

# Procesos estocásticos y ecuaciones diferenciales: una introducción

Fabrizio Macià y Gerardo Oleaga  
2007

Disponible en: <http://dcain.etsin.upm.es/fabrizio/Docencia.html>

# Índice

<b>1. Repaso de la teoría básica de probabilidades</b>	<b>4</b>
1.1. Probabilidad . . . . .	4
1.2. Propiedades básicas. . . . .	4
1.3. Probabilidad condicional e independencia . . . . .	5
1.4. Variables aleatorias . . . . .	7
1.5. Vectores aleatorios . . . . .	9
1.6. Transformación de variables aleatorias . . . . .	10
1.7. Valor esperado y momentos . . . . .	11
1.8. Distribuciones condicionales . . . . .	13
1.9. Funciones generatrices . . . . .	14
<b>2. La caminata aleatoria.</b>	<b>16</b>
2.1. El proceso aleatorio más simple en una dimensión. . . . .	16
2.2. Solución explícita discreta y su comportamiento asintótico . . . . .	16
2.3. La función densidad de probabilidad . . . . .	17
2.3.1. Pared reflejante . . . . .	18
2.3.2. Pared absorbente . . . . .	19
2.4. Una ecuación en diferencias para la densidad de probabilidad . . . . .	21
2.5. La caminata aleatoria como límite de pasos independientes: El teorema de Laplace-De Moivre . . . . .	23
<b>3. El proceso de Wiener.</b>	<b>25</b>
3.1. Definición y propiedades básicas del proceso de Wiener. . . . .	25
3.2. La medida de Wiener . . . . .	27
<b>4. Interpretación estocástica de esquemas numéricos para EDP'S.</b>	<b>29</b>
4.1. La ecuación del calor: solución numérica. . . . .	29
4.2. La ecuación del calor con una fuente . . . . .	31
4.3. La ecuación del calor con potencial . . . . .	33
<b>5. Introducción</b>	<b>35</b>
<b>6. Construcción de una caminata aleatoria infinitesimal</b>	<b>35</b>
<b>7. El movimiento Browniano y la medida de Wiener</b>	<b>37</b>
7.1. Sobre el espacio muestral . . . . .	38
7.2. El teorema de Riesz-Markov . . . . .	40
7.3. La construcción de la medida de Wiener . . . . .	41
7.4. La continuidad de las trayectorias del movimiento Browniano . . . . .	42
<b>8. Fórmula de Feynman-Kac y ecuaciones en derivadas parciales</b>	<b>46</b>
8.1. Introducción . . . . .	46
8.2. La fórmula de Feynman-Kac para $-\frac{1}{2}\Delta + V$ . . . . .	46
8.3. La fórmula de Feynman para la ecuación de Schrödinger . . . . .	48
<b>9. Aplicación al estudio de los autovalores de un operador elíptico</b>	<b>50</b>
9.1. Los autovalores del operador de Schrödinger $-\frac{1}{2}\Delta + V(x)$ . . . . .	50
9.2. Aplicación de Feynman-Kac al estudio del comportamiento asintótico de los autovalores . . . . .	52
9.2.1. Aplicación de la fórmula de Feynman-Kac . . . . .	54

9.2.2. El puente Browniano . . . . .	55
9.2.3. Estimaciones de $\text{tr}(e^{-t(\Delta/2-V)})$ . . . . .	56
<b>A. La ecuación del calor: solución analítica.</b>	<b>57</b>

# 1. Repaso de la teoría básica de probabilidades

## 1.1. Probabilidad

Al construir un modelo probabilístico tenemos en mente la idea de la *frecuencia* con la que ocurren ciertos sucesos cuya naturaleza aleatoria nos impide conocer con detalle las leyes que los gobiernan. Ejemplos de este tipo de fenómenos son el resultado de arrojar una moneda, el ganador de una carrera de caballos o el comportamiento del precio de una acción. La abstracción y generalización de estas situaciones y conceptos, nos llevan a las siguientes definiciones:

1) **Experimento.** Cualquier procedimiento que nos proporcione un conjunto bien definido de resultados posibles se llama *experimento*.

2) **Espacio muestral.** El conjunto de todos los resultados posibles de un experimento se llama el *espacio muestral* y se denota por  $\Omega$ . Un resultado particular (pero no especificado) se suele denotar  $\omega$ .

3) **Eventos.** Un *evento* es un subconjunto del espacio muestral  $\Omega$ . En particular,  $\Omega$  es llamado el *evento seguro* y su complemento, denotado por  $\Omega^c = \emptyset$ , se llama *evento imposible*.

4) **Espacio de eventos** (o sigma álgebra). Una colección de eventos se llama *espacio de eventos*, denotada por  $\mathcal{F}$ , si satisface:

1.  $\emptyset$  está en  $\mathcal{F}$ ;
2. si  $A$  está en  $\mathcal{F}$ , entonces  $A^c$  está en  $\mathcal{F}$ ;
3. si  $A_n$  están en  $\mathcal{F}$  para  $n = 1, 2, \dots$ , entonces  $\cup_{n=1}^{\infty} A_n$  está en  $\mathcal{F}$ .

