada de asesorarme y guiarme durante el proceso de escribir este trabajo, y por haberme dado ánimos durante mis momentos de frustración, así como consejos.

A mi hermano y en particular a mis padres les quiero agradecer por su inigualable apoyo no sólo desde que me trajeron al mundo sino durante toda mi experiencia en la universidad, donde siempre se preocuparon de que nunca me faltara nada y me dejaron claro que jamás se dejarán de preocupar por mí, y eso me es de infinito valor.

Agradezco a Martha Liliana Espinosa Tavares, pues su amor, amistad, apoyo, compañía, cariño y comprensión me fueron indispensables para ser feliz durante este proceso, donde siempre tuvo un hombro para que yo me recargara. Gracias por sacarme del algoritmo de la vida silenciosa y de pocas sorpresas y ocurrencias.

Contenido

Capítulo 1

Introducción

1.1. Motivación

El siguiente trabajo tiene por objetivo el estudio de algunos resultados de grandes desviaciones y cambios de medidas en procesos estocásticos, para al final del trabajo mostrar una aplicación de estos resultados al resolver un problema de hallar un cambio de medida que minimice a una funcional de interés en una aplicación particular.

En grandes desviaciones se pueden estudiar eventos que son en cierta medida "extraños" al estudiarse bajo una medida de probabilidad, y en particular se puede estudiar qué pasa con eventos raros conforme va convergiendo en cierto sentido una familia de medidas de probabilidad, como se verá en el capítulo 3.

En particular, el estudio de la teoría expuesta aquí está motivado por una aplicación en finanzas dada en el artículo de P. Guasoni y S. Robertson [20]. El énfasis dado aquí es en la parte teórica de las matemáticas empleadas en [20], es decir, se estudia teoría que puede ser utilizada para resolver un problema de matemáticas aplicadas. Esta teoría se estudia en los capítulos 2 y 3 de este trabajo, y en el capítulo 4 se muestra cómo se pueden emplear estas herramientas en un problema de finanzas que se expone en [20].

Para dar una dirección y motivación para la teoría que se estudia aquí, se explica en términos generales en qué consiste la aplicación de la teoría de grandes desviaciones y cambios de medida dada en el artículo de P. Guasoni y S. Robertson [20]. En la siguiente sección se explicará brevemente el problema que tratan en dicho artículo, pero en aún menos palabras, en [20] se propone un método para buscar minimizar dentro de una familia específica de funciones a *una aproximación* de la varianza de una variable aleatoria al estudiar una versión perturbada de la varianza y su comportamiento asintótico conforme hacemos la perturbación más pequeña (tender a cero). En este trabajo no se da una expresión general cerrada para esta aproximación de varianza minimizada, sino que se estudia cómo cambiar el problema que involucra un proceso aleatorio como lo es el movimiento browniano, a un problema determinista, como se verá en el capítulo 4 de este trabajo, cuya solución minimiza a una cantidad de interés. Para ver expresiones concretas de estas aproximaciones minimizadas de la varianza, véase [20] p. 8-11 donde se obtienen estas expresiones al utilizar variables aleatorias particulares (funciones particulares del movimiento browniano) que se dan en las aplicaciones.

En este trabajo nos limitamos a dar una aplicación de la teoría estudiada en los capítulos 2 y 3, al resolver en el capítulo 4 el problema de hallar un cambio de medida que minimice a una funcional relacionada con el comportamiento asintótico de una aproximación perturbada de una varianza.

1.2. Sobre una aplicación

El siguiente planteamiento es heurístico y servirá para dar un panorama de la aplicación en la cual aplicaremos la teoría de cambios de medida y grandes desviaciones que se desarrollará en los capítulos 2 y 3. En el capítulo 4 se da una manera de abordar esta aplicación desde la perspectiva de las grandes desviaciones.

En la práctica, con frecuencia se quiere estimar el valor esperado de una variable aleatoria G, que en el caso que interesa al artículo de P. Guasoni y S. Robertson [20] en el que nos basamos para este trabajo, puede ser el precio de un producto financiero que depende del valor de un proceso estocástico en un tiempo determinado. En finanzas, un modelo aleatorio importante y utilizado es el proporcionado por el movimiento browniano, del cual dependerá la variable aleatoria G, o mejor dicho, $G(B)$, donde $B := \{B_t\}_{t\geq 0}$ es un movimiento browniano estándar.[1] Consideremos entonces el caso cuando G depende de las trayectorias de un movimiento browniano, que pueden ser vistas como funciones de $C_0[0,T]$, el espacio de trayectorias continuas que en cero valen cero. Una técnica muy común para calcular este valor esperado $\mu = E_P[G(B)]$, donde P denota que estamos midiendo a los eventos bajo la probabilidad P, es utilizar la simulación Monte Carlo[2] para calcular el estimador

$$\hat{G}_n = \frac{1}{n}\sum_{i=1}^{n} G_i, \tag{1.1}$$

donde $\{G_i\}_{i=1}^{n}$ es una muestra de tamaño n (es decir, n variables aleatorias independientes idénticamente distribuidas), que pueden ser pagos asociados a $G(B)$, y dependen de las trayectorias del movimiento browniano. El problema es que si un evento tiene relativamente muy baja probabilidad de ocurrir, al tomar una muestra, por muy grande que sea ésta, es poco probable que demos con estos valores de G.

Como desconocemos tanto a $\mu = E_P[G]$ como a la varianza, usamos al estimador de varianza

$$\hat{\sigma}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(G_i - \hat{G}_n)^2 \tag{1.2}$$

[1] Para una definición de movimiento browniano, véase el apéndice, definición A.11.

[2] No se estudia a fondo el método Monte Carlo, pero véase la sección A.2 del apéndice para conocer en qué consiste este método en términos generales. Para más detalles véase el texto de P. Glasserman [4]. También damos una definición de estimadores y estimadores de intervalo en la sección A.2.

y usamos a este último estimador para obtener un estimador de intervalo para el parámetro μ:

$$\left(\hat{G}_n - z_{\alpha/2}\frac{\hat{\sigma}_n}{\sqrt{n}},\ \hat{G}_n + z_{\alpha/2}\frac{\hat{\sigma}_n}{\sqrt{n}}\right). \tag{1.3}$$

(véase G. Casella y R. Berger [19] p. 429 para una derivación de este intervalo de confianza), donde $z_{\alpha/2}$ es el valor de una variable aleatoria de distribución normal $N(0,1)$ tal que la función de acumulación de distribución de ha acumulado una probabilidad de $1-\alpha/2$.[3] Dentro de este último intervalo de confianza se busca a un candidato para $\mu = E_P[G]$.

Los autores de [20] explican la siguiente problemática que exponemos a continuación. El asunto se complica cuando estos valores poco probables tienen grandes magnitudes, por lo cual tienen contribuciones significativas al valor esperado $\mu = E_P[G]$. Peor aún, dado que estos valores poco probables tienen grandes magnitudes en comparación con los eventos "de mayor probabilidad", entonces contribuyen de manera aún más crítica a la varianza de la variable aleatoria G, porque la varianza involucra elevar al cuadrado estas grandes magnitudes. Intuitivamente, notamos que esto hace que nuestro intervalo de confianza de $E_P[G]$ dado por (1.3) sea poco eficiente en el sentido que nuestras muestras no están "dando" con eventos importantes, es decir, valores grandes que de dar con ellos con nuestras muestras, contribuirían significativamente al valor tanto de $\hat{G}_n$ como de $\hat{\sigma}_n$, por lo cual el intervalo calculado (1.3) es más estrecho y a la vez está un tanto descentralizado en valores pequeños de $\hat{G}_n$. Aquí se ve la necesidad de "centralizar" nuestras muestras en valores más grandes, y esto haremos con un término de deriva agregado a las trayectorias del movimiento browniano. El teorema de Girsanov jugará un rol fundamental en este aspecto, pues nos ayuda a tomar nuestras muestras en regiones más relevantes para calcular el valor esperado buscado, al usar medidas de probabilidad que dan mayor peso a estas regiones, y a la vez cumpliendo con necesidades impuestas por el mundo de las finanzas.

En el artículo [20], se propone un método para encontrar un estimador para la media $E_P[G]$ y a su vez reduciendo la varianza en el mismo. La manera en que en dicho artículo abordan este problema es utilizando una técnica de muestreo de importancia, que básicamente consiste en realizar un cambio de medida que mantenga el mismo valor teórico de $E_P[G]$, pero que disminuya la varianza, para que la muestra realizada bajo esta nueva medida esté más cercana al valor de $E_P[G]$, pues recordemos que la varianza es una medida de qué tan lejos están los valores que toma una variable aleatoria de su esperanza. Intuitivamente, al comparar dos variables aleatorias con el mismo valor esperado, si en la práctica son iguales de difíciles de calcular (en términos de esfuerzo computacional, tiempo, costo, etc.), entonces puede ser preferible, o más confiable usar el que tiene menor varianza.

En este trabajo nos limitamos a buscar desde la perspectiva de grandes desviaciones un cambio de medida que minimice a una funcional de interés en el artículo de P. Guasoni

[3] En los capítulos 8 y 9 del texto de G. Casella y R. Berger [19] se estudian estos intervalos de confianza.

y S. Robertson [20]. La manera de razonar se presenta a continuación. Sea Q una medida de probabilidad equivalente a P, es decir, los conjuntos de medida cero bajo Q coinciden con los conjuntos de medida cero con respecto a P. Definimos a la variable aleatoria H como el producto de la variable aleatoria original G y el factor $\frac{dP}{dQ}$ (la derivada de Radon-Nikodym véase Williams [14] p. 145-149) de P con respecto de Q, que existe por ser estas medidas equivalentes), es decir:

$$H := G\frac{dP}{dQ}.$$

Entonces se cumple lo siguiente:

$$E_P[G] = \int_\Omega GdP = \int_\Omega G\frac{dP}{dQ}dQ = \int_\Omega HdQ = E_Q[H],$$

donde hemos utilizado que

$$\int fd\mu = \int f\frac{d\mu}{d\nu}d\nu$$

siempre que exista $\int fd\mu$ y $\mu \ll \nu$, con f una función medible, este resultado puede consultarse en P. Halmos [5] p. 134 teorema B. Así que H corresponde al precio del producto financiero G multiplicado por un factor correspondiente a un cambio de medida, y tiene el mismo valor esperado que G. El objetivo es disminuir la varianza, que depende de la medida utilizada Q, por lo cual lo ideal sería hallar alguna Q óptima en el sentido que minimice la varianza de H:

$$\mathrm{Var}_Q(H) = E_Q[H^2] - \underbrace{\big(E_Q[H]\big)^2}_{\left(E_P[G]\right)^2} = E_Q[H^2] - \big(E_P[G]\big)^2.$$

Como vemos de la expresión anterior para $\mathrm{Var}_Q(H)$, basta con hallar una medida Q tal que se minimice el valor de $E_Q[H^2]$, pero por definición de H,

$$\begin{aligned} E_Q[H^2] = \int_\Omega H^2 dQ = \int_\Omega \left(G\frac{dP}{dQ}\right)^2 dQ &= \int_\Omega \left(G^2\frac{dP}{dQ}\right)\cdot\frac{dP}{dQ}dQ \\ &= \int_\Omega G^2\frac{dP}{dQ}dP \\ &= E_P\left[G^2\frac{dP}{dQ}\right] \end{aligned} \tag{1.4}$$

donde en la cuarta igualdad hemos vuelto a aplicar el teorema B del texto de Halmos [5] mencionado anteriormente, pero ahora a la variable aleatoria $G^2\frac{dP}{dQ}$, pues recordemos que $\frac{dP}{dQ}$ es una variable aleatoria positiva.

Recordando que G es una funcional del movimiento browniano, por (1.4) vemos que lo ideal sería minimizar a la expresión $E_P\Big[G^2(B)\frac{dP}{dQ}\Big]$, que depende de Q. Sin embargo, en las páginas 2 y 6 de [20], los autores consideran una *aproximación asintótica* de (1.4), al reescalar el movimiento browniano B por $\sqrt{\epsilon}$, usando $\sqrt{\epsilon}B$ y compensando este reescalamiento al multiplicar a la función de pago G por $1/\epsilon$, de tal manera que terminan minimizando en una familia de medidas Q a la aproximación

$$\epsilon \log E_P \left[\frac{1}{\epsilon} G^2(\sqrt{\epsilon}B) \frac{dP}{dQ} \right] \tag{1.5}$$

conforme la perturbación del movimiento browniano converge a cero, es decir, conforme ϵ tiende a cero. Al analizar el comportamiento asintótico de (1.5) conforme ϵ tiende a cero, se utilizan resultados de grandes desviaciones que estudiaremos en este trabajo.

Respecto a la medida Q implícita en el factor dP/dQ, no es cualquier medida la que debemos usar, sino que debe cumplir con ciertas condiciones que se presentan en finanzas, y por razones de esta área, se requiere que el cambio de medida dado por la derivada de Radon-Nikodym $\frac{dP}{dQ}$ nos dé trayectorias que sean martingalas bajo Q, por razones de un concepto conocido como *arbitraje*.[4] Estas restricciones de finanzas, que no se abordan en este trabajo, pero para cuyos detalles referimos a [20], motivan el uso del teorema de Girsanov.

La teoría de Girsanov nos dice que Q va a depender de un término de deriva, que es una función que cumple las hipótesis del teorema de Girsanov. Por razones explicadas en el artículo de P. Guasoni y S. Robertson en [20] p. 3, se utilizan funciones deterministas como términos de deriva, que serán tomadas dentro del espacio *clásico de Cameron-Martin*, denotado por $\mathbb{H}_T$, y que se introducirá en el capítulo 3, en la definición 3.3. De este modo identificaremos a los cambios de medida con este espacio $\mathbb{H}_T$. La razón principal de usar estas derivas deterministas es la velocidad con que se pueden obtener estas derivas y por ello calcular a estos intervalos de confianza se hace con ahorro de tiempo. Entonces se debe hallar un término de deriva determinista que minimice a la funcional implícita en (1.5), y esto es donde usaremos un método variacional básico.

En el artículo de P. Guasoni y S. Robertson [20], se trabaja con una *aproximación* de (1.4) y minimizan a dicha aproximación. Obtienen resultados numéricos para probar la efectividad de sus aproximaciones para minimizar a la varianza original al analizar funciones de pago particulares que se presentan en finanzas, y estos resultados pueden revisarse en p. 8-11 de [20]. *Al final nuestra aplicación de grandes desviaciones y cambios de medida será proponer un método para minimizar dentro de una familia específica de medidas a dicha aproximación desde la perspectiva de grandes desviaciones.*

Para dejar en claro el problema específico que abordamos en este trabajo, que más bien se concentra en grandes desviaciones y cambios de medida, y no en finanzas, se indica a continuación el problema a tratar en el capítulo 4:

[4] Estas condiciones y conceptos de finanzas no se estudian aquí, pero para más detalles ver el texto de P. Glasserman [10] sección 1.2.2 y las referencias ahí contenidas.

Problema 1.1. *Dentro de una familia de medidas identificadas por el teorema de Girsanov con $\mathbb{H}_T$, proponer un método para hallar un cambio de medida dado por $\frac{dP}{dQ}$ tal que minimice a la expresión*

$$\limsup_{\epsilon\downarrow 0} \epsilon \log E_P \left[\frac{1}{\epsilon} G^2(\sqrt{\epsilon} B) \frac{dP}{dQ} \right]. \tag{1.6}$$

1.3. Organización del Trabajo

En este trabajo se proseguirá del modo que describimos a continuación.

Primero, en el capítulo 2 se estudia la *integral de Wiener* para funciones de $L^2[0,T]$, donde $[0,T]$ es un intervalo finito. Esta integral se realiza con respecto al movimiento browniano, que tiene trayectorias de variación infinita casi seguramente[5], por lo cual no se puede definir para funciones de $L^2[0,T]$ como una integral de Stieltjes. Dentro de $L^2[0,T]$ se buscará a un término de deriva, que como se verá, este término de deriva se puede emplear para cambiar la zona de donde tomamos nuestras muestras, lo cual en particular puede usarse para mejorar aproximaciones Monte Carlo. Pero veremos que introducir un término de deriva se puede compensar con un cambio de medida, que justificaremos por un resultado de Girsanov, el teorema 2.3. De hecho, se usa un caso particular del llamado teorema de Girsanov, pues no usaremos la integral estocástica en la generalidad con la que comúnmente se da el teorema de Girsanov, pues sólo requerimos un caso particular de la integral con respecto a movimiento browniano. Este capítulo 2 está basado en teoría expuesta en el libro de Michael Steele [12] y el texto de H. Kuo [7].

En el capítulo 3 se exponen resultados básicos de grandes desviaciones, pues obtendremos expresiones que nos auxiliarán en el proceso de obtener valores extremos, por ejemplo, para hallar una función que minimice a una funcional, como vimos en la primera sección que se quiere hacer. Usaremos resultados conocidos de grandes desviaciones, como lo son el lema de Schilder y el lema integral de Varadhan. Debido a que en el artículo de P. Guasoni y Robertson [20] se presenta un problema en las aplicaciones que requiere una hipótesis adicional al lema de Varadhan, tendremos que hacer una pequeña extensión de este resultado. Este capítulo 3 está basado principalmente en el texto de los autores Mark Freidlin y Alexander Wenztell [15] y el libro de A. Dembo y O. Zeitouni [17]. La extensión del lema de Varadhan está basada en el artículo de P. Guasoni y S. Robertson en [20].

En el capítulo 4 se aborda el Problema 1.1, relacionado con minimizar a una aproximación de varianza, como dijimos al final de la sección anterior, desde la perspectiva ofrecida por el teorema de Girsanov que daremos en el capítulo 2 y los resultados de grandes desviaciones del capítulo 3. Este capítulo 4 está basado principalmente en el artículo de P. Guasoni y S. Robertson en [20]. Veremos que aplicando el teorema de Girsanov, la

[5] Véase el teorema A.5 del apéndice para una prueba de esto.

cantidad a minimizar dependerá de términos de deriva h y estará dada por

$$L(h) := \limsup_{\epsilon\to 0} \epsilon \, \log E_P\Big[\exp\Big(\frac{1}{\epsilon}\Big(2F(\sqrt{\epsilon}B) - \int_0^T \sqrt{\epsilon}\dot{h}_t dB_t + \frac{1}{2}\int_0^T \dot{h}_t^2 dt \Big)\Big)\Big], \qquad (1.7)$$

donde $F = \log(G)$ y B es un movimiento browniano estándar. Se prueba que al problema de minimizar (1.7) se puede aplicar la extensión del lema de Varadhan. También se usa un método variacional para probar la existencia de extremos en un espacio de Hilbert que introduciremos en el capítulo 3 en la definición 3.3, llamado *espacio clásico de Cameron-Martin*, que se denotará por $\mathbb{H}_T$. Al final de este capítulo 4 se obtiene el teorema 4.3, que nos proporciona una manera de abordar el problema de minimizar (1.7) desde una perspectiva determinista, pues como vemos, $L(h)$ en (1.7) tiene un aspecto probabilístico al incorporar un movimiento browniano. Para probarlo, nos auxiliamos de los resultados de grandes desviaciones del capítulo 3, como el lema de Schilder para obtener una expresión cerrada que minimiza (1.7). Concretamente, obtendremos que con $L(h)$ definido por (1.7), para los $h \in \mathbb{H}_T$ tales que su derivada $\dot{h}$ tenga variación finita, tendremos que

$$L(h) = \sup_{x\in\mathbb{H}_T} \Big\{ 2F(x) + \frac{1}{2}\int_0^T (\dot{x}_t - \dot{h}_t)^2 dt - \int_0^T \dot{x}_t^2 dt \Big\}. \qquad (1.8)$$

Notemos que esta expresión alternativa de $L(h)$ es la que nos lleva a una expresión determinista en el sentido que la expresión original de $L(h)$ en (1.7) depende de las trayectorias de un movimiento browniano, mientras que la expresión (1.8) depende de funciones en $\mathbb{H}_T$, que veremos que son deterministas. Por otra parte, usando un principio de grandes desviaciones, probaremos que para hallar h que minimice a $L(h)$ en (1.7), es suficiente hallar a $\hat{h}$ tal que

$$L(\hat{h}) = \sup_{x\in\mathbb{H}_T} \Big\{ 2F(x) - \int_0^T \dot{x}_t^2 dt \Big\}. \qquad (1.9)$$

Esto nos ayuda a plantear el problema desde el punto de vista del cálculo de variaciones, pues queremos buscar $\hat{h} \in \mathbb{H}_T$ tal que minimice a $L(h)$, que como veremos, basta con ser el valor extremo dado por el lado derecho de (1.9). De acuerdo a los autores P. Guasoni y S. Robertson en [20], obtener el supremo en (1.8) y (1.9) se reduce a resolver las ecuaciones diferenciales ordinarias unidimensionales de Euler-Lagrange asociadas, donde se obtiene una ecuación de Euler-Lagrange particular al utilizar funciones particulares G de pago en las aplicaciones, que describen el pago del producto financiero G. Ejemplos de funciones particulares de pago utilizadas en la práctica se dan en p. 8-11 de P. Guasoni y S. Robertson [20].

Es importante destacar que en este trabajo no se abarca el estudio de funciones de pago G particulares utilizadas en la práctica, por lo cual no llegamos a resolver la ecuación diferencial de Euler-Lagrange, sino que el objetivo es proponer un método para resolver el problema de hallar el mínimo de la expresión (1.7) (al hallar una función h que minimice (1.7)) al relacionar $L(h)$ dado por (1.7) con una expresión determinista que al meter la función de pago G de interés particular puede resolverse, por ejemplo, al resolver la ecuación de Euler-Lagrange. Con un método variacional básico se probará que los problemas

(1.8) y (1.9) tienen solución en el espacio $\mathbb{H}_T$; al final la solución es una función que nos determina un cambio de medida o un término de deriva.

Cabe destacar que en la prueba del teorema 4.3, basado en el resultado principal del artículo [20] (el teorema 3.6, p.7), se tuvo que hacer una modificación de la prueba proporcionada en dicho artículo[6], misma para la cual fui auxiliado por uno de los autores del artículo, Scott Robertson.

En resumen, el resultado principal del teorema 4.3 nos dice que la solución del Problema 1.1 es $\hat{h}$ tal que cumpla las dos condiciones siguientes: que sea el supremo del lado derecho de la ecuación (1.9), y que al evaluar $L(\hat{h})$, resulte que se cumpla la igualdad (1.9). Más aún, si $\hat{h}$ es tal que su derivada $\dot{\hat{h}}$ tiene variación finita, entonces podemos evaluar a $L(\hat{h})$ por medio de la expresión determinista (1.8), y la solución será única. Según Scott Robertson, uno de los autores de [20], en las aplicaciones suele ser el caso que $\dot{\hat{h}}$ tiene variación finita, por lo cual verificamos que se cumple la igualdad (1.9) al utilizar (1.8) para evaluar a $L(\hat{h})$.

Al final, con el teorema 4.3 en mano, se propone un algoritmo para resolver el problema de minimizar (1.7) para funciones de pago G generales.

[6]Esto se hace por razones explicadas después del teorema 4.3.

Capítulo 2

Teoría de Girsanov y la integral de Wiener

2.1. Introducción

Como vimos en el capítulo anterior, podría ser útil recorrer el lugar donde se toma una muestra hacia información que nos interesa, para mejorar la capacidad de estimación del valor esperado de una variable aleatoria por el método de Monte Carlo. Este capítulo está basado en teoría expuesta en el libro de Michael Steele [12], H. Kuo [7], y también por el texto de los autores I. Karatzas y S. Shreve [21]. El producto final de este capítulo es el teorema 2.3.

El siguiente ejemplo simplificado nos da una idea de cómo recorrer el lugar de donde se toma una muestra, al hacer un cambio de medida. Supongamos que tenemos una variable aleatoria X normal con valor esperado cero y varianza 1, denotado $X \sim N(0,1)$. Vamos a denotar a la ley de probabilidad de X como P_0, donde el subíndice 0 nos indica que X tiene la media cero bajo P_0. De este modo denotamos por E_0 al operador de valor esperado (que es una integral bajo la medida P_0). Si $X \sim N(\mu, 1)$, entonces denotaríamos a la ley de probabilidad de X con media μ como P_μ, y a su correspondiente operador de esperanza como E_μ.

Sea f una función acotada y Borel medible. Entonces

$$\begin{aligned} E_0[f(X)] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-(x-\mu)^2/2} e^{\mu^2/2-\mu x} dx \\ &= E_\mu[f(X) e^{-\mu X+\mu^2/2}], \end{aligned} \tag{2.1}$$

donde la segunda igualdad se obtuvo sumando y restando $\mu^2/2 - \mu x$. Esta última ecuación, por simple que se ve, nos está diciendo algo importante:

Nos está proporcionando una familia paramétrica de operadores E_μ que nos auxiliarán para calcular un mismo valor esperado $E_0[f(X)]$.

Cambiar el valor esperado μ de una variable aleatoria normal, es como darle más peso a otras regiones de $\mathbb{R}$. Entonces la fórmula dada por (2.1) nos está dando cierta flexibilidad acerca del lugar donde tomamos nuestras muestras, a cambio de una compensación (que depende del recorrimiento μ) por el factor $e^{-\mu X+\mu^2/2}$. Si usáramos el método de Monte Carlo para estimar a $E_0[f(X)]$, tomaríamos muestras de tamaño n de la forma $X_1,\ldots,X_n$, entonces de acuerdo a la ley fuerte de grandes números, para valores de n suficientemente grandes:

$$E_0[f(X)] \approx \frac{1}{n}\sum_{i=1}^{n} f(X_i).$$

Pero si f depende de eventos poco probables (como valores muy grandes de X_i) entonces dado que aún para n's muy grandes, es poca la cantidad de veces que observamos estos eventos[1] por lo cual nuestros estimadores para $E_0[f(X)]$ nos dan valores de cero. La forma en la que se aborda este problema es que, apoyándonos en la identidad (2.1), recorremos el lugar donde tomamos la muestra (hacia los valores que nos importan más) y compensamos este recorrimiento al multiplicar por el factor $e^{-\mu X+\mu^2/2}$, de tal manera que mejor formamos las sumas:

$$E_0[f(X)] \approx \frac{1}{n}\sum_{i=1}^{n} f(Y_i)e^{-\mu Y_i+\mu^2/2}, \qquad \text{con } Y_i \sim N(\mu,1). \tag{2.2}$$

Por ejemplo, si $f(x) = \mathbf{1}_{\{x>30\}}$ y $X \sim N(0,1)$, entonces debido al rápido decaimiento (exponencial) de la densidad de X (ver el siguiente capítulo para una medida cuantitativa de este decaimiento exponencial), es muy poco probable dar con estos valores $x > 30$. Así que tomando $\mu = 30$, se mejora mucho el cálculo de $E_0[f(X)]$. Para un ejemplo numérico de cómo mejoran este tipo de aproximaciones, véase p. 11 del trabajo [8] de César Almenara Martínez.

2.2. Motivación de teorema de Girsanov

Sabemos que nos puede ayudar recorrer el lugar donde tomamos una muestra, pero eso lo hicimos para una variable aleatoria X con valores en $\mathbb{R}$. Debemos hacerlo para un proceso estocástico en tiempo continuo, es decir, un *conjunto de variables aleatorias*

$$X := \{X_t : 0 \le t \le T\}$$

indizadas por $t \in [0,T]$, un intervalo de longitud finita.

[1] Incluso, como indican los autores P. Guasoni y S. Robertson en la p.2 de [20] o M. Steele p. 213-215 en [12], en las aplicaciones, puede ocurrir que *nunca* se dé con estos eventos, con los valores de n disponibles en la práctica.

Es importante destacar que podemos ver a un proceso estocástico con trayectorias casi seguramente continuas (como lo es el movimiento browniano y el movimiento browniano on deriva) $\{X_t : 0 \leq t \leq T\}$ como una variable aleatoria que toma valores en $C_0[0,T]$, el espacio de funciones continuas en el intervalo $[0,T]$ tales que $f(0) = 0$. No abundaremos mucho en el procedimiento formal de definir medidas en espacios de dimensión infinita como lo es $C_0[0,T]$, esto se puede revisar en I. Karatzas y S. Shreve [21] en las secciones 2.1-2.5. Daremos aquí los conceptos necesarios para entender los resultados que se exponen en este trabajo.

Una vez que tenemos a un proceso estocástico recorrido por un término de deriva, a veces para facilitarnos las cuentas nos conviene en cierto sentido poder ver a este proceso con deriva como uno sin deriva. Se conoce por teoría de Girsanov a resultados relacionados con este tipo de intentos, aunque de acuerdo al autor Michael Steele en [12], el matemático I.V. Girsanov no fue el primero en intentar resolver este tipo de problemas.

Se da un ejemplo que motivará la idea central tras el teorema de Girsanov, y que a su vez nos evoca el concepto de integración con respecto al movimiento browniano. La teoría de integración con respecto al movimiento browniano, o integración de Itô, como suele conocerse, es una teoría vasta. Podremos trabajar con un ejemplo muy particular de integral estocástica, menos general que la integral de Itô, conocida como la integral de Wiener.

Ahora empezamos, en cierta manera, al revés de como se inició en la introducción de este capítulo, pues ahora tendremos el movimiento browniano recorrido por el término de deriva μ, es decir, tenemos al proceso estocástico $\{X_t : 0 \leq t \leq T\}$ dado por

$$X_t = B_t + \mu t \qquad \text{para todo } \ 0 \leq t \leq T \tag{2.3}$$

donde B_t corresponde a un movimiento browniano estándar[2], es decir, sin deriva. Supongamos que tomamos un conjunto finito de tiempos $0 = t_0 < t_1 < \cdots < t_n \leq T$, y que en estos tiempos obtenemos $X_{t_1}, X_{t_2}, \ldots, X_{t_n}$. Sea $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ una función acotada y Borel medible, y supongamos que se quiere calcular

$$E[f(X_{t_1}, X_{t_2}, \ldots, X_{t_n})]. \tag{2.4}$$

Dado que $\{X_t\}$ definido por (2.3) es un proceso con incrementos independientes, y queremos aprovechar esto en el cálculo de (2.4), suponemos que podemos hacer $f(x_1, x_2 - x_1, \ldots, x_n - x_{n-1})$, de tal manera que nos fijamos en la distribución del vector de incrementos independientes

$$(X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}). \tag{2.5}$$

Entonces la densidad de este vector, dada por la distribución gaussiana, evaluada en $(x_1, x_2 - x_1, \ldots, x_n - x_{n-1}) \in \mathbb{R}^n$ es, dados los incrementos independientes y estacionarios,

[2]Para una definición del movimiento browniano, véase la definición A.11 del apéndice.

con $x_0 = 0$:

$$(2\pi)^{-n/2}\big((t_1(t_2-t_1)\cdots(t_n-t_{n-1})\big)^{-1/2}\exp\left(-\sum_{i=1}^{n}\frac{\big\{(x_i-x_{i-1})-\mu(t_i-t_{i-1})\big\}^2}{2(t_i-t_{i-1})}\right).$$

Definiendo $\Delta x_i = x_i - x_{i-1}$ y $\Delta t_i = t_i - t_{i-1}$, desarrollamos

$$\begin{aligned}\exp&\left(-\sum_{i=1}^{n}\frac{\big\{\Delta x_i-\mu\Delta t_i\big\}^2}{2\Delta t_i}\right)\\ &=\exp\left(-\sum_{i=1}^{n}\frac{\Delta x_i^2-2\mu\Delta t_i\Delta x_i+\mu^2\Delta t_i^2}{2\Delta t_i}\right)\\ &=\exp\left(-\sum_{i=1}^{n}\frac{\Delta x_i^2}{2\Delta t_i}\right)\exp\left(-\sum_{i=1}^{n}\frac{-2\mu\Delta t_i\Delta x_i+\mu^2\Delta t_i^2}{2\Delta t_i}\right)\\ &=\exp\left(-\sum_{i=1}^{n}\frac{\Delta x_i^2}{2\Delta t_i}\right)\exp\left(\sum_{i=1}^{n}\Big(\mu\Delta x_i-\frac{1}{2}\mu^2\Delta t_i\Big)\right)\end{aligned}\tag{2.6}$$

y notamos que en el segundo producto exponencial de la última igualdad (2.6) tenemos una suma telescópica, de tal forma que la densidad del vector aleatorio tiene el factor

$$\begin{aligned}\exp&\left(-\sum_{i=1}^{n}\frac{\big\{\Delta x_i-\mu\Delta t_i\big\}^2}{2\Delta t_i}\right)\\ &=\exp\left(-\sum_{i=1}^{n}\frac{(x_i-x_{i-1})^2}{2(t_i-t_{i-1})}\right)\exp\Big(\mu x_n-\frac{1}{2}\mu^2 t_n\Big).\end{aligned}\tag{2.7}$$

Nótese que la primera exponencial de (2.7) está relacionada con la distribución de los incrementos de un movimiento browniano *sin deriva* $\{B_t\}$. Esto quiere decir que para calcular el valor esperado $E[f(X_{t_1}, X_{t_2}, \ldots, X_{t_n})]$, tenemos que

$$E[f(X_{t_1},X_{t_2},\ldots,X_{t_n})] = E\Big[f(B_{t_1},B_{t_2},\ldots,B_{t_n})e^{\mu B_{t_n}-\frac{1}{2}\mu^2 t_n}\Big].\tag{2.8}$$

Esta última ecuación nos está indicando que podemos mejor evaluar a f en $(B_{t_1}, B_{t_2}, \ldots, B_{t_n})$, que es un movimiento browniano sin deriva, pues de (2.7) vemos que la densidad del vector (2.5) es como la densidad de un movimiento browniano estándar, sólo que multiplicado por el factor

$$\exp\left(\mu B_{t_n}-\frac{1}{2}\mu^2 t_n\right).$$

Aquí vemos por qué esto es el inverso análogo a la ecuación de (2.1), pues esta vez empezamos con un movimiento browniano con deriva, y terminamos viéndolo como un movimiento browniano sin deriva, con un factor de compensación dado por (2.8). Este factor de compensación es la derivada de Radon-Nikodym de la ley de probabilidad P_μ del proceso $\{X_t\}$ definido por (2.3) con respecto a la ley del proceso sin deriva $\{B_t\}$, que denotaremos por P, es decir:

$$\left.\frac{dP_\mu}{dP}\right|_{\mathcal{F}_t} = \exp\left(\mu B_t - \frac{1}{2}\mu^2 t\right). \tag{2.9}$$

La derivada de Radon-Nikodym de esta última ecuación nos habla de medidas equivalentes P_μ y P, pues el lado derecho de (2.9) es siempre positivo, por lo cual ambas medidas tienen los mismos conjuntos de medida cero. Por otra parte, la medida está bien definida, gracias a que $e^{\mu B_t - \frac{1}{2}\mu^2 t}$ nos define un proceso que es *martingala*[3] bajo P, pues si $s < t$, si denotamos al valor esperado con respecto a P como $E[\cdot]$

$$\begin{aligned}
E\left[e^{\mu B_t - \frac{1}{2}\mu^2 t}\middle|\mathcal{F}_s\right] &= E\left[e^{\mu(B_t - B_s) + \mu B_s - \frac{1}{2}\mu^2(t-s) - \frac{1}{2}\mu^2 s}\middle|\mathcal{F}_s\right] \\
&= e^{\mu B_s - \frac{1}{2}\mu^2 s} E\left[e^{\mu(B_t - B_s) - \frac{1}{2}\mu^2(t-s)}\middle|\mathcal{F}_s\right] \\
&= e^{\mu B_s - \frac{1}{2}\mu^2 s} E\left[e^{\mu(B_t - B_s) - \frac{1}{2}\mu^2(t-s)}\right] \quad \text{(por incrementos independientes)} \\
&= e^{\mu B_s - \frac{1}{2}\mu^2 s} \quad \text{(por la función generadora de momentos).}
\end{aligned}$$

Entonces si definimos $M_t := \exp\left(\mu B_{t_n} - \frac{1}{2}\mu^2 t_n\right)$, usando (2.9) tenemos que para $A \in \mathcal{F}_s$, con $s < t$:

$$\begin{aligned}
P_\mu(A) :&= E\left[\mathbf{1}_A M_t\right] \\
&= E\big[E\left[\mathbf{1}_A M_t | \mathcal{F}_s\right]\big] \\
&= E\big[\mathbf{1}_A E\left[M_t | \mathcal{F}_s\right]\big] \qquad (A \in \mathcal{F}_s) \\
&= E\left[\mathbf{1}_A M_s\right] \qquad (\{M_t\} \;\; \text{es martingala}),
\end{aligned}$$

de modo que $\{M_t\}$ genera una medida.