5) **Función de probabilidad.** Una función  $P(\cdot)$ , definida sobre los eventos en un espacio  $\mathcal{F}$ , es una probabilidad, si:

1.  $P(\Omega) = 1$ ;
2.  $0 \leq P(A) \leq 1$  para cualquier evento  $A \in \mathcal{F}$ .
3. (Axioma de aditividad numerable) si  $A_j \cap A_k = \emptyset$  para  $j \neq k$ , entonces

$$P\left(\cup_{n \in I} A_n\right) = \sum_{n \in I} P(A_n),$$

donde  $(A_n; n \in I)$  puede ser una colección finita o infinita numerable de eventos.

6) **Espacio de probabilidad.** La función  $P(\cdot)$  puede ser llamada una medida de probabilidad, o una distribución de probabilidad. La estructura completa consistente en  $\Omega$  (espacio muestral),  $\mathcal{F}$  (sigma álgebra de eventos) y  $P$  (función de probabilidad), es identificada como un *espacio de probabilidad*.

## 1.2. Propiedades básicas.

Listamos a continuación algunas propiedades elementales que pueden demostrarse como ejercicio a partir de las definiciones.

P1:

$$P(A^c) = 1 - P(A).$$

P2:

$$P(\emptyset) = 0.$$

P3:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

P4:

$$P(A) \leq P(B), \quad \text{si } A \subset B.$$

P5 (*Identidad de inclusión-exclusión*):

$$P(\cup_{r=1}^n A_r) = \sum_{r=1}^n P(A_r) - \sum_{r < s} P(A_r \cap A_s) + \dots + (-1)^{n+1} P(\cap_{r=1}^n A_r).$$

P6 (*Propiedad de continuidad*): supongamos que  $A_n$  es una sucesión de conjuntos tales que  $A_n \subset A_{n+1}$ , entonces se verifica que:

$$\text{Si } \lim_{n \rightarrow \infty} A_n = A, \text{ entonces } \lim_{n \rightarrow \infty} P(A_n) = P(A).$$

Aquí el límite de los conjuntos debe entenderse del siguiente modo: dado  $a \in A$ , existe un  $n$  suficientemente grande tal que  $a \in \cup_{r=1}^n A_r$ .

P7 (*Desigualdad de Boole*) Para eventos cualesquiera  $A_1, \dots, A_n$  tenemos que:

$$P(\cup_{r=1}^n A_r) \leq \sum_{r=1}^n P(A_r).$$

**Ejercicio 1.** Pruébense las propiedades P1-P7.

### 1.3. Probabilidad condicional e independencia

La probabilidad condicional cambia el espacio muestral  $\Omega$  a un evento  $B$ , de modo que las proporciones del número de veces que ocurre  $A$  se refieren al evento  $B$ . Es decir, medimos el número de veces que ocurren los eventos  $A$  y  $B$  al mismo tiempo y hallamos la proporción respecto del número de veces que ocurrió  $B$ .

**Probabilidad condicional.** Si  $P(B) > 0$ , la *probabilidad condicional* de que  $A$  ocurra *dado que*  $B$  ocurre se denota por  $P(A|B)$ , y se define por:

$$P(A|B) = P(A \cap B) / P(B).$$

Para entender intuitivamente la definición podemos identificar la probabilidad con la frecuencia relativa con la que ocurren los eventos al realizar un número  $N$  de pruebas. Si  $f_X$  es la frecuencia con la que ocurre el evento  $X$ , entonces el número de veces que ocurren  $A$  y  $B$  al mismo tiempo es:

$$N f_{A \cap B}$$

Mientras que el número de veces que ocurre  $B$  es

$$N f_B$$

y el cociente entre ambas es la definición de probabilidad condicional desde el punto de vista intuitivo de las frecuencias:

$$f_{A|B} = \frac{f_{A \cap B}}{f_B}.$$

Con más generalidad, aún cuando  $P(B) = 0$  admitimos la llamada *ley de multiplicación*:

$$P(A \cap B) = P(A|B)P(B).$$

Es importante tener en cuenta que la correspondencia

$$A \mapsto P(A|B)$$

es una función de probabilidad para los eventos en  $\mathcal{F}$  y satisface todas las propiedades de la definición.

**Identidades importantes.** La definición de probabilidad condicional, junto con las reglas básicas de la probabilidad, proveen varias identidades útiles.

**Regla de partición.** La forma más simple es:

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c).$$

Con más generalidad, si  $B_j \cap B_k = \emptyset$  para  $j \neq k$ , y  $A \subseteq \cup_r B_r$ , luego

$$P(A) = \sum_r P(A|B_r)P(B_r).$$

**Regla de multiplicación.** La forma elemental  $P(A \cap B) = P(A|B)P(B)$  puede extenderse fácilmente a la forma más general

$$P\left(\bigcap_{r=1}^{n+1} A_r\right) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \dots P(A_{n+1}|\bigcap_{r=1}^n A_r).$$

Finalmente, si combinamos la regla de partición con la definición de probabilidad condicional, obtenemos la conocida:

**Regla de Bayes.** Si  $A_j \cap A_k = \emptyset$  para  $j \neq k$ ,  $B \subseteq \cup_{k=1}^n A_k$  y  $P(B) > 0$ , entonces

$$P(A_r|B) = \frac{P(B|A_r)P(A_r)}{\sum_{k=1}^n P(B|A_k)P(A_k)}.$$

**Ejercicio 2.** (Uso de la regla de Bayes) Un test médico detecta una enfermedad con probabilidad 0,99 cuando se le aplica a una persona enferma y con probabilidad 0,01 cuando se lo aplica a una persona sana. ¿Cuál es la probabilidad de tener efectivamente la enfermedad si el test da positivo y la enfermedad afecta a una persona cada 100? ¿y si la enfermedad afectara a una persona cada  $10^5$ ?

La idea de condicionar nos lleva a un concepto muy importante: la *independencia* de sucesos.

**Independencia.**

a)] Los eventos  $A$  y  $B$  son *independientes* si

$$P(A \cap B) = P(A)P(B).$$

Una *familia* de eventos  $(A_i, 1 \leq i \leq n)$  es independiente si

$$P(A_{i_1} \cap A_{i_2} \dots \cap A_{i_r}) = P(A_{i_1})P(A_{i_2}) \dots P(A_{i_r})$$

para cualquier selección  $1 \leq i_1 < i_2 < \dots < i_r \leq n$ .

**Ejercicio 3.** Se arrojan 3 dados. Se definen los eventos

$A :=$  El primero y el segundo muestran el mismo resultado.

$B :=$  El segundo y el tercero muestran el mismo resultado.

$C :=$  El primero y el tercero muestran el mismo resultado.

Muéstrese que los eventos son dos a dos independientes pero que no son independientes como familia de eventos.

## 1.4. Variables aleatorias

Usualmente no estamos interesados en los resultados  $\omega$  del espacio muestral  $\Omega$  de un experimento aleatorio sino en algún resultado numérico que dependa de  $\omega$ . Por ejemplo, en una apuesta uno está más interesado en la ganancia o pérdida resultante y no tanto en el resultado en sí. Al asignar un número  $X(\omega)$  a cada resultado, transferimos en forma natural la definición de probabilidad a subconjuntos de la recta real, ya que estamos más interesados en la probabilidad de que  $X(\omega)$  tome valores en distintos subconjuntos de  $\mathbb{R}$ . Denotaremos a las variables aleatorias por letras mayúsculas, mientras que sus valores posibles serán escritos con letras minúsculas. Pasamos a las definiciones.

a) **Variable aleatoria.** Dado  $\Omega$  y una función de probabilidad  $P$ , una *variable aleatoria* es una función  $X(\omega)$ , definida sobre  $\Omega$ , tomando valores en  $\mathbb{R}$ .

2. **Función de distribución.** La *función de distribución*  $F_X(x)$  de la variable  $X$  está definida como

$$F_X := P(\{\omega : X(\omega) \leq x\})$$

y también se denota directamente por  $P(X \leq x)$ . Nótese que estamos asumiendo que el conjunto  $\{\omega : X(\omega) \leq x\} \in \mathcal{F}$  para todo valor de  $x \in \mathbb{R}$ . Esto es equivalente a que la función  $X$  sea *medible* en la  $\sigma$ -álgebra  $\mathcal{F}$ . Asumiremos que este es el caso para las variables consideradas.

3. **Variables aleatorias discretas.** Una *variable discreta* toma valores en un subconjunto numerable  $D$  de  $\mathbb{R}$ . Comúnmente es un subconjunto de números enteros. Luego, la probabilidad de que  $X$  asuma un valor  $x$  determinado se denota como

$$f_X(x) = P(X = x) := P(\{\omega : X(\omega) = x\}) .$$

La función  $f_X$  se llama *función densidad de probabilidad* de la variable  $X$ .

Ejemplos de variables discretas:

a) **Distribución binomial.** Si se realizan  $n$  pruebas independientes de un experimento cuyos resultados son sólo dos, E (de éxito) con probabilidad  $p$  y F (de fracaso) con probabilidad  $(1-p)$ , entonces la variable  $X$  definida por el número de éxitos obtenidos en las  $n$  pruebas tiene la distribución:

$$f_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad 0 \leq k \leq n .$$

Nos referiremos a ella como la distribución Binomial con parámetros  $n$  y  $p$  y la denotaremos por  $B(n, p)$ . Para indicar que una variable  $X$  tiene esta distribución se usa la notación

$$X \sim B(n, p) .$$

b) **Distribución de Poisson.** Cuando  $n$  es grande y  $p$  pequeño de modo que  $\lim_{n \rightarrow \infty} np = \lambda$ , los valores de la distribución binomial convergen a

$$f_X(k) = \frac{e^{-\lambda} \lambda^k}{k!} .$$

Nos referiremos a ella como Poisson( $\lambda$ ). Esta distribución está asociada al número de éxitos que pueden ocurrir en un intervalo temporal dado, cuando los mismos pueden ocurrir en cualquier instante del intervalo. Ejemplos pueden obtenerse en el número de llamadas recibidas en un intervalo dado; en este caso el parámetro  $\lambda$  representa el número esperado de llamadas en ese intervalo.