La idea central del teorema de Girsanov está visible en la expresión (2.8), con la derivada de Radon-Nikodym $M_t = e^{\mu B_t - \frac{1}{2}\mu^2 t}$ auxiliándonos para poder manipular a un movimiento browniano con deriva μ como en (2.3) como si fuera un movimiento browniano sin deriva. Sin embargo queremos algo aún un poco más ambicioso. Recordemos al segundo término exponencial de (2.6) que nos dio $M_t = e^{\mu B_t - \frac{1}{2}\mu^2 t}$:

$$\exp\left(\sum_{i=1}^{n}\left(\mu \Delta x_i - \frac{1}{2}\mu^2 \Delta t_i\right)\right). \tag{2.10}$$

Notemos que este término depende del vector aleatorio

$$(B_{t_1}, B_{t_2}, \ldots, B_{t_n}) \in \mathbb{R}^n$$

[3] Véase el apéndice para una definición de martingala.

que corresponde a tomar un número finito de tiempos $0 = t_0 < t_1 < \cdots < t_n \leq T$ y registrar el valor de $\{B_t\}$ en estos instantes. Queremos involucrar a un número infinito de tiempos, con mayor precisión, a toda la historia del proceso $\{B_t\}$ en el intervalo $[0, t] \subset [0, T]$ (con $t \geq 0$), y además supondremos que la deriva μ *no es necesariamente constante*, es decir, es una función $\mu(t)$. Pediremos además que $\mu(t)$ satisfaga el criterio de integrabilidad

$$\int_0^T \mu^2(s)ds < \infty.$$

Para capturar toda la información de $\{B_t\}$ en el intervalo $[0, t]$, haremos que los tiempos t_i se vayan acercando más y más, haciendo más chicos los intervalos $[t_{i-1}, t_i]$ conforme n crece. Entonces el análogo al argumento del segundo factor exponencial de (2.6), pero ahora con $\mu(s)$ es

$$\sum_{i=1}^{n} \mu(t_i^*)\big(B_{t_i}(\omega) - B_{t_{i-1}}(\omega)\big) - \frac{1}{2}\sum_{i=1}^{n} \mu^2(t_i^*)\Delta t_i, \tag{2.11}$$

donde $t_{i-1} \leq t_i^* \leq t_i$. Nótese que a diferencia de (2.6), no se cancelan los términos como una suma telescópica. Sabemos que para cada n tenemos una suma de Riemann $\frac{1}{2}\sum_{i=1}^{n} \mu^2(t_i^*)\Delta t_i$, por lo cual conforme $\Delta t_i = t_i - t_{i-1} \to 0$ haciendo a $n \to \infty$, tenemos que

$$\frac{1}{2}\sum_{i=1}^{n} \mu^2(t_i^*)\Delta t_i \to \frac{1}{2}\int_0^t \mu^2(s)ds.$$

¿Qué sucede con el término $\sum_{i=1}^{n} \mu(t_i^*)\big(B_{t_i}(\omega) - B_{t_{i-1}}(\omega)\big)$? [4]

Para esto, nos ayudaría poder definir para cada $\omega \in \Omega$ la integral de $\mu(t)$ con respecto al movimiento browniano $\{B_t\}$ de la forma

$$\int_0^t \mu(s)dB_s(\omega), \tag{2.12}$$

para poder obtener el análogo en tiempo continuo de la expresión (2.10):

$$M_t = \exp\left(\int_0^t \mu(s)dB_s - \frac{1}{2}\int_0^t \mu^2(s)ds\right).$$

Pero si intentáramos obtener a la integral de (2.12) en el sentido de Riemann-Stieltjes por medio de la fórmula de integración por partes, dado que el movimiento browniano tiene trayectorias casi seguramente de variación infinita en intervalos finitos,[5] entonces sólo

[4]Pudimos haber usado una función $\mu(\omega, t)$ y haber pedido que $\mu(\omega, t)$ cumpliera otro criterio de integrabilidad distinto al que dimos aquí, y esto induciría el concepto de integral de Itô. Nos basta con funciones μ deterministas para el Problema 1.1 a resolver en esta tesis.

[5]Esto se prueba en el apéndice en el teorema A.5.

podríamos obtener la integral de Stieltjes para funciones $\mu(s)$ continuas y de variación acotada, y nosotros queremos a la integral para $\mu \in L^2[0,T]$, como dijimos anteriormente. Esto nos lleva al concepto de la integral de Wiener, que desarrollamos a continuación.

2.3. La integral de Wiener

Se comienza en el espiritu de los cursos básicos de integración, es decir, primero se define la integral para una familia de integrandos simples (escalonados) y posteriormente, a partir de esta integral simple se extiende el proceso de integración a $L^2[a,b]$, donde $[a,b]$ es un intervalo finito arbitrario.
Al momento de definir la integral con respecto al movimiento browniano, nuestra experiencia con integración en cursos básicos de análisis matemático nos dice que deberíamos de esperar que esta integral cumpla que

$$\int_a^b \mathbf{1}_{[c,d)} dB_s = B_d - B_c$$

para cada intervalo $[c,d) \subset [a,b]$. Además, debería cumplir una condición de la forma

$$\int_a^b \alpha \, dB_s = \alpha \int_a^b dB_s \qquad \text{para toda constante } \alpha \in \mathbb{R}.$$

y si $a \leq c \leq b$ entonces

$$\int_a^c dB_s + \int_c^b dB_s = \int_a^b dB_s.$$

Esto motiva la siguiente definición.

Definición 2.1. (La integral de Wiener para funciones escalonadas) *Para las **funciones escalonadas** f de la forma*

$$f = \sum_{i=1}^{n} a_i \mathbf{1}_{[t_{i-1},t_i)} \tag{2.13}$$

con $a =: t_0 < \ldots < t_n := b$ definimos a la integral de f con respecto a $\{B_t\}$ en el intervalo $[a,b]$, denotada por $I[f]$ como

$$I[f] = \sum_{i=1}^{n} a_i (B_{t_i} - B_{t_{i-1}}). \tag{2.14}$$

Esto quiere decir que $I[f]$ es la variable aleatoria $I[f] : \Omega \longrightarrow \mathbb{R}$ tal que para cada $\omega \in \Omega$ tenemos una suma de la forma (2.14).

Es fácil probar que hemos obtenido un operador lineal I, es decir que para f, g escalonadas y $\alpha, \beta \in \mathbb{R}$ se cumple que $I[\alpha f + \beta g] = \alpha I[f] + \beta I[g]$. Tenemos el siguiente resultado sencillo.

Teorema 2.1. (La integral de f escalonada es una v.a. normal) *Para toda función escalonada f de la forma* (2.13), *la variable aleatoria $I[f]$ es gaussiana de valor esperado 0 y de varianza*

$$\mathrm{Var}(I[f]) = E\big[I[f]^2\big] = \int_a^b f^2(t)dt$$

Demostración. Por definición tenemos que $I[f] = \sum_{i=1}^n a_i(B_{t_i} - B_{t_{i-1}})$, donde cada sumando $a_i(B_{t_i} - B_{t_{i-1}})$ es una variable aleatoria normal independiente (por los incrementos del movimiento browniano) $N(0, a_i^2(t_i - t_{i-1}))$. Esto quiere decir que $I[f]$ es una combinación lineal de variables aleatorias normales independientes de media cero, por lo cual también $I[f]$ es una variable aleatoria normal de media cero.[6]

Calculemos la varianza de $I[f]$:

$$\begin{aligned} E\big[I[f]^2\big] &= E\left[\left(\sum_{i=1}^n a_i(B_{t_i} - B_{t_{i-1}})\right)^2\right] \\ &= E\left[\sum_{i=1}^n a_i^2(B_{t_i} - B_{t_{i-1}})^2 + 2\sum_{i<j}^n a_i a_j(B_{t_i} - B_{t_{i-1}})(B_{t_j} - B_{t_{j-1}})\right]. \end{aligned} \tag{2.15}$$

Ahora, debido a la distribución de los incrementos del movimiento browniano, tenemos que

$$E[(B_{t_i} - B_{t_{i-1}})^2] = t_i - t_{i-1},$$

y por la independencia de los incrementos, para $i \neq j$:

$$E[(B_{t_i} - B_{t_{i-1}})(B_{t_j} - B_{t_{j-1}})] = E[(B_{t_i} - B_{t_{i-1}})]E[(B_{t_j} - B_{t_{j-1}})] = 0.$$

De esta manera, tenemos en (2.15) que

$$E\big[I[f]^2\big] = \sum_{i=1}^n a_i^2(t_i - t_{i-1}) = \int_a^b f^2(t)dt.$$

□

Si $(\Omega, \mathcal{F}, P)$ denota a nuestro espacio de probabilidad de donde tomamos a ω y $L^2(\Omega)$ es el espacio asociado de funciones cuadrado-integrables con respecto a la medida P en la σ-álgebra $\mathcal{F}$ de Ω, nótese que este último resultado implica que el operador I definido para $f \in L^2[a,b]$ escalonada:

$$f \mapsto I[f] \in L^2(\Omega)$$

es tal que preserva la norma de $f \in L^2[a,b]$, pues

$$\big\|I[f]\big\|^2_{L^2(\Omega)} := E\big[I[f]^2\big] = \int_a^b f^2(t)dt =: \big\|f\big\|^2_{L^2[a,b]}.$$

[6] Por el corolario 4.6.10 del texto de G. Casella y L. Berger [19] p. 184, que habla de la distribución de combinaciones lineales de v.a. normales.

Esto implica que I manda sucesiones de Cauchy en $L^2[a,b]$ a sucesiones de Cauchy en el espacio de Hilbert $L^2(\Omega)$, donde a $L^2(\Omega)$ le asignamos el producto interior

$$\langle X, Y \rangle_{L^2(\Omega)} = E[XY] \qquad \text{para } \; X, Y \in L^2(\Omega).$$

Este hecho nos ayudará a definir a la integral de Wiener para integrandos más generales, es decir para toda $f \in L^2[a,b]$. Para esto, sabemos que en para toda $f \in L^2[a,b]$ existe una sucesión $\{f_n\}_n \subset L^2[a,b]$ de funciones escalonadas tal que $f_n \to f$ en la norma de $L^2[a,b]$. Por el teorema 2.1 tenemos que $\{I[f_n]\}_n$ es una sucesión de Cauchy en $L^2(\Omega)$, por lo cual converge a un límite en $L^2(\Omega)$, y definimos

$$I[f] := \lim_{n\to\infty} I[f_n]. \tag{2.16}$$

Esta integral está bien definida para toda $f \in L^2[a,b]$, pues si $\{f'_n\}_n \subset L^2[a,b]$ es otra sucesión tal que $f'_n \to f$ en $L^2[a,b]$, entonces por la desigualdad del triángulo tendríamos que

$$\|f_n - f'_n\| \leq \|f_n - f\| + \|f - f'_n\| \to 0 \quad \text{si } n \to \infty.$$

Pero por el teorema 2.2, se sigue que

$$\|I(f_n) - I(f'_n)\|_{L^2(\Omega)} = \|f_n - f'_n\|_{L^2[a,b]} \to 0 \quad \text{si } n \to \infty,$$

por lo cual $\lim_{n\to\infty} I[f_n] = \lim_{n\to\infty} I[f'_n]$ en el sentido $L^2[a,b]$, lo cual prueba que $I[f]$ está bien definida. Esto motiva la siguiente definición:

Definición 2.2. (La integral de Wiener de $f \in \mathbf{L^2[a,b]}$) ***La integral de Wiener*** *de $f \in L^2[a,b]$ será definida por el límite* (2.16)*, y usaremos la notación*

$$I[f](\omega) = \int_a^b f(t)dB_t(\omega) \qquad \omega \in \Omega, \quad \text{casi seguramente}$$

donde ω denota que escogimos como integradora a una trayectoria $t \mapsto B_t(\omega)$ del movimiento browniano, que es continua casi seguramente.

El siguiente resultado es tomado del texto de H. Kuo [7].

Teorema 2.2. (Las integrales de Wiener son v.a. normales) *Para cada $f \in L^2[a,b]$, la integral de Wiener*

$$\int_a^b f(s)dB_s$$

es una variable aleatoria de distribución gaussiana tal que

$$E\left[\int_a^b f(s)dB_s\right] = 0 \quad \text{y} \quad \operatorname{Var}\left(\int_a^b f(s)dB_s\right) = \int_a^b f^2(s)ds$$

Demostración. Por el teorema 2.1 sabemos que el resultado es cierto para funciones escalonadas. Para $f \in L^2[a,b]$ arbitraria, usemos una sucesión $\{f_n\}_n \subset L^2[a,b]$ tal que $f_n \to f$ en $L^2[a,b]$. Cada $I[f_n]$ es una variable aleatoria normal $N(0,\sigma_n^2)$, con $\sigma_n^2 = \int_a^b f_n^2(s)ds$.

El teorema 17.2 del texto de J. Jacod y Philip Protter en [18] nos dice que la convergencia en $L^2(\Omega)$ de $I[f_n]$ a $I[f]$ implica la convergencia en probabilidad, y el teorema 18.2 del mismo [18] nos dice que la convergencia en probabilidad implica la convergencia en distribución, es decir, la *distribución*[7] de $I[f]$ es el límite débil de las distribuciones de $I[f_n] \sim N(0,\sigma_n^2)$. Por el teorema de continuidad de Lévy, que puede consultarse en P. Billingsley [2] p. 46 teorema 7.6, tenemos que las funciones características convergen para cada $t \in \mathbb{R}$:

$$e^{-\sigma_n^2 t^2/2} = E\big[e^{itI[f_n]}\big] \to E\big[e^{itI[f]}\big] \qquad \text{si } \ n \to \infty,$$

y dado que $f_n \to f$ en $L^2[a,b]$, entonces $\sigma_n^2 = \big\|f_n\big\|^2_{L^2[a,b]} \to \big\|f\big\|^2_{L^2[a,b]}$. La continuidad de la función exponencial implica que

$$E\big[e^{itI[f]}\big] = \exp\left(-\frac{t^2\big\|f\big\|^2_{L^2[a,b]}}{2}\right),$$

lo cual es la función característica de una normal $N\big(0, \big\|f\big\|^2_{L^2[a,b]}\big)$, es decir, $I[f]$ se distribuye como una variable aleatoria normal de varianza $\int_a^b f^2(s)ds$.

□

Al último enunciado del teorema anterior: $\mathrm{Var}\left(\int_a^b f(s)dB_s\right) = \int_a^b f^2(s)ds$ es un caso particular de un resultado conocido como la *isometría de Itô*, y puede ser generalizada a otros integrandos, ver el teorema 6.1 p. 82 del texto de M. Steele [12].

2.4. Un teorema de Girsanov

Con la integral de Wiener a nuestra disposición, enunciamos un caso particular del llamado teorema de Girsanov. Este es el análogo en tiempo continuo de lo que se hizo en la sección 2.2.

A continuación, $C_0[0,T]$ es el espacio de funciones continuas tales que $f(0) = 0$ y $\mathcal{B}\big(C_0[0,T]\big)$ es su σ-álgebra de Borel correspondiente. E_P y E_Q denotan los operadores de valor esperado bajos las leyes de probabilidad P y Q, respectivamente. Como la filtración usada en el siguiente resultado, se utiliza la *filtración asociada* al movimiento browniano, que es explicada en el apéndice en (A.2).

[7] Véase el apéndice para una definición de distribución. La convergencia débil se define de acuerdo a P. Billingsley en [2] p.7 como: *Decimos que la leyes de probabilidad P_n convergen débilmente a la medida de probabilidad P si $\int_S f dP_n \to \int_S f dP$ para toda función real f continua y acotada definida en S.*

Por otra parte, las variables aleatorias X_n convergen a X en distribución si las distribuciones P_n de X_n convergen débilmente a la distribución P de X. La convergencia en distribución puede verse en [2] p. 23.

Teorema 2.3. Teorema de Girsanov (Caso particular) *Sean* $\mu(s) \in L^2[0,T]$ *y* $\{B_t : 0 \leq t \leq T\}$ *un movimiento browniano estándar con respecto a la ley de probabilidad* P. *Entonces el proceso exponencial, (o exponencial estocástica)*

$$M_t = \exp\left(\int_0^t \mu(u)dB_u - \frac{1}{2}\int_0^t \mu^2(u)du\right) \tag{2.17}$$

es una martingala bajo P. *Finalmente, sea* Q *la medida en* $C_0[0,T]$ *definida por la ecuación*

$$\left.\frac{dQ}{dP}\right|_{\mathcal{F}_t} = M_t, \tag{2.18}$$

es decir, Q *es la medida de probabilidad*

$$Q(A) = E_P[\mathbf{1}_A\, M_T] \tag{2.19}$$

con $A \in \mathcal{F}_T$. *Entonces el proceso recorrido*

$$X_t := B_t - \int_0^t \mu(u)du \tag{2.20}$$

es un movimiento browniano estándar bajo Q*, en* $[0,T]$.

Demostración. Probemos que $E_P[M_t|\mathcal{F}_s] = M_s$ para todo $s < t$. Para esto, sumamos y restamos la expresión $\int_0^s \mu(u)dB_u - \frac{1}{2}\int_0^s \mu^2(u)du$ en la expresión para M_t:

$$\begin{aligned} E_P[M_t|\mathcal{F}_s] &= e^{\int_0^s \mu(u)dB_u - \frac{1}{2}\int_0^s \mu^2(u)du} E\left[\exp\left(\int_s^t \mu(u)dB_u - \frac{1}{2}\int_s^t \mu^2(u)du\right)\Big|\mathcal{F}_s\right] \\ &= M_s E\left[\exp\left(\int_s^t \mu(u)dB_u - \frac{1}{2}\int_s^t \mu^2(u)du\right)\right], \end{aligned}$$

donde la primera igualdad se sigue porque una vez que tenemos la información de $\mathcal{F}_s$, conocemos a $\int_0^s \mu(u)dB_u$, y la segunda igualdad se sigue porque la integral $\int_s^t \mu(u)dB_u$ no depende de tiempos menores a s, por los incrementos independientes del movimiento browniano que son usados para definirla. Pero por el teorema 2.2, tenemos que

$$E\left[\exp\left(\int_s^t \mu(u)dB_u\right)\right] = \exp\left(\frac{1}{2}\int_s^t \mu^2(u)du\right),$$

pues es la generadora de momentos de una variable aleatoria normal $N\big(0, \int_s^t \mu^2(u)du\big)$. Esto implica que

$$E\left[\exp\left(\int_s^t \mu(u)dB_u - \frac{1}{2}\int_s^t \mu^2(u)du\right)\right] = 1,$$

de donde se sigue que en la expresión de arriba tenemos que $E_P[M_t|\mathcal{F}_s] = M_s$, es decir $\{M_t\}$ es una martingala bajo P.

Para probar que $\{X_t\}$ dado por (2.20) es un movimiento browniano estándar bajo Q, es claro que $\{X_t\}$ tiene trayectorias casi seguramente continuas y que sus incrementos son independientes, por ser independiendes los incrementos de $\{B_t\}$. Probemos que la función característica de sus incrementos bajo Q, es igual que la que correponde al movimiento browniano estándar, para concluir que es un movimento browniano. Esto se seguirá de la fórmula siguiente:

Con $\{X_t\}$ y Q como definidos en (2.20) y (2.19) respectivamente, para toda $f : [0,T] \longrightarrow \mathbb{C}$ acotada y determinista se sigue que

$$E_Q\left[\exp\left(\int_0^T f(s)dX_s\right)\right] = \exp\left(\frac{1}{2}\int_0^T f^2(s)ds\right). \tag{2.21}$$

Para esto, por definición $E_Q\left[\exp\left(\int_0^T f(s)dX_s\right)\right] = E_P\left[M_T\exp\left(\int_0^T f(s)dX_s\right)\right]$, y tenemos que

$$\begin{aligned}
&E_P\left[M_T\exp\left(\int_0^T f(s)dX_s\right)\right] = \\
&E_P\left[\exp\left(\int_0^T \mu(s)dB_s - \frac{1}{2}\int_0^T \mu^2(s)ds + \int_0^T f(s)dB_s - \int_0^T f(s)\mu(s)ds\right)\right] \\
&= E_P\left[\exp\left(\int_0^T \big(\mu(s)+f(s)\big)dB_s - \int_0^T f(s)\mu(s)ds - \frac{1}{2}\int_0^T \mu^2(s)ds\right)\right],
\end{aligned}$$

que sumando y restando $\frac{1}{2}\int_0^T f^2(s)ds$ de convierte en:

$$\begin{aligned}
&E_P\left[M_T\exp\left(\int_0^T f(s)dX_s\right)\right] \\
&= \exp\left(\frac{1}{2}\int_0^T f^2(s)ds\right) E_P\left[\exp\left(\int_0^T \big(\mu(s)+f(s)\big)dB_s - \frac{1}{2}\int_0^T \big(\mu(s)+f(s)\big)^2 ds\right)\right].
\end{aligned}$$

Pero nótese en esta última expresión, que dado que f es acotada y determinista en $[0,T]$ entonces $\mu + f \in L^2[0,T]$, de tal forma que el lado derecho de la última igualdad tiene implícita a una martingala $\hat{M}_t$ bajo P similar a M_t, lo cual quiere decir que

$$E_P\left[\exp\left(\int_0^T \big(\mu(s)+f(s)\big)dB_s - \frac{1}{2}\int_0^T \big(\mu(s)+f(s)\big)^2 ds\right)\right] = E[\hat{M}_0] = 1$$

y se sigue (2.21). Esta formula implica que para todo conjunto finito de tiempos $0 = t_0 < \ldots < t_n = T$ los incrementos $(X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}})$ de $\{X_t\}$ cumplen que si $i = \sqrt{-1}$ y $\theta_k \in \mathbb{R}$ para todo $1 \leq k \leq n$ y definimos:

$$f(s) := \sum_{k=1}^{n} i\,\theta_k \mathbf{1}_{[t_{k-1},t_k)}(s) \qquad s \in [0,T]$$

entonces tenemos que

$$\int_0^T f(s)dX_s = i\sum_{k=1}^{n}\theta_k \int_{t_{k-1}}^{t_k} dX_s$$

y que

$$\int_0^T f^2(s)ds = -\sum_{k=1}^{n}\theta_k^2 \int_{t_{k-1}}^{t_k} ds,$$

de tal forma que en la ecuación (2.21) tenemos que

$$E_Q\left[\exp\left(i\sum_{k=1}^{n}\theta_k\big(X_{t_k}-X_{t_{k-1}}\big)\right)\right] = \exp\left(-\frac{1}{2}\sum_{k=1}^{n}\theta_k^2(t_k-t_{k-1})\right).$$

Esta última igualdad nos indica que la función característica con respecto de la medida Q del vector $(X_{t_1}, X_{t_2}-X_{t_1}, \ldots, X_{t_n}-X_{t_{n-1}})$ es igual al de un movimiento browniano estándar, y esto es suficiente para probar que $\{X_t\}$ es un movimiento browniano estándar bajo[8] Q. □

Un ejemplo que ilustra la importancia del teorema de Girsanov se puede ver en finanzas, que es el área donde interesa resolver el problema de este trabajo. Como vimos en el Problema 1.1 del capítulo 1, queremos minimizar a la integral $E_P\left[G^2(\sqrt{\epsilon}B)\frac{dP}{dQ}\right]$, que depende de una medida dada por

$$\frac{dP}{dQ}.$$

El teorema de Girsanov nos da una familia de medidas indizadas por los términos de deriva $h \in L^2[0,T]$, y nos proporciona trayectorias $\{X_t\}$ como en (2.20) tales que bajo el cambio de medida dado por dQ^h/dP, donde usamos el superíndice h para indicar la dependencia de la medida del término de deriva, $\{X_t\}$ se ve como un movimiento browniano sin deriva, es decir, es martingala.

Debido a condiciones y restricciones que se presentan en aplicaciones en finanzas[9], en el capítulo 4 se empleará el teorema de Girsanov para abordar el Problema 1.1 de este trabajo, al buscar una función h que usaremos como término de deriva y que minimice la cantidad (1.6) del Problema 1.1.

[8] Véase el apéndice Definición A.11 para la definición de movimiento browniano utilizada aquí.

[9] Para una explicación de estas condiciones de finanzas, véase la sección 1.2.2 del texto de P. Glasserman [4] y las referencias ahí contenidas.

Capítulo 3

Grandes Desviaciones

3.1. Introducción

En este capítulo se exponen algunas ideas básicas de grandes desviaciones. Este área de las matemáticas está ligada a la ley de grandes números. En lo que respecta a la llamada "teoría de grandes desviaciones", tomamos las siguientes palabras de Jean-Dominique Deuschel y Daniel Stroock en el prefacio de su libro [16]:

"*...no hay una "teoría" de grandes desviaciones, así como no hay una "teoría" de ecuaciones diferenciales parciales, y lo que pasamos por "teoría" es en realidad el conjunto de técnicas que han sido aplicadas exitosamente a situaciones especiales y que por lo tanto valen la pena intentar aplicar en situaciones cercanamente relacionadas.*"

Este capítulo está basado principalmente en el texto de los autores Mark Freidlin y Alexander Wenztell [15] y el libro de A. Dembo y O. Zeitouni [17]. La extensión del lema de Varadhan está basada en el artículo de P. Guasoni y S. Robertson en [20].

Así pues, veamos el siguiente ejemplo, que es bastante ilustrativo de lo que se intenta hacer en grandes desviaciones.

Supongamos que tenemos n variables aleatorias independientes e idénticamente distribuidas $X_1, X_2, \ldots, X_n$. Supongamos que son variables aleatorias normales $N(0,1)$ y consideremos su promedio

$$\hat{S}_n := \frac{1}{n}\sum_{k=1}^{n} X_k. \tag{3.1}$$

La ley débil de los grandes números (véase Casella y Berger [19]), nos dice que para toda $\delta > 0$:

$$\lim_{n\to\infty} P\{|\hat{S}_n| \geq \delta\} = 0$$

y el teorema central del límite nos dice que

$$\lim_{n\to\infty} P\{\sqrt{n}\hat{S}_n \geq \delta\} = \frac{1}{\sqrt{2\pi}}\int_{\delta}^{\infty} e^{-x^2/2}dx.$$

Estos dos resultados nos hablan de una convergencia en distribución. Pero podemos decir aún un poco más que lo anterior, pues para cada $n \in \mathbb{N}$, tenemos que $\hat{S}_n \sim N(0, 1/n)$, de tal forma que

$$P\{\hat{S}_n \geq \delta\} = \frac{\sqrt{n}}{\sqrt{2\pi}} \int_\delta^\infty e^{-nx^2/2} dx,$$

que con el cambio de variable $y = \sqrt{n}x - \sqrt{n}\delta$, esto se convierte en

$$P\{\hat{S}_n \geq \delta\} = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-(y+\sqrt{n}\delta)^2/2} dy,$$

y tenemos que:

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-(y+\sqrt{n}\delta)^2/2} dy \leq & \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-(y^2+n\delta^2)/2} dy \\ = & e^{-n\delta^2/2} \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-y^2/2} dy \\ = & \frac{1}{2} e^{-n\delta^2/2}. \end{aligned}$$

Es decir, que tenemos la siguiente cota superior

$$P\{\hat{S}_n \geq \delta\} \leq \frac{1}{2} e^{-n\delta^2/2},$$

que nos indica que la probabilidad de que $\hat{S}_n$ supere a $\delta > 0$ decae exponencialmente conforme crece δ, y la *tasa* a la cual decae exponencialmente está acotada por

$$\log P\{\hat{S}_n \geq \delta\} \leq \log\left(\frac{1}{2} e^{-n\delta^2/2}\right) = \log\left(\frac{1}{2}\right) - \frac{n\delta^2}{2}. \tag{3.2}$$

Por otra parte, de un modo análogo tenemos que

$$\begin{aligned} P\{\hat{S}_n \geq \delta\} = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-(y+\sqrt{n}\delta)^2/2} dy &\geq \frac{1}{\sqrt{2\pi}} \int_0^1 e^{-(y+\sqrt{n}\delta)^2/2} dy \\ &\geq \frac{1}{\sqrt{2\pi}} e^{-(1+\sqrt{n}\delta)^2/2} \\ &= \frac{1}{\sqrt{2\pi}} e^{-(1+2\sqrt{n}\delta)/2} e^{-n\delta^2/2} \end{aligned}$$

y esto implica la cota inferior de la “tasa de decaimiento” de la probabilidad

$$\log\left(\frac{1}{\sqrt{2\pi}} e^{-(1+2\sqrt{n}\delta)/2}\right) - \frac{n\delta^2}{2} \leq \log P\{\hat{S}_n \geq \delta\}. \tag{3.3}$$

Las cotas (3.2) y (3.3) nos están diciendo que

$$\lim_{n\to\infty} \frac{1}{n} \log P\{\hat{S}_n \geq \delta\} = -\frac{\delta^2}{2}. \tag{3.4}$$

Este último límite nos está dando una afirmación cuantitativa de un dato adicional al proporcionado por el teorema central del límite y la ley débil de los grandes números: el teorema central del límite nos habla de cómo converge la distribución de la caminata aleatoria $\hat{S}_n$, la ley débil de grandes números nos habla de cómo las probabilidades de alejarse del valor esperado de una v.a. normal de media 0, convergen a cero conforme n crece, y finalmente la ecuación (3.4) nos habla de la tasa a la cual decae a cero la probabilidad de estar fuera de una vecindad que contenga a 0. Conforme $n \to \infty$, de acuerdo a la ley débil de grandes números los valores típicos relativo a P se van concentrando alrededor de la media 0. Hallar esta tasa de decadencia exponencial es uno de los objetivos que se persiguen en grandes desviaciones.

Por lo que acabamos de decir, eventualmente, conforme n crece, quedar *fuera* de un intervalo que contenga al cero es un evento raro, y conforme hacemos los pasos de la caminata aleatoria pequeños haciendo a n tender a infinito, obtenemos una tasa de este ritmo exponencial de decadencia de las probabilidades conforme nos alejamos del valor esperado del límite de la caminata aleatoria. Es decir, tenemos una medida de qué tan "desviado" está un evento de lo que es en cierto sentido normal ver para las medidas o distribuciones conforme convergen en algún sentido. En el caso de $\hat{S}_n$ aquí tratado, lo "normal" o "no desviado" es ver con mucha frecuencia al valor esperado conforme los pasos de la caminata aleatoria se van haciendo más pequeños. ¿Qué tan desviado está de lo "regular" ver un valor de $\hat{S}_n > x > 0$ conforme n tiende a infinito? La medida de eso nos lo da (3.4), es decir, $-x^2/2$.

3.2. El Principio de Grandes Desviaciones

Hablando con un poco más de generalidad, vamos a querer buscar algo análogo a lo que hicimos para la caminata aleatoria, pero para otros procesos o variables aleatorias diferentes. Esto se formula en términos generales: dado una familia $\{X_\epsilon\}_{\epsilon>0}$ de variables aleatorias con valores en un espacio métrico $\mathcal{X}$ dotado de su σ-álgebra de Borel $\mathcal{B}$, si $\{X_\epsilon\}_{\epsilon>0}$ induce en $\mathcal{B}$ a la familia de medidas $\{\mu_\epsilon\}_{\epsilon>0}$ de la forma

$$\mu_\epsilon(A) := P\{X_\epsilon \in A\} \quad \text{para todo}\, A \in \mathcal{B},$$

y la familia de medidas es tal que μ_ϵ converge en algún sentido (por ejemplo, débilmente) a una medida delta de Dirac δ_q (con $q \in \mathcal{X}$) conforme ϵ tiende a cero, entonces es claro que en cierto sentido, eventualmente el punto q (o la ocurrencia del evento de dar con $q \in \mathcal{X}$) es un valor típico, de acuerdo a como se va concentrando el peso en q conforme $\epsilon \to 0$. Si $U \in \mathcal{B}$ es una vecindad abierta de q, entonces es razonable averiguar qué tan rápido decae a cero la probabilidad

$$P\{X_\epsilon \notin U\}$$

conforme ϵ tiende a cero. En el caso de $\hat{S}_n$ tratado arriba, el punto q en el que se concentra la medida es el valor esperado. A continuación se da una manera eficiente de abordar estas cuestiones para situaciones más generales. Como nos va a interesar el caso continuo, pues usaremos el análogo de las caminatas pero en tiempo continuo, es decir, el movimiento browniano, entonces damos una definición que se nos acomodará bien.

Definición 3.1. (Principio de Grandes Desviaciones) *Sea $(\mathcal{X},\mathcal{B})$ un espacio medible, donde $\mathcal{X}$ es un espacio métrico y $\mathcal{B}$ es su σ-álgebra de Borel. Sea una función semicontinua inferiormente $I:\mathcal{X}\longrightarrow[0,\infty]$. Se dice que una familia de medidas $\{\mu_\epsilon\}_{\epsilon\in(0,\delta)}$ definidas en el espacio medible $(\mathcal{X},\mathcal{B})$ satisface un **principio de grandes desviaciones** (PGD) con **función de tasa buena** I si*

*(i) Los **conjuntos de nivel** $\Phi_I(\alpha):=\big\{x\in\mathcal{X}:I(x)\leq\alpha\big\}$ son compactos para todo $\alpha\in\mathbb{R}$.*

(ii) Para todo boreliano $A\in\mathcal{B}$,

$$-\inf_{x\in\overset{\circ}{A}} I(x)\leq \liminf_{\epsilon\to 0}\epsilon\log\mu_\epsilon(A)\leq\limsup_{\epsilon\to 0}\epsilon\log\mu_\epsilon(A)\leq-\inf_{x\in\overline{A}}I(x).$$

Como vemos en la definición anterior, si la aplicamos al caso discreto[1] de la caminata aleatoria con $\epsilon:=1/n$ y con μ_n la distribución de cada caminata $\hat{S}_n=\frac{1}{n}\sum_{k=1}^n X_k$, entonces tenemos que $\mathcal{X}=\mathbb{R}$, la función de tasa buena es $I(x)=x^2/2$, como se podrá fácilmente verificar a partir de las cuentas hechas anteriormente.

Nosotros ocuparemos principalmente en este trabajo dos resultados de grandes desviaciones: el lema de Schilder y el lema integral de Varadhan. El lema de Schilder es un análogo en tiempo continuo al resultado de la caminata aleatoria que obtuvimos en la introducción a este capítulo. Respecto al lema de Varadhan, en realidad emplearemos una pequeña extensión de este.

En lo que sigue, $(\mathcal{X},\mathcal{B})$ será un espacio métrico con su σ-álgebra de Borel y $\{Z_\epsilon\}$ será una familia de variables aleatorias con valores en $\mathcal{X}$. La familia de medidas inducidas por las variables aleatorias $\{Z_\epsilon\}$ será denotada por $\{\mu_\epsilon\}$. En particular nos interesará cuando

$$\mathcal{X}=C_0[0,T]:=\big\{\omega\in C[0,T]:\omega_0=0\big\}$$

donde $\mathcal{X}$ posee la norma uniforme y $Z_\epsilon=\sqrt{\epsilon}B$, donde B es un movimiento browniano estándar. Entonces la familia de medidas $\{\mu_\epsilon\}$ estarán definidas en $\mathcal{B}\big(C_0[0,T]\big)$ y serán tales que

$$\mu_\epsilon(A):=P\{Z_\epsilon\in A\}\quad\text{para todo}\ \ A\in\mathcal{B},$$

es decir, en el caso que nos interesará, la familia de medidas es la inducida en $C_0[0,T]$ por $\sqrt{\epsilon}B$. Como escoger una $\omega\in\Omega$ es equivalente a escoger una trayectoria $t\mapsto B_t(\omega)$ en $C_0[0,T]$, y la probabilidad de

$$P\Big\{\omega\in\Omega:t\mapsto B_t(\omega)\in C[0,T]\Big\}=1,$$

[1] Existe una definición semejante para caminatas aleatorias, con $\epsilon=1/n$ y tomando los límites conforme $n\to\infty$, pero a nosotros nos interesa el caso continuo en el problema a resolver en este trabajo.

entonces usaremos la notación $\omega_t := B_t(\omega)$ para indicar que se ha escogido una trayectoria de $C_0[0,T]$ (al escoger a $\omega \in \Omega$) y se ha evaluado al instante t.

Definición 3.2. (El espacio Clásico de Wiener) *Al espacio métrico $\mathcal{X} = C_0[0,T]$ con su norma uniforme y la medida inducida en $\mathcal{B}(\mathcal{X})$ por un movimiento browniano estándar $\{B_t\}_{t\geq 0}$ lo denotamos como **el espacio clásico de Wiener**, denotado $\mathbb{W}_T$.*

3.3. El Lema de Schilder

En esta sección se probará el lema de Schilder, herramienta muy útil en la teoría de grandes desviaciones aplicada al movimiento browniano. La exposición del lema de Schilder que se toma aquí está basada en Freidlin y Wenztell [15].

Como trabajaremos con movimientos brownianos definidos en un intervalo finito $[0,T]$, con término de deriva determinista de la forma $h(t)$ (queremos que esté recorrido el movimiento browniano por una función h, no por su integral como en el teorema de Girsanov) de tal forma que tomamos $h \in C[0,T]$ tal que

$$h(t) = \int_0^t \dot{h}(s)ds, \quad \text{para cada } \ t \in [0,T],$$

de tal manera que también pediremos que $h(0) = 0$, y vamos a hacer uso de la integral de Wiener, que definimos en el capítulo anterior en el espacio $L^2[0,T]$. Entonces requeriremos que

$$\int_0^T \dot{h}^2(t)dt < \infty.$$

Esto induce la siguiente definición.

Definición 3.3. (El espacio Clásico de Cameron-Martin) *El espacio clásico de Cameron-Martin en el intervalo $[0,T]$ es el espacio funcional de las funciones absolutamente continuas[2] tales que:*

$$\mathbb{H}_T := \left\{ h \in AC[0,t] : h_0 = 0, \int_0^T \dot{h}^2(t)dt < \infty \right\}.$$

[2]Una función f es *absolutamente continua* en un intervalo $[a,b]$ si dado $\epsilon > 0$ existe $\delta > 0$ tal que

$$\sum_{i=1}^{n} |f(x_i) - f(y_i)| < \epsilon$$

para toda partición de $[a,b]$ de subintervalos de la forma $\{(x_i, y_i)\}$ tales que no se intersectan (es decir, $(x_i,y_i) \cap (x_j,y_j) = \emptyset$ para todo $i \neq j$) y que cumple

$$\sum_{i=1}^{n} |x_i - y_i| < \delta.$$

Equivalentemente, por el teorema 14 y el corolario 15 p. 110 del texto de H.L. Royden [11], la derivada $\dot{f}$ de f existe c.d. en $[a,b]$ y $f(t) = f(a) + \int_a^t \dot{f}(s)ds$, para $t \in [a,b]$. Al conjunto de funciones absolutamente continuas en $[a,b]$ se les denotará por $AC[a,b]$.

Haremos uso pues, del espacio clásico de Cameron-Martin. Nótese que el requerimiento de integrabilidad $\dot{h} \in L^2[0,T]$ nos permite usar el teorema de Girsanov 2.3.

En lo que sigue, Z_ϵ será el movimiento browniano escalado dado por $Z_\epsilon := \sqrt{\epsilon}B$ y la familia de medidas inducidas por las variables aleatorias $\{Z_\epsilon\}$ en $C_0[0,T]$ será denotada por $\{\mu_\epsilon\}$.

Como estaremos usando el proceso escalado, junto con el teorema de Girsanov, entonces en la expresión de la exponencial estocástica de dicho teorema, el término de deriva queda escalado como $\frac{1}{\sqrt{\epsilon}}h(t)$, pues el movimiento browniano escalado y recorrido queda como

$$\sqrt{\epsilon}B_t - \int_0^t \dot{h}(s)ds =: Y_{\epsilon,t},$$

y esto puede ser visto como

$$\frac{1}{\sqrt{\epsilon}}Y_{\epsilon,t} = B_t - \frac{1}{\sqrt{\epsilon}}h(t).$$

De esta manera, si denotamos por ν_ϵ a la distribución del proceso Y_ϵ definido arriba, por el teorema de Girsanov obtenemos que ν_ϵ es una medida absolutamente continua con respecto a μ_ϵ, y la derivada de Radon-Nikodym de ν_ϵ con respecto a μ_ϵ es

$$\frac{d\nu_\epsilon}{d\mu_\epsilon} = \exp\left(\frac{1}{\sqrt{\epsilon}}\int_0^T \dot{h}(t)dB_t - \frac{1}{2\epsilon}\int_0^T \dot{h}^2(t)dt\right). \tag{3.5}$$

Usaremos esta expresión de la exponencial estocástica para probar el siguiente lema, que nos habla de la probabilidad de que el proceso Z_ϵ esté alejado en el sentido de la norma uniforme en $C[0,T]$ de la trayectoria determinada por el término de deriva $h \in \mathbb{H}_T$. Nos ayudará para obtener una cota inferior de un principio de grandes desviaciones para el movimiento browniano. Sabemos que para el movimiento browniano Z_ϵ recorrido por la deriva h, el valor esperado de $Z_{\epsilon,t} - h(t)$ es $h(t)$, pero queremos saber qué tan rápido terminará la trayectoria de $Z_\epsilon - h$ en un tubo de radio $\delta > 0$ alrededor de h.

Lema 3.1. *Para cualesquiera $\delta, \gamma > 0$, si definimos*

$$I(x) = \begin{cases} \frac{1}{2}\int_0^T \dot{x_t}^2 dt & si\ x \in \mathbb{H}_T \\ \infty & si\ x \in \mathbb{W}_T \setminus \mathbb{H}_T, \end{cases}$$

entonces para cada $h \in \mathbb{H}_T$

$$P\left\{\|Z_\epsilon - h\|_\infty < \delta\right\} \geq \exp\left(-\frac{1}{\epsilon}(I(h)+\gamma)\right).$$

Demostración. Por el teorema de Girsanov probado en el capítulo anterior, el hecho de que $h \in \mathbb{H}_T$ nos garantiza que h induce un cambio de medida en $C_0[0,T]$ (estamos identificando a $\omega \in \Omega$ con una trayectoria en $C_0[0,T]$)

$$\frac{d\nu_\epsilon}{d\mu_\epsilon}(\sqrt{\epsilon}\omega) = \exp\left(\frac{1}{\sqrt{\epsilon}}\int_0^T \dot{h}(t)d\omega_t - \frac{1}{2\epsilon}\int_0^T \dot{h}^2(t)dt\right).$$

Por el teorema de Girsanov, $Y_{\epsilon,t} := \sqrt{\epsilon}B_t - h(t)$ *con respecto a la medida* ν_ϵ se ve como un movimiento browniano centrado en cero (de valor esperado cero), de tal manera que

$$\begin{aligned} P\left\{\|Z_\epsilon - h\|_\infty < \delta\right\} &= P\left\{\|Y_\epsilon\|_\infty < \delta\right\} \\ &= \int_{\|\sqrt{\epsilon}\omega\|_\infty < \delta} \frac{d\nu_\epsilon}{d\mu_\epsilon}(\sqrt{\epsilon}\omega) \cdot d\mu_\epsilon(\sqrt{\epsilon}\omega) \\ &= e^{-\frac{I(h)}{\epsilon}} \int_{\|\sqrt{\epsilon}\omega\|_\infty < \delta} \exp\left(\frac{1}{\sqrt{\epsilon}} \int_0^T \dot{h}(t) d\omega_t\right) d\mu_\epsilon(\sqrt{\epsilon}\omega). \end{aligned} \tag{3.6}$$

Aquí usamos que $\frac{1}{\sqrt{\epsilon}} Y_{\epsilon,t}$ con respecto a ν_ϵ se ve como un movimiento browniano estándar: por el principio de reflexión para el movimiento browniano A.2 del apéndice:

$$\begin{aligned} P\left\{\|Y_\epsilon\|_\infty \geq \delta\right\} &= P\left\{\left\|\frac{Y_\epsilon}{\sqrt{\epsilon}}\right\|_\infty \geq \delta/\sqrt{\epsilon}\right\} \\ &= P\left\{\|B\|_\infty \geq \delta/\sqrt{\epsilon}\right\} \\ &= 2P\{B_T \geq \delta/\sqrt{\epsilon}\} \\ &= \sqrt{\frac{2}{\pi T}} \int_{\delta/\sqrt{\epsilon}}^\infty e^{-y^2/2T} dy, \end{aligned}$$

de tal manera que si hacemos tender ϵ a cero, obtenemos que

$$\lim_{\epsilon \to 0} P\left\{\|Y_\epsilon\|_\infty < \delta\right\} = \lim_{\epsilon \to 0} \left(1 - P\left\{\|B\|_\infty \geq \delta/\sqrt{\epsilon}\right\}\right) = 1. \tag{3.7}$$

Por esto último, para ϵ suficientemente pequeña tenemos que $P\left\{\|Y_\epsilon\|_\infty < \delta\right\} \geq 3/4$.

Por otra parte, aplicando la desigualdad de Chebyshev junto con la isometría de Itô (Teorema 2.2) obtenemos que

$$\begin{aligned} P\left\{\frac{1}{\sqrt{\epsilon}} \int_0^T \dot{h}(t) d\omega_t \leq -2\sqrt{\frac{2}{\epsilon} \mathrm{I}(h)}\right\} &\leq P\left\{\left|\frac{1}{\sqrt{\epsilon}} \int_0^T \dot{h}(t) d\omega_t\right| \geq 2\sqrt{\frac{2}{\epsilon}} \sqrt{\mathrm{I}(h)}\right\} \\ &\leq \frac{1}{8\mathrm{I}(h)} \mathrm{Var}\left(\int_0^T \dot{h}(t) d\omega_t\right) \\ &= \frac{1}{8\mathrm{I}(h)} E\left[\left(\int_0^T \dot{h}(t) d\omega_t\right)^2\right] = \frac{1}{4} \end{aligned}$$

pues la isometría de Itô nos dice que $E\left[\left(\int_0^T \dot{h}(t) d\omega_t\right)^2\right] = \int_0^T \dot{h}(t)^2 dt = 2\mathrm{I}(h)$. De esta manera concluimos que

$$P\left\{\exp\left(\frac{1}{\sqrt{\epsilon}} \int_0^T \dot{h}(t) d\omega_t\right) \geq \exp\left(-2\sqrt{\frac{2}{\epsilon} \mathrm{I}(h)}\right)\right\} \geq \frac{3}{4}. \tag{3.8}$$

Sea $D := \left\{ e^{\frac{1}{\sqrt{\epsilon}} \int_0^T \dot{h}(t) d\omega_t} \geq e^{-2\sqrt{\frac{2}{\epsilon}\mathrm{I}(h)}} \right\}$. Entonces debido a (3.7), tenemos que para $\epsilon > 0$ suficientemente pequeña entonces $P\left\{\|Y_\epsilon\|_\infty < \delta\right\} \geq 3/4$, por lo cual se sigue que

$$P\Big(D \cap \{\|Y_\epsilon\|_\infty < \delta\}\Big) \geq \frac{1}{2}.$$

Esto es así porque en la ecuación (3.8) tenemos que $P(D) \geq 3/4$ y para $\epsilon > 0$ pequeña $P\left\{\|Y_\epsilon\|_\infty < \delta\right\} \geq 3/4$, por lo cual

$$\frac{6}{4} \leq P(D) + P\left\{\|Y_\epsilon\|_\infty < \delta\right\},$$

y para que se cumpla la desigualdad

$$P\big(D \cup \{\|Y_\epsilon\|_\infty < \delta\}\big) = P(D) + P\left\{\|Y_\epsilon\|_\infty < \delta\right\} - P\big(D \cap \{\|Y_\epsilon\|_\infty < \delta\}\big) \leq 1,$$

entonces es necesario que $P\left(D \cap \{\|Y_\epsilon\|_\infty < \delta\}\right) \geq \frac{1}{2}$. De esto último se sigue que

$$\begin{aligned}
&\int_{\|\sqrt{\epsilon}\omega\|_\infty < \delta} \exp\left(\frac{1}{\sqrt{\epsilon}} \int_0^T \dot{h}(t) d\omega_t\right) d\mu_\epsilon(\sqrt{\epsilon}\omega) \\
&\qquad \geq \int_{\{\|\sqrt{\epsilon}\omega\|_\infty < \delta\} \cap D} \exp\left(\frac{1}{\sqrt{\epsilon}} \int_0^T \dot{h}(t) d\omega_t\right) d\mu_\epsilon(\sqrt{\epsilon}\omega) \\
&\qquad \geq \int_{\{\|\sqrt{\epsilon}\omega\|_\infty < \delta\} \cap D} e^{-2\sqrt{\frac{2}{\epsilon}\mathrm{I}(h)}} d\mu_\epsilon(\sqrt{\epsilon}\omega) \\
&\qquad = e^{-2\sqrt{\frac{2}{\epsilon}\mathrm{I}(h)}} P\left(D \cap \{\|Y_\epsilon\|_\infty < \delta\}\right) \geq \frac{1}{2} e^{-2\sqrt{\frac{2}{\epsilon}\mathrm{I}(h)}}.
\end{aligned}$$

Entonces por (3.6), vemos que

$$P\left\{\|Z_\epsilon - h\|_\infty < \delta\right\} \geq \frac{1}{2} e^{-\frac{\mathrm{I}(h)}{\epsilon} - 2\sqrt{\frac{2}{\epsilon}\mathrm{I}(h)}}$$

Pero para $\epsilon > 0$ suficientemente pequeña y cualquier $\gamma > 0$ es posible hacer que

$$\frac{1}{2} e^{-2\sqrt{2\mathrm{I}(h)/\epsilon}} \geq e^{-\gamma/\epsilon}$$

de tal forma que $P\left\{\|Z_\epsilon - h\|_\infty < \delta\right\} \geq \frac{1}{2} e^{-\frac{\mathrm{I}(h)}{\epsilon} - 2\sqrt{\frac{2}{\epsilon}\mathrm{I}(h)}} \geq e^{-\frac{1}{\epsilon}(\mathrm{I}(h) + \gamma)}$. $\square$

El siguiente resultado es solo una cota superior para la probabilidad de un evento que nos interesará para poder obtener una cota superior para un principio de grandes desviaciones.

Lema 3.2. *Sea $\Delta > 0$ y W_Δ una variable aleatoria de distribución normal $N(0, \Delta)$. Entonces para toda $z > 0$ se cumple que*

$$P\{W_\Delta > z\} \leq \frac{\sqrt{\Delta}}{z\sqrt{2\pi}} \exp\left(-z^2/2\Delta\right).$$

Demostración. Tenemos que mediante el cambio de variable $y = x/\sqrt{\Delta}$, entonces

$$P\{W_\Delta > z\} = \frac{1}{\sqrt{2\pi\Delta}} \int_z^\infty \exp\left(-x^2/2\Delta\right) dx = \frac{1}{\sqrt{2\pi}} \int_{z/\sqrt{\Delta}}^\infty \exp\left(-y^2/2\right) dy.$$

Pero en la región en que estamos integrando, $\frac{\sqrt{\Delta}y}{z} \geq 1$, de donde se sigue que

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_{z/\sqrt{\Delta}}^\infty \exp\left(-y^2/2\right) dy &\leq \frac{\sqrt{\Delta}}{\sqrt{2\pi}} \int_{z/\sqrt{\Delta}}^\infty \frac{y}{z} \exp\left(-y^2/2\right) dy \\ &= -\frac{\sqrt{\Delta}}{\sqrt{2\pi}} e^{-y^2/2} \Big|_{z/\sqrt{\Delta}}^\infty = \frac{\sqrt{\Delta}}{z\sqrt{2\pi}} e^{-z^2/2\Delta} \end{aligned}$$

□

El siguiente lema nos dará la cota superior para poder probar el lema de Schilder. Denotamos por d_∞ a la métrica inducida por la norma en $C[0,T]$, y si $x \in C[0,T]$ y $K \subset C[0,T]$ es un subconjunto, entonces $d_\infty(x,K) := \inf_{\theta \in K} d_\infty(x,\theta)$. Como siempre, $Z_{\epsilon,t} = \sqrt{\epsilon} B_t$.

Lema 3.3. *Para $s > 0$, definamos*

$$\Phi(s) := \left\{ x \in C[0,T] : x(0) = 0 \ \text{ y } \ I(x) = \frac{1}{2} \int_0^T |\dot{x}(t)|^2 dt \leq s \right\}.$$

Entonces para cualquier $\delta, \gamma > 0$ y $s > 0$ se cumple que

$$P\left\{ d_\infty\big(Z_\epsilon\,, \Phi(s)\big) \geq \delta \right\} \leq \exp\left(-\frac{s-\gamma}{\epsilon} \right) \tag{3.9}$$

Demostración. Básicamente aproximaremos las trayectorias de $\sqrt{\epsilon}B$ por una versión discretizada Y_ϵ^Δ de $\sqrt{\epsilon}B$ tal que las trayectorias de Y_ϵ sean más suaves (derivables salvo conjuntos discretos de puntos). Para esto tomamos subintervalos de $[0,T]$ de longitud $\Delta > 0$, donde por el momento solo pediremos que Δ sea tal que T/Δ sea un número entero. Con estas Δ definamos a Y_ϵ^Δ tal que su gráfica sea el polígono formado por interpolar linealmente a los vértices de la gráfica de Z_ϵ en las coordenadas $(0,0), (\Delta, Z_{\epsilon,\Delta}), (2\Delta, Z_{\epsilon,2\Delta}) \ldots, (T, Z_{\epsilon,T})$. El objetivo es usar esta trayectoria que es más manejable con respecto a la funcional I en el sentido que es finita I en Y_ϵ^Δ, pues en las trayectorias Z_ϵ sabemos que $I(\sqrt{\epsilon}B) = \infty$. Se harán cálculos de la probabilidad del evento de la expresión (3.9) con la discretización de $\sqrt{\epsilon}B$ y se probará que para una aproximación suficientemente buena de $\sqrt{\epsilon}B$, se puede probar (3.9).

Nótese que si $d_\infty(Z_\epsilon, Y_\epsilon^\Delta) < \delta$ y también ocurre que $d_\infty\big(Z_\epsilon\,, \Phi(s)\big) \geq \delta$, entonces $Y_\epsilon^\Delta \notin \Phi(s)$, por lo cual $I(Y_\epsilon^\Delta) > s$. Usando esto último y condicionando sobre el evento $\{d_\infty(Z_\epsilon, Y_\epsilon^\Delta) < \delta\}$ tenemos que

$$\begin{aligned} P\left\{ d_\infty\big(Z_\epsilon\,, \Phi(s)\big) \geq \delta \right\} &= P\left\{ d_\infty\big(Z_\epsilon\,, \Phi(s)\big) \geq \delta\,,\, d_\infty(Z_\epsilon, Y_\epsilon^\Delta) < \delta \right\} \\ &\quad + P\left\{ d_\infty\big(Z_\epsilon\,, \Phi(s)\big) \geq \delta\,,\, d_\infty(Z_\epsilon, Y_\epsilon^\Delta) \geq \delta \right\} \\ &\leq P\left\{ I(Y_\epsilon^\Delta) > s \right\} + P\left\{ d_\infty(Z_\epsilon, Y_\epsilon^\Delta) \geq \delta \right\}. \end{aligned} \tag{3.10}$$

Acotaremos la primera probabilidad de (3.10).

Debido a que $Y_\epsilon^\Delta(s)$ con $s \in [(k-1)\Delta, k\Delta]$ $(k = 0, 1, \ldots, T/\Delta)$ es el segmento de recta entre los dos puntos de $\mathbb{R}^2$: $((k-1)\Delta, \sqrt{\epsilon}B_{(k-1)\Delta})$ y $(k\Delta, \sqrt{\epsilon}B_{k\Delta})$, entonces tenemos que

$$\frac{d}{dt}Y_\epsilon^\Delta(s) = \sqrt{\epsilon}\,\frac{(B_{k\Delta} - B_{(k-1)\Delta})}{\Delta} \qquad \text{si } s \in \big((k-1)\Delta, k\Delta\big),$$

por lo cual:

$$\begin{aligned} I(Y_\epsilon^\Delta) =& \frac{1}{2}\int_0^T |\dot{Y}_\epsilon^\Delta(t)|^2 dt \\ =& \frac{1}{2}\sum_{k=1}^{T/\Delta}\int_{(k-1)\Delta}^{k\Delta} |\dot{Y}_\epsilon^\Delta(t)|^2 dt \\ =& \frac{\Delta\epsilon}{2}\sum_{k=1}^{T/\Delta}\frac{|B_{k\Delta} - B_{(k-1)\Delta}|^2}{\Delta^2} = \frac{\epsilon}{2}\sum_{k=1}^{T/\Delta}\frac{|B_{k\Delta} - B_{(k-1)\Delta}|^2}{\Delta}. \end{aligned}$$

Observemos que por incrementos estacionarios:

$$\xi_k = \frac{B_{k\Delta} - B_{(k-1)\Delta}}{\sqrt{\Delta}} \sim N(0,1),$$

y por ser independientes los incrementos, cada ξ_k es una variable aleatoria $N(0,1)$ independiente, de tal forma que

$$\sum_{k=1}^{T/\Delta}\frac{|B_{k\Delta} - B_{(k-1)\Delta}|^2}{\Delta} = \sum_{k=1}^{T/\Delta}\xi_k^2 \sim \chi^2_{T/\Delta}$$

donde $\chi^2_{T/\Delta}$ indica una variable aleatoria χ^2 con T/Δ grados de libertad. De este modo,

$$P\left\{I(Y_\epsilon^\Delta) > s\right\} = P\left\{\frac{\epsilon}{2}\sum_{k=1}^{T/\Delta}\xi_k^2 > s\right\} = P\left\{\sum_{k=1}^{T/\Delta}\xi_k^2 > \frac{2s}{\epsilon}\right\}, \tag{3.11}$$

y podemos acotar esta última probabilidad, usando la desigualdad exponencial de Chebyshev que nos dice[3] que para todo $\lambda > 0$ entonces $P\{X > \lambda\} \leq e^{-a\lambda}E[e^{aX}]$ para todo $a > 0$ tal que esté definido $E[e^{aX}]$. En este caso, $X \sim \chi^2_{T/\Delta}$, por lo cual

$$E[e^{aX}] = \left(\frac{1}{1-2a}\right)^{T/2\Delta} \qquad \text{para } a < 1/2 \tag{3.12}$$

(véase [Casella-Berger] p. 623 para la función generadora de momentos de una v.a. χ^2). Sea $\alpha > 0$. Usando $\lambda = \frac{2s}{\epsilon}$ y $a := \frac{1-\alpha}{2} < \frac{1}{2}$ en la desigualdad exponencial de Chebyshev

[3] Este resultado es probado en el lema A.3 del apéndice.

junto con la fórmula anterior para $E[e^{aX}]$, obtenemos que

$$\begin{aligned}P\left\{\sum_{k=1}^{T/\Delta}\xi_k^2>\frac{2s}{\epsilon}\right\}&\le\exp\left(-\frac{2s(1-\alpha)}{2\epsilon}\right)E\left[\exp\left(\frac{1-\alpha}{2}\sum_{k=1}^{T/\Delta}\xi_k^2\right)\right]\\&=\frac{1}{\left(1-2\left(\frac{1-\alpha}{2}\right)\right)^{T/2\Delta}}\exp\left(-\frac{s}{\epsilon}(1-\alpha)\right)=\alpha^{-T/2\Delta}e^{-(1-\alpha)s/\epsilon}.\end{aligned}$$

Dado que $\lim_{\Delta\to 0}\alpha^{-T/2\Delta}=0$, entonces para Δ suficientemente pequeño tenemos que $\alpha^{-T/2\Delta}<1/2$ (basta con $\Delta<T/2\log_\alpha(2)$). De esta manera, si $\alpha=\gamma/s>0$ obtenemos que si $\Delta<T/2\log_\alpha(2)$:

$$P\left\{I(Y_\epsilon^\Delta)>s\right\}=P\left\{\sum_{k=1}^{T/\Delta}\xi_k^2>\frac{2s}{\epsilon}\right\}\le\frac{1}{2}e^{-(1-\alpha)s/\epsilon}=\frac{1}{2}e^{-(s-\gamma)/\epsilon}. \tag{3.13}$$

Ahora acotamos al segundo sumando de (3.10). Como

$$\left\{\|Z_\epsilon-Y_\epsilon^\Delta\|_\infty\ge\delta\right\}\subseteq\bigcup_{k=1}^{T/\Delta}\left\{\max_{(k-1)\Delta\le t\le k\Delta}|Z_{\epsilon,t}-Y_{\epsilon,t}^\Delta|\ge\delta\right\},$$

entonces

$$\begin{aligned}P\left\{\|Z_\epsilon-Y_\epsilon^\Delta\|\ge\delta\right\}&\le\sum_{k=1}^{T/\Delta}P\left\{\max_{(k-1)\Delta\le t\le k\Delta}|Z_{\epsilon,t}-Y_{\epsilon,t}^\Delta|\ge\delta\right\}\\&=\frac{T}{\Delta}P\left\{\max_{0\le t\le\Delta}|Z_{\epsilon,t}-Y_{\epsilon,t}^\Delta|\ge\delta\right\},\end{aligned}$$

donde la última igualdad se sigue de los incrementos independientes y estacionarios del movimiento browniano. Por construcción de Y_ϵ^Δ, para $t\in[0,\Delta]$ tenemos que

$$Y_\epsilon^\Delta(t)=\frac{t}{\Delta}\sqrt{\epsilon}\omega_\Delta,$$

por lo cual

$$P\left\{\|Z_\epsilon-Y_\epsilon^\Delta\|_\infty\ge\delta\right\}\le\frac{T}{\Delta}P\left\{\max_{0\le t\le\Delta}|\sqrt{\epsilon}\omega_t-\frac{t}{\Delta}\sqrt{\epsilon}\omega_\Delta|\ge\delta\right\} \tag{3.14}$$

Nótese que si ocurre $\left\{\max_{0\le t\le\Delta}|\sqrt{\epsilon}\omega_t-\frac{t}{\Delta}\sqrt{\epsilon}\omega_\Delta|\ge\delta\right\}$, entonces por la desigualdad del triángulo es claro que debería ser que

$$\max_{0\le t\le\Delta}\left|\frac{t}{\Delta}\sqrt{\epsilon}\omega_\Delta\right|+\max_{0\le t\le\Delta}|\sqrt{\epsilon}\omega_t|=\sqrt{\epsilon}|\omega_\Delta|+\max_{0\le t\le\Delta}|\sqrt{\epsilon}\omega_t|\ge\delta,$$

y aquí se hace evidente que debe suceder el evento $\{\max_{0\leq t\leq\Delta}|\sqrt{\epsilon}\omega_t| \geq \delta/2\}$, pues en caso contrario no se cumpiría la desigualdad anterior, pues ocurriría que $|\omega_\Delta| \leq \max_{0\leq t\leq\Delta}|\sqrt{\epsilon}\omega_t| \leq \delta/2$. Así que hemos obtenido que

$$\left\{\max_{0\leq t\leq\Delta}|\sqrt{\epsilon}\omega_t - \frac{t}{\Delta}\sqrt{\epsilon}\omega_\Delta| \geq \delta\right\} \subset \left\{\max_{0\leq t\leq\Delta}|\sqrt{\epsilon}\omega_t| \geq \delta/2\right\}. \tag{3.15}$$

Pero si ocurre $\{\max_{0\leq t\leq\Delta}|\sqrt{\epsilon}\omega_t| \geq \delta/2\}$, y si usamos que

$$\min_{0\leq t\leq\Delta}\sqrt{\epsilon}\omega_t = -\max_{0\leq t\leq\Delta}\{-\sqrt{\epsilon}\omega_t\}$$

y que la distribución de $\sqrt{\epsilon}\omega_t$ es igual que la de $-\sqrt{\epsilon}\omega_t$, se sigue que por simetría del movimiento browniano junto con el principio de reflexión A.2 del apéndice:

$$\begin{aligned} P\left\{\max_{0\leq t\leq\Delta}|\sqrt{\epsilon}\omega_t| \geq \frac{\delta}{2}\right\} &\leq P\left\{\max_{0\leq t\leq\Delta}\sqrt{\epsilon}\omega_t \geq \frac{\delta}{2}\right\} + P\left\{\min_{0\leq t\leq\Delta}\sqrt{\epsilon}\omega_t \leq -\frac{\delta}{2}\right\} \\ &= 2P\left\{\max_{0\leq t\leq\Delta}\sqrt{\epsilon}\omega_t \geq \frac{\delta}{2}\right\} = 4P\left\{\sqrt{\epsilon}\omega_\Delta \geq \frac{\delta}{2}\right\}. \end{aligned}$$

De (3.15) y lo anterior se sigue en la ecuación (3.14) que

$$\begin{aligned} P\left\{\|Z_\epsilon - Y_\epsilon^\Delta\|_\infty \geq \delta\right\} &\leq \frac{T}{\Delta}P\left\{\max_{0\leq t\leq\Delta}|\sqrt{\epsilon}\omega_t - \frac{t}{\Delta}\sqrt{\epsilon}\omega_\Delta| \geq \delta\right\} \\ &\leq \frac{T}{\Delta}P\left\{\max_{0\leq t\leq\Delta}|\sqrt{\epsilon}\omega_t| \geq \frac{\delta}{2}\right\} \leq \frac{4T}{\Delta}P\left\{\sqrt{\epsilon}\omega_\Delta \geq \frac{\delta}{2}\right\}. \end{aligned} \tag{3.16}$$

Usando el lema 3.2 con $z := \delta/2\sqrt{\epsilon}$ en (3.16) obtenemos que

$$\begin{aligned} P\left\{\|Z_\epsilon - Y_\epsilon^\Delta\|_\infty \geq \delta\right\} &\leq \frac{4T}{\Delta}P\left\{\omega_\Delta \geq \frac{\delta}{2\sqrt{\epsilon}}\right\} \\ &\leq \frac{4T}{\Delta}\cdot\frac{\sqrt{\Delta}}{(\delta/2\sqrt{\epsilon})\sqrt{2\pi}}e^{-\delta^2/8\epsilon\Delta} = \sqrt{\epsilon}\frac{8T}{\delta\sqrt{2\pi\Delta}}e^{-\delta^2/8\epsilon\Delta}. \end{aligned} \tag{3.17}$$

Si escogemos Δ tal que $\Delta < \delta^2/4s$, entonces $1/\Delta > 4s/\delta^2$ y para $\gamma > 0$ es cierto que $e^{\gamma/2\epsilon} \geq 1$, de tal forma que

$$e^{-\delta^2/8\epsilon\Delta} < e^{-4s\delta^2/8\epsilon\delta^2} = e^{-s/2\epsilon} \leq e^{-s/2\epsilon}e^{\gamma/2\epsilon} = e^{-(s-\gamma)/2\epsilon}.$$

Escogiendo ϵ suficientemente chica, podemos hacer que $\sqrt{\epsilon}\frac{8T}{\delta\sqrt{2\pi\Delta}} < \frac{1}{2}$, de tal manera que en (3.17) obtenemos que

$$P\left\{\|Z_\epsilon - Y_\epsilon^\Delta\|_\infty \geq \delta\right\} \leq \frac{1}{2}e^{-(s-\gamma)/2\epsilon}.$$

Esto prueba el resultado, pues ambos sumandos de (3.10) son menores que $\frac{1}{2}e^{-(s-\gamma)/2\epsilon}$. $\square$

Ahora probamos que la funcional $I(x)$ definida en el lema 3.1 tiene conjuntos de nivel compactos.