- c) **Distribución geométrica.** En una sucesión de pruebas independientes con probabilidad de éxito  $p$  en cada intento y probabilidad  $(1 - p)$  de fallo, sea  $X$  el número de pruebas que se realizan hasta obtener un éxito (incluyendo esta última prueba). Luego  $X$  tiene distribución

$$f_X(k) = (1 - p)^{k-1}p, \quad k = 1, 2, \dots$$

Nos referiremos a esta distribución como  $\text{Geom}(p)$ .

**Variabes aleatorias continuas.** Una variable  $X$  que no sea discreta y cuya función de distribución pueda escribirse como:

$$F_X(x) = \int_{-\infty}^x f_X(s) ds$$

para alguna función no-negativa  $f_X(x)$  definida para todo valor de  $x$ , se llama *variable aleatoria continua*. En este caso  $f_X(x)$  es llamada la *función de densidad* (o *densidad*) de  $X$ . Es análoga a la función de probabilidad de una variable discreta. Cuando no haya confusión denotaremos la densidad por  $f(x)$  omitiendo mencionar la variable.

**Ejemplos:**

1. **Densidad uniforme.** Si elegimos un punto al azar en el intervalo  $(a, b)$  con la idea intuitiva de que “los intervalos de igual longitud tienen la misma probabilidad de ser elegidos”, es fácil ver que la distribución de probabilidad viene dada por:

$$F(x) = P(X \leq x) = \int_a^x \frac{dx}{b-a} = \frac{x-a}{b-a}.$$

La densidad de  $X$  queda definida por:

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{para } a < x < b, \\ 0, & \text{en otro caso.} \end{cases}$$

2. **Densidad exponencial.** Para cualquier constante  $\lambda > 0$ ,

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

Nos referiremos a la distribución generada como  $\text{Exponencial}(\lambda)$ .

3. **Densidad normal.** Para constantes  $\mu \in \mathbb{R}$ ,  $\sigma^2 > 0$ ,

$$f(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad -\infty < x < \infty.$$

Nos referiremos a esta distribución con  $N(\mu, \sigma^2)$ . Si  $\mu = 0$  y  $\sigma^2 = 1$ , esta será la *Densidad Normal estándar*. Para ésta usaremos la notación especial

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right),$$

$$\Phi(x) = \int_{-\infty}^x \phi(s) ds.$$

**Ejercicio 4.** a) Sea  $X \sim N(\mu, \sigma^2)$  y definamos la variable  $Y = aX + b$ , con  $a$  y  $b$  constantes. Muéstrase que  $Y \sim N(a\mu + b, a^2\sigma^2)$ . b) Sea  $U$  una variable aleatoria con distribución uniforme en  $(0, 1)$ , y definamos  $Y := -(1/\lambda) \log U$ . Muéstrase que si  $\lambda > 0$ ,  $Y \sim \text{Exponencial}(\lambda)$ .



## 1.5. Vectores aleatorios

En muchos modelos es necesario tener en cuenta los efectos conjuntos de varias variables al mismo tiempo. Dichas variables deben estar definidas en el mismo espacio muestral  $\Omega$  con su correspondiente función de probabilidad  $P$  y su espacio de eventos  $\mathcal{F}$ . Definimos algunas cantidades útiles para tratar estos casos distinguiendo entre el caso discreto y el continuo.

1. **Función de probabilidad conjunta.** Para variables discretas  $X_1, X_2, \dots, X_n$ , tenemos:

$$f(x_1, x_2, \dots, x_n) = P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n),$$

donde cada  $x_i$  recorre los valores posibles de cada variable  $X_i$ . Notación:

$$\mathbf{x} = (x_1, x_2, \dots, x_n) \quad \mathbf{X} = (X_1 = x_1, X_2 = x_2, \dots, X_n = x_n).$$

De este modo, si  $D$  es un subconjunto de  $\mathbb{R}^n$ , tenemos que:

$$P(\mathbf{X} \in D) = \sum_{\mathbf{x} \in D} f(\mathbf{x}).$$

**Distribuciones Marginales.** Podemos recuperar las distribuciones de cada una de las variables  $X_i$  si conocemos su distribución conjunta. Consideremos la distribución de la variable  $X_1$ :

$$f_{X_1}(x) = \sum_{x_2, \dots, x_n} f(x, x_2, \dots, x_n).$$

Es decir, sumamos sobre todos los valores posibles del resto de las variables, dejando fijo al valor  $x$ . En términos de eventos podemos interpretarlo de la siguiente forma:

$$\begin{aligned} \{X_1 = x\} &= \{X_1 = x\} \cap \Omega = \{X_1 = x\} \cap \cup_{x_2, \dots, x_n} \{X_2 = x_2, \dots, X_n = x_n\} \\ &= \cup_{x_2, \dots, x_n} \{X_1 = x\} \cap \{X_2 = x_2, \dots, X_n = x_n\} \end{aligned}$$

Las distribuciones así obtenidas se denominan *distribuciones marginales*.