Lema 3.4. *La funcional*

$$I(x) = \begin{cases} \frac{1}{2}\int_0^T \dot{x}_t{}^2 dt & si\ x \in \mathbb{H}_T \\ \infty & si\ x \in \mathbb{W}_T \setminus \mathbb{H}_T. \end{cases}$$

es tal que sus conjuntos de nivel $\Phi_I(\alpha) = \{x \in \mathbb{W}_T : I(x) \leq \alpha\}$ *son compactos.*

Demostración. Como $I : \mathbb{W}_T \longrightarrow [0,\infty]$, entonces para $\alpha < 0$ tenemos que $\Phi_I(\alpha) = \emptyset$, por lo cual tomamos $\alpha \geq 0$.

Sean $\epsilon > 0$ y $x \in \Phi_I(\alpha)$. Entonces debe ser que $x \in \mathbb{H}_T$, pues en $\mathbb{H}_T^c$ tenemos que $I \equiv \infty$. De este modo tenemos que $\Phi_I(\alpha) \subset \mathbb{H}_T$, y además, para toda $x \in \Phi_I(\alpha)$:

$$\|\dot{x}\|_{L^2} \leq \sqrt{2\alpha}.$$

Si $\delta = \epsilon^2/2\alpha$, entonces para $s,t \in [0,T]$ tales que $|s-t| < \delta$ y usando que $x \in \mathbb{H}_T$:

$$|x(t) - x(s)| = \left|\int_s^t \dot{x}(u)du\right| \leq \|\dot{x}\|_{L^2}\left(\int_s^t du\right)^{1/2} \leq \sqrt{2\alpha}|s-t|^{1/2} < \sqrt{2\alpha}\delta^{1/2} = \epsilon,$$

de tal forma que $\Phi_I(\alpha)$ es un conjunto de funciones equicontinuo. Del mismo modo, dado que

$$|x(t)| = \left|x(0) + \int_0^t \dot{x}(u)du\right| \leq |x(0)| + \|\dot{x}\|_{L^2}T^{1/2} \leq |x(0)| + \sqrt{2\alpha T},$$

entonces $\Phi_I(\alpha)$ es un conjunto de funciones puntualmente acotada. Por el teorema de Arzelá-Ascoli, esto implica que $\Phi_I(\alpha)$ es compacto en el espacio métrico $\mathbb{W}_T$.

□

Ahora probamos el lema de Schilder, que es un teorema clásico de grandes desviaciones.

Teorema 3.1. (Lema de Schilder) *La familia de medidas* $\{\mu_\epsilon\}_\epsilon$ *inducidas por el movimiento browniano escalado* $\{\sqrt{\epsilon}B\}$ *satisfacen en* $\mathbb{W}_T$ *un principio de grandes desviaciones con función de tasa buena*

$$I(x) = \begin{cases} \frac{1}{2}\int_0^T \dot{x}_t{}^2 dt & si\ x \in \mathbb{H}_T \\ \infty & si\ x \in \mathbb{W}_T \setminus \mathbb{H}_T \end{cases}$$

Demostración. Sea A en $\mathcal{B}(\mathbb{W}_T)$. Probemos la desigualdad

$$-\inf_{x \in \mathring{A}} I(x) \leq \liminf_{\epsilon \to 0} \epsilon \log \mu_\epsilon(A).$$

Para esto, usemos que $-\operatorname{ínf}_{x\in\mathring{A}} I(x) = \sup_{x\in\mathring{A}}\{-I(x)\}$. A su vez, como buscamos el supremo en $\mathring{A}$ y ocurre que $-I \equiv -\infty$ en $\mathbb{W}_T \setminus \mathbb{H}_T$, usando que $\mathbb{H}_T$ es denso[4] en $\mathbb{W}_T$ (como $\mathbb{W}_T$ es un espacio métrico, la densidad de $\mathbb{H}_T$ implica que toda vecindad de un $x \in \mathbb{W}_T$ contiene a un elemento de $\mathbb{H}_T$), entonces nos limitamos a $x \in \mathbb{H}_T \cap \mathring{A}$. Como $x \in \mathring{A}$, entonces existe una bola de radio $\delta > 0$ alrededor de x, totalmente contenida en $\mathring{A}$, denotado $B_\delta(x) \subset \mathring{A}$. Entonces para estos $x \in \mathbb{H}_T \cap \mathring{A}$, gracias al lema 3.1 tenemos que dada $\epsilon > 0$ y para toda $\gamma > 0$:

$$e^{-(I(x)+\gamma)/\epsilon} \leq P\{Z_\epsilon \in B_\delta(x)\} \leq P\{Z_\epsilon \in A\},$$

si y sólo si

$$-(I(x) + \gamma) \leq \epsilon \log P\{Z_\epsilon \in B_\delta(x)\} \leq \epsilon \log P\{Z_\epsilon \in A\}.$$

De esta última expresión, que se cumple para toda $x \in \mathbb{H}_T \cap \mathring{A}$, deducimos que $\epsilon \log P\{Z_\epsilon \in A\}$ es una cota superior para $-(I(x) + \gamma)$, para toda $\gamma > 0$. De aquí deducimos que

$$-\operatorname*{ínf}_{x\in\mathring{A}} I(x) = \sup_{x\in\mathring{A}}\{-I(x)\} \leq \epsilon \log P\{Z_\epsilon \in B_\delta(x)\} \leq \epsilon \log P\{Z_\epsilon \in A\},$$

y tomando el límite inferior conforme ϵ tiende a cero, tenemos la primera desigualdad del principio de grandes desviaciones para I.

Para la segunda parte, usamos $\overline{A}$ cerrado, y $\gamma > 0$. Sea $s := \operatorname{ínf}_{x\in\overline{A}} I(x) - \gamma$. Supongamos primero que existe un $h \in \mathbb{H}_T \cap \overline{A}$, para garantizar que $\operatorname{ínf}_{x\in\overline{A}} I(x) < \infty$. Por lo que probamos en el lema anterior, el conjunto de nivel $K := \{x : I(x) \leq s\}$ es compacto. Por construcción $\overline{A} \cap K = \emptyset$, pues $K := \{x : I(x) \leq \operatorname{ínf}_{x\in\overline{A}} I(x) - \gamma < \operatorname{ínf}_{x\in\overline{A}} I(x)\}$. Es un resultado estándar de análisis que en un espacio métrico la distancia entre un cerrado y un compacto se alcanza, por lo cual está bien definido $\delta = \operatorname{ínf}_{x\in\overline{A}, y\in K} d_\infty(x, y) =: d_\infty(\overline{A}, K) > 0$. Dado que

$$Z_\epsilon \in \overline{A} \quad \text{implica} \;\; d_\infty(Z_\epsilon, K) \geq \delta,$$

entonces por esto último y el lema 3.3 tenemos que

$$\begin{aligned} P\{Z_\epsilon \in \overline{A}\} &\leq P\{d_\infty(Z_\epsilon, K) \geq \delta\} \\ &\leq e^{-(s-\gamma)/\epsilon} = \exp\left(-\frac{\operatorname{ínf}_{x\in\overline{A}} I(x) - 2\gamma}{\epsilon}\right) \end{aligned}$$

de tal suerte que

$$\epsilon \log P\{Z_\epsilon \in \overline{A}\} \leq -\operatorname*{ínf}_{x\in\overline{A}} I(x) + 2\gamma$$

tomando el límite superior conforme ϵ tiende a cero, y dado que $\gamma > 0$ fue arbitrario, se sigue el resultado.

[4] $\mathbb{H}_T$ es denso en $\mathbb{W}_T$ porque los polinomios están en $\mathbb{H}_T$, y estos son densos en $C[0, T]$.

Si no existe un h en $\mathbb{H}_T \cap \overline{A}$, entonces $I(x) \equiv \infty$ en $\overline{A}$, y entonces $\inf_{x\in\overline{A}} I(x) = \infty$. Necesitamos entonces probar que

$$\limsup_{\epsilon\to 0} \epsilon \log P\{Z_\epsilon \in \overline{A}\} = -\infty.$$

Para esto, para cada $n \in \mathbb{N}$, tenemos que

$$K_n := \{x : I(x) \leq n\}$$

es compacto. Como $I(x) \equiv \infty$ para todo $x \in \overline{A}$, entonces es claro que $K_n \cap \overline{A} = \emptyset$. Procediendo análogamente al caso donde había un $h \in \mathbb{H}_T \cap \overline{A}$, definimos $\delta := \inf_{x\in\overline{A}, y\in K_n} d_\infty(x,y)$. Si $Z_\epsilon \in \overline{A}$, entonces $d_\infty(Z_\epsilon, K_n) \geq \delta$, de tal suerte que nuevamente aplicando el lema 3.3

$$\begin{aligned} P\{Z_\epsilon \in \overline{A}\} &\leq P\{d_\infty(Z_\epsilon, K_n) \geq \delta\} \\ &\leq e^{-(n-\gamma)/\epsilon} \end{aligned}$$

de tal suerte que

$$\epsilon \log P\{Z_\epsilon \in \overline{A}\} \leq -n + \gamma$$

tomando el límite superior conforme ϵ tiende a cero, y dado que $\gamma > 0$ y $n \in \mathbb{N}$ fueron arbitrarios, se sigue el resultado.

□

3.4. El Lema Integral de Varadhan

El siguiente resultado, conocido como el lema integral de Varadhan, jugará un rol fundamental en la solución del problema de este trabajo. La manera de exponer el lema de Varadhan aquí, está basada en el libro de A. Dembo y O. Zeitouni en [17].

Teorema 3.2. (Varadhan) *Sea $\{\mu_\epsilon\}$ una familia de medidas que satisface un PGD con función de tasa buena $I : \mathcal{X} \longrightarrow [0,\infty]$. Sea $\phi : \mathcal{X} \longrightarrow \mathbb{R}$ una función continua tal que*

$$\lim_{M\to\infty} \limsup_{\epsilon\to 0} \epsilon \log E\left[\exp\left(\frac{1}{\epsilon}\phi(Z_\epsilon)\right)\mathbf{1}_{\{\phi(Z_\epsilon)\geq M\}}\right] = -\infty. \tag{3.18}$$

Entonces

$$\lim_{\epsilon\to 0} \epsilon \log E\left[\exp\left(\frac{\phi(Z_\epsilon)}{\epsilon}\right)\right] = \sup_{x\in X}\{\phi(x) - I(x)\}.$$

Finalmente, la conclusión sigue válida si se cumple la condición alternativa

$$\limsup_{\epsilon\to 0} \epsilon \log E\left[\exp\left(\frac{\gamma}{\epsilon}\phi(Z_\epsilon)\right)\right] < \infty. \tag{3.19}$$

para alguna $\gamma > 1$.

La prueba del lema de Varadhan requerirá de unos resultados previos, que se dan a continuación.

Lema 3.5. *Sea N un número entero fijo. Entonces para toda colección de números $a_i^\epsilon \geq 0$ con $1 \leq i \leq N$ y $\epsilon > 0$:*

$$\limsup_{\epsilon \to 0} \epsilon \log \left(\sum_{i=1}^{N} a_\epsilon^i \right) = \max_{1 \leq i \leq N} \left\{ \limsup_{\epsilon \to 0} \epsilon \log \left(a_\epsilon^i \right) \right\}$$

Demostración. Debido a que la función logaritmo natural es creciente, se sigue que

$$\log \left(\max_{1 \leq i \leq N} a_\epsilon^i \right) = \max_{1 \leq i \leq N} \log(a_\epsilon^i).$$

Por otra parte, dado que $a_i^\epsilon \geq 0$ implica que

$$\sum_{i=1}^{N} a_\epsilon^i \geq a_\epsilon^j$$

para toda $j \leq N$, de la primera igualdad junto con esta última desigualdad tenemos que:

$$0 \leq \epsilon \log \left(\sum_{i=1}^{N} a_\epsilon^i \right) - \epsilon \max_{1 \leq i \leq N} \log(a_\epsilon^i) = \epsilon \log \left(\frac{\sum_{i=1}^{N} a_\epsilon^i}{\max_{1 \leq j \leq N} \{a_\epsilon^j\}} \right) \leq \epsilon \log N.$$

Dado que $\epsilon \log N$ tiende a cero conforme ϵ tiende a cero, entonces dado que el límite superior de una función siempre existe, se sigue de la ecuación anterior que

$$\limsup_{\epsilon \to 0} \epsilon \log \left(\sum_{i=1}^{N} a_\epsilon^i \right) = \limsup_{\epsilon \to 0} \max_{1 \leq i \leq N} \left\{ \epsilon \log \left(a_\epsilon^i \right) \right\}.$$

Como $N < \infty$ es fijo entonces

$$\limsup_{\epsilon \to 0} \max_{1 \leq i \leq N} \left\{ \epsilon \log \left(a_\epsilon^i \right) \right\} = \max_{1 \leq i \leq N} \left\{ \limsup_{\epsilon \to 0} \epsilon \log \left(a_\epsilon^i \right) \right\}.$$

□

En los lemas siguientes, $I : \mathcal{X} \longrightarrow [0, \infty]$ es la función de tasa buena asociada a la familia de medidas $\{\mu_\epsilon\}$ mencionadas al inicio de esta sección.

Lema 3.6. *Sea $\phi : \mathcal{X} \longrightarrow \mathbb{R}$ una función semicontinua inferiormente e $I : \mathcal{X} \longrightarrow [0, \infty]$ tal que cumple la desigualdad*

$$-\inf_{x \in \mathring{A}} I(x) \leq \liminf_{\epsilon \longrightarrow 0} \epsilon \log \mu_\epsilon(A). \tag{3.20}$$

para todo $A \in \mathcal{B}$. Entonces

$$\liminf_{\epsilon \to 0} \epsilon \log E \left[\exp \left(\frac{1}{\epsilon} \, \phi(Z_\epsilon) \right) \right] \geq \sup_{x \in \mathcal{X}} \{ \phi(x) - I(x) \}.$$

Demostración. Sean $x \in \mathcal{X}$ fijo y $\delta > 0$. Dado que ϕ es semicontinua inferiormente, entonces existe una vecindad abierta V de x tal que $\operatorname{ínf}_{y\in V} \phi(y) \geq \phi(x) - \delta$ (véase apéndice para detalles sobre semicontinuidad). Usando esta vecindad V tenemos que

$$\begin{aligned} \liminf_{\epsilon\to 0} \epsilon \log E\left[\exp\left(\frac{\phi(Z_\epsilon)}{\epsilon}\right)\right] &\geq \liminf_{\epsilon\to 0} \epsilon \log E\left[\exp\left(\frac{\phi(Z_\epsilon)}{\epsilon}\right)\mathbf{1}_{\{\phi(Z_\epsilon)\in V\}}\right] \\ &\geq \liminf_{\epsilon\to 0} \epsilon \log E\left[\exp\left(\frac{1}{\epsilon}\cdot \inf_{y\in V}\phi(y)\right)\mathbf{1}_{\{\phi(Z_\epsilon)\in V\}}\right] \end{aligned} \tag{3.21}$$

Pero

$$\begin{aligned} \epsilon \log E\left[\exp\left(\frac{1}{\epsilon}\cdot \inf_{y\in V}\phi(y)\right)\mathbf{1}_{\{\phi(Z_\epsilon)\in V\}}\right] &= \epsilon \log\left(\exp\left(\frac{1}{\epsilon}\cdot \inf_{y\in V}\phi(y)\right)\cdot E\left[\mathbf{1}_{\{\phi(Z_\epsilon)\in V\}}\right]\right) \\ &= \inf_{y\in V}\phi(y) + \epsilon\log\mu_\epsilon(V). \end{aligned}$$

Sustituyendo esto último en (3.21) tenemos que

$$\liminf_{\epsilon\to 0} \epsilon \log E\left[\exp\left(\frac{\phi(Z_\epsilon)}{\epsilon}\right)\right] \geq \inf_{y\in V}\phi(y) + \liminf_{\epsilon\to 0}\epsilon\log\mu_\epsilon(V).$$

Dado que se cumple la desigualdad (3.20) y la vecindad abierta V es tal que $\operatorname{ínf}_{y\in V} \phi(y) \geq \phi(x) - \delta$, se sigue que

$$\inf_{y\in V}\phi(y) + \liminf_{\epsilon\to 0}\epsilon\log\mu_\epsilon(V) \geq \inf_{y\in V}\phi(y) - \underbrace{\inf_{y\in V} I(y)}_{\geq -I(x)} \geq \phi(x) - I(x) - \delta.$$

Dado que $x \in \mathcal{X}$ y $\delta > 0$ fueron arbitrarios, se sigue el resultado.

□

Ahora se prueba un resultado análogo al anterior.

Lema 3.7. *Sea $\phi : \mathcal{X} \longrightarrow \mathbb{R}$ una función semicontinua superiormente que cumple la condición de integrabilidad* (3.18) *y la desigualdad*

$$\limsup_{\epsilon\to 0}\epsilon\log\mu_\epsilon(A) \leq -\inf_{x\in\overline{A}} I(x), \tag{3.22}$$

para todo $A \in \mathcal{B}$. Entonces

$$\limsup_{\epsilon\to 0} \epsilon \log E\left[\exp\left(\frac{1}{\epsilon}\,\phi(Z_\epsilon)\right)\right] \leq \sup_{x\in\mathcal{X}}\{\phi(x) - I(x)\}.$$

Demostración. Primero supongamos que ϕ es acotada por $K > 0$ constante. ϕ cumple la condición de integrabilidad (3.18), pues para $M > K$, se cumple que $\mathbf{1}_{\{\phi(Z_\epsilon)\geq M\}} = 0$. Sea $a < \infty$ fijo y $\delta > 0$. Sea el conjunto de nivel

$$\Phi_I(a) := \left\{x \in \mathcal{X} : I(x) \leq a\right\}.$$

Dado que I es una función de tasa buena, este conjunto es compacto para cada $a \geq 0$. Para cada $x \in \Phi_I(a)$, dado que I es semicontinua inferiormente y ϕ es semicontinua superiormente, existe una vecindad V_x de x tal que

$$\inf_{y \in V_x} I(y) \geq I(x) - \delta \quad \text{y} \quad \sup_{y \in V_x} \phi(y) \leq \phi(x) + \delta.$$

Dado que estamos en un espacio métrico $\mathcal{X}$, existe una bola abierta A_x centrada en x tal que $\overline{A}_x \subset V_x$ (en un espacio métrico, como V_x es abierto, existe una bola de radio $\epsilon > 0$ alrededor de x, achicando el radio a la mitad tenemos un conjunto abierto cuya cerradura está totalmente contenida en V_x.)

De esta forma, para cada $x \in \Phi_I(a)$ existe una vecindad A_x de x tal que su cerradura $\overline{A}_x$ cumple

$$\inf_{y \in \overline{A}_x} I(y) \geq I(x) - \delta \quad , \sup_{y \in \overline{A}_x} \phi(y) \leq \phi(x) + \delta. \tag{3.23}$$

De esta manera obtenemos una cubierta abierta $\bigcup_{x \in \Phi_I(a)} A_x$ de $\Phi_I(a)$. Dado que los conjuntos de nivel de I son compactos, podemos extraer una cubierta finita $\bigcup_{i=1}^N A_{x_i} =: A$. Usando esta cubierta, el hecho de que $\phi(y) \leq K$ para todo $y \in \mathcal{X}$ y la segunda desigualdad de (3.23), tenemos que

$$\begin{aligned}
E\left[\exp\left(\frac{1}{\epsilon} \cdot \phi(Z_\epsilon)\right)\right] &= E\left[\mathbf{1}_{A \cup A^c} \cdot \exp\left(\frac{1}{\epsilon} \cdot \phi(Z_\epsilon)\right)\right] \\
&\leq e^{K/\epsilon} E\left[\mathbf{1}_{(\bigcup_{i=1}^N A_{x_i})^c}\right] + \sum_{i=1}^N E\left[\mathbf{1}_{\{Z_\epsilon \in A_{x_i}\}} \cdot \exp\left(\frac{1}{\epsilon} \cdot \phi(Z_\epsilon)\right)\right] \\
&\leq e^{K/\epsilon} \mu_\epsilon\left(\left(\bigcup_{i=1}^N A_{x_i}\right)^c\right) + \sum_{i=1}^N \exp\left(\frac{1}{\epsilon}\big(\phi(x_i) + \delta\big)\right) \mu_\epsilon(\overline{A}_{x_i})
\end{aligned}$$

Se sigue que

$$\limsup_{\epsilon \downarrow 0} \epsilon \log E\left[e^{\phi(Z_\epsilon)/\epsilon}\right] \leq \limsup_{\epsilon \downarrow 0} \epsilon \log \left\{ e^{K/\epsilon} \mu_\epsilon\left(\left(\bigcup_{i=1}^N A_{x_i}\right)^c\right) + \sum_{i=1}^N e^{(\phi(x_i)+\delta)/\epsilon} \mu_\epsilon(\overline{A}_{x_i}) \right\}$$

Pero por el lema 3.5 tenemos que (sustituyendo $A := \bigcup_{i=1}^N A_{x_i}$)

$$\begin{aligned}
&\limsup_{\epsilon \downarrow 0} \epsilon \log \left\{ e^{K/\epsilon} \mu_\epsilon\left(\left(\bigcup_{i=1}^N A_{x_i}\right)^c\right) + \sum_{i=1}^N e^{(\phi(x_i)+\delta)/\epsilon} \mu_\epsilon(\overline{A}_{x_i}) \right\} \\
&= \max\left\{ \max_{1 \leq i \leq N} \left\{ \limsup_{\epsilon \downarrow 0} \epsilon \log \left(e^{(\phi(x_i)+\delta)/\epsilon} \mu_\epsilon(\overline{A}_{x_i})\right)\right\}, \limsup_{\epsilon \downarrow 0} \epsilon \log \left(e^{K/\epsilon} \cdot \mu_\epsilon(A^c)\right)\right\} \\
&= \max\left\{ \max_{1 \leq i \leq N} \left\{ \phi(x_i) + \delta + \limsup_{\epsilon \downarrow 0} \epsilon \log \mu_\epsilon(\overline{A}_{x_i})\right)\right\}, K + \limsup_{\epsilon \downarrow 0} \epsilon \log \mu_\epsilon(A^c) \right\}
\end{aligned}$$

Recordemos la hipótesis $\limsup_{\epsilon\to 0} \epsilon \log \mu_\epsilon(A) \leq -\inf_{x\in\overline{A}} I(x)$ para todo $A \in \mathcal{B}$. Nótese que $A^c \subset \Phi_I(a)^c$ es un conjunto cerrado tal que por definición de $\Phi_I(a)$, $I(y) > a$ para todo $y \in A^c$, por lo cual se cumplen:

$$\limsup_{\epsilon\to 0} \epsilon \log \mu_\epsilon(A^c) \leq -\inf_{y\in A^c} I(y) \quad \text{y} \quad -\inf_{y\in A^c} I(y) < -a.$$

Usando estos hechos, tenemos que

$$\begin{aligned}
&\limsup_{\epsilon\downarrow 0} \epsilon \log E\left[e^{\phi(Z_\epsilon)/\epsilon}\right] \\
&\qquad \leq \max\left\{ \max_{1\leq i\leq N} \left\{\phi(x_i) + \delta + \limsup_{\epsilon\downarrow 0} \epsilon \log \mu_\epsilon(\overline{A}_{x_i})\right\},\, K + \limsup_{\epsilon\downarrow 0} \epsilon \log \mu_\epsilon(A^c) \right\} \\
&\qquad \leq \max\left\{ \max_{1\leq i\leq N} \left\{\phi(x_i) + \delta - \inf_{y\in\overline{A}_{x_i}} I(y)\right\},\, K - \inf_{y\in A^c} I(y) \right\} \\
&\qquad \leq \max\left\{ \max_{1\leq i\leq N} \left\{\phi(x_i) - I(x_i) + 2\delta\right\},\, K - a\right\} \\
&\qquad \leq \max\left\{ \sup_{x\in\mathcal{X}} \left\{\phi(x) - I(x)\right\},\, K - a\right\} + 2\delta
\end{aligned}$$

donde en la penúltima igualdad hemos utilizado la primera desigualdad de (3.23) para decir que $-\inf_{y\in\overline{A}_{x_i}} I(y) \leq \delta - I(x_i)$ para todo $1 \leq i \leq N$. Como $a \geq 0$ y $\delta > 0$ fueron arbitrarios, dejando $a \to \infty$, $\delta \to 0$ se sigue el resultado para ϕ acotada.

Si ϕ no está necesariamente acotada, sea $M > 0$, entonces aplicando el lema 3.5

$$\begin{aligned}
&\limsup_{\epsilon\downarrow 0} \epsilon \log E\left[e^{\phi(Z_\epsilon)/\epsilon}\right] \\
&\qquad = \limsup_{\epsilon\downarrow 0} \epsilon \log \left(E\left[e^{\phi(Z_\epsilon)/\epsilon}\mathbf{1}_{\{\phi(Z_\epsilon)<M\}}\right] + E\left[e^{\phi(Z_\epsilon)/\epsilon}\mathbf{1}_{\{\phi(Z_\epsilon)\geq M\}}\right]\right) \\
&\qquad = \max\left\{ \limsup_{\epsilon\downarrow 0} \epsilon \log E\left[e^{\phi(Z_\epsilon)/\epsilon}\mathbf{1}_{\{\phi(Z_\epsilon)<M\}}\right],\, \limsup_{\epsilon\downarrow 0} \epsilon \log E\left[e^{\phi(Z_\epsilon)/\epsilon}\mathbf{1}_{\{\phi(Z_\epsilon)\geq M\}}\right]\right\} \\
&\qquad \leq \max\left\{ \sup_{x\in X} \left\{\phi(x) - I(x)\right\},\, \limsup_{\epsilon\downarrow 0} \epsilon \log E\left[e^{\phi(Z_\epsilon)/\epsilon}\mathbf{1}_{\{\phi(Z_\epsilon)\geq M\}}\right]\right\}
\end{aligned}$$

Donde en la desigualdad hemos utilizado el resultado para ϕ acotada obtenido anteriormente en esta prueba. Usando la hipótesis de integrabilidad (3.18) tenemos que si $M \to \infty$, se sigue el resultado. □

Ahora se prueba que la condición (3.19) implica la condición (3.18).

Lema 3.8. *Si para alguna $\gamma > 1$ se cumple que*

$$\limsup_{\epsilon\to 0} \epsilon \log E\left[\exp\left(\frac{\gamma}{\epsilon}\phi(Z_\epsilon)\right)\right] < \infty,$$

entonces

$$\lim_{M\to\infty} \limsup_{\epsilon\to 0} \epsilon \log E\left[\exp\left(\frac{1}{\epsilon}\phi(Z_\epsilon)\right) \mathbf{1}_{\{\phi(Z_\epsilon)\geq M\}} \right] = -\infty.$$

Demostración. Sea $X_\epsilon := e^{(\phi(Z_\epsilon)-M)/\epsilon}$. Nótese que $\phi(Z_\epsilon) \geq M$ si y sólo si

$$\frac{\phi(Z_\epsilon) - M}{\epsilon} \geq 0 \quad \text{si y sólo si} \quad X_\epsilon \geq 1.$$

Como $\gamma > 1$, entonces

$$\begin{aligned} e^{-M/\epsilon} E\big[e^{\phi(Z_\epsilon)/\epsilon} \mathbf{1}_{\{\phi(Z_\epsilon)\geq M\}}\big] = E\big[X_\epsilon \mathbf{1}_{\{X_\epsilon \geq 1\}}\big] &\leq E\big[(X_\epsilon)^\gamma\big] \\ &= e^{-\gamma M/\epsilon} E\big[e^{\gamma\phi(Z_\epsilon)/\epsilon}\big], \end{aligned}$$

es decir:

$$E\big[e^{\phi(Z_\epsilon)/\epsilon} \mathbf{1}_{\{\phi(Z_\epsilon)\geq M\}}\big] \leq e^{M(1-\lambda)/\epsilon} E\big[e^{\gamma\phi(Z_\epsilon)/\epsilon}\big].$$

De esto último se sigue que

$$\begin{aligned} \limsup_{\epsilon\downarrow 0} \epsilon \log E\big[e^{\phi(Z_\epsilon)/\epsilon} \mathbf{1}_{\{\phi(Z_\epsilon)\geq M\}}\big] &\leq \limsup_{\epsilon\downarrow 0} \epsilon \log \left(e^{M(1-\gamma)/\epsilon} E\big[e^{\gamma\phi(Z_\epsilon)/\epsilon}\big] \right) \\ &= \limsup_{\epsilon\downarrow 0} \epsilon \log \left(e^{M(1-\gamma)/\epsilon} \right) + \limsup_{\epsilon\downarrow 0} \epsilon \log E\big[e^{\gamma\phi(Z_\epsilon)/\epsilon}\big] \\ &= M(1-\gamma) + \limsup_{\epsilon\downarrow 0} \epsilon \log E\big[e^{\gamma\phi(Z_\epsilon)/\epsilon}\big] \end{aligned}$$

Por hipótesis el último sumando de la desigualdad anterior es finito. Tomando $M \to \infty$ y usando que $\gamma > 1$, tenemos que el lado derecho de la desigualdad anterior diverge a $-\infty$, de donde se sigue el resultado.

□

Con estos resultados a la mano, podemos probar el lema de Varadhan:

Demostración. **(del lema de Varahan 3.2)** Debido a que ϕ es continua, en particular es semicontinua inferiormente y a su vez semicontinua superiormente. Como la familia $\{\mu_\epsilon\}$ cumple el PGD con función de tasa buena I, en particular cumple la conclusión del lema 3.6, es decir

$$\liminf_{\epsilon\to 0} \epsilon \log E\left[\exp\left(\frac{1}{\epsilon}\, \phi(Z_\epsilon)\right) \right] \geq \sup_{x\in\mathcal{X}} \{\phi(x) - I(x)\}.$$

A su vez, cumple la conlusión del lema 3.7:

$$\limsup_{\epsilon\to 0} \epsilon \log E\left[\exp\left(\frac{1}{\epsilon}\, \phi(Z_\epsilon)\right) \right] \leq \sup_{x\in\mathcal{X}} \{\phi(x) - I(x)\}.$$

De aquí se sigue inmediatamente que

$$\liminf_{\epsilon\to 0} \epsilon \log E\left[\exp\left(\frac{1}{\epsilon}\,\phi(Z_\epsilon)\right)\right] = \limsup_{\epsilon\to 0} \epsilon \log E\left[\exp\left(\frac{1}{\epsilon}\,\phi(Z_\epsilon)\right)\right],$$

es decir, $\lim_{\epsilon\to 0} \epsilon \log E\left[\exp\left(\frac{1}{\epsilon}\,\phi(Z_\epsilon)\right)\right]$ existe y es igual a

$$\sup_{x\in\mathcal{X}}\{\phi(x) - I(x)\}.$$

Esto concluye la primera parte del lema de Varadhan, la segunda parte se sigue inmediatamente de una aplicación del lema 3.8.

□

Necesitaremos extender el lema de Varadhan a un caso un poco más general, pues como vemos, $\phi : \mathcal{X} \longrightarrow \mathbb{R}$, y a nosotros nos va a interesar cuando ϕ toma el valor adicional extendido $-\infty$, es decir, cuando tiene imagen en $\mathbb{R}\cup\{-\infty\}$. Para esto, necesitaremos del siguiente resultado previo, que es parte del ejercicio 4.3.11 del texto de Dembo y Zeitouni [17].

Lema 3.9. *Supongamos que la familia de medidas $\{\mu_\epsilon\}$ satisface un principio de grandes desviaciones con función de tasa buena $I : \mathcal{X} \longrightarrow [0,\infty]$. Sea $\phi : \mathcal{X} \longrightarrow \mathbb{R}$ una función continua que cumple la condición*

$$\limsup_{\epsilon\to 0} \epsilon \log E\left[\exp\left(\frac{\gamma}{\epsilon}\phi(Z_\epsilon)\right)\right] < \infty.$$

para alguna $\gamma > 1$ (por lo cual cumple la condición (3.18)). Entonces para todo conjunto cerrado F de $\mathcal{X}$ se cumple que

$$\limsup_{\epsilon\to 0} \epsilon \log\left(\int_F e^{\phi(x)/\epsilon} d\mu_\epsilon\right) \leq \sup_{x\in F}\{\phi(x) - I(x)\}.$$

Demostración. Definiendo a

$$\tilde{\phi}(x) = \begin{cases} \phi(x) & \text{si } x \in F \\ -\infty & \text{si } x \notin F, \end{cases} \tag{3.24}$$

se probará que $\tilde{\phi}$ cumple las hipótesis del lema 3.7.

Empezaremos probando que es semicontinua superiormente. Si $x \notin F$, entonces $\tilde{\phi}(x) = -\infty$. Dado que F es un cerrado, entonces $X \setminus F$ es abierto, por lo cual existe una vecindad V_x de x totalmente contenida en $X \setminus F$ tal que $\tilde{\phi}(y) = -\infty$. Por otra parte, si $x \in F$, entonces como en F tenemos que $\tilde{\phi}(x) = \phi(x)$, y ϕ es continua, entonces existen dos opciones:

(i) Existe una vecindad V_x de x *totalmente contenida en* F, en cuyo caso por continuidad de ϕ, se sigue que $\tilde{\phi}$ es continua en esta vecindad, en particular es semicontinua superiormente.

(ii) Si no existe una vecindad de x totalmente contenida en F, entonces x es un punto frontera de F, por ser $\mathcal{X}$ un espacio métrico. Como ϕ no puede tomar el valor $-\infty$ en F, $\tilde{\phi}(x)$ tampoco, por lo cual para una vecindad suficientemente pequeña V_x de x, se cumple que

$$\tilde{\phi}(y) \leq \tilde{\phi}(x) + \delta,$$

pues si $y \notin F$ entonces $\tilde{\phi}(y) = -\infty \leq \tilde{\phi}(x) + \delta$. Si $(V_x \cap F) \setminus \{x\} \neq \emptyset$, entonces para los $y \in V_x \cap F$ se cumple la misma desigualdad, por ser ϕ continua en $\mathcal{X}$.

De este modo $\tilde{\phi}$ es semicontinua superiormente. Nótese ahora que por definición de $\tilde{\phi}$, si $x \in \mathcal{X} \setminus F$ entonces

$$\exp\left(\frac{\gamma}{\epsilon}\,\tilde{\phi}(x)\right) = 0,$$

por lo cual sólo $x \in F$ constribuyen a la integral de esta función, de donde se sigue que

$$\begin{aligned}
\limsup_{\epsilon\to 0} \epsilon \log E\left[\exp\left(\frac{\gamma}{\epsilon}\tilde{\phi}(Z_\epsilon)\right)\right] &= \limsup_{\epsilon\to 0} \epsilon \log\left(\int_F e^{\gamma\phi(x)/\epsilon} d\mu_\epsilon\right) \\
&= \limsup_{\epsilon\to 0} \epsilon \log E\left[\exp\left(\frac{\gamma}{\epsilon}\phi(Z_\epsilon)\right)\mathbf{1}_F\right],
\end{aligned}$$

por lo cual

$$\limsup_{\epsilon\to 0} \epsilon \log E\left[\exp\left(\frac{\gamma}{\epsilon}\tilde{\phi}(Z_\epsilon)\right)\right] \leq \limsup_{\epsilon\to 0} \epsilon \log E\left[\exp\left(\frac{\gamma}{\epsilon}\phi(Z_\epsilon)\right)\right] < \infty.$$

Entonces $\tilde{\phi}$ cumple las condiciones del lema 3.7, por lo cual

$$\limsup_{\epsilon\to 0} \epsilon \log E\left[\exp\left(\frac{1}{\epsilon}\tilde{\phi}(Z_\epsilon)\right)\right] \leq \sup_{x\in\mathcal{X}}\{\phi(x) - I(x)\}.$$

Pero en el lado derecho de la ecuación anterior, podemos descartar $x \in \mathcal{X} \setminus F$, pues en estos $\tilde{\phi} = -\infty$, por lo cual podemos limitarnos a $x \in F$. De este modo obtenemos

$$\limsup_{\epsilon\to 0} \epsilon \log\left(\int_F e^{\phi(x)/\epsilon} d\mu_\epsilon\right) \leq \sup_{x\in F}\{\phi(x) - I(x)\}$$

□

Ahora probamos la extensión del lema de Varadhan. Como se dijo al principio de esta sección, $(\mathcal{X}, \mathcal{B})$ un espacio métrico con $\{\mu_\epsilon\}$ una familia de medidas inducidas en $(\mathcal{X}, \mathcal{B})$ por el conjunto de variables aleatorias $\{Z_\epsilon\}$ con valores en $\mathcal{X}$. Esta extensión está basada en el artículo de P. Guasoni y S. Robertson [20].

Teorema 3.3. (Extensión del Lema de Varadhan) *Sea $\{Z_\epsilon\}$ una familia de variables aleatorias con valores en $\mathcal{X}$ cuyas leyes de probabilidad asociadas $\{\mu_\epsilon\}$ satisfacen un PGD con función de tasa buena $I : \mathcal{X} \longrightarrow [0, \infty]$. Sea $\phi : \mathcal{X} \longrightarrow \mathbb{R} \cup \{-\infty\}$ una función continua tal que para algún $\gamma > 1$ se cumple que*

$$\limsup_{\epsilon \downarrow 0} \epsilon \log E\left[\exp\left(\frac{\gamma}{\epsilon} \phi(Z_\epsilon) \right) \right] < \infty.$$

Entonces para cualquier $A \in \mathcal{B}$ se cumple que:

$$\begin{aligned} \sup_{x \in \mathring{A}} \{\phi(x) - I(x)\} &\leq \liminf_{\epsilon \downarrow 0} \epsilon \log \left(\int_{\mathring{A}} e^{\phi(x)/\epsilon} d\mu_\epsilon \right) \\ &\leq \limsup_{\epsilon \downarrow 0} \epsilon \log \left(\int_{\overline{A}} e^{\phi(x)/\epsilon} d\mu_\epsilon \right) \leq \sup_{x \in \overline{A}} \{\phi(x) - I(x)\} \end{aligned} \tag{3.25}$$

Demostración. Para la primera desigualdad, nótese que aunque ϕ ahora puede tomar el valor $-\infty$, esto no afecta la demostración del lema 3.6, pues usando que ϕ es semicontinua inferiormente, entonces si $\phi(x) \in \mathbb{R} \cup \{-\infty\}$, sigue existiendo una vecindad abierta V de x tal que $\inf_{y \in V} \phi(y) \geq \phi(x) - \delta$, y se siguen los mismos pasos de la demostración del lema 3.6, sólo que tomando $x \in \mathring{A}$ en vez de $\mathcal{X}$. De aquí que podemos obtener una conclusión como la del lema 3.6, es decir:

$$\sup_{x \in \mathring{A}} \{\phi(x) - I(x)\} \leq \liminf_{\epsilon \downarrow 0} \epsilon \log \left(\int_{\mathring{A}} e^{\phi(x)/\epsilon} d\mu_\epsilon \right)$$

Es de notar que para la última desigualdad de (3.25), no podemos seguir el mismo procedimiento que en la prueba del lema 3.7, pues en dicho lema, la hipótesis que $\phi(x) \neq -\infty$ nos pudo garantizar que se cumplía la segunda desigualdad de (3.23), es decir, que para todo x existe un abierto A_x tal que $\sup_{y \in \overline{A}_x} \phi(y) \leq \phi(x) + \delta$. Esta desigualdad fue importante para la prueba del lema 3.7, pero ahora que $\phi(x)$ podría ser $-\infty$, no podemos garantizar la existencia de dicha vecindad.

Entonces proseguiremos de otra manera. Es claro que si $\phi = -\infty$ en todo $\mathcal{X}$, entonces el resultado es cierto. Supongamos que ϕ no es $-\infty$ en todo $\mathcal{X}$. Sea $M > 0$ y definimos

$$C_M := \overline{A} \cap \{x : \phi(x) \geq -M\}.$$

Como $\{x : \phi(x) \geq -M\} = \phi^{-1}\big([-M, \infty)\big)$, por continuidad de ϕ tenemos que este conjunto es cerrado, por lo cual C_M es cerrado. Usaremos a C_M para controlar lo suficiente a ϕ como para poder aplicar el lema 3.9, pues $\phi|_{C_m} > -\infty$. Para esto nótese que

$$\begin{aligned} \int_{\overline{A}} e^{\phi(x)/\epsilon} d\mu_\epsilon &= \int_{\overline{A} \cap C_M} e^{\phi(x)/\epsilon} d\mu_\epsilon + \int_{\overline{A} \setminus C_M} e^{\phi(x)/\epsilon} d\mu_\epsilon \\ &\leq \int_{\overline{A} \cap C_M} e^{\phi(x)/\epsilon} d\mu_\epsilon + e^{-M/\epsilon} \mu_\epsilon(\overline{A} \setminus C_M). \end{aligned} \tag{3.26}$$

Usando que las medidas $\{\mu_\epsilon\}$ satisfacen un principio de grandes desviaciones con función de tasa buena I, tenemos que

$$\begin{aligned}\limsup_{\epsilon\downarrow 0}\epsilon\log\left(e^{-M/\epsilon}\mu_\epsilon(\overline{A}\setminus C_M)\right) &= -M+\limsup_{\epsilon\downarrow 0}\epsilon\log\mu_\epsilon(\overline{A}\setminus C_M)\\ &\leq -M-\inf_{x\in\overline{A\setminus C_M}}I(x).\end{aligned}\tag{3.27}$$

Por otra parte, usando que C_M es cerrado y que en este conjunto, $\phi|_{C_m}>-\infty$, podemos aplicar el lema 3.9 para obtener que

$$\limsup_{\epsilon\to 0}\epsilon\log\left(\int_{C_M}e^{\phi(x)/\epsilon}d\mu_\epsilon\right)\leq\sup_{x\in C_M}\{\phi(x)-I(x)\}.\tag{3.28}$$

Usando el lema 3.5 y después aplicando (3.27) y (3.28) en (3.26) obtenemos que

$$\begin{aligned}&\limsup_{\epsilon\downarrow 0}\epsilon\log\left(\int_{\overline{A}\cap C_M}e^{\phi(x)/\epsilon}d\mu_\epsilon+e^{-M/\epsilon}\mu_\epsilon(\overline{A}\setminus C_M)\right)\\ =&\max\left\{\limsup_{\epsilon\downarrow 0}\epsilon\log\left(\int_{\overline{A}\cap C_M}e^{\phi(x)/\epsilon}d\mu_\epsilon\right),\limsup_{\epsilon\downarrow 0}\epsilon\log\left(e^{-M/\epsilon}\mu_\epsilon(\overline{A}\setminus C_M)\right)\right\}\\ \leq&\max\left\{\sup_{x\in C_M}\{\phi(x)-I(x)\},-M-\inf_{x\in\overline{A\setminus C_M}}I(x)\right\}\\ \leq&\max\left\{\sup_{x\in\overline{A}}\{\phi(x)-I(x)\},-M\right\}.\end{aligned}$$

Dejando $M\to\infty$ en la última ecuación, se sigue que

$$\limsup_{\epsilon\downarrow 0}\epsilon\log\left(\int_{\overline{A}}e^{\phi(x)/\epsilon}d\mu_\epsilon\right)\leq\sup_{x\in\overline{A}}\{\phi(x)-I(x)\}.$$

Esto prueba la tercera desigualdad de (3.25). La segunda desigualdad se sigue de la definición de límite superior y límite inferior. □

Capítulo 4

Aplicación a un problema de optimización

4.1. Replanteamiento del Problema 1.1

En este capítulo se da una aplicación de la teoría estudiada en los capítulos anteriores 2 y 3 al problema propuesto en el capítulo 1 de este trabajo, en particular al Problema 1.1.

Como se dijo en el capítulo 1, en el artículo de P. Guasoni y S. Robertson [20], se busca minimizar a la integral

$$E_P\left[G^2(B)\frac{dP}{dQ^h}(B)\right]. \tag{4.1}$$

Vamos a estudiar el problema de minimizar una aproximación de (4.1) desde el punto de vista de las herramientas desarrolladas en los capítulos 2 y 3. La manera de proseguir será basada en el artículo de P. Guasoni y S. Robertson en [20].

Debido a que el problema a resolver se presenta en finanzas, y por razones de esta área, se requiere que el cambio de medida dado por la derivada de Radon-Nikodym $\frac{dP}{dQ}$ nos mantenga trayectorias que sean martingalas[1] bajo Q. Esto es logrado por el teorema de Girsanov 2.3, por lo cual escoger un cambio de medida $\frac{dP}{dQ}$ es análogo a escoger un término de deriva adecuado h, por lo cual denotamos la dependencia de Q de h al escribir Q^h. Comentamos en el capítulo 1 que se usarían términos de deriva deterministas. Usaremos un principio de grandes desviaciones para el movimiento browniano, por lo cual, como vimos en el lema de Schilder del capítulo anterior, usaremos el espacio clásico de Cameron-Martin $\mathbb{H}_T$. De acuerdo a P. Guasoni y S. Robertson en p. 3,4,7 y 8 de [20], la velocidad de simulación de estimadores con las trayectorias en este espacio es buena, y en

[1] Por razones del concepto conocido como *arbitraje*. Estas condiciones y conceptos de finanzas no se estudian aquí, pero para más detalles ver el texto de P. Glasserman [10] sección 1.2.2 y las referencias ahí contenidas.

la práctica este espacio da soluciones eficientes. En las secciones 4 y 5 de [20], los autores trabajan con funciones particulares de interés en finanzas, y en particular en la sección 5 dan ejemplos numéricos de la eficiencia de estas soluciones.

Pongamos a la expresión (4.1) en términos de una derivada de Radon-Nikodym dada por el teorema 2.3 del capítulo 2, y usando $h \in \mathbb{H}_T$. Esto lo hacemos del modo siguiente: para $h \in \mathbb{H}_T$ formamos su exponencial estocástica

$$\mathcal{E}\Big(\int \dot{h}_s dB_s\Big)_t = \exp\left(\int_0^t \dot{h}_s dB_s - \frac{1}{2}\int_0^t \dot{h}_s^2 ds\right).$$

Como sabemos, por el teorema de Girsanov 2.3 esta exponencial estocástica es una P-martingala y es tal que bajo el cambio de medida $\frac{dQ^h}{dP} = \mathcal{E}\big(\int \dot{h}_s dB_s\big)_t$, el proceso

$$X_t := B_t - h_t$$

es un movimiento browniano estándar, donde se ha utilizado que $h_t - h_0 = h_t = \int_0^t \dot{h}_s ds$. Tenemos que

$$\frac{dP}{dQ^h} = \frac{1}{\frac{dQ^h}{dP}} = \frac{1}{\mathcal{E}\left(\int \dot{h}_s dB_s\right)_T} = \exp\left(-\int_0^T \dot{h}_t dB_t + \frac{1}{2}\int_0^T \dot{h}_t^2 dt\right)$$

donde la primera igualdad de la expresión anterior se sigue del resultado en Halmos [5] p. 136. Así que la expresión (4.1) toma la forma

$$E_P\Big[G^2(B)\frac{dP}{dQ^h}\Big] = E_P\left[G^2(B)\exp\left(-\int_0^T \dot{h}_t dB_t + \frac{1}{2}\int_0^T \dot{h}_t^2 dt\right)\right].$$

Usando que $G^2 = \exp\big(\log(G^2)\big) = \exp\big(2\log(G)\big)$, definiendo $F := \log(G)$, tenemos en la ecuación anterior que

$$E_P\left[G^2(B)\frac{dP}{dQ^h}\right] = E_P\left[\exp\left(2F(B) - \int_0^T \dot{h}_t dB_t + \frac{1}{2}\int_0^T \dot{h}_t^2 dt\right)\right]. \tag{4.2}$$

Sin embargo, como dijimos en el capítulo 1, los autores P. Guasoni y S. Robertson indican en la p. 6 de [20] que la integral (4.1) es difícil de calcular con exactitud, por lo cual trabajan con una *aproximación* perturbada de $E_P\left[G^2(B)\frac{dP}{dQ^h}\right]$, y en la p. 6 de [20] indican que con esta aproximación, les interesa minimizar

$$\epsilon \log E_P\left[\frac{1}{\epsilon}G^2(\sqrt{\epsilon}B)\frac{dP}{dQ^h}(\sqrt{\epsilon}B)\right] \tag{4.3}$$

al dejar tender la perturbación ϵ a cero. Es decir, se busca minimizar con respecto a $h \in \mathbb{H}_T$ a

$$\limsup_{\epsilon\downarrow 0} \epsilon \log E_P\left[\frac{1}{\epsilon}G^2(\sqrt{\epsilon}B)\frac{dP}{dQ^h}(\sqrt{\epsilon}B)\right].$$

Así que en este trabajo nos limitaremos a un método para hallar

$$\min_{h\in\mathbb{H}_T}\left\{\limsup_{\epsilon\downarrow 0}\epsilon\,\log E_P\left[\frac{1}{\epsilon}G^2(\sqrt{\epsilon}B)\frac{dP}{dQ^h}(\sqrt{\epsilon}B)\right]\right\}. \tag{4.4}$$

Multiplicando (4.2) por $1/\epsilon$ y sustituyendo B por $\sqrt{\epsilon}B$, entonces al tomar en cuenta a la aproximación de (4.1), obtenemos

$$\begin{aligned}\limsup_{\epsilon\downarrow 0}\epsilon\,\log E_P&\left[\frac{1}{\epsilon}G^2(\sqrt{\epsilon}B)\frac{dP}{dQ^h}(\sqrt{\epsilon}B)\right]\\ &=\limsup_{\epsilon\to 0}\epsilon\,\log E_P\left[\exp\left(\frac{1}{\epsilon}\Big(2F(\sqrt{\epsilon}B)-\int_0^T\sqrt{\epsilon}\dot{h}_t dB_t+\frac{1}{2}\int_0^T\dot{h}_t^2dt\Big)\right)\right].\end{aligned}$$

Entonces el objetivo es minimizar a esta última expresión, y de este modo tenemos el siguiente problema, que es el replanteamiento del Problema 1.1 del capítulo 1.

Problema 4.1. *Con L definido para $h\in\mathbb{H}_T$ como*

$$L(h):=\limsup_{\epsilon\to 0}\epsilon\,\log E_P\left[\exp\left(\frac{1}{\epsilon}\Big(2F(\sqrt{\epsilon}B)-\int_0^T\sqrt{\epsilon}\dot{h}_t dB_t+\frac{1}{2}\int_0^T\dot{h}_t^2dt\Big)\right)\right], \tag{4.5}$$

proponer un método para hallar

$$\min_{h\in\mathbb{H}_T}L(h). \tag{4.6}$$

Una solución al problema (4.6) *será denominada una* ***solución asintóticamente óptima***.

La siguiente suposición jugará un papel esencial en el principal resultado a probar en este capítulo. En nuestro contexto, $F:=\log(G)$.

Suposición 4.1. $F:\mathbb{W}_T\longrightarrow\mathbb{R}\cup\{-\infty\}$ *es continua y satisface la desigualdad*

$$F(x)\leq K_1+K_2\left(\sup_{t\in[0,T]}|x_t|\right)^{\alpha} \tag{4.7}$$

para algunas $K_1,K_2>0$ *y para alguna* $0<\alpha<2$.

Según P. Guasoni y S. Robertson en [20] p. 6, para prácticamente todas las funciones de pago G de interés en las aplicaciones, se cumple la suposición anterior al hacer $F:=\log(G)$.

Vamos a obtener dos lemas auxiliares cuya única finalidad es probar que la expresión (4.5) a minimizar, bajo hipótesis suficientes podemos relacionarlo con un problema variacional por medio de la extensión del lema integral de Varadhan dada en el teorema 3.3 y el lema de Schilder que se dio en el teorema 3.1 del capítulo anterior. Después obtenemos el lema 4.3, cuya finalidad es probar que los problemas variacionales que hallemos tienen una solución. El producto final de lo anterior será el teorema 4.3, que nos da implícitamente un algoritmo para resolver el problema de minimizar (4.5), que proponemos al final de este capítulo.

4.2. Resultados Auxiliares

Básicamente se compilan resultados de los capítulos anteriores para probar lemas que nos llevan directamente a la solución del problema. Se usarán unos principios básicos de cálculo de variaciones. El siguiente es únicamente un resultado auxiliar, tomado de P. Guasoni y S. Robertson [20].

Lema 4.1. *Sean $A, B > 0$ y $\alpha \in (0,2)$. Sea $b_0 := \left(\frac{\alpha A}{2B}\right)^{1/(2-\alpha)}$. Entonces $f : [0,\infty) \longrightarrow \mathbb{R}$ definida por $f(b) := Ab^\alpha - Bb^2$ satisface*

$$\int_0^\infty \exp\big(f(b)\big)db \leq \exp(Ab_0^\alpha - Bb_0^2)\left(b_0 + \sqrt{\frac{2\pi}{\text{mín}\{2B, 2B(2-\alpha)\}}}\right) \tag{4.8}$$

Demostración. Nótese que

$$\begin{aligned} f'(b) &= \alpha A b^{\alpha-1} - 2Bb \\ f''(b) &= \alpha(\alpha-1)Ab^{\alpha-2} - 2B \\ f'''(b) &= \alpha(\alpha-1)(\alpha-2)Ab^{\alpha-3} \end{aligned} \tag{4.9}$$

Entonces

$$\begin{aligned} f'(b_0) &= \alpha A\Big(\frac{\alpha A}{2B}\Big)^{\frac{1}{2-\alpha}\cdot(\alpha-1)} - 2B\Big(\frac{\alpha A}{2B}\Big)^{\frac{1}{2-\alpha}} \\ &= \frac{(\alpha A)^{\frac{\alpha-1}{2-\alpha}+1}}{(2B)^{\frac{\alpha-1}{2-\alpha}}} - \frac{(\alpha A)^{\frac{1}{2-\alpha}}}{(2B)^{\frac{1}{2-\alpha}-1}} \\ &= \frac{(\alpha A)^{\frac{1}{2-\alpha}}}{(2B)^{\frac{\alpha-1}{2-\alpha}}} - \frac{(\alpha A)^{\frac{1}{2-\alpha}}}{(2B)^{\frac{\alpha-1}{2-\alpha}}} = 0, \end{aligned}$$

por lo cual se sigue que b_0 es un punto crítico de f. Observemos lo siguiente:

(i) Es inmediato ver en (4.9) que $f'(b) < 0$ si y sólo si $\alpha Ab^{\alpha-1} < 2Bb$, y esto último ocurre si y sólo si $\frac{\alpha A}{2B} < b^{2-\alpha}$, es decir, si y sólo si $\left(\frac{\alpha A}{2B}\right)^{1/(2-\alpha)} = b_0 < b$. Por otra parte $f'(b) > 0$ si y sólo si $b < b_0$. Esto nos indica que $f'(b)$ es positiva antes de b_0, toma el valor de cero sólo una vez en b_0, y una vez que pasa este valor es mantiene negativa la derivada, lo cual significa que b_0 es un máximo global de f.

(ii) Por otra parte, observemos que

$$f''(b) = \alpha(\alpha-1)Ab^{\alpha-2} - 2B \leq -2B \quad \text{si y sólo si} \quad \alpha(\alpha-1)Ab^{\alpha-2} \leq 0$$
$$\text{si y sólo si} \quad \alpha - 1 \leq 0,$$

donde pasamos de la segunda a la tercera desigualdad porque integraremos en valores $b \geq 0$ y las constantes son tales que $A, \alpha > 0$. Entonces si $\alpha \in (0,1]$ tenemos que $f''(b) \leq -2B$.

(iii) Por otra parte, $f'''(b) < 0$ si $\alpha(\alpha-1)(\alpha-2) < 0$, lo cual ocurre si $\alpha \in (1,2)$. Entonces si $\alpha \in (1,2)$, tendríamos que $f''' < 0$, por lo cual f'' sería decreciente. Entonces si $b \geq b_0$:

$$\begin{aligned} f''(b) \leq f''(b_0) =& \alpha(\alpha-1)A\Big(\frac{\alpha A}{2B}\Big)^{(\alpha-2)/(2-\alpha)} - 2B \\ =& A\Big(\frac{2B}{\alpha A}\Big) - 2B \\ =& 2B(\alpha-2) = -2B(2-\alpha). \end{aligned}$$

De (ii) y (iii) vemos que si $\alpha \in (0,2)$ y $b \geq b_0$, se cumple que

$$f''(b) \leq \text{máx}\{-2B, -2B(2-\alpha)\}. \tag{4.10}$$

(No necesitamos $b \geq b_0$ para la condición (ii), pero sí la ocupamos para la condición (iii).) Haciendo una expansión de Taylor de primer orden alrededor de b_0 y usando que $f'(b_0) = 0$ obtenemos que:

$$\begin{aligned} f(b) =& Ab_0^\alpha - Bb_0^2 + f'(b_0)(b-b_0) + \frac{1}{2}f''\big(\xi(b)\big)(b-b_0)^2 \\ =& Ab_0^\alpha - Bb_0^2 + \frac{1}{2}f''\big(\xi(b)\big)(b-b_0)^2 \end{aligned}$$

donde $\xi(b) \in [b, b_0]$ si $b < b_0$ y $\xi(b) \in [b_0, b]$ si $b > b_0$. Pero si $b > b_0$, entonces $\xi(b) \in [b_0, b]$, en particular $\xi(b) \geq b_0$ por lo cual, como ya hemos visto en (4.10),

$$f''(\xi(b)) \leq \text{máx}\{-2B, -2B(2-\alpha)\}.$$

De esto último combinado con la expansión de Taylor obtenida anteriormente se sigue que

$$\begin{aligned} \int_{b_0}^{\infty} \exp\big(f(b)\big)db \leq& \int_{b_0}^{\infty} \exp\Big(f(b_0) + \frac{1}{2}(b-b_0)^2\,\text{máx}\{-2B, -2B(2-\alpha)\}\Big)db \\ =& \exp\big(f(b_0)\big)\int_{b_0}^{\infty} \exp\Big(\frac{1}{2}(b-b_0)^2\,\text{máx}\{-2B, -2B(2-\alpha)\}\Big)db \end{aligned}$$

Entonces tenemos lo siguiente:

$$\begin{aligned}
&\int_0^\infty \exp\big(f(b)\big)db \\
=&\int_0^{b_0} \exp\big(f(b)\big)db + \int_{b_0}^\infty \exp\big(f(b)\big)db \\
\leq& b_0 e^{f(b_0)} + e^{f(b_0)} \int_{b_0}^\infty \exp\Big(\frac{1}{2}(b-b_0)^2 \text{máx}\{-2B, -2B(2-\alpha)\}\Big)db \\
=&e^{f(b_0)}\Big(b_0 + \int_{b_0}^\infty \exp\Big(\frac{1}{2}(b-b_0)^2 \text{máx}\{-2B, -2B(2-\alpha)\}\Big)db\Big) \qquad (4.11)
\end{aligned}$$

Usando que $\text{máx}\, A = -\text{mín}(-A)$ para todo conjunto de números y definiendo $C := \text{mín}\{2B, 2B(2-\alpha)\}$, obtenemos que la última integral de (4.11) es igual a

$$\begin{aligned}
&\int_{b_0}^\infty \exp\Big(-\frac{1}{2}(b-b_0)^2 \text{mín}\{2B, 2B(2-\alpha)\}\Big)db \\
=&\int_{b_0}^\infty \exp\Big(-\frac{(b-b_0)^2}{2(1/C)}\Big)db \leq \int_{-\infty}^\infty \exp\Big(-\frac{(b-b_0)^2}{2(1/C)}\Big)db.
\end{aligned} \qquad (4.12)$$

Haciendo el cambio de variable $w := \dfrac{b-b_0}{\sqrt{2/C}}$, la última integral de la expresión anterior es

$$\sqrt{\frac{2}{C}}\int_{-\infty}^\infty e^{-w^2}dw = \sqrt{\frac{2\pi}{C}},$$

donde hemos utilizado que $\int_{-\infty}^\infty e^{-w^2}dw = \sqrt{\pi}$.

De aquí se sigue que al sustituir en (4.12) obtenemos que

$$\int_{b_0}^\infty \exp\Big(-\frac{1}{2}(b-b_0)^2 \text{mín}\{2B, 2B(2-\alpha)\}\Big) \leq \sqrt{\frac{2\pi}{C}},$$

y sustituyendo esto último en la integral de (4.11) obtenemos que

$$\begin{aligned}
\int_0^\infty \exp\big(f(b)\big)db \leq& e^{f(b_0)}\Big(b_0 + \sqrt{\frac{2\pi}{C}}\Big) \\
=& e^{f(b_0)}\Big(b_0 + \sqrt{\frac{2\pi}{\text{mín}\{2B, 2B(2-\alpha)\}}}\Big)
\end{aligned}$$

□

Para el siguiente resultado, ocuparemos de la suposición 4.1 dada anteriormente:

Suposición 4.2. $F : \mathbb{W}_T \longrightarrow \mathbb{R} \cup \{-\infty\}$ *es continua y satisface la desigualdad*

$$F(x) \leq K_1 + K_2 \Big(\sup_{t\in[0,T]} |x_t|\Big)^\alpha \qquad (4.13)$$

para algunas $K_1, K_2 > 0$ *y para alguna* $0 < \alpha < 2$.

Lema 4.2. *Supongamos que $F : \mathbb{W}_T \longrightarrow \mathbb{R} \cup \{-\infty\}$ satisface la suposición 4.2. Sea $h \in \mathbb{H}_T$ tal que $\dot{h}$ es de variación acotada y definamos $F_h : \mathbb{W}_T \longrightarrow \mathbb{R} \cup \{-\infty\}$ como*

$$F_h(x) := 2F(x) - \int_0^T \dot{h}_t dx_t + \frac{1}{2}\int_0^T \dot{h}_t^2 dt.$$

Entonces F_h está bien definida, es continua y si $Z_\epsilon = \sqrt{\epsilon}B$, entonce F_h también satisface el criterio de integrabilidad

$$\limsup_{\epsilon\downarrow 0} \epsilon \log E\Big[\exp\Big(\frac{\gamma}{\epsilon} F_h(Z_\epsilon)\Big)\Big] < \infty$$

para toda $\gamma > 1$ y $h \in \mathbb{H}_T$ tal que $\dot{h}$ es de variación acotada.

Demostración. La definición de $\mathbb{H}_T$ implica que $\int_0^T \dot{h}_t^2 dt < \infty$, así que para verificar que F_h está bien definida, sólo necesitamos verificar que $\int_0^T \dot{h}_t dx_t$ está bien definida para toda $x \in \mathbb{W}_T$. Pero esto se sigue de la fórmula de integración por partes, que nos dice que existe $\int_0^T \dot{h}_t dx_t$ como una integral ordinaria de Lebesgue-Stieltjes, porque existe $\int_0^T x_t d\dot{h}_t$ (véase el teorema de integración por partes en el texto de D. Stroock [13] teorema 1.2.3 p. 9) gracias a que $\dot{h}$ tiene variación acotada en $[0,T]$ y x es continua en $[0,T]$, y más aún

$$\int_0^T \dot{h}_t dx_t = x_T\dot{h}_T - x_0\dot{h}_0 - \int_0^T x_t d\dot{h}_t = x_T\dot{h}_T - \int_0^T x_t d\dot{h}_t,$$

donde hemos usado que $x_0 = 0$. De aquí obtenemos que para toda $h \in \mathbb{H}_T$ con $\dot{h}$ de variación acotada, F_h está bien definida. Para verificar la continuidad, nótese que para $x, y \in \mathbb{W}_T$,

$$\begin{aligned} |F_h(x) - F_h(y)| =& \Big|2\big(F(x) - F(y)\big) - \Big(\int_0^T \dot{h}_t dx_t - \int_0^T \dot{h}_t dy_t\Big)\Big| \\ \leq& 2\big|F(x) - F(y)\big| + \Big|\int_0^T \dot{h}_t dx_t - \int_0^T \dot{h}_t dy_t\Big|. \end{aligned}$$

Por continuidad de F, vemos que la continuidad de F_h se seguirá de la continuidad del mapeo $x \mapsto \int_0^T \dot{h}_t dx_t$. Por la fórmula de integración por partes usada anteriormente, tenemos que si $\mathbf{V}(\dot{h})$ denota la variación de $\dot{h}$ en $[0,T]$, entonces:

$$\begin{aligned} \Big|\int_0^T \dot{h}_t dx_t\Big| = \Big|x_T\dot{h}_T - \int_0^T x_t d\dot{h}_t\Big| &\leq \|x\|_\infty |\dot{h}_T| + \|x\|_\infty \Big|\int_0^T d\dot{h}_t\Big| \\ &\leq \|x\|_\infty\big(|\dot{h}_T| + \mathbf{V}(\dot{h})\big), \end{aligned} \tag{4.14}$$

donde en la última desigualdad hemos utilizado que $\Big|\int_0^T d\dot{h}_t\Big| \leq \mathbf{V}(\dot{h})$, pues para toda partición $\Pi = \{0 = t_0 < t_1 < \ldots < t_n = T\}$ de $[0,T]$ se cumple que

$$\sum_{i=1}^n (\dot{h}_{t_i} - \dot{h}_{t_{i-1}}) \leq \sum_{i=1}^n \big|\dot{h}_{t_i} - \dot{h}_{t_{i-1}}\big| < M < \infty,$$

donde $M > 0$ es una cota para la variación de $\dot{h}$, que hemos supuesto que es finita. Tomando particiones Π tales que $\|\Pi\| \to 0$ obtenemos que $\left|\int_0^T d\dot{h}_t\right| \leq \mathbf{V}(\dot{h}) < M$. De aquí se sigue que en la ecuación (4.14) se tiene que

$$\left|\int_0^T \dot{h}_t dx_t\right| \leq \|x\|_\infty \big(|\dot{h}_T| + M\big) = M^* \|x\|_\infty,$$

con $0 < M^* = |\dot{h}_T| + M < \infty$. Como en $\mathbb{W}_T$ usamos la norma uniforme, esto prueba que $x \mapsto \int_0^T \dot{h}_t dx_t$ es un mapeo continuo de $\mathbb{W}_T$ a $\mathbb{R}$, y con ello se sigue la continuidad de F_h.