2. **Función de densidad conjunta.** Un par de variables aleatorias  $(X_1, X_2)$  es (conjuntamente) continuo con densidad  $f(x_1, x_2)$  si, para todo  $a, b, c, d$ ,

$$P(\{a < X_1 \leq b\} \cap \{c < X_2 \leq d\}) = \int_a^b \int_c^d f(x_1, x_2) dx_2 dx_1$$

En particular, esto define la función de distribución conjunta de  $X_1, X_2$  denotada por  $F(x_1, x_2)$ , por:

$$F(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f(u, v) dv du.$$

La densidad tiene la misma propiedad que en el caso discreto. Si  $D$  es un subconjunto del plano, tenemos que:

$$P((X_1, X_2) \in D) = \int \int_{(x_1, x_2) \in D} f(x_1, x_2) dx_1 dx_2.$$

También podemos recuperar la distribución de cada variable a partir de la conjunta:

$$P(a < X_1 \leq b) = P(\{a < X_1 \leq b\} \cap \{-\infty < X_2 < \infty\}) = \int_a^b \int_{-\infty}^{\infty} f(x_1, x_2) dx_2 dx_1.$$

De modo que

$$f_{X_1}(x) = \int_{-\infty}^{\infty} f(x, x_2) dx_2.$$

3. **Independencia de variables aleatorias.** El par de variables  $X_1$  y  $X_2$  son independientes si, para todo  $x_1$  y  $x_2$ ,

$$P(X_1 \leq x_1, X_2 \leq x_2) := F(x_1, x_2) = F_{X_1}(x_1)F_{X_2}(x_2) =: P(X_1 \leq x_1)P(X_2 \leq x_2).$$

Para conjuntos arbitrarios podemos deducir la independencia a partir de las densidades. Derivando la expresión para  $F$  podemos obtener la factorización:

$$f(x_1, x_2) = f_{X_1}(x_1)f_{X_2}(x_2)$$

y probar que:

$$\begin{aligned} P(X_1 \in A, X_2 \in B) &= \int_A \int_B f(x_1, x_2) dx_1 dx_2 \\ &= \left( \int_A f_{X_1}(x_1) dx_1 \right) \left( \int_B f_{X_2}(x_2) dx_2 \right) = P(A)P(B) \end{aligned}$$

y recuperamos la noción previa de independencia.

**Ejercicio 5.** a) Sean  $X_1$  y  $X_2$  con distribución conjunta  $F(x_1, x_2)$ , y supongamos que  $a < b$  y  $c < d$ . Muéstrase que

$$P(a < X_1 \leq b, c < X_2 \leq d) = F(b, d) + F(a, c) - F(a, d) - F(b, c).$$

b) Sean  $X_1, X_2$  variables independientes con distribución exponencial con parámetros  $\lambda$  y  $\mu$  respectivamente. Muéstrase que

$$\mu P(X_1 \leq t < X_1 + X_2) = \lambda P(X_2 \leq t < X_1 + X_2).$$

## 1.6. Transformación de variables aleatorias

Consideremos el caso de variables continuas; buscamos la distribución de probabilidad de una función  $Z = g(X, Y)$  del vector aleatorio  $(X, Y)$ . Usando la definición de la función de probabilidad:

$$F_Z(z) := P(Z \leq z) = P(g(X, Y) \leq z) = \int \int_{x, y: g(x, y) \leq z} f(x, y) dx dy.$$

**Suma de variables aleatorias continuas.** Si  $X, Y$ , tienen densidad conjunta  $f(x, y)$  y su suma es  $Z = X + Y$  luego:

$$\begin{aligned} F_Z(z) &= P(X + Y \leq z) = \int \int_{x, y: x+y \leq z} f(x, y) dx dy \\ &= \int_{x=-\infty}^{\infty} \int_{y=-\infty}^{z-x} f(x, y) dx dy \\ &= \int_{u=-\infty}^z \int_{-\infty}^{\infty} f(v, u-v) dv du \end{aligned}$$

donde  $u = x+y$  y  $v = x$ . De esta fórmula vemos inmediatamente que la densidad de probabilidad

$$f_Z(z) = \int_{-\infty}^{\infty} f(v, z-v) dv.$$

Si  $X, Y$  son independientes tenemos que  $f(x, y) = f_X(x)f_Y(y)$ , y nos queda:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(v) f_Y(z-v) dv, \tag{1}$$

que representa la *convolución* de las densidades de  $X$  y de  $Y$ .

**Ejemplo: Suma de normales estándar** Si  $X, Y$  son variables  $N(0, 1)$ , luego usando (1):

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} \frac{1}{2\pi} \exp\left(-\frac{1}{2}v^2 - \frac{1}{2}(z-v)^2\right) dv \\ &= \int_{-\infty}^{\infty} \frac{1}{2\pi} \exp\left(-\frac{1}{4}z^2 - \left(v - \frac{z}{2}\right)^2\right) dv = \frac{1}{2\sqrt{\pi}} e^{-(1/4)z^2}, \end{aligned}$$

que es una normal de media cero y varianza 2.