Ahora verificamos la condición de integrabilidad. Antes calculamos lo siguiente.

$$\begin{aligned}
&\epsilon \log E\Big[\exp\Big(\frac{\gamma}{\epsilon} F_h(\sqrt{\epsilon}B)\Big)\Big] \\
&= \epsilon \log E\Bigg[\exp\Bigg(\frac{\gamma}{\epsilon}\Bigg(2F(\sqrt{\epsilon}B) - \int_0^T \sqrt{\epsilon}\dot{h}_t dB_t + \frac{1}{2}\underbrace{\int_0^T \dot{h}_t^2 dt}_{\text{determinista}}\Bigg)\Bigg)\Bigg] \\
&= \epsilon \log\Bigg\{\exp\Bigg(\frac{\gamma}{2\epsilon}\int_0^T \dot{h}_t^2 dt\Bigg) \times E\Bigg[\exp\Bigg(\frac{\gamma}{\epsilon}\Bigg(2F(\sqrt{\epsilon}B) - \int_0^T \sqrt{\epsilon}\dot{h}_t dB_t\Bigg)\Bigg)\Bigg]\Bigg\} \\
&= \epsilon \log\exp\Bigg(\frac{\gamma}{2\epsilon}\int_0^T \dot{h}_t^2 dt\Bigg) + \epsilon \log E\Bigg[\exp\Bigg(\frac{\gamma}{\epsilon}\Bigg(2F(\sqrt{\epsilon}B) - \int_0^T \sqrt{\epsilon}\dot{h}_t dB_t\Bigg)\Bigg)\Bigg] \\
&= \frac{\gamma}{2}\int_0^T \dot{h}_t^2 dt + \epsilon \log E\Bigg[\exp\Bigg(\frac{\gamma}{\epsilon}\Bigg(2F(\sqrt{\epsilon}B) - \int_0^T \sqrt{\epsilon}\dot{h}_t dB_t\Bigg)\Bigg)\Bigg] \qquad (4.15)
\end{aligned}$$

Aplicando la desigualdad de Cauchy-Schwarz $E[e^a e^b] \leq E[e^{2a}]^{1/2} E[e^{2b}]^{1/2}$ al último sumando de (4.15) obtenemos que

$$\begin{aligned}
&\epsilon \log E\Bigg[\exp\Bigg(\frac{\gamma}{\epsilon}\Bigg(2F(\sqrt{\epsilon}B) - \int_0^T \sqrt{\epsilon}\dot{h}_t dB_t\Bigg)\Bigg)\Bigg] \\
&= \epsilon \log E\Bigg[\underbrace{\exp\Bigg(\frac{\gamma}{\epsilon}2F(\sqrt{\epsilon}B)\Bigg)}_{e^a} \times \underbrace{\exp\Bigg(-\frac{\gamma}{\epsilon}\int_0^T \sqrt{\epsilon}\dot{h}_t dB_t\Bigg)}_{e^b}\Bigg] \\
&\leq \epsilon \log\Bigg\{E\Big[\exp\Big(\frac{4\gamma}{\epsilon}F(\sqrt{\epsilon}B)\Big)\Big]^{1/2} \times E\Bigg[\exp\Bigg(-\frac{2\gamma}{\epsilon}\int_0^T \sqrt{\epsilon}\dot{h}_t dB_t\Bigg)\Bigg]^{1/2}\Bigg\} \\
&= \frac{\epsilon}{2}\log E\Big[\exp\Big(\frac{4\gamma}{\epsilon}F(\sqrt{\epsilon}B)\Big)\Big] + \frac{\epsilon}{2}\log E\Bigg[\exp\Bigg(-\frac{2\gamma}{\epsilon}\sqrt{\epsilon}\int_0^T \dot{h}_t dB_t\Bigg)\Bigg].
\end{aligned}$$

Usando esta última desigualdad en la ecuación (4.15), obtenemos que

$$\begin{aligned}
&\epsilon \log E\Big[\exp\Big(\frac{\gamma}{\epsilon} F_h(\sqrt{\epsilon}B)\Big)\Big] \qquad (4.16)\\
&\leq \frac{\gamma}{2}\int_0^T \dot{h}_t^2 dt + \frac{\epsilon}{2}\log E\Bigg[\exp\Bigg(\frac{4\gamma}{\epsilon}F(\sqrt{\epsilon}B)\Bigg)\Bigg] + \frac{\epsilon}{2}\log E\Bigg[\exp\Bigg(-\frac{2\gamma}{\sqrt{\epsilon}}\int_0^T \dot{h}_t dB_t\Bigg)\Bigg].
\end{aligned}$$

Probemos entonces que el término del lado derecho de la desigualdad anterior es finito al tomar el límite superior. Es claro que el primer término es finito, por ser $h \in \mathbb{H}_T$. Para el tercer sumando, usaremos que

$$\int_0^T \dot{h}_t dB_t \sim N\Big(0, \int_0^T \dot{h}_t^2 dt\Big)$$

(véase el teorema 2.2 del capítulo 2) y que $E[e^{\theta X}] = e^{\theta^2\sigma^2/2}$ si $X \sim N(0,\sigma^2)$. De esta manera obtenemos que si $\sigma^2 = \int_0^T \dot{h}_t^2 dt$ y $\theta = 2\gamma/\sqrt{\epsilon}$, entonces la tercera integral del lado derecho de (4.16) cumple que

$$\begin{aligned}
\limsup_{\epsilon\downarrow 0} \frac{\epsilon}{2} \log E\Big[\exp\Big(-\frac{2\gamma}{\sqrt{\epsilon}}\int_0^T \dot{h}_t dB_t\Big)\Big] &= \limsup_{\epsilon\downarrow 0} \frac{\epsilon}{2} \log \exp\left(\frac{1}{2}\left(\frac{2\gamma}{\sqrt{\epsilon}}\right)^2 \int_0^T \dot{h}_t^2 dt\right)\\
&= \gamma^2 \int_0^T \dot{h}_t^2 dt < \infty.
\end{aligned}$$

Falta ver que el segundo término del lado derecho de (4.16) es finito. Para esto, usemos la suposición 4.2 para obtener que:

$$\begin{aligned}
&\frac{\epsilon}{2} \log E\Big[\exp\Big(\frac{4\gamma}{\epsilon} F(\sqrt{\epsilon}B)\Big)\Big]\\
&\quad\leq \frac{\epsilon}{2} \log E\left[\exp\left(\frac{4\gamma}{\epsilon}\Big(K_1 + K_2\Big(\sup_{t\in[0,T]} |\sqrt{\epsilon}B_t|\Big)^\alpha\Big)\right)\right]\\
&\quad= \frac{\epsilon}{2} \log\left\{\exp\Big(\frac{4\gamma}{\epsilon}K_1\Big) \cdot E\left[\exp\Big(\frac{4\gamma}{\epsilon} K_2\Big(\sup_{t\in[0,T]} |\sqrt{\epsilon}B_t|\Big)^\alpha\Big)\right]\right\}\\
&\quad= 2\gamma K_1 + \frac{\epsilon}{2}\log E\left[\exp\Big(\frac{4\gamma}{\epsilon} K_2 \sqrt{\epsilon}^{\,\alpha}\Big(\sup_{t\in[0,T]} |B_t|\Big)^\alpha\Big)\right]\\
&\quad= 2\gamma K_1 + \frac{\epsilon}{2}\log E\left[\exp\Big(\frac{4\gamma K_2}{\epsilon^{1-\alpha/2}}\Big(\sup_{t\in[0,T]} |B_t|\Big)^\alpha\Big)\right],
\end{aligned}$$

es decir,

$$\begin{aligned}
&\limsup_{\epsilon\downarrow 0} \frac{\epsilon}{2} \log E\Big[\exp\Big(\frac{4\gamma}{\epsilon} F(\sqrt{\epsilon}B)\Big)\Big]\\
&\qquad\leq 2\gamma K_1 + \limsup_{\epsilon\downarrow 0}\frac{\epsilon}{2}\log E\left[\exp\Big(\frac{4\gamma K_2}{\epsilon^{1-\alpha/2}}\Big(\sup_{t\in[0,T]} |B_t|\Big)^\alpha\Big)\right].
\end{aligned} \tag{4.17}$$

Si podemos probar que el último sumando del lado derecho de (4.17) es finito, habremos terminado. Nótese que

$$\Big\{\sup_{t\in[0,T]} |B_t| \geq b\Big\} \subset \Big\{\sup_{t\in[0,T]} B_t \geq b\Big\} \bigcup \Big\{\inf_{t\in[0,T]} B_t \leq -b\Big\}.$$

Usemos que $\inf_{t\in[0,T]} B_t = -\sup_{t\in[0,T]}(-B_t)$. Pero como sabemos (basta con ver la distribución normal, que es simétrica respecto al valor esperado), si B es un movimiento browniano estándar, entonces $-B$ también lo es. De estas observaciones, se sigue que

$$P\Big\{\sup_{t\in[0,T]} |B_t| \geq b\Big\} \leq 2P\Big\{\sup_{t\in[0,T]} B_t \geq b\Big\}. \tag{4.18}$$

Definamos a $A := \dfrac{4\gamma K_2}{\epsilon^{1-\alpha/2}}$ y a la variable aleatoria

$$X := \exp\left(A\cdot\Big(\sup_{t\in[0,T]}|B_t|\Big)^{\alpha}\right) \geq 0.$$

Nótese que podemos aplicar el lema A.1 del apéndice[2] a X para calcular su valor esperado. Usando dicho lema junto con la ecuación (4.18), tenemos que

$$\begin{aligned}
E[X] &= \int_0^\infty P\{X \geq b\}db \\
&= \int_0^\infty P\Big\{\exp\left(A\cdot\Big(\sup_{t\in[0,T]}|B_t|\Big)^{\alpha}\right) \geq b\Big\}db \\
&= \int_0^\infty P\Big\{\sup_{t\in[0,T]}|B_t| \geq \Big(\frac{\log b}{A}\Big)^{1/\alpha}\Big\}db \\
&\leq 2\int_0^\infty P\Big\{\sup_{t\in[0,T]}B_t \geq \Big(\frac{\log b}{A}\Big)^{1/\alpha}\Big\}db \\
&= 2\int_0^\infty P\Big\{\exp\left(A\Big(\sup_{t\in[0,T]}B_t\Big)^{\alpha}\right) \geq b\Big\}db \\
&= 2E\left[\exp\left(A\Big(\sup_{t\in[0,T]}B_t\Big)^{\alpha}\right)\right]
\end{aligned}$$

Usando esto último y el hecho de que por el principio de reflexión A.2 del apéndice, la distribución de $\sup_{t\in[0,T]} B_t$ está dada por

$$P\Big\{\sup_{t\in[0,T]} B_t \in db\Big\} = \sqrt{\frac{2}{\pi T}}\exp\big(-b^2/2T\big)db.$$

[2] Dicho lema nos dice que para $X \geq 0$, $E[X] = \int_0^\infty P\{X \geq b\}db$.

Usando esto último, obtenemos que

$$
\begin{aligned}
E\left[\exp\left(\frac{4\gamma K_2}{\epsilon^{1-\alpha/2}}\Big(\sup_{t\in[0,T]}|B_t|\Big)^{\alpha}\right)\right] &\le 2E\left[\exp\left(\frac{4\gamma K_2}{\epsilon^{1-\alpha/2}}\Big(\sup_{t\in[0,T]}B_t\Big)^{\alpha}\right)\right] \\
&= 2\int_0^{\infty}\exp\left(\frac{4\gamma K_2}{\epsilon^{1-\alpha/2}}b^{\alpha}\right)P\Big\{\sup_{t\in[0,T]}B_t\in db\Big\} \\
&= 2\sqrt{\frac{2}{\pi T}}\int_0^{\infty}\exp\left(\frac{4\gamma K_2}{\epsilon^{1-\alpha/2}}b^{\alpha}-\frac{b^2}{2T}\right)db \\
&= 2\sqrt{\frac{2}{\pi T}}\int_0^{\infty}\exp\big(Ab^{\alpha}-Bb^2\big)db, \qquad (4.19)
\end{aligned}
$$

donde hemos definido $B := 1/2T$ y A como definimos arriba. Podemos aplicar el lema 4.1, con

$$
\begin{aligned}
b_0 :&= \Big(\frac{\alpha A}{2B}\Big)^{1/(2-\alpha)} \\
&= \Big(\frac{4\alpha\gamma K_2}{2\epsilon^{1-\alpha/2}/2T}\Big)^{\frac{1}{2-\alpha}} \quad = (4\alpha T\gamma K_2)^{\frac{1}{2-\alpha}}\Big(\frac{1}{\epsilon^{(2-\alpha)/2}}\Big)^{\frac{1}{2-\alpha}} = (4\alpha T\gamma K_2)^{1/(2-\alpha)}\frac{1}{\sqrt{\epsilon}}.
\end{aligned}
$$

De este modo, con el lema 4.1 obtenemos que la integral de (4.19) cumple que

$$
\begin{aligned}
\int_0^{\infty}\exp\big(Ab^{\alpha}-Bb^2\big)db &\le \exp(Ab_0^{\alpha}-Bb_0^2)\left(b_0+\sqrt{\frac{2\pi}{\text{mín}\{2B,2B(2-\alpha)\}}}\right) \\
&= \exp\left(\frac{4\gamma K_2}{\epsilon^{1-\alpha/2}}(4\alpha T\gamma K_2)^{\alpha/(2-\alpha)}\epsilon^{-\alpha/2}-\frac{1}{2T}(4\alpha T\gamma K_2)^{2/(2-\alpha)}\epsilon^{-1}\right) \\
&\times\left((4\alpha T\gamma K_2)^{1/(2-\alpha)}\epsilon^{-1/2}+\sqrt{\frac{2\pi}{\text{mín}\{1/T\,,\,1/T(2-\alpha)\}}}\right).
\end{aligned}
$$

Nótese que si definimos $N := 4\alpha T\gamma K_2$ y a $M := \text{mín}\{\frac{1}{T},\frac{1}{T}(2-\alpha)\}$, entonces al simplificar al argumento de la función exponencial en la última ecuación obtenida:

$$
\begin{aligned}
\frac{4\gamma K_2}{\epsilon^{1-\alpha/2}}(4\alpha T\gamma K_2)^{\alpha/(2-\alpha)}\epsilon^{-\alpha/2}-\frac{1}{2T\epsilon}(4\alpha T\gamma K_2)^{2/(2-\alpha)} &= \frac{4\gamma K_2}{\epsilon^{1-\alpha/2}}\epsilon^{-\alpha/2}N^{\alpha/(2-\alpha)}-\frac{1}{2T\epsilon}N^{2/(2-\alpha)} \\
&= \frac{\alpha T}{\alpha T}\cdot\frac{4\gamma K_2}{\epsilon}N^{\alpha/(2-\alpha)}-\frac{1}{2T\epsilon}N^{2/(2-\alpha)} \\
&= \frac{N}{\epsilon\alpha T}N^{\alpha/(2-\alpha)}-\frac{1}{2T\epsilon}N^{2/(2-\alpha)} \\
&= \frac{N^{2/(2-\alpha)}}{\epsilon T}\Big(\frac{1}{\alpha}-\frac{1}{2}\Big),
\end{aligned}
$$

de tal forma que obtenemos que

$$
\int_0^{\infty}\exp\big(Ab^{\alpha}-Bb^2\big)db \le \exp\left(\frac{N^{2/(2-\alpha)}}{\epsilon T}\Big(\frac{1}{\alpha}-\frac{1}{2}\Big)\right)\times\underbrace{\left(N^{1/(2-\alpha)}\epsilon^{-1/2}+\sqrt{\frac{2\pi}{M}}\right)}_{\text{Sea esto}=:K_\epsilon}.
$$

Usando esto último en la ecuación (4.19) obtenemos que

$$\limsup_{\epsilon\downarrow 0}\frac{\epsilon}{2}\log E\left[\exp\left(\frac{4\gamma K_2}{\epsilon^{1-\alpha/2}}\Big(\sup_{t\in[0,T]}|B_t|\Big)^{\alpha}\right)\right]$$

$$\leq \limsup_{\epsilon\downarrow 0}\frac{\epsilon}{2}\log\left(2K_\epsilon\sqrt{\frac{2}{\pi T}}\exp\left(\frac{N^{\frac{2}{2-\alpha}}}{\epsilon T}\Big(\frac{1}{\alpha}-\frac{1}{2}\Big)\right)\right)$$

$$=\underbrace{\limsup_{\epsilon\downarrow 0}\frac{\epsilon}{2}\log\left(2K_\epsilon\sqrt{\frac{2}{\pi T}}\right)}_{0}+\limsup_{\epsilon\downarrow 0}\frac{\epsilon}{2}\log\left(\exp\left(\frac{N^{\frac{2}{2-\alpha}}}{\epsilon T}\Big(\frac{1}{\alpha}-\frac{1}{2}\Big)\right)\right)$$

$$=\limsup_{\epsilon\downarrow 0}\frac{\epsilon}{2}\frac{N^{2/(2-\alpha)}}{\epsilon T}\Big(\frac{1}{\alpha}-\frac{1}{2}\Big)=\frac{N^{2/(2-\alpha)}}{2T}\Big(\frac{1}{\alpha}-\frac{1}{2}\Big)<\infty,$$

donde usamos que $\alpha<2$ y en la primera igualdad que

$$\limsup_{\epsilon\downarrow 0}\frac{\epsilon}{2}\log\left(2K_\epsilon\sqrt{\frac{2}{\pi T}}\right)$$

$$=\limsup_{\epsilon\downarrow 0}\frac{\epsilon}{2}\log\left(N^{1/(2-\alpha)}\epsilon^{-1/2}+\sqrt{\frac{2\pi}{M}}\right)+\limsup_{\epsilon\downarrow 0}\frac{\epsilon}{2}\log\left(2\sqrt{\frac{2}{\pi T}}\right)=0,$$

donde se ha utilizado que $\lim_{x\to 0}x\log\left(\frac{1}{\sqrt{x}}\right)=0$, por una sencilla aplicación de la regla de L'Hospital. Como se dijo anteriormente en la ecuación (4.17), esto es suficiente para probar el resultado, pues esto prueba que el lado derecho de (4.16) es finito.

$\square$

OBSERVACIÓN: Nótese que gracias a que por definición, F es continua, y probamos en las ecuaciones (4.16) y (4.17) que $\limsup_{\epsilon\downarrow 0}\frac{\epsilon}{2}\log E\big[\exp\Big(\frac{4\gamma}{\epsilon}F(\sqrt{\epsilon}B)\Big)\big]<\infty$, entonces se sigue que F también cumple las hipótesis requeridas para aplicar la extensión del lema de Varadhan.

Antes de continuar, probaremos que el espacio clásico de Cameron-Martin $\mathbb{H}_T$ con el producto interior

$$\langle x,y\rangle_{\mathbb{H}_T}:=\int_0^T\dot{x}(t)\dot{y}(t)dt$$

es un espacio de Hilbert.

Teorema 4.1. $\mathbb{H}_T$ *es un espacio de Hilbert.*

Demostración. Sea (g_n) una sucesión de Cauchy en $\mathbb{H}_T$. Por definición de la métrica de $\mathbb{H}_T$, esto es equivalente a decir que $(\dot{g}_n)$ es una sucesión de Cauchy en $L^2[0,T]$. Pero

$L^2[0,T]$ es un espacio de Hilbert, por lo cual $\dot{g}_n \to \gamma$ si $n \to \infty$, para alguna $\gamma \in L^2[0,T]$. Nótese que

$$g(t) := \int_0^t \gamma(s)ds \in AC[0,T],$$

por el Teorema 14 p. 110 en H.L. Royden [11]. Claramente, $g(0) = 0$, y dado que γ es Lebesgue integrable, se sigue que para casi toda $t \in [0,T]$: $\dot{g}(t) = \gamma(t)$, por el teorema fundamental del cálculo de Lebesgue. Finalmente gracias a que $\gamma \in L^2[0,T]$ se sigue que $\dot{g} \in L^2[0,T]$. Esto se traduce a que $g_n \to g$ si $n \to \infty$ en $\mathbb{H}_T$, por lo cual $\mathbb{H}_T$ es completo en la métrica inducida por su producto interior.

$\square$

Nótese que $\mathbb{H}_T$ puede verse como un subespacio de Hilbert continuamente encajado en $\mathbb{W}_T$, el espacio clásico de Wiener. Esto es así porque para cada $h \in \mathbb{H}_T$ y $t \in [0,T]$ se cumple que

$$|h(t)| = \left|\int_0^t \dot{h}(s)ds\right| \leq T^{1/2}\left(\int_0^T \dot{h}^2(s)ds\right)^{1/2} = T^{1/2}\|h\|_{\mathbb{H}_T},$$

es decir que la función de inclusión $i : \mathbb{H}_T \longrightarrow \mathbb{W}_T$ tal que $h \overset{i}{\hookrightarrow} h$ es continua. $\mathbb{H}_T$ es un subespacio de Hilbert importante para $\mathbb{W}_T$, y explotaremos sus propiedades para hallar una solución a nuestro problema principal a resolver en el trabajo.

Diremos que una función definida en un espacio métrico $F : \mathcal{X} \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ es *coerciva* si cumple que

$$F(x) \longrightarrow \infty \quad \text{conforme} \quad \|x\| \to \infty.$$

El siguiente resultado es inspirado de [22] de J. Jost y X. Li-Jost, p. 186 Teorema 4.2.1, sólo que usamos hipótesis más restringidas para nuestro caso particular, por lo cual no ocupamos tantos resultados como en [22], y la prueba se facilita.

Teorema 4.2. (Existencia de mínimo en espacio de Hilbert) *Sea $\mathcal{H}$ un espacio de Hilbert, $F : \mathcal{H} \longrightarrow \mathbb{R} \cup \{\infty\}$ tal que F no es idénticamente ∞, es coerciva y débilmente semicontinua inferiormente. Entonces existe un minimizador $x_0 \in \mathcal{H}$, es decir, un elemento x_0 tal que*

$$F(x_0) = \inf_{x \in \mathcal{H}} F(x) > -\infty.$$

Demostración. Sea $\{x_n\}$ una sucesión minimizadora para F, es decir, una sucesión tal que

$$\lim_{n\to\infty} F(x_n) = \inf_{x \in \mathcal{H}} F(x).$$

Nótese que como $F \neq \infty$, entonces $\inf_{x\in\mathcal{H}} F(x) < \infty$. Como F es coerciva, si $\|x_n\| \to \infty$ conforme $n \to \infty$, entonces $F(x_n) \xrightarrow{n} \infty$. Pero como $\{x_n\}_n$ es una sucesión minimizadora, esto no puede suceder, pues $\lim_{n\to\infty} F(x_n) = \inf_{x\in\mathcal{H}} F(x) < \infty$. Entonces se sigue que $\{x_n\}_n$ está acotada en $\mathcal{H}$. Como toda sucesión acotada en un espacio de Hilbert tiene una subsucesión débilmente convergente (véase teorema 21.8 p. 293 del texto de J. Jost [6]), entonces existe $x_0 \in \mathcal{H}$ tal que $x_n \xrightarrow{n} x_0$ débilmente. Pero como F es semicontinua inferiormente en la topología débil (véase la sección A.1 de semicontinuidad en el apéndice):

$$F(x_0) \leq \liminf_{n\to\infty} F(x_n) = \inf_{x\in\mathcal{H}} F(x).$$

Dado que $x_0 \in \mathcal{H}$, debe ser que $F(x_0) = \inf_{x\in\mathcal{H}} F(x) > -\infty$, pues F no toma el valor extendido $-\infty$ en $\mathcal{H}$.

$\square$

Nótese que si $g_n \xrightarrow{n} g$ débilmente en $\mathbb{H}_T$, entonces la sucesión $\{\dot{g}_n\}$ está acotada en $L^2[0,T]$, es decir, existe $M > 0$ tal que

$$\|\dot{g}_n\|_{L^2} \leq M \quad \text{para toda } n \in \mathbb{N}.$$

(lema 21.11 p. 297 del texto de J. Jost [6]). Sea $\epsilon > 0$. Nótese que si $\delta = \epsilon^2/M^2$, entonces para $s, t \in [0,T]$ tales que $|s-t| < \delta$:

$$|g_n(t) - g_n(s)| = \left|\int_s^t \dot{g}_n(u)du\right| \leq \|\dot{g}_n\|_{L^2}\left(\int_s^t du\right)^{1/2} \leq |s-t|^{1/2}M < \delta^{1/2}M = \epsilon,$$

de tal modo que $\{g_n\}$ es equicontinua. Por otro lado, $\{g_n\}$ converge a g puntualmente, porque si para cada $t \in [0,T]$ definimos

$$h^t(s) = \begin{cases} s & \text{si } s \in [0,t) \\ t & \text{si } s \in [t,T], \end{cases}$$

entonces podemos usar

$$\dot{h}^t(s) = \begin{cases} 1 & \text{si } s \in (0,t) \\ 0 & \text{si } s \in (t,T), \end{cases}$$

de tal manera que es cierto que $h^t(s) = \int_0^s \dot{h}^t(u)du$ y $h^t \in \mathbb{H}_T$, entonces por convergencia débil en $\mathbb{H}_T$:

$$\langle g_n, h^t\rangle_{\mathbb{H}_T} \to \langle g, h^t\rangle_{\mathbb{H}_T} \quad \text{si } n \to \infty.$$

Esto se traduce en

$$\langle \dot{g}_n, \dot{h}^t\rangle_{L^2[0,T]} \to \langle \dot{g}_n, \dot{h}^t\rangle_{L^2[0,T]} \quad \text{si } n \to \infty,$$

de donde obtenemos que

$$\left|\int_0^T \big(\dot{g}_n(s) - \dot{g}(s)\big)\dot{h}^t(s)ds\right| = \left|\int_0^t \big(\dot{g}_n(s) - \dot{g}(s)\big)ds\right| \to 0 \quad \text{si } n \to \infty.$$

Pero por definición de $\mathbb{H}_T$, esto implica que

$$\left|\int_0^t \big(\dot{g}_n(s) - \dot{g}(s)\big)ds\right| = |g_n(t) - g(t)| \to 0 \quad \text{si } n \to \infty.$$

Como esto ocurre para cada $t \in [0,T]$, entonces concluimos que $\{g_n\}$ converge a g puntualmente. Por el lema A.4 del apéndice, se sigue que $\{g_n\}$ converge a g uniformemente en $[0,T]$, es decir $g_n \xrightarrow{n} g$ en $C[0,T]$. la observación anterior nos auxilia en la prueba del siguiente resultado, tomado de P. Guasoni y S. Robertson [20].

Lema 4.3. *Sea $F : \mathbb{W}_T \longrightarrow \mathbb{R} \cup \{-\infty\}$ continua tal que satisface la suposición 4.2. Entonces para cualquier $h \in \mathbb{H}_T$ y $M > 0$, existe un maximizador para el problema*

$$\max_{x \in \mathbb{H}_T} \left\{ 2F(x) + M \int_0^T (\dot{x}_t - \dot{h}_t)^2 dt - 2M \int_0^T \dot{x}_t^2 dt + (1-2M) \int_0^T \dot{h}_t^2 dt \right\} \tag{4.20}$$

Demostración. Nótese que (4.20) es igual a

$$\max_{x \in \mathbb{H}_T} \left\{ 2F(x) - M\|x+h\|^2_{\mathbb{H}_T} + \|h\|^2_{\mathbb{H}_T} \right\}.$$

De esta última ecuación, definimos $G(x) := 2F(x) - M\|x+h\|^2_{\mathbb{H}_T} + \|h\|^2_{\mathbb{H}_T}$. Aplicaremos el teorema 4.2 a $-G$ restringido a $\mathbb{H}_T$, pues

$$-\min_{x \in \mathbb{H}_T} \left\{ -G(x) \right\} = \max_{x \in \mathbb{H}_T} G(x), \tag{4.21}$$

por lo cual el mínimo de $-G$ nos proporciona el máximo buscado. Nótese que por definición de F tenemos que $-G : \mathbb{W}_T \longrightarrow \mathbb{R} \cup \{\infty\}$.

Como buscaremos el máximo en $\mathbb{H}_T$, entonces probaremos que $-G$ restringida a $\mathbb{H}_T$ es semicontinua inferiormente en la topología débil de $\mathbb{H}_T$. Para esto, como dijimos en la observación justo antes de este teorema, si $g_n \xrightarrow{n} g$ débilmente en $\mathbb{H}_T$, entonces $g_n \xrightarrow{n} g$ en $C[0,T]$. Pero F por hipótesis es continua en $\mathbb{W}_T$, por lo cual $g_n \xrightarrow{n} g$ en $C[0,T]$ implica que $F(g_n) \xrightarrow{n} F(g)$, es decir, entonces F es débilmente continua *con el producto interior de* $\mathbb{H}_T$.

Por otra parte, el mapeo $f : \mathbb{H}_T \longrightarrow \mathbb{R}$ definido por $f(x) := \sqrt{M}\|x+h\|_{\mathbb{H}_T}$ es continuo, pues si $\epsilon > 0$, tomando $\delta = \epsilon/\sqrt{M}$ se sigue que si $\|x-y\|_{\mathbb{H}_T} < \delta$, entonces por la desigualdad triangular

$$|f(x) - f(y)| = \sqrt{M}\left| \|x+h\|_{\mathbb{H}_T} - \|y+h\|_{\mathbb{H}_T} \right| \le \sqrt{M}\|x-y\|_{\mathbb{H}_T} < \epsilon,$$

es decir, f es continua, por lo cual f^2 también es continua y con ello $x \mapsto M\|x+h\|^2_{\mathbb{H}_T}$ es continua, lo cual implica que es también $\mathbb{H}_T$ débilmente semicontinua inferiormente. Con todo esto, obtenemos que

$$-G(x) = M\|x+h\|^2_{\mathbb{H}_T} - 2F(x) - \|h\|^2_{\mathbb{H}_T}$$

es débilmente semicontinua inferiormente.

Ahora verifiquemos la condición de coercividad. Para esto, nótese que gracias a la suposición 4.2, existen $K_1, K_2 > 0$ tales que

$$M\|x+h\|^2_{\mathbb{H}_T} - \|h\|^2_{\mathbb{H}_T} - 2F(x) \ge M\|x+h\|^2_{\mathbb{H}_T} - \|h\|^2_{\mathbb{H}_T} - 2K_1 - 2K_2\|x\|^{\alpha}_{\infty}. \tag{4.22}$$

Pero por la desigualdad de Cauchy-Schwarz:

$$\begin{aligned}
\|x\|_\infty^\alpha &= \left(\max_{t\in[0,T]} |x_t|\right)^\alpha \\
&= \left(\max_{t\in[0,T]} \left|\int_0^t \dot{x}_s ds\right|\right)^\alpha \\
&\leq \left(\max_{t\in[0,T]} \int_0^t |\dot{x}_s| ds\right)^\alpha \\
&= \left(\int_0^T |\dot{x}_s| ds\right)^\alpha \leq T^{\alpha/2}\left(\int_0^T |\dot{x}_s|^2 ds\right)^{\alpha/2},
\end{aligned}$$

de donde se sigue que $-2K_2\|x\|_\infty^\alpha \geq -2K_2T^{\alpha/2}\|x\|_{\mathbb{H}_T}^\alpha$. Como $\alpha < 2$, si $\|x\|_{\mathbb{H}_T} \to \infty$ entonces

$$M\|x+h\|_{\mathbb{H}_T}^2 - 2K_2T^{\alpha/2}\|x\|_{\mathbb{H}_T}^\alpha \to \infty.$$

Usando esto en la desigualdad (4.22), obtenemos que

$$-G(x) \geq M\|x+h\|_{\mathbb{H}_T}^2 - \|h\|_{\mathbb{H}_T}^2 - 2K_1 - 2K_2T^{\alpha/2}\|x\|_{\mathbb{H}_T}^\alpha \to \infty \qquad \text{si } \|x\|_{\mathbb{H}_T} \to \infty.$$

De aquí se sigue la coercividad de $-G$. Como $-G$ también es $\mathbb{H}_T$ semicontinua inferiormente, y $\mathbb{H}_T$ es un espacio de Hilbert, aplicando el teorema 4.2 obtenemos que existe un mínimo $x_0 \in \mathbb{H}_T$ para $-G$, y concluimos que existe x_0 tal que

$$G(x_0) = \max_{x\in\mathbb{H}_T} G(x) = \max_{x\in\mathbb{H}_T}\Big\{2F(x) - M\|x+h\|_{\mathbb{H}_T}^2 + \|h\|_{\mathbb{H}_T}^2\Big\}. \tag{4.23}$$

Nótese además que gracias a que $\min_{x\in\mathbb{H}_T}\big\{-G(x)\big\} > -\infty$, entonces

$$-\min_{x\in\mathbb{H}_T}\big\{-G(x)\big\} = \max_{x\in\mathbb{H}_T} G(x) < \infty.$$

□

4.3. Solución al Problema

Es hora es compilar todos los resultados obtenidos hasta ahora, para solucionar por fin el Problema 4.1 dado al inicio de este capítulo. Recordemos que queremos minimizar a una aproximación de una integral, lo cual de acuerdo a la ecuación (4.5) es equivalente a minimizar

$$L(h) := \limsup_{\epsilon\to 0} \epsilon \log E_P\left[\exp\left(\frac{1}{\epsilon}\Big(2F(\sqrt{\epsilon}B) - \int_0^T \sqrt{\epsilon}\dot{h}_t dB_t + \frac{1}{2}\int_0^T \dot{h}_t^2 dt\Big)\right)\right]. \tag{4.24}$$

Entonces el objetivo es minimizar esta última expresión y una solución al problema

$$\min_{h\in\mathbb{H}_T} L(h)$$

será denominada una **solución asintóticamente óptima**.

Auxiliándonos de todos los resultados anteriores que hemos probado, se demuestra el resultado principal.

Teorema 4.3. *Supóngase que $F = \log(G)$ satisface la suposición 4.2. Entonces se cumple lo siguiente:*

(i) Si $h \in \mathbb{H}_T$ tal que su derivada $\dot{h}$ tiene variación finita en $[0,T]$ y L está definido como en (4.24), entonces

$$L(h) = \sup_{x \in \mathbb{H}_T} \left\{ 2F(x) + \frac{1}{2}\int_0^T (\dot{x}_t - \dot{h}_t)^2 dt - \int_0^T \dot{x}_t^2 dt \right\} \tag{4.25}$$

(ii) Para todo $h \in \mathbb{H}_T$, existen maximizadores tanto para la expresión (4.25) como para

$$\sup_{x \in \mathbb{H}_T} \left\{ 2F(x) - \int_0^T \dot{x}_t^2 dt \right\}. \tag{4.26}$$

(iii) Si $\hat{h}$ es una solución al problema (4.26), entonces $\hat{h}$ cumple ser asintóticamente óptimo si

$$L(\hat{h}) = 2F(\hat{h}) - \int_0^T \dot{\hat{h}}_t^2 dt. \tag{4.27}$$

Más aún, si se cumple (4.27) y $\dot{\hat{h}}$ tiene variación finita, entonces $\hat{h}$ es solución única de (4.26).

Demostración. Para probar (i), nos fijamos en el conjunto indizado por $\epsilon \geq 0$ de movimientos brownianos escalados de la forma

$$Z_\epsilon := \{\sqrt{\epsilon} B_t : t \in [0,T]\}.$$

Es decir, para cada ϵ tenemos un proceso estocástico, o equivalentemente, una variable aleatoria Z_ϵ con valores en $\mathcal{X} = \mathbb{W}_T$. Por el lema de Schilder (Teorema 3.1), las medidas inducidas $\{\mu_\epsilon\}_\epsilon$ por el conjunto $\{Z_\epsilon\}_\epsilon$ satisfacen un principio de grandes desviaciones con función de tasa buena

$$I(x) = \begin{cases} \dfrac{1}{2}\displaystyle\int_0^T \dot{x_t}^2 dt & \text{si } x \in \mathbb{H}_T \\ \infty & \text{si } x \in \mathbb{W}_T \setminus \mathbb{H}_T. \end{cases} \tag{4.28}$$

Gracias a esto último en conjunto con el lema 4.2, para cada $h \in \mathbb{H}_T$ tal que $\dot{h}_t$ tiene variación finita, la función

$$F_h(x) := 2F(x) - \int_0^T \dot{h}_t dx_t + \frac{1}{2}\int_0^T \dot{h}_t^2 dt$$

satisface las hipótesis de la extensión del lema de Varadhan (teorema 3.3), pues es continua y cumple el criterio de integrabilidad requerido. De aquí se sigue que si en la desigualdad (3.25) de la página 44 que nos da el teorema 3.3 usamos $A = \mathcal{X} = \mathbb{W}_T = \mathring{\mathbb{W}}_{\mathbb{T}} = \overline{\mathbb{W}}_T$, esta toma la forma[3]

$$\sup_{x\in\mathbb{W}_T}\{F_h(x) - I(x)\} \leq \liminf_{\epsilon\downarrow 0} \epsilon \log E\left[\exp\left(\frac{1}{\epsilon}F_h(\sqrt{\epsilon}B)\right)\right]$$
$$\leq \limsup_{\epsilon\downarrow 0} \epsilon \log E\left[\exp\left(\frac{1}{\epsilon}F_h(\sqrt{\epsilon}B)\right)\right] \leq \sup_{x\in\mathbb{W}_T}\{F_h(x) - I(x)\}$$

de tal suerte que

$$\limsup_{\epsilon\downarrow 0} \epsilon \log E\left[\exp\left(\frac{1}{\epsilon}F_h(\sqrt{\epsilon}B)\right)\right] = \sup_{x\in\mathbb{W}_T}\{F_h(x) - I(x)\},$$

lo cual por definición de F_h quiere decir:

$$\begin{aligned}\limsup_{\epsilon\downarrow 0} \epsilon \log E\left[\exp\left(\frac{1}{\epsilon}\left(2F(\sqrt{\epsilon}B) - \int_0^T \sqrt{\epsilon}\dot{h}_t dB_t + \frac{1}{2}\int_0^T \dot{h}_t^2 dt\right)\right)\right]& \\ = \sup_{x\in\mathbb{W}_T}\left\{2F(x) - \int_0^T \dot{h}_t dx_t + \frac{1}{2}\int_0^T \dot{h}_t^2 dt - I(x)\right\}.&\end{aligned} \tag{4.29}$$

Pero como sabemos, F puede tomar el valor extendido $-\infty$, y por el lema de Schilder, la función de tasa buena cumple que

$$-I(x) = -\infty \qquad \text{si } x \in \mathbb{W}_T \setminus \mathbb{H}_T.$$

Como nos interesa el supremo en $\mathbb{W}_T$, descartamos las $x \in \mathbb{W}_T \setminus \mathbb{H}_T$ y nos fijamos únicamente en $\mathbb{H}_T$, donde ocurre que $I(x) = \frac{1}{2}\int_0^T \dot{x_t}^2 dt$ si $x \in \mathbb{H}_T$. De esta manera, por la expresión (4.29) obtenemos que

$$\begin{aligned}\limsup_{\epsilon\downarrow 0} \epsilon \log E\left[\exp\left(\frac{1}{\epsilon}\left(2F(\sqrt{\epsilon}B) - \int_0^T \sqrt{\epsilon}\dot{h}_t dB_t + \frac{1}{2}\int_0^T \dot{h}_t^2 dt\right)\right)\right]& \\ = \sup_{x\in\mathbb{H}_T}\left\{2F(x) + \frac{1}{2}\int_0^T (\dot{x}_t - \dot{h}_t)^2 dt - \int_0^T \dot{x}_t^2 dt\right\}.&\end{aligned}$$

Para probar (ii), simplemente usemos $M = 1/2$ en el lema 4.3, y así obtenemos que existe una solución para (4.25). Usando el mismo lema, también existe el maximizador del problema (4.26) si usamos a la función constante $h := 0$ y $M = 1$.

[3] Por el teorema de cambio de variable, $\int_{\mathbb{W}_T} e^{F_h(x)/\epsilon} d\mu_\epsilon$ es como integrar sobre todo Ω, por lo cual es un valor esperado.

Para probar (iii), observamos que por la desigualdad de Cauchy-Schwarz, si X es una variable aleatoria entonces

$$E_P\left[e^{2X}\frac{dP}{dQ}\right]=E_Q\left[\left(e^X\frac{dP}{dQ}\right)^2\right]\geq E_Q\left[e^X\frac{dP}{dQ}\right]^2=(E_P[e^X])^2, \tag{4.30}$$

por lo cual si $X:=F(\sqrt{\epsilon}B)/\epsilon$, y

$$\frac{dP}{dQ^h}(\sqrt{\epsilon}B)=\exp\left(-\frac{1}{\sqrt{\epsilon}}\int_0^T \dot{h}_t dB_t+\frac{1}{2\epsilon}\int_0^T \dot{h}_t^2 dt\right),$$

entonces $L(h)$ como definido es $L(h)=\limsup_{\epsilon\to 0}\epsilon\log E_P\left[e^{2X}\frac{dP}{dQ^h}(\sqrt{\epsilon}B)\right]$. Pero por la desigualdad (4.30) tenemos que

$$L(h)\geq 2\limsup_{\epsilon\to 0}\epsilon\log E_P[e^{F(\sqrt{\epsilon}B)/\epsilon}]. \tag{4.31}$$

Como dijimos en la observación después de la prueba del lema 4.2, F también cumple las hipótesis para poder aplicar la extensión del lema de Varadhan dada en el teorema 3.3, por lo cual tenemos que

$$\limsup_{\epsilon\to 0}\epsilon\log E_P[e^{F(\sqrt{\epsilon}B)/\epsilon}]=\sup_{x\in\mathbb{H}_T}\{F(x)-I(x)\}=\sup_{x\in\mathbb{H}_T}\left\{F(x)-\frac{1}{2}\int_0^T \dot{x}_t^2 dt\right\}.$$

Sustituyendo esto en (4.31) obtenemos que

$$L(h)\geq\sup_{x\in\mathbb{H}_T}\left\{2F(x)-\int_0^T \dot{x}_t^2 dt\right\}.$$

Si $\hat{h}$ es una solución al problema (4.26) y cumple $2F(\hat{h})-\int_0^T \dot{\hat{h}}_t^2 dt=L(\hat{h})$, entonces tomando el ínfimo en $\mathbb{H}_T$ se sigue que

$$\inf_{h\in\mathbb{H}_T}L(h)\geq 2F(\hat{h})-\int_0^T \dot{\hat{h}}_t^2 dt=L(\hat{h}), \tag{4.32}$$

de donde se sigue que debe ser que $L(\hat{h})=\inf_{h\in\mathbb{H}_T}L(h)$, y $\hat{h}$ es asintóticamente óptimo.

Para verificar unicidad, supongamos que existen dos soluciones $\hat{h}$ y g a (4.26), y que $\hat{h}\neq g$. Entonces $\|g-\hat{h}\|^2_{\mathbb{H}_T}>0$, y por (i), tenemos que

$$\begin{aligned}L(\hat{h})&\geq 2F(g)+\frac{1}{2}\int_0^T(\dot{g}_t-\dot{\hat{h}}_t)^2dt-\int_0^T \dot{g}_t^2 dt\\&>2F(g)-\int_0^T \dot{g}_t^2 dt=2F(\hat{h})-\int_0^T \dot{\hat{h}}_t^2 dt=L(\hat{h}),\end{aligned}$$

y esto último contradice la optimalidad de $\hat{h}$ que es implicada por (4.32). De esta contradicción se sigue la unicidad.

□

OBSERVACIÓN: Es de notar que en la p. 18 del artículo de Guasoni y Robertson [20], en la prueba del inciso (*iii*) para probar que $\hat{h}$ es asintóticamente óptimo, comienzan con la expresión

$$\begin{aligned} L(h) &= \sup_{x\in\mathbb{H}_T}\left\{2F(x)+\frac{1}{2}\int_0^T(\dot{x}_t-\dot{h}_t)^2dt-\int_0^T\dot{x}_t^2dt\right\} \\ &\geq \sup_{x\in\mathbb{H}_T}\left\{2F(x)-\int_0^T\dot{x}_t^2dt\right\}. \end{aligned}$$

Pero para poder dar la primera igualdad, es necesario que la derivada $\dot{\hat{h}}$ sea de variación finita, por el inciso (*i*) de este teorema 4.3 que acabamos de probar, y era necesaria la hipótesis de que $\dot{h}$ sea de variación finita para lograr la igualdad del inciso (*i*), como puede verse en una revisión detallada del lema auxiliar 4.2. Pero los autores no dijeron como hipótesis del inciso (*iii*) que $\dot{\hat{h}}$ sea necesariamente de variación acotada. Una forma correcta de lograr la desigualdad

$$L(h)\geq \sup_{x\in\mathbb{H}_T}\left\{2F(x)-\int_0^T\dot{x}_t^2dt\right\}$$

es usar el principio de grandes desviaciones para F, como hicimos en la prueba de este teorema 4.3. De este modo, no es necesario que $\hat{h}$ tenga derivada de variación finita para lograr la desigualdad deseada. Usar el principio de grandes desviaciones para lograr la desigualdad de arriba, me fue sugerido por Scott Robertson, uno de los autores del artículo en el cual fue basado este trabajo.

Para probar la unicidad, se ve en la prueba del inciso (*iii*) cómo se necesitó que $\dot{\hat{h}}$ sea de variación acotada. Scott Robertson me comentó que en las aplicaciones, al resolver la ecuación de Euler-Lagrange para hallar a $\hat{h}$ óptima, *suele ser el caso que la derivada de* $\hat{h}$ *tiene variación finita.*

Conclusión del Problema 4.1

Para resolver el Problema 4.1 de minimizar a la cantidad de interés, buscamos $\hat{h}$ tal que resuelva el problema de optimización de la expresión (4.26). De acuerdo a P. Guasoni y S. Robertson en p. 3 y 7 de [20], se puede hallar $\hat{h}$ al resolver una ecuación diferencial de Euler-Lagrange. Por el inciso (*iii*) del teorema anterior 4.3, $\hat{h}$ es asintóticamente óptimo si satisface (4.27). En el caso que $\dot{\hat{h}}$ sea de variación finita, podemos usar el inciso (*i*) para evaluar a $L(\hat{h})$, que nuevamente, los autores P. Guasoni y S. Robertson indican que se puede resolver el lado derecho de (4.25) al resolver la ecuación de Euler-Lagrange asociada.

Como vemos, por los resultados de este capítulo, con $\hat{h}$ hemos minimizado a la funcional de interés $L(h)$ del Problema 4.1.

Apéndice A

Apéndice

A.1. Preliminares de Probabilidad

Con el objetivo de indicar la notación y las convenciones usadas en este trabajo, se da aquí un repaso de definiciones de probabilidad y procesos estocásticos, misma que puede ser vista en libros como el de los autores Jacod y Protter [18], los textos de P. Billingsley [2], [3] y el de I. Karatzas y S. Shreve [21].

A.1.1. Conceptos de Probabilidad

Los conceptos de probabilidad manejados en este caso, son en gran parte conceptos de teoría de la medida.

Definición A.1. *Sea Ω un conjunto dado. Una **σ-álgebra** $\mathcal{F}$ para Ω es una familia $\mathcal{F}$ de subconjuntos de Ω que cumple las siguientes propiedades:*

(i) $\Omega \in \mathcal{F}, \emptyset \in \mathcal{F}$.

(ii) *Si* $A \in \mathcal{F}$, *entonces* $A^c \in \mathcal{F}$.

(iii) *Si* $\{A_n\}_{n\in\mathbb{N}}$ *es una sucesión de conjuntos tal que* $A_n \in \mathcal{F}$ $\forall n \geq 1$, *entonces* $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

*A la pareja $(\Omega, \mathcal{F})$ se le llama un **espacio medible**, y a los subconjuntos F tales que $F \in \mathcal{F}$ se les conoce como conjuntos $\mathcal{F}$-medibles. Llamaremos también a los subconjuntos medibles de $\mathcal{F}$ **eventos**.*

Se le recuerda al lector que cuando se tiene dado una clase de subconjuntos $\mathcal{C}$ de Ω, existe la σ-álgebra más chica que contiene a $\mathcal{C}$, que denotaremos por $\sigma(\mathcal{C})$. Es la σ-álgebra más chica en el sentido que cualquier otra σ-álgebra que contenga a $\mathcal{C}$ contiene a $\sigma(\mathcal{C})$. Esta σ-álgebra, conocida como la **σ-álgebra generada** por $\mathcal{C}$, la obtenemos como

$$\sigma(\mathcal{C}) = \bigcap \{\mathcal{G} : \mathcal{G} \text{ es } \sigma\text{-álgebra para } \Omega,\ \mathcal{G} \supset \mathcal{C}\}.$$

En particular, cuando Ω es un espacio métrico, a la σ-álgebra generada por los conjuntos abiertos en la topología de Ω se le conoce como la **σ-álgebra de Borel de** Ω, denotada por $\mathcal{B}(\Omega)$. A los conjuntos $B \in \mathcal{B}(\Omega)$ se les conoce como *conjuntos borelianos*, o simplemente como borelianos. A continuación se da la definición de una medida de probabilidad.

Definición A.2. *Sea $(\Omega, \mathcal{F})$ un espacio medible. Una **medida de probabilidad** en $(\Omega, \mathcal{F})$ es una función conjuntista $P : \mathcal{F} \longrightarrow [0, 1]$ con las siguientes propiedades:*

(i) $P(\Omega) = 1$ *y* $P(\emptyset) = 0$.

(ii) Si $A_n \in \mathcal{F}$ para toda $n \geq 1$ y $A_i \cap A_j = \emptyset$ $\forall i \neq j$, entonces

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

*A la terna $(\Omega, \mathcal{F}, P)$ se le llama un **espacio de probabilidad.***

Dado que estamos trabajando con aspectos que se comparan con respecto a nuestra medida de probabilidad, es decir, el peso de un evento lo da nuestra medida P, entonces es suficiente garantizar una propiedad salvo conjuntos que bajo P son "despreciables".

Definición A.3. *Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad. Si N es un subconjunto de Ω tal que $N \subset F$ donde $F \in \mathcal{F}$ y $P(F) = 0$, entonces se dice que N es un **conjunto nulo** con respecto a P, o P-nulo. A la familia de conjuntos que cumplen esto se les denota por $\mathcal{N}$ y se les llaman los conjuntos P-**nulos**.*

Diremos que una propiedad se cumple **casi seguramente con respecto a** P (o P-c.s.) si se cumple en $\Omega \setminus N$, donde N es un conjunto P-nulo. Cuando en el contexto sea claro que nos referimos únicamente a la medida probabilidad P, diremos simplemente que la propiedad se cumple casi seguramente (c.s.). Nótese que la definición de $\mathcal{N}$ no necesariamente implica que $\mathcal{N} \subset \mathcal{F}$. Cuando un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ cumple que todo $N \in \mathcal{N}$ también está en $\mathcal{F}$, es decir, $\mathcal{N} \subset \mathcal{F}$, se dice que $(\Omega, \mathcal{F}, P)$ es un **espacio de probabilidad completo**.

Dado un espacio de probabilidad $(\Omega, \mathcal{F}, P)$, se puede completar la σ-álgebra con los conjuntos P-nulos para obtener una nueva σ-álgebra de la forma $\mathcal{F}' = \{F \cup N : F \in \mathcal{F}, N \in \mathcal{N}\}$ y extender la función de probabilidad P a esta nueva σ-álgebra, que llamaremos la P-completación de $\mathcal{F}$. Esta nueva σ-álgebra $\mathcal{F}'$ tiene la propiedad de ser la más chica que contiene a $\mathcal{F}$ y a $\mathcal{N}$, en el sentido de que cualquier otra σ-álgebra que contenga a $\mathcal{F}$ y a $\mathcal{N}$ contiene a la P-completación de $\mathcal{F}$. Véase p. 37 Teorema 6.4 de J. Jacod y P. Protter [18]. Usaremos seguido este resultado para garantizar que se puede completar un espacio de probabilidad dado.

Si se tienen dos espacios medibles $(\Omega, \mathcal{F})$ y $(\Gamma, \mathcal{E})$, recuérdese que una función $f : \Omega \longrightarrow \Gamma$ se dice que es *$\mathcal{F}/\mathcal{E}$-medible* si $f^{-1}(E) \in \mathcal{F}$ para todo $E \in \mathcal{E}$.[1] En el caso particular de $\Gamma = \mathbb{R}^d$ tenemos la siguiente definición.

Definición A.4. *Sea un espacio de probabilidad $(\Omega, \mathcal{F}, P)$. Un mapeo $X : \Omega \longrightarrow \mathbb{R}^d$ es una **variable aleatoria** si X es $\mathcal{F}/\mathcal{B}(\mathbb{R}^d)$-medible, donde $\mathcal{B}(\mathbb{R}^d)$ denota la σ-álgebra de Borel de $\mathbb{R}^d$. Esto se traduce a que $\{\omega \in \Omega : X(\omega) \in B\} = X^{-1}(B) \in \mathcal{F}$, para todo $B \in \mathcal{B}(\mathbb{R}^d)$.*

Si tenemos un espacio de probabilidad $(\Omega, \mathcal{F}, \mathrm{P})$, un espacio medible $(\Omega', \mathcal{F}')$ y una función medible $X : \Omega \longrightarrow \Omega'$, entonces podemos inducir una medida de probabilidad en Ω' dada por

$$(P \circ X^{-1})(B') = P\big(X^{-1}(B')\big) = P\big\{\omega \in \Omega : X(\omega) \in B'\big\} \qquad B' \in \mathcal{F}'.$$

Si tenemos una función $f : \Omega' \longrightarrow \mathbb{R}$ Borel-medible, entonces la función

$$f \circ X : \Omega \longrightarrow \mathbb{R}$$

también es Borel-medible con respecto a la σ-álgebra $\mathcal{F}$. A nosotros nos interesará cuando Ω' tiene la estructura de espacio métrico con su respectiva σ-álgebra de Borel, y a esta medida *inducida en el espacio donde toma valores la función medible X* se le conoce como la **distribución** de X o **la medida inducida por** X, o **la ley de probabilidad de** X, que algunos denotan por un subíndice P_X. Esto quiere decir que

$$P_X(B') = P\{X \in B'\} \qquad B' \in \mathcal{F}'.$$

Esta medida nos asigna pesos a eventos *desde la perspectiva de X*, es decir, desde el punto de vista de los valores posibles que puede tomar X.

En general el espacio de probabilidad $(\Omega, \mathcal{F}, P)$ es bastante arbitrario y un tanto abstracto, y en general nos interesa trabajar en contextos donde tenemos un poco más de estructura matemática que únicamente la de un espacio medible. En este trabajo, la estructura adicional en que nos concentraremos será el caso cuando Ω' es un espacio métrico, es decir cuando tenemos un mapeo medible $X : \Omega \longrightarrow E$, donde E es un espacio métrico. Este mapeo es como nuestro contacto con el espacio de probabilidad $(\Omega, \mathcal{F}, P)$, que usamos básicamente para dar un ambiente de aleatoriedad a nuestro espacio métrico E, al poder asignar probabilidades a los conjuntos borelianos $\mathcal{B}(E)$. En el trabajo a desarrollar, nos interesará seguido cuando $E = \mathbb{R}$ o incluso si E es el espacio de dimensión infinita $C[0, T]$, el espacio de trayectorias continuas en $[0, T]$. Es común decir que una variable aleatoria es una función medible con dominio un espacio de probabilidad e imagen un espacio métrico, por lo cual no se limita este concepto únicamente al de funciones con valores en $\mathbb{R}^d$.

El siguiente concepto es de suma importancia en probabilidad.

[1] Una condición suficiente y necesaria para que $f : \Omega \longrightarrow \Gamma$ sea medible es que $f^{-1}(\mathcal{C}) \subset \mathcal{F}$, donde $\mathcal{C}$ es un subconjunto de Γ tal que $\sigma(\mathcal{C}) = \mathcal{E}$ y $f^{-1}(\mathcal{C})$ denota a la familia $\{f^{-1}(C) : C \in \mathcal{C}\}$. Véase J. Jacod y P. Protter [18] p. 47 Teorema 8.1 para una demostración de esto.

Definición A.5. (Valor Esperado) *Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad, $(E, \mathcal{B}(E))$ un espacio métrico con su σ-álgebra de Borel y $X : \Omega \longrightarrow E$ una variable aleatoria. Si $f : E \longrightarrow \mathbb{R}$ es una función Borel-medible, el **valor esperado** o **esperanza matemática** de la variable aleatoria $f(X)$ es la integral*

$$E[f(X)] := \int_\Omega f(X(\omega))dP(\omega),$$

siempre que exista esta integral.

Algunos denotan por E_P al operador de valor esperado, para hacer énfasis en la medida P, pues es común calcular valores esperados bajo diferentes medidas de probabilidad. Es común decir simplemente *la esperanza de* $f(X)$. También se le dice *media* al valor esperado, pues nos da una idea de lo que ocurre "en promedio" con los valores de una variable aleatoria.

En general no calcularemos la integral $\int_\Omega f(X(\omega))dP(\omega)$ en Ω, y usaremos la estructura del espacio métrico para auxiliarnos en el cálculo de esta integral (como cuando X toma valores en $\mathbb{R}$), y mejor calculamos

$$\begin{aligned}
&\int_E f(x)dP\big(X^{-1}(x)\big) \\
&\qquad = \int_E f(x)dP_X(x) \\
&\qquad = \int_\Omega f(X(\omega))dP(\omega) = E[f(X)],
\end{aligned}$$

donde la segunda igualdad se da por el teorema de cambio de variable, que puede consultarse en el texto de P. Halmos [5] p. 163 teorema C:

Teorema A.1. (Teorema de Cambio de Variable) *Sea un espacio de probabilidad $(\Omega, \mathcal{F}, P)$, un espacio medible $(\Omega', \mathcal{F}')$ y un mapeo medible $X : \Omega \longrightarrow \Omega'$, entonces si $f : \Omega' \longrightarrow \mathbb{R}$ es Borel-medible, entonces para todo $B' \in \mathcal{F}'$:*

$$\int_{f^{-1}(B')} f(X(\omega))dP(\omega) = \int_{B'} f(\omega')dP\big(f^{-1}(\omega')\big),$$

donde la existencia de cualquiera de las integrales de arriba implica la existencia de la otra.

Cuando decimos que "un evento A es independiente de B", queremos decir que el hecho de que ocurra uno no implica que ocurre el otro, por lo cual, desde el punto de vista de probabilidad, la ocurrencia de uno no afecta en absoluto la ocurrencia del otro, por lo cual la probabilidad de que ocurra uno de los eventos, en cierto sentido no es afectada si ocurre o no ocurre el otro. Algo similar ocurre para dos variables aleatorias X y Y definidas en un mismo espacio de probabilidad. Intuitivamente, X es independiente de Y si los eventos $\{X \in A\}$ y $\{Y \in B\}$ son independientes para toda pareja de conjuntos A, B en el conjunto imagen de estas variables aleatorias. Nótese que X y Y pueden tomar valores en distintos espacios de medida.

Recordemos que si $X : (\Omega, \mathcal{F}) \longrightarrow (E, \mathcal{E})$ es medible, entonces

$$X^{-1}(\mathcal{E}) := \{X^{-1}(A) : A \in \mathcal{E}\}$$

es una sub σ-álgebra de $\mathcal{F}$, llamada la *σ-álgebra inducida* por X. Recalcamos que no es necesario que las variables aleatorias en cuestión tomen valores en los mismos espacios, pues las sub σ-álgebras que inducen estas variables aleatorias son inducidas en su dominio.

Definición A.6. *Sea $\{\mathcal{F}_i\}_{i\in I}$ una colección de sub σ-álgebras de $\mathcal{F}$. Diremos que $\{\mathcal{F}_i\}_{i\in I}$ es una colección de sub σ-álgebras* ***independientes*** *si para toda subcolección $\{A_i\}_{i\in J}$ de eventos tales que $A_i \in \mathcal{F}_i$, para cada $i \in J$, con $J \subset I$ un subconjunto de índices arbitrario, ocurre:*

$$P\left(\bigcap_{i\in J} A_i\right) = \prod_{i\in J} P(A_i).$$

Sea un conjunto de variables aleatorias $\{X_i\}_{i\in I}$ donde cada $X_i : (\Omega, \mathcal{F}) \longrightarrow (E, \mathcal{E}_i)$ toma valores en un espacio de medida $(E, \mathcal{E}_i)$. Diremos que las variables aleatorias $\{X_i\}_{i\in I}$ son independientes si las sub σ-álgebras $X_i^{-1}(\mathcal{E}_i)$ de $\mathcal{F}$ son independientes.

Las variables aleatorias independientes pueden ser más fáciles para manejar, en términos de calcular probabilidades y valores esperados, como podemos ver del siguiente resultado, cuya prueba se puede hallar en J. Jacod y P. Protter [18] capítulo 10 teorema 10.1.

Teorema A.2. *Sean el espacio de probabilidad $(\Omega, \mathcal{F}, P)$ y dos variables aleatorias definidas en él: $X : \Omega \longrightarrow (E, \mathcal{E})$ y $Y : \Omega \longrightarrow (S, \mathcal{S})$, donde $(E, \mathcal{E})$ y $(S, \mathcal{S})$ son espacios medibles. Entonces X y Y son independientes si y sólo si se cumple cualquiera de las condiciones siguientes:*

(i) $P\{X \in A, Y \in B\} = P\{X \in A\}P\{Y \in B\}$ *para todo $A \in \mathcal{E}$, $B \in \mathcal{S}$.*

(ii) Para todo par f, g de funciones medibles (en sus respectivos espacios) $f(X)$ es independiente de $g(Y)$.

(iii) Para todo par f, g de funciones medibles y acotadas, o medibles y positivas, se cumple que $E[f(X)g(Y)] = E[f(X)]E[g(Y)]$.

(iv) Sean $(E, \mathcal{E})$ y $(S, \mathcal{S})$ espacios métricos con sus respectivas σ-álgebras de Borel. Para todo par f, g de funciones continuas y acotadas se cumple que $E[f(X)g(Y)] = E[f(X)]E[g(Y)]$.

A.1.2. Esperanza Condicional

El concepto de esperanza condicional juega un papel importante en la teoría de procesos estocásticos.

Diremos que una función medible (o variable aleatoria) f definida en un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ que cumple una propiedad $\mathcal{P}$ es casi seguramente única (con respecto a la medida de probabilidad P) si existe la posibilidad de que haya otra función que también cumple la propiedad $\mathcal{P}$ y $f = g$ casi seguramente. Ahora se dará la noción de esperanza condicional.

Definición A.7. (Esperanza Condicional) *Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y X una variable aleatoria integrable, es decir $E[|X|] < \infty$. Sea $\mathcal{G}$ una sub σ-álgebra de $\mathcal{F}$. La esperanza condicional de X con respecto a $\mathcal{G}$ (o también conocida como la esperanza condicional de X dada $\mathcal{G}$) es la variable aleatoria $E[X|\mathcal{G}] : \Omega \longrightarrow \mathbb{R}$ c.s. única tal que*

(i) $E[X|\mathcal{G}]$ *es $\mathcal{G}$-medible.*

(ii) Para todo $H \in \mathcal{G}$ se cumple que

$$\int_H E(X|\mathcal{G})(\omega) dP(\omega) = \int_H X(\omega) dP(\omega),$$

es decir $E\big[E[X|\mathcal{G}]\mathbf{1}_H\big] = E[X\mathbf{1}_H]$ *para todo* $H \in \mathcal{G}$.

Es importante recordar siempre que la esperanza condicional con respecto a una sigma álgebra $\mathcal{G}$ es una variable aleatoria medible con respecto a $\mathcal{G}$. La existencia de la esperanza condicional se sigue del teorema de Radon-Nikodym.[2] Dado que la esperanza condicional de X dada $\mathcal{G}$ sólo es única salvo conjuntos de medida cero, entonces cuando trabajemos con $E[X|\mathcal{G}]$, diremos que tenemos *una versión* de esta esperanza condicional. Usaremos las siguientes propiedades de la esperanza condicional, la prueba de estas puede verse en D. Williams [14] secciones 9.7 y 9.8, p. 88-90.

Teorema A.3. (Propiedades de la esperanza condicional) *Sean X, Y variables aleatorias integrables y $\mathcal{G}$ una sub σ-álgebra de $\mathcal{F}$. Sea $E[X|\mathcal{G}]$ una versión de la esperanza condicional de X dada $\mathcal{G}$. La esperanza condicional cumple que:*

(i) Linealidad: $E[aX + bY|\mathcal{G}] = aE[X|\mathcal{G}] + bE[Y|\mathcal{G}]$, *con* $a, b \in \mathbb{R}$.

(ii) $E\big[E[X|\mathcal{G}]\big] = E[X]$

(iii) Si X es $\mathcal{G}$-medible, entonces $E[X|\mathcal{G}] = X$.

(iv) Si X es $\mathcal{G}$-medible, entonces $E[XY|\mathcal{G}] = XE[Y|\mathcal{G}]$.

(v) Si X es independiente de $\mathcal{G}$, entonces $E[X|\mathcal{G}] = E[X]$.

[2] Véase Halmos [5] p. 128. Para el caso particular donde la medida μ es una medida de probabilidad, véase Williams [14] p. 145-149.

La esperanza condicional puede verse intuitivamente como el resultado de calcular la esperanza *dado* que conocemos la información contenida en la σ-álgebra $\mathcal{G}$. Si X es cuadrado-integrable, la esperanza condicional incluso puede verse como una proyección de la variable aleatoria $X \in L^2(\mathcal{F})$ sobre el subespacio de Hilbert $L^2(\mathcal{G})$ de $L^2(\mathcal{F})$, es decir, $E[X|\mathcal{G}]$ es la mejor aproximación en $L^2(\mathcal{G})$ a X, donde recordemos que $E[X]$ es el valor que se espera que tome X, cuando sabemos *toda* la información posible, es decir, la contenida en $\mathcal{F}$. Véase el capítulo 23 de J. Jacod y P. Protter en [18] para más detalles sobre esta interpretación de la esperanza condicional en L^2. De este modo, las propiedades de la esperanza condicional son intuitivas. Por ejemplo, si X es $\mathcal{G}$-medible, entonces el inciso (*iii*) dice que toda la información que podríamos buscar de X está contenida en $\mathcal{G}$, por lo cual es claro que la mejor aproximación a X dado que sabemos $\mathcal{G}$, es X misma. En términos de proyecciones en L^2, como $X \in L^2(\mathcal{G})$, entonces al proyectar sobre el espacio donde "vive" X, obtenemos a X nuevamente. Algo similar puede razonarse con las demás propiedades.

A.1.3. Procesos Estocásticos

Un papel central en este trabajo lo juegan los procesos estocásticos, donde:

Definición A.8. *Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad. Un proceso estocástico en tiempo continuo*

$$X = \{X_t\}_{t \geq 0}$$

definido en el espacio de probabilidad dado es un conjunto de variables aleatorias indizadas por $t \in [0, \infty)$ de la forma $X_t : \Omega \longrightarrow \mathbb{R}$.

Dotaremos siempre a $\mathbb{R}$ de su σ-álgebra de Borel $\mathcal{B}(\mathbb{R})$. El proceso $X = \{X_t\}_t$ también puede ser visto como un mapeo $X : \Omega \times [0, \infty) \longrightarrow \mathbb{R}$ tal que

$$X_t := X(\cdot, t) : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

es $\mathcal{F}$-medible para cada $t \geq 0$, y t tiene una interpretación útil como el tiempo, por lo cual podemos pensar en un proceso estocástico como la evolución de un sistema aleatorio en el tiempo.

En este trabajo, como es obviamente lo práctico, nos fijaremos en intervalos finitos $[0, T]$. Recordemos que una *trayectoria* del proceso X está definida ω por ω de tal manera que $t \mapsto X_t(\omega) \in \mathbb{R}$ es una función, o trayectoria. Siempre asumiremos en este trabajo que los procesos estocásticos toman valores en $\mathbb{R}$ a menos que se indique lo contrario. Muchos autores identifican a $\omega \in \Omega$ con una función en el producto cartesiano $\mathbb{R}^{[0,\infty)}$. Cuando las trayectorias son casi seguramente continuas, es decir, que

$$P\Big\{\omega : t \mapsto X_t(\omega) \quad \text{es continuo}\Big\} = 1, \tag{A.1}$$

podemos decir que escoger una $\omega \in \Omega$ es como escoger una trayectoria en $C[0, T] \subset \mathbb{R}^{[0,\infty)}$. Existe una manera formal de definir medidas en $\mathbb{R}^{[0,\infty)}$ y en $C[0, T]$, de tal manera que podemos ver a un proceso estocástico

$$X = \{X_t\}_{t \geq 0}$$

como *una variable aleatoria* con valores en $\mathbb{R}^{[0,\infty)}$, o $C[0,T]$, si es el caso. Esto es una manera elegante y concisa de mejor tratar a todas las variables aleatorias X_t con $t \in [0,\infty)$ (o un subintervalo $[0,T]$ de $\mathbb{R}$) como *una sola* variable aleatoria. No abordaremos demasiado en el procedimiento de definir estas medidas, pero esto puede consultarse en el texto de P. Billinsgley [3] capítulo 7, o I. Karatzas y S. Shreve [21] secciones 2.1-2.5, donde estos últimos autores consideran el caso particular de proceso estocástico que nos interesará en este trabajo.

Como dijimos, podemos pensar en un proceso estocástico como la evolución de un sistema aleatorio, y como sabemos de nuestra experiencia en el mundo de los hechos, conforme evoluciona un sistema acumulamos información. Por ejemplo, si tenemos una serie de juegos que dependen del resultado de echar un volado con una moneda, supongamos que después de n juegos tenemos los resultados $X_1, X_2, \ldots, X_n$ de estos juegos, que toman valores reales. Entonces al momento n hemos acumulado información de cada juego. Nótese que $X_i^{-1}(A) = \{\omega \in \Omega : X_i(\omega) \in A\}$, es decir, este conjunto indica el subconjunto de Ω que tal que para ω en este subconjunto $X_i(\omega)$ termine en A. Intuitivamente, la información que tenemos despues de n juegos, nos lleva a pensar en

$$\mathcal{F}_n := \sigma(X_1, X_2, \ldots, X_n),$$

la σ-álgebra generada por las variables aleatorias $X_1, X_2, \ldots, X_n$, pues recordemos que esta σ-álgebra $\mathcal{F}_n$ es generada por los eventos de la forma $\{X_i^{-1}(A) : A \in \mathcal{B}(\mathbb{R})\}$, $i = 1, 2, \ldots, n$, por lo cual nos proporciona una idea de "información hasta el juego n".
La definición formal que viene de este ejemplo y nos auxilia en el estudio de los procesos estocásticos es:

Definición A.9. (Filtración) *Sea un espacio de probabilidad $(\Omega, \mathcal{F}, P)$. Una* ***filtración*** *$\{\mathcal{F}_t\}_{t \in [0,T]}$ es una familia de sub σ-álgebras de $\mathcal{F}$ tales que $s < t$ implica que $\mathcal{F}_s \subset \mathcal{F}_t$.*

Si tenemos un proceso estocástico $X = \{X_t\}_{t \geq 0}$ definido en el espacio de probabilidad $(\Omega, \mathcal{F}, P)$ entonces este proceso define para cada $t \geq 0$ una σ-álgebra

$$\mathcal{F}_t^X := \sigma(X_s : 0 \leq s \leq t)$$

que denota a la σ-álgebra más pequeña tal que el conjunto de variables aleatorias $\{X_s : s \leq t\}$ es medible. El conjunto de σ-álgebras definidas de este modo nos da una filtración $\{\mathcal{F}_t^X\}_{t \geq 0}$, y se le llama la ***filtración natural generada por el proceso estocástico*** *$X = \{X_t\}$.*

A un espacio de probabilidad con una filtración asociada, se le conoce como *espacio de probabilidad filtrado*. Decimos que un proceso estocástico $\{X_t\}$ en un espacio de probabilidad filtrado tal que X_t es $\mathcal{F}_t$-medible para cada $t \geq 0$ (denotado seguido $X_t \in \mathcal{F}_t$)

es un *proceso adaptado* a la filtración $\{\mathcal{F}_t\}$.

El concepto de *martingala* es muy importante en procesos estocásticos. Jugará un papel implícitamente importante en este trabajo, pues casi no usaremos explícitamente las propiedades que normalmente se explotan de las martingalas en el área de procesos aleatorios. De todos modos damos su definición.

Definición A.10. (Martingala) *Sea un espacio de probabilidad* $(\Omega, \mathcal{F}, P)$ *con una filtración* $\{\mathcal{F}_t\}_{t\geq 0}$ *y un proceso estocástico* $\{X_t\}_{t\geq 0}$ *adaptado a esta filtración. Decimos que* $\{X_t\}$ *es* ***martingala*** *si cumple que*

1. $E(|X_t|) < \infty$ *para todo* $t \geq 0$.
2. $E[X_t|\mathcal{F}_s] = X_s$ *para todo* $0 \leq s \leq t$.