**Ejercicio 6.** Muestra que si  $X_1 \sim N(\mu_1, \sigma_1^2)$  y  $X_2 \sim N(\mu_2, \sigma_2^2)$  con  $X_1, X_2$  independientes, luego  $X_1 + X_2$  es  $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ .

## 1.7. Valor esperado y momentos

Para hacer el promedio de una cierta cantidad de datos  $N$ , solemos hacer una lista  $d_n$ , los sumamos y dividimos por la cantidad total de datos:

$$w = \frac{1}{N} \sum_n d_n.$$

Si los datos  $d_n$  corresponden a una variable discreta, podemos reescribir esto agrupando los datos repetidos. Pongamos por ejemplo que  $d_n \in \{x_i : 1 \leq i < \infty\}$ . Entonces, podemos reescribir el promedio del siguiente modo:

$$w = \frac{1}{N} \sum_{i=1}^{\infty} n(i)x_i,$$

donde  $n(i)$  es el número de veces que aparece el dato  $x_i$ . Llamando  $f_i := n(i)/N$  a la *frecuencia* con que aparece el dato  $x_i$ , nos queda la siguiente expresión:

$$w = \sum_{i=1}^{\infty} f_i x_i.$$

**Valor esperado.** Sea  $X$  una variable discreta con distribución de probabilidad  $f(k)$ ; el *valor esperado* (o *media*) de  $X$  está denotado por  $E[X]$  y definido por

$$E[X] := \sum_k x_k f(k),$$

siempre que  $\sum_k |x_k| f(k) < \infty$ . Si esta condición falla diremos que la variable no tiene media finita. En el caso de una variable continua la definición es análoga:

$$E[X] := \int_{-\infty}^{\infty} x f(x) dx,$$

siempre que  $\int_{-\infty}^{\infty} |x| f(x) dx < \infty$ .

Consideremos ahora dos variables vinculadas por  $Y = g(X)$ . Para calcular el valor esperado debemos conocer (en principio) la distribución de  $g(X)$ . Podemos evitar esta dificultad por medio del siguiente resultado.

**Teorema** Sean las variables aleatorias  $X$  e  $Y$  tales que  $Y = g(X)$ , donde  $g(\cdot)$  es una función real definida sobre  $\mathbb{R}$ .

a) Si  $X$  es discreta, luego

$$E[Y] = \sum_k g(x_k) f(k) ,$$

siempre que  $\sum_k |g(x_k)| f(k) < \infty$ .

b) Si  $X$  es continua, luego:

$$E[X] := \int_{-\infty}^{\infty} g(x) f(x) dx ,$$

siempre que  $\int_{-\infty}^{\infty} |g(x)| f(x) dx < \infty$ .

*Demostración:* vemos solamente la parte a), ya que la b) es análoga.

$$\begin{aligned} E[Y] &:= \sum_y y P(Y = y) \\ &= \sum_y y \left( \sum_{x:g(x)=y} P(X = x) \right) \\ &= \sum_y \sum_{x:g(x)=y} g(x) P(X = x) \\ &= \sum_x g(x) P(X = x) , \end{aligned}$$

donde la última igualdad se obtiene teniendo en cuenta que  $\cup_y \{x : g(x) = y\}$  son todos los valores posibles de  $X$ .

**Propiedad** Si  $X$  e  $Y$  son independientes, entonces para funciones reales  $g$  y  $h$  tenemos que:

$$E[g(X)h(Y)] = E[g(X)]E[h(Y)] .$$

**Ejercicio 7.** Demuéstrese esta propiedad.

**Momentos.** Son cantidades que se definen a partir del valor esperado.

a) El *momento*  $k$ -ésimo de  $X$  es  $\mu_k := E[X^k]$ .

b) El *momento central*  $k$ -ésimo de  $X$  es  $\sigma_k := E[(X - E[X])^k]$ .

En particular  $\mu_1$  es la media y  $\sigma_2$  es llamada la varianza y denotada por  $\sigma^2$  o  $\text{Var}[X]$ .

**Ejercicio 8.** Muestra que la varianza de la suma de dos variables independientes es igual a la suma de las varianzas.

**Covarianza y correlación** Para variables  $X, Y$  con distribución conjunta debemos considerar usualmente los siguientes parámetros:

La covarianza de  $X$  e  $Y$  es:

$$\text{cov}(X, Y) = E[(X - E[X])(Y - E[Y])] .$$

El coeficiente de correlación (o correlación) es:

$$\rho(X, Y) := \frac{\text{cov}(X, Y)}{(\text{var}[X] \text{var}[Y])^{1/2}} .$$

## 1.8. Distribuciones condicionales

Las distribuciones condicionales de una variable se obtienen al considerar una restricción en el valor de otra variable aleatoria. Consideramos el caso discreto, ya que las extensiones son obvias.

**Función de probabilidad condicional** Consideremos dos variables discretas  $X$  e  $Y$ , distribuidas en forma conjunta. El evento  $Y = y$ , permite considerar las probabilidades condicionales de la variable  $X$  *dado que*  $Y = y$ . Esto define una nueva función de probabilidad para  $X$  que la denotamos por:

$$\begin{aligned} f_{X|Y}(x|y) &:= P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} \\ &= \frac{f_{X,Y}(x, y)}{f_Y(y)}, \quad f_Y(y) > 0. \end{aligned}$$

La función  $x \rightarrow f_{X|Y}(\cdot|y)$  define una nueva función de probabilidad para la variable  $X$ :

$$\sum_x f_{X|Y}(x|y) = \sum_x \frac{f_{X,Y}(x, y)}{f_Y(y)} = \frac{1}{f_Y(y)} \sum_x f_{X,Y}(x, y) = \frac{f_Y(y)}{f_Y(y)} = 1.$$

**Esperanza condicional.** La esperanza condicional de  $X$  dado  $Y = y$  se obtiene calculando el valor esperado cuando fijamos la variable  $Y$  y aplicamos la distribución condicional:

$$E(X|Y = y) = \sum_x x f_{X|Y}(x|y).$$

Cuando no se especifica un valor concreto de  $y$ , la esperanza condicional de  $X$  dado  $Y$  es una nueva variable aleatoria cuyos valores dependen del resultado de la variable  $Y$ , es decir:

$$E(X|Y) : \omega \mapsto E(X|Y = Y(\omega)).$$

**Caso continuo.** Cuando las variables son continuas no se puede hacer la derivación elemental usando la probabilidad condicional. En este caso nos guiamos heurísticamente por analogía con el caso discreto, aunque estas definiciones pueden justificarse en un contexto más general. Si  $X$  e  $Y$  son continuas y tienen densidad de distribución conjunta  $f$ , entonces la densidad condicional de  $X$  dado  $Y = y$  está definida por:

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}, \quad \text{cuando } f_Y(y) > 0.$$

La esperanza condicional de  $X$  dado  $Y = y$  viene dada por:

$$E(X|Y = y) = \int_{\mathbb{R}} x f_{X|Y}(x|y) dx,$$

siempre que la integral converja absolutamente. La función  $f_{X|Y}(\cdot, y)$  es una densidad de probabilidad y  $E(X|Y = Y(\cdot))$  es una variable aleatoria.

**El lema de partición.** Sean  $X$  e  $Y$  variables distribuidas en forma conjunta.

a) Si  $X$  e

$Y$  son discretas, entonces  $f_X(x) = \sum_y f_{X|Y}(x|y) f_Y(y)$ .

b) Si  $X$  e  $Y$  son continuas, entonces:

$$f_X(x) = \int_{\mathbb{R}} f_{X|Y}(x|y) f_Y(y) dy.$$

**Ejercicio 9.** Prueba el lema de partición.

A partir de este lema podemos mostrar algunas identidades útiles. Consideremos el caso discreto:

$$\begin{aligned} E[X] &= \sum_x x f_X(x) = \sum_x \sum_y x f_{X|Y}(x|y) f_Y(y) \\ &= \sum_y \left( \sum_x x f_{X|Y}(x|y) \right) f_Y(y) \\ &= \sum_y E(X|Y=y) f_Y(y) \\ &= E[E(X|Y)]. \end{aligned}$$

Y hay una identidad análoga para el caso continuo.

**Ejemplo de aplicación del lema de partición.** Sean  $X$  e  $Y$  variables aleatorias continuas e independientes. Entonces:

$$P(X < Y) = \int_{-\infty}^{\infty} P(X < Y|Y=y) f_Y(y) dy$$

Al ser  $X$  e  $Y$  independientes, tenemos que:

$$P(X < Y|Y=y) f_Y(y) = P(X < y) f_Y(y) = F_X(y) f_Y(y)$$

y finalmente

$$P(X < Y) = \int_{-\infty}^{\infty} F_X(y) f_Y(y) dy.$$

Es costumbre utilizar la siguiente notación, que es más intuitiva que la anterior:

$$P(X < Y) = \int_{-\infty}^{\infty} P(X < Y|y < Y < y + dy) f_Y(y) dy$$

resaltando el hecho de que  $f_Y(y) dy$  es la probabilidad de encontrar a la variable  $Y$  en el intervalo infinitesimal  $(y, y + dy)$ .

## 1.9. Funciones generatrices

**Función generatriz de momentos** La *función generatriz de momentos* (fgm) de la variable aleatoria  $X$  está dada por:

$$M_X(t) := E[e^{tX}],$$

para todos los valores reales  $t$  en que esté definido el valor esperado. En particular, cuando la función está definida en un intervalo real  $(-a, a)$ , podemos escribir, dentro del círculo de convergencia:

$$E[e^{tX}] = E\left[\sum_{r=0}^{\infty} \frac{t^r X^r}{r!}\right] = \sum_{r=0}^{\infty} \frac{t^r}{r!} E[X^r].$$

De modo que los coeficientes del desarrollo, que podemos recuperar derivando  $E[e^{tX}]$  nos indican el momento  $r$ -ésimo de la variable  $X$ .