A.1.4. Movimiento Browniano

El movimiento browniano es quizás el proceso estocástico más estudiado, y en el contexto de muchas aplicaciones, uno de los procesos estocásticos más importantes.

En 1828, el botánico escocés Robert Brown observó que partículas de polen suspendidas en un fluido realizaban un movimiento irregular, es decir, sus trayectorias eran un tanto erráticas. Este fenómeno después fue explicado por Albert Einstein en 1905 en términos de las colisiones aleatorias entre las moléculas del fluido en que estaban suspendidas estas partículas. El físico Jean Baptiste Perrin confirmó experimentalmente la explicación de A. Einstein, y con ello se confirmó la teoría atómica de la materia, pues no se tenía aún suficiente evidencia de que la materia estaba constituida por átomos. J. B. Perrin recibió el premio Nobel de física en 1926 por estos trabajos.

Por otra parte, se acepta en general que las matemáticas financieras comenzaron en 1900 con la tesis del matemático francés Louis Bachelier, tesis titulada *Théorie de la spéculation*. En dicho trabajo, Bachelier propuso la modelación de la evolución (o más bien, de la fluctuación) de precios de activos riesgosos por medio del movimiento browniano. Bachelier es considerado el primero en realizar trabajos cuantitativos con este modelo estocástico. Aunque actualmente se sabe que estas fluctuaciones de precios pueden ser discontinuos, el movimiento browniano se sigue empleando por ser una buena aproximación.

La primera persona en dar una descripción matemáticamente rigurosa del movimiento browniano, y de probar la existencia del movimiento browniano como un objeto matemático (y con ello, probar que no es sólo un fenómeno para el cual se pueden dar diferentes modelos que lo aproximan), fue el matemático Nobert Wiener en 1923.

Daremos aquí la descripción matemática de este proceso estocástico. Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad. Recordemos que decimos que una variable aleatoria $X : \Omega \longrightarrow \mathbb{R}$

tiene *distribución normal* o *gaussiana* de parámetros μ, σ si su función de densidad es de la forma

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

Denotamos la dependencia de estos parámetros μ y σ al escribir $X \sim N(\mu, \sigma^2)$, y con esto entenderemos que X tiene una distribución normal de parámetros μ y σ.

La *función característica* $\Psi(\theta) := E[e^{i\theta X}]$ de una v.a. $X \sim N(\mu, \sigma^2)$ es

$$E[e^{i\theta X}] = \exp\left(i\mu\theta - \frac{1}{2}\sigma^2\theta^2\right)$$

y su *función generadora de momentos* $\psi(\theta) = E[e^{\theta X}]$ es

$$E[e^{\theta X}] = \exp\left(\mu\theta + \frac{1}{2}\sigma^2\theta^2\right).$$

Si X es una variable aleatoria normal $N(\mu, \sigma^2)$, entonces su valor esperado es $E[X] = \mu$ y su varianza es $\mathrm{Var}(X) = \sigma^2$.

A continuación presentamos la definición del movimiento browniano.

Definición A.11. (Movimiento browniano) *Sean $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y un proceso estocástico en tiempo continuo $\{B_t\}_{t\geq 0}$ definido en Ω. Decimos que $\{B_t\}_{t\geq 0}$ es un* ***movimiento browniano*** *si cumple las siguientes propiedades:*

(i) $B_0 = 0$ con probabilidad 1.

(ii) (Incrementos independientes) Para cualquier cantidad finita de tiempos $0 \leq t_1 < t_2 < \ldots < t_n < \infty$ las variables aleatorias $B_{t_2} - B_{t_1}, B_{t_3} - B_{t_2}, \ldots, B_{t_n} - B_{t_{n-1}}$ son independientes.

(iii) (Incrementos estacionarios) Para todo $0 \leq s \leq t < \infty$ el incremento $B_t - B_s$ tiene distribución normal $N(0, t-s)$.

(iv) (Continuidad de las trayectorias) Con probabilidad 1 las trayectorias $t \mapsto B_t(\omega)$ son continuas.

La existencia y construcción del movimiento browniano se puede probar ya sea usando el Teorema de Extensión de Kolmogorov o usando una construcción por medio de *onduletes*, ambos métodos se ilustran en I. Karatzas y S. Shreve [21] Capítulo 2.

Supongamos que tenemos un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ y $\{B_t\}_{t\geq 0}$ un movimiento browniano definido en él. Si $\mathcal{N}$ denota la clase de conjuntos nulos de la σ-álgebra $\mathcal{F}$ con respecto a P y $\{\mathcal{F}_t^B\}$ con $\mathcal{F}_t^B = \sigma(B_s : s \leq t)$ es la filtración natural generada por el movimiento browniano $\{B_t\}_{t\geq 0}$, entonces definiendo

$$\mathcal{F}_t = \sigma(\mathcal{F}_t^B, \mathcal{N}) \text{ para cada } t \geq 0, \tag{A.2}$$

obtenemos una filtración $\{\mathcal{F}_t\}_{t\geq 0}$ para nuestro espacio de probabilidad $(\Omega, \mathcal{F}, P)$, que denotaremos en este trabajo como la *filtración asociada al movimiento browniano*. **En este trabajo, siempre usaremos la filtración asociada al movimiento browniano, cuando hablemos de este proceso estocástico.**

Observación. Por la propiedad (ii) de la definición de movimiento browniano, se sigue que para todo $0 \leq s \leq t < \infty$, el incremento $B_t - B_s$ es independiente de $\mathcal{F}_s$, denotado $B_t - B_s \perp \mathcal{F}_s$. Este hecho se aplicará cuando se trabaje con martingalas relacionadas a movimientos brownianos.

Definición A.12. *Sean $t \in [0, \infty)$ fijo, $\Pi = \{0 = t_0 < t_1 < \ldots < t_n = t\}$ una partición de $[0, t]$ y $\omega \in \Omega$ fijo. Tomando $p > 1$ fijo, se define la **p-variación de la trayectoria** $s \mapsto X_s(\omega)$ en $[0, t]$ con respecto a la partición Π como el número*

$$V_t^{\Pi,p}\big[X(\omega)\big] = \sum_{k=1}^{n} |X_{t_k}(\omega) - X_{t_{k-1}}(\omega)|^p$$

*Se dice que la trayectoria $X(\omega)$ es de **p-variación finita** en $[0, t]$ si*

$$\lim_{\|\Pi\|\to 0} V_t^{\Pi,p}\big[X(\omega)\big] < \infty.$$

En particular, tomando $p = 1$ en la definición anterior obtenemos la *variación de la trayectoria* en $[0, t]$ y si $p = 2$ denotaremos a esta 2-variación como la *variación cuadrática* en el intervalo $[0, t]$. En general tomaremos diferentes modos de convergencia como la convergencia en probabilidad o la convergencia en L^2 cuando hablemos de estos procesos de variación.

Para un movimiento browniano, si $X := B_t \sim N(0, t)$, entonces al calcular sus primeros cuatro momentos tenemos que si $E[e^{\theta X}] = \exp(\frac{1}{2}\theta^2 t)$ entonces:

$$\begin{aligned}
\frac{d}{d\theta}E[e^{\theta X}] &= \frac{d}{d\theta}e^{\frac{1}{2}\theta^2 t} = \theta t e^{\frac{1}{2}\theta^2 t}.\\
\frac{d^2}{d\theta^2}E[e^{\theta X}] &= e^{\frac{1}{2}\theta^2 t}(\theta^2 t^2 + t).\\
\frac{d^3}{d\theta^3}E[e^{\theta X}] &= e^{\frac{1}{2}\theta^2 t}(3\theta t^2 + \theta^3 t^3).
\end{aligned}$$

de donde se sigue que

$$\frac{d^4}{d\theta^4}E[e^{\theta X}] = E\left[\frac{d^4}{d\theta^4}e^{\theta X}\right] = e^{\frac{1}{2}\theta^2 t}(3t^2 + 6\theta^2 t^3 + \theta^4 t^4)$$

por lo cual evaluando en $\theta = 0$ se obtiene que

$$E[X^4] = 3t^2. \tag{A.3}$$

Usaremos el cuarto momento del movimiento browniano en el siguiente resultado.

Teorema A.4. *La variación cuadrática de un movimiento browniano en un intervalo de tiempo finito* $[0,t]$ *converge en* $L^2(\Omega)$ *a la longitud del intervalo de tiempo, es decir, para cualquier sucesión* $\{\Pi_n\}_n$ *de particiones de* $[0,t]$ *tales que* $\|\Pi_n\| \to 0$ *cuando* $n \to \infty$, *entonces*

$$V_t^{\Pi_n,2}(B) \to t \quad \text{en } L^2(\Omega) \quad \text{si } n \to \infty.$$

Demostración. Para ahorrar un poco de espacio denotaremos por ΔB_i a la diferencia $B_{t_i} - B_{t_{i-1}}$ y $\Delta t_i = t_i - t_{i-1}$. Usaremos que $\sum_{i=1}^n \Delta t_i = t$.

Si $Z_n := \sum_{i=1}^n (B_{t_i} - B_{t_{i-1}})^2 - t = \sum_{i=1}^n (\Delta B_i^2 - \Delta t_i)$, entonces queremos probar que $Z_n \to 0$ en $L^2(dP)$, es decir, que $E[Z_n^2] \to 0$.

Pero

$$\begin{aligned} E[Z_n^2] =& E\left[\left(\sum_{i=1}^n (\Delta B_i^2 - \Delta t_i)\right)^2\right] \\ =& E\left[\sum_{i=1}^n (\Delta B_i^2 - \Delta t_i)^2 + 2\sum_{i<j} (\Delta B_i^2 - \Delta t_i)(\Delta B_j^2 - \Delta t_j)\right]. \end{aligned} \tag{A.4}$$

Sea la filtración $(\mathcal{F}_t)_{t\geq 0}$ definida como en (A.2). Para la última suma de (A.4), usando la propiedad (ii) de la esperanza condicional $E\big[E[X|\mathcal{F}]\big] = E[X]$ se sigue que

$$E[(\Delta B_i^2 - \Delta t_i)(\Delta B_j^2 - \Delta t_j)] = E\left[E[(\Delta B_i^2 - \Delta t_i)(\Delta B_j^2 - \Delta t_j)|\mathcal{F}_{t_{j-1}}]\right]$$

pero si $i < j$ entonces $(\Delta B_i^2 - \Delta t_i)$ es $\mathcal{F}_{t_{j-1}}$-medible, por lo cual, usando la propiedad (iv) de la esperanza condicional, esto último es igual a

$$E\left[(\Delta B_i^2 - \Delta t_i)E[(\Delta B_j^2 - \Delta t_j)|\mathcal{F}_{t_{j-1}}]\right],$$

y dado que $(B_{t_j} - B_{t_{j-1}}) \perp \mathcal{F}_{t_{j-1}}$ (por incrementos independientes) entonces se sigue de la propiedad (v) de la esperanza condicional que

$$E[(\Delta B_j^2 - \Delta t_j)|\mathcal{F}_{t_{j-1}}] = E[\Delta B_j^2 - \Delta t_j] = 0,$$

puesto que $\Delta B_j \sim N(0, \Delta t_j)$. Por esto el último sumando de la expresión (A.4) es cero.

Si nos fijamos en los términos del primer sumando de la segunda igualdad de (A.4), obtenemos que

$$\begin{aligned} E[(\Delta B_i^2 - \Delta t_i)^2] =& E[\Delta B_i^4 - 2\Delta t_i \Delta B_i^2 + \Delta t_i^2] \\ =& 3\Delta t_i^2 - 2\Delta t_i \cdot \Delta t_i + \Delta t_i^2 = 2\Delta t_i^2, \end{aligned} \tag{A.5}$$

donde en la última igualdad hemos utilizado que $E[\Delta B_i^4] = 3\Delta t_i^2$, como habíamos obtenido anteriormente en Ec. (A.3). Insertando la Ec. (A.5) en (A.4) tenemos que

$$\begin{aligned} E[Z_n^2] = 2\sum_{i=1}^{n} \Delta t_i^2 \leq 2\Big(\max_{1\leq i\leq n} \Delta t_i\Big) \sum_{i=1}^{n} \Delta t_i \\ =2t\|\Pi_n\| \overset{n\to\infty}{\to} 0. \end{aligned}$$

De esta forma se concluye que $V_t^{\Pi_n,2}(B) \to t$ en $L^2(\Omega)$, si $n \to \infty$. □

Con base en este último resultado, probaremos ahora que la variación del movimiento browniano en un intervalo finito $[0, t]$ es infinita con probabilidad uno.

Teorema A.5. *Sea $t \geq 0$ finito. La variación de un movimiento browniano en el intervalo $[0, t]$ es infinita casi seguramente.*

Demostración. Sea cualquier sucesión $\{\Pi_n\}_n$ de particiones de $[0, t]$ tales que $\|\Pi_n\| \to 0$ cuando $n \to \infty$. Formamos las sumas

$$V_t^{\Pi_n,2}(B) = \sum_{t_{k-1},t_k\in\Pi_n} |B_{t_k} - B_{t_{k-1}}|^2.$$

Por el Teorema A.4 recién probado, tenemos que $V_t^{\Pi_n,2}(B) \to t$ en $L^2(\Omega)$. La convergencia en $L^2(\Omega)$ implica convergencia en probabilidad de la sucesión

$$\{V_t^{\Pi_n,2}(B)\}_{n\geq 1}$$

y esta convergencia implica la convergencia casi segura de una subsucesión de $\{V_t^{\Pi_n,2}(B)\}_{n\geq 1}$.[3] Esto significa que existe una subsucesión $\{\Pi_{n_k}\}_k$ de particiones de $[0, t]$ tales que

$$\sum_{t_{j-1},t_j\in\Pi_{n_k}} |B_{t_j} - B_{t_{j-1}}|^2 \to t$$

con probabilidad uno. Pero

$$\sum_{t_{j-1},t_j\in\Pi_{n_k}} |B_{t_j} - B_{t_{j-1}}|^2 \leq \Big(\max_j |B_{t_j} - B_{t_{j-1}}|\Big) \sum_{t_{j-1},t_j} |B_{t_j} - B_{t_{j-1}}|. \tag{A.6}$$

Dado que conforme $n_k \to \infty$ tenemos que $\|\Pi_{n_k}\| \to 0$, entonces $\max_j |B_{t_j} - B_{t_{j-1}}| \to 0$ porque el movimiento browniano es c.s. continuo en su variable temporal, por lo cual en intervalos compactos es uniformemente continuo.

Se sigue que si hacemos $n_k \to \infty$ en Ec. (A.6), tendríamos que con probabilidad uno:

$$t \leq 0 \cdot \left(\lim_{n_k\to\infty} \sum_{t_{j-1},t_j\in\Pi_{n_k}} |B_{t_j} - B_{t_{j-1}}| \right),$$

por lo cual debe ser que $\lim_{n_k\to\infty} \sum_{t_{j-1},t_j\in\Pi_{n_k}} |B_{t_j} - B_{t_{j-1}}| = \infty$ casi seguramente. □

[3] Estos resultados son estándares de teoría de la medida, y pueden consultarse, por ejemplo, en p. 144-145 de J. Jacod y P. Protter en [18].

A.1.5. Resultados de Probabilidad

El siguiente resultado que se utiliza en el trabajo se prueba en el texto de P. Billingsley [3] p. 282.

Lema A.1. *SI X es una variable aleatoria no negativa, entonces*

$$E[X] = \int_0^\infty P(X \geq x)dx.$$

El resultado a continuación se prueba en p. 79-80 del texto de I. Karatzas y S. Shreve [21].

Lema A.2. (El principio de reflexión) *Sea $\{B_t\}$ un movimiento browniano estándar y $b > 0$. Entonces para cualquier $T > 0$ se cumple que*

$$P\Big\{ \sup_{t\in[0,T]} B_t \geq b \Big\} = 2P\{B_T > b\} = \sqrt{\frac{2}{\pi T}} \int_b^\infty e^{-w^2/2T} dw.$$

Finalmente damos el siguiente resultado, consecuencia sencilla de la famosa desigualdad de Markov: $P\{X \geq \epsilon\} \leq E[X]/\epsilon$ para $X \geq 0$.

Lema A.3. (Desigualdad Exponencial de Chebyshev) *Para todo $\lambda > 0$ entonces $P\{X > \lambda\} \leq e^{-a\lambda} E[e^{aX}]$ para todo $a > 0$ tal que esté definido $E[e^{aX}]$.*

Demostración. Si f es no-decreciente positiva y $E[f(X)] < \infty$, entonces $X \geq \epsilon$ si y sólo si $f(X) \geq f(\epsilon)$, de tal manera que

$$P\{X \geq \epsilon\} = P\{f(X) \geq f(\epsilon)\} \leq E[f(X)]/f(\epsilon).$$

Usando $f(x) = e^{ax}$ para $a > 0$ en esta última desigualdad, se sigue el resultado para toda $a > 0$ tal que $E[e^{aX}] < \infty$. □

A.2. Algunos conceptos de Estadística.

Se hablará superficialmente del método de Monte Carlo para el cálculo de integrales, o en palabras de la probabilidad, del cálculo de un valor esperado. No se abordará demasiado en los métodos de Monte Carlo, pero hablar un poco de en qué consiste este método nos explicará por qué se usa la expresión (1.1) para aproximar a $E_P[G]$ en el problema a resolver en esta tesis.

La idea básica se ve en el siguiente ejemplo. Supongamos que queremos obtener la integral de una función G integrable en el intervalo $[0, 1]$. Por definición de valor esperado de una variable aleatoria uniforme, esto lo podemos representar como el valor esperado de $G(X)$, donde X es una variable aleatoria uniforme en el intervalo $[0, 1]$, es decir:

$$\mu = E[G(X)] = \int_0^1 G(x)dx.$$

Sea una medida de probabilidad p definida en $\mathbb{R}$, es decir, una medida positiva cuya integral en $\mathbb{R}$ es uno. Recordemos que una *muestra aleatoria* de tamaño $n \in \mathbb{N}$ de una población p es un conjunto de variables aleatorias $X_1, X_2, \ldots, X_n$ independientes e idénticamente distribuidas de acuerdo a la distribución p. Intuitivamente, esto es lo análogo a suponer que una variable aleatoria en teoría se distribuye de acuerdo a la ley de probabilidad p, y aleatoriamente escogemos n eventos (o hacemos n experimentos) $X_1, X_2, \ldots, X_n$, donde cada X_i se obtiene de acuerdo a la ley de probabilidad p. El proceso de experimentar n veces y almacenar los resultados $X_1, X_2, \ldots, X_n$ es la *obtención de una muestra.* Un *estimador* es cualquier función $F(X_1, \ldots, X_n)$ de una muestra. Por ejemplo, la media aritmética $\hat{\mu}_n := \frac{1}{n}\sum_{i=1}^{n} X_i$ es un estimador.

Citamos la ley fuerte de grandes números, cuya prueba puede verse en p. 119 de D. Williams [14], teorema 12.10.

Teorema A.6. (La ley fuerte de grandes números) *Sea $X_1, X_2, \ldots, X_n$ un conjunto de variables aleatorias idénticamente distribuidas tales que $E[|X_i|] < \infty$ para toda i. Si $E[X_i] = \mu$, entonces*

$$\hat{\mu}_n := \frac{1}{n}\sum_{i=1}^{n} X_i \longrightarrow \mu \quad \text{casi seguramente, conforme } \ n \to \infty.$$

La independencia de $X_1, X_2, \ldots, X_n$ implica la independencia de las n variables aleatorias $G(X_1), G(X_2), \ldots, G(X_n)$. Nótese que cada $G(X_i)$ cumple que $E[G(X_i)] = E[G(X)] = \mu$. Si calculamos la media de $G(X_1), G(X_2), \ldots, G(X_n)$

$$\hat{\mu}_n := \frac{1}{n}\sum_{i=1}^{n} G(X_i), \tag{A.7}$$

la integrabilidad de G en $[0,1]$ implica, por la ley fuerte de grandes números que

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} G(X_i) = \mu = \int_0^1 G(x)dx \quad \text{casi seguramente.} \tag{A.8}$$

Esto quiere decir que para aproximar a la integral de G en $[0,1]$, podemos obtener muestras aleatorias suficientemente grandes $X_1, X_2, \ldots, X_n$ (por ejemplo, usando simulación para obtener números pseudoaleatorios) y calculamos su media $\hat{\mu}_n$. A este método se le conoce como el método de Monte Carlo para el cálculo de integrales, nombrado así porque supuestamente el tío de Stanislaw Ulam, uno de los pioneros de este método, apostaba dinero pidiendo prestado en un casino en Monte Carlo.[4] En general, se conoce como método de Monte Carlo a una amplia gama de métodos computacionales que dependen del uso repetido de toma de muestras aleatorias para obtener un valor numérico. Estos métodos son muy útiles cuando no se conoce un método exacto para calcular cantidades de interés.

[4] Según el físico Nicholas Metropolis en el artículo [9].

Con frecuencia, al tener un estimador de un parámetro, como lo es el estimador $\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ del valor esperado $\mu = E[X]$ (donde cada X_i se distribuye como X), queremos una medida de qué tan buen estimador es. Esto nos lleva a la siguiente definición.

Definición A.13. (Estimador Insesgado) *Decimos que un estimador F del parámetro θ es **insesgado** si $E[F] = \theta$.*

El estimador $\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ del parámetro $\mu = E[X]$ es insesgado, pues

$$E[\hat{\mu}_n] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n} nE[X] = E[X].$$

Recordemos que la varianza de una variable aleatoria X es

$$\sigma^2 = \mathrm{Var}(X) := E\left[(X - E[X])^2\right] = E[X^2] - (E[X])^2\,,$$

y nos da una medida de qué tan lejos está una variable aleatoria de su valor esperado. Un estimador comúnmente usado para la varianza σ^2 de una variable aleatoria involucra al estimador para la media $\hat{\mu}_n$, y es

$$\hat{\sigma}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \hat{\mu}_n)^2,$$

y este estimador también es insesgado.

Se puede ir aún más lejos y buscar intervalos de $\mathbb{R}$ como estimadores para un parámetro θ, es decir, conjuntos de potenciales candidatos para θ.

Definición A.14. (Intervalo de Confianza) *Un intervalo de confianza para un parámetro $\theta \in \mathbb{R}$ es el conjunto determinado por un par de estimadores L y U tales que $L(X_1, \ldots, X_n) \leq U(X_1, \ldots, X_n)$ para toda muestra $(X_1, \ldots, X_n) = \mathbf{X}$. El intervalo aleatorio $(L(\mathbf{X}), U(\mathbf{X}))$ es denominado **intervalo de confianza** o **estimador de intervalo**, y es el conjunto de candidatos para hallar al parámetro θ.*

La importancia del concepto de intervalo de confianza, es que en estos estimadores buscaremos candidatos para el valor esperado μ de una variable aleatoria.

A.3. Semicontinuidad y equicontinuidad

En este trabajo, si X es un espacio topológico, una *vecindad* de un punto $x_0 \in X$ es un conjunto V tal que x_0 es un punto interior de V, es decir, existe un abierto $A \subset V$ tal que $x_0 \in A$. Para este trabajo, no se ocupará usar de la generalidad de un espacio topológico, pues en particular para el tema de interés a tratar, X será un espacio métrico, por lo cual los conjuntos abiertos estarán caracterizados por la métrica de X.

Recordemos que una función con valores reales definida en un espacio métrico, $f : X \longrightarrow \mathbb{R}$ es continua en x_0 si para todo $\epsilon > 0$ existe $\delta > 0$ tal que si $d(x, x_0) < \delta$, entonces

$$|f(x) - f(x_0)| < \epsilon,$$

es decir

$$f(x_0) - \epsilon < f(x) < f(x_0) + \epsilon.$$

A veces es necesario relajar alguna de las desigualdades anteriores y pedir que f solamente cumpla la primera o la segunda de las desigualdades de la expresión anterior. Esto motiva los siguientes conceptos.

Definición A.15. *Sea X un espacio topológico y f una función $f : X \longrightarrow \mathbb{R} \cup \{\pm\infty\}$. Sea $x_0 \in X$ un punto tal que $f(x_0) < \infty$. Diremos que f es **semicontinua inferiormente** en x_0 si para toda $\epsilon > 0$ existe una vecindad V_{x_0} de x_0 tal para todo $x \in V_{x_0}$ se cumple la desigualdad*

$$f(x) > f(x_0) - \epsilon.$$

*Para x_0 tal que $f(x_0) = \infty$, pediremos que se cumpla que $\lim_{x \to x_0} f(x) = \infty$, es decir que f tome valores arbitrariamente grandes al acercarse a x_0. Análogamente, f es **semicontinua superiormente** en x_0 tal que $f(x_0) > -\infty$ si para toda $\epsilon > 0$ existe una vecindad de x_0 tal que*

$$f(x) < f(x_0) + \epsilon$$

para toda x en dicha vecindad. Para x_0 tal que $f(x_0) = -\infty$, pediremos que se cumpla que $\lim_{x \to x_0} f(x) = -\infty$. Diremos que f es semicontinua inferiormente (superiormente) si es semicontinua inferiormente (superiormente) en todo punto de su dominio.

Probaremos a continuación un resultado que nos proporciona una equivalencia entre esta definición de semicontinuidad inferior y otra definición que se suele dar para semicontinuidad inferior.

Teorema A.7. *Sea X un espacio métrico[5] y f una función $f : X \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ definida en X. Entonces f es semicontinua inferiormente si y sólo si para todo $a \in \mathbb{R}$ se cumple que los conjuntos de nivel*

$$\{x \in X : f(x) \leq a\}$$

son cerrados en X.

Demostración. $\Rightarrow$ Usando la definición de semicontinuidad dada, probemos que si f es semicontinua en X, entonces para cualquier $a \in \mathbb{R}$ el conjunto de nivel $\mathfrak{N}_a := \{x \in X : f(x) \leq a\}$ es cerrado, o equivalentemente, basta probar que el complemento $\mathfrak{N}_a^c = \{x \in X : f(x) > a\}$ es abierto en X. Supongamos que $\mathfrak{N}_a \neq \emptyset, X$, pues estos dos son abiertos y cerrados en X. Sea $x_0 \in \{x \in X : f(x) > a\}$. Tomemos $\epsilon = f(x_0) - a > 0$. Entonces por ser f semicontinua inferiormente en X, tenemos que existe una vecindad V_{x_0} de x_0 en X tal que

$$f(x_0) - \epsilon = a < f(x)$$

[5] Basta con que X sea un espacio topológico.

para todo $x \in V_{x_0}$. Esto implica que $V_{x_0} \subset \mathfrak{N}_a^c$. Como $x_0 \in \mathfrak{N}_a^c$ fue arbitrario, esto prueba que cualquier punto de $\mathfrak{N}_a^c$ tiene una vecindad totalmente contenida en $\mathfrak{N}_a^c$, por lo cual este conjunto es vecindad de todos sus puntos, es decir, es abierto.

$\Leftarrow$ Recíprocamente, si $\mathfrak{N}_a^c$ es abierto para todo $a \in \mathbb{R}$, entonces para todo $\epsilon > 0$ y $x_0 \in X$ tal que $f(x_0) \in \mathbb{R}$ tenemos que $a := f(x_0) - \epsilon \in \mathbb{R}$. entonces $\{x \in X : f(x) > a\}$ es un abierto que contiene a x_0, es decir, es una vecindad de x_0 tal que

$$f(x_0) - \epsilon < f(x)$$

para todo elemento de $\{x \in X : f(x) > a\}$. Esto prueba que f es semicontinua inferiormente en X. $\square$

El resultado análogo del teorema anterior es el siguiente, cuya prueba es totalmente análoga a la del resultado anterior, o se sigue de aplicar el resultado anterior a $-f$, pues f es semicontinua inferiormente si y sólo si $-f$ es semicontinua superiormente.

Teorema A.8. *Sea X un espacio métrico y f una función $f : X \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ definida en X. Entonces f es semicontinua superiormente si y sólo si para todo $a \in \mathbb{R}$ se cumple que los conjuntos de nivel*

$$\{x \in X : f(x) \geq a\}$$

son cerrados en X.

Por los resultados anteriores, es común definir que f es semicontinua inferiormente en x_0 si para todo $a \in \mathbb{R}$ tal que $f(x_0) > a$ se cumple que $\{x \in X : f(x) > a\}$ es una vecindad abierta de x_0. Análogamente, f es semicontinua superiormente en x_0 si para todo $a \in \mathbb{R}$ tal que $f(x_0) < a$ se cumple que $\{x \in X : f(x) < a\}$ es una vecindad abierta de x_0. Entonces en esta convención se dice que una función f es semicontinua inferiormente (superiormente) si $\{x \in X : f(x) \leq a\}$ (respectivamente $\{x \in X : f(x) \geq a\}$) es cerrado para todo $a \in \mathbb{R}$.

Otra definición de semicontinuidad que suele darse está justificada por el siguiente resultado, cuya prueba puede verse en el teorema 5.5.17 p. 234 del texto de S. Berberian [1].

Teorema A.9. *Sea X un espacio métrico y f una función $f : X \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ definida en X. Entonces f es semicontinua inferiormente en el punto $x \in X$ si se cumple que para toda sucesión $\{x_n\}$ que converge a x en X se cumple que*

$$f(x) \leq \liminf_{n\to\infty} f(x_n).$$

Diremos que una familia $\{g_\alpha\}_{\alpha\in\mathcal{A}}$ de funciones entre espacios métricos $\mathcal{X}$ e $\mathcal{Y}$ es *equicontinua* en el punto $x_0 \in \mathcal{X}$ si para toda $\epsilon > 0$ existe una $\delta > 0$ tal que

$$d_{\mathcal{X}}(x, x_0) < \delta \quad \text{implica que}$$

$$d_{\mathcal{Y}}(g(x), g(x_0)) < \epsilon \quad \text{para toda } \ g \in \{g_\alpha\}_{\alpha\in\mathcal{A}},$$

es decir, que exista una misma $\delta > 0$ tal que funcione para verificar continuidad de toda $g \in \{g_\alpha\}_{\alpha\in\mathcal{A}}$.

Lema A.4. *Sea $\{g_n\}$ una sucesión de funciones reales equicontinuas definidas en un espacio métrico. Si en un conjunto compacto K, $\{g_n\}$ converge puntualmente a una función g, entonces $\{g_n\}$ converge uniformemente en K a g.*

Demostración. Sea $\epsilon > 0$. Por equicontinuidad, existe $\delta > 0$ tal que si $d(x, y) < \delta$, entonces $|g_n(x) - g_n(y)| < \epsilon$ para toda $n \in \mathbb{N}$. Sea $B_\delta(x)$ una bola centrada en x y de radio δ. Como K es compacto, podemos extraer una subcolección finita $\{x_1, \ldots, x_M\}$ tal que

$$K \subset \bigcup_{k=1}^{M} B_\delta(x_k).$$

Por convergencia puntual, $\{g_n\}$ converge uniformemente en el conjunto finito $\{x_1, \ldots, x_M\}$, pues podemos tomar $N := \max_{1\leq k\leq M} N(\epsilon, x_k)$ para obtener que si $n \geq N$, entonces

$$|g_n(x) - g(x)| < \epsilon \quad \text{para todo } x \in \{x_1, \ldots, x_M\}.$$

De este modo, para toda $1 \leq k \leq M$ existe N tal que si $m, n \geq N$ entonces $|g_n(x_k) - g_m(x_k)| < \epsilon$. Para cada $x \in K$, existe $k \leq M$ tal que $x \in B_\delta(x_k)$. Entonces por esto último y equicontinuidad del conjunto de funciones $\{g_n\}$, para esta x ocurre que si $m, n \geq N$ entonces

$$|g_n(x) - g_m(x)| \leq |g_n(x) - g_n(x_k)| + |g_n(x_k) - g_m(x_k)| + |g_m(x_k) - g_m(x)| < 3\epsilon.$$

Como toda $x \in B_\delta(x_k)$ para alguna $k \leq M$, la desigualdad anterior ocurre para toda $x \in K$, por lo cual $\{g_n\}$ es una sucesión uniformemente de Cauchy, por lo cual converge a g en $C(K)$. Esto implica que $g_n \xrightarrow{n} g$ uniformemente en K, pues la convergencia en $C(K)$ es la de la convergencia uniforme.

□

Bibliografía

[1] Sterling K. Berberian. *Fundamentals of Real Analysis.* Springer-Verlag, 1999.

[2] Patrick Billingsley. *Convergence of Probability Measures.* John Wiley & Sons, 1968.

[3] Patrick Billingsley. *Probability and Measure.* John Wiley & Sons, second edition, 1986.

[4] Paul Glasserman. *Monte Carlo Methods in Financial Engineering.* Springer-Verlag, 2003.

[5] Paul R. Halmos. *Measure Theory.* Springer-Verlag, 1974.

[6] J. Jost. *Postmodern Analysis.* Springer-Verlag, third edition, 2005.

[7] Hui-Hsiung Kuo. *Introduction to Stochastic Integration.* Springer-Verlag, 2006.

[8] César Almenara Martínez. *Muestreo de Importancia: Una Aplicación Estadística de las Grandes Desviaciones.* UNAM, tesis de licenciatura.

[9] Nicholas Metropolis. The beginning of the monte-carlo method. *Los Alamos Science*, Special Issue:125–130, 1989.

[10] P. Heidelberger y P. Shahabuddin P. Glasserman. Asymptotically optimal importance sampling and stratification for price-dependent options. *Math.Finance*, 9:117–152, 1999.

[11] H.L. Royden. *Real Analysis.* Macmillan Publishing Company, third edition, 1988.

[12] J. Michael Steele. *Stochastic Calculus and Financial Applications.* Applications of Mathematics. Springer-Verlag, New York, 2001.

[13] Daniel W. Stroock. *Essentials of Integration Theory for Analysis.* Springer-Verlag, 2011.

[14] David Williams. *Probability with Martingales.* Cambridge University Press, 1991.

[15] Mark I. Freidlin y Alexander D. Wentzell. *Random Perturbations of Dynamical Systems.* Springer-Verlag, third edition, 2012.

[16] Jean-Dominique Deuschel y Daniel W. Stroock. *Large Deviations.* Academic Press, 1989.

[17] Amir Dembo y Ofer Zeitouni. *Large Deviations Techniques and Applications.* Springer-Verlag, 1998.

[18] Jean Jacod y Philip Protter. *Probability Essentials.* Springer-Verlag, second edition, 2000.

[19] G. Casella y Roger L. Berger. *Statistical Inference.* Duxbury Advanced Series, second edition, 2002.

[20] Paolo Guasoni y Scott Robertson. Optimal importance sampling with explicit formulas in continuous time. *Finance and Stochastics*, 12:1–19, 2008.

[21] Ioannis Karatzas y Steven E. Shreve. *Brownian motion and Stochastic Calculus.* Springer-Verlag, Berlin Heidelberg, second edition, 1991.

[22] J. Jost y X. Li-Jost. *Calculus of Variations.* Cambridge University Press, 1998.

Índice alfabético