Hay tres propiedades importantes que nos interesan de la fgm.

**Unicidad** Si  $M_X(t)$  es convergente en un intervalo  $(-a, a)$  con  $a > 0$ , entonces los momentos de la variable  $X$  determinan en forma única la distribución  $F_X(x)$ . Más aún, todos los momentos de  $X$  son finitos. Si  $X$  e  $Y$  son independientes, luego tenemos que la función generatriz de momentos *conjunta*

$$M(s, t) := E[e^{sX+tY}] = M_X(s)M_Y(t).$$

Más aún, puede mostrarse que si la fgm conjunta factoriza de esta manera, entonces  $X$  e  $Y$  son independientes. Sea  $M_n(t)$  una sucesión de fgms, tal que, cuando  $n \rightarrow \infty$ ,  $M_n(t) \rightarrow M(t)$ , donde  $M(t)$  es una fgm de una distribución  $F(x)$ . Además, si  $M_n(t)$  es la fgm de una distribución  $F_n(x)$ , tenemos que

$$F_n(x) \rightarrow F(x), \quad n \rightarrow \infty.$$

**Ejemplo.** Para una variable  $X \sim N(0, 1)$  tenemos que:

$$\begin{aligned} M_X(t) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} e^{tx} dx \\ &= e^{t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(x-t)^2/2} dx \\ &= e^{t^2/2}. \end{aligned}$$

Del mismo modo si  $Y \sim N(\mu, \sigma^2)$ , obtenemos:

$$M_Y(t) = E[e^{(\mu+\sigma X)t}] = e^{\mu t} M_X(\sigma t) = e^{\left(\mu t + \frac{\sigma^2}{2} t^2\right)}.$$

**La función característica** La función definida por

$$\phi_X(t) := E[e^{itX}] = E[\cos(tX) + i \operatorname{sen}(tX)], \quad t \in \mathbb{R}, \quad i^2 = -1,$$

se llama la *función característica* de la variable  $X$  o también fc. Esta función, a diferencia de la fgm, existe siempre, ya que:

$$|E[e^{itX}]| \leq E[|e^{itX}|] = 1.$$

Cuando la función generatriz de momentos existe en un intervalo  $(-a, a)$  tenemos que:

$$\phi_X(t) = M_X(it), \quad t \in \mathbb{R}.$$

Esto implica que la normal estándar tiene función característica:

$$\phi_X(t) = e^{-t^2/2}.$$

## 2. La caminata aleatoria.

### 2.1. El proceso aleatorio más simple en una dimensión.

Nos planteamos el siguiente problema elemental de probabilidades:

**UnicidadFactorizaciónContinuidad** Una partícula puede moverse sobre la recta real dando saltos de longitud 1 hacia la derecha o la izquierda de modo aleatorio. La probabilidad de desplazarse hacia uno u otro lado es de  $1/2$ . Después de  $N$  movimientos puede ocupar alguna de las posiciones:

$$X_N = -N, -N + 1, \dots, -1, 0, 1, \dots, N - 1, N$$

Si su posición inicial es  $X_0 = 0$  deseamos calcular la probabilidad  $w(m, N) := P(X_N = m)$  de encontrarla en la posición  $X_N = m$  al dar un número de pasos  $N$ .

Haremos en primer lugar un abordaje directo calculando una expresión explícita para estas probabilidades. Como veremos más adelante, éste también puede resolverse desde puntos de vista distintos: a través de la solución explícita de una ecuación en derivadas parciales, como una ecuación en diferencias, o como el valor esperado de una variable estocástica.

### 2.2. Solución explícita discreta y su comportamiento asintótico

Dadas las condiciones del problema, está claro que  $w(m, N) = 0$  para  $|m| > N$ . Debemos entonces contar de cuántas maneras diferentes podemos llegar a un valor  $m$ , ya que todos los caminos posibles son igualmente probables, con probabilidad  $1/2^N$ .

Sean  $i, d \geq 0$  el número de pasos realizado hacia la izquierda y a la derecha respectivamente. Tenemos entonces que:

$$\begin{aligned}d + i &= N \\d - i &= m\end{aligned}$$

Los posibles valores de  $i, d$  son  $0, \dots, N$  de modo que *no todas* las posiciones  $x = m$  son accesibles para un  $N$  dado, sólo lo son aquellas para las cuales  $d = \frac{N+m}{2}$ ,  $i = \frac{N-m}{2}$  son enteros positivos. En caso de que lo sean el problema es equivalente a contar las palabras de  $N$  letras, entre las que hay  $d$  letras A e  $i$  letras B:

$$w(m, N) = \frac{N!}{d!i!} \frac{1}{2^N} = \frac{N!}{\left(\frac{N+m}{2}\right)! \left(\frac{N-m}{2}\right)!} \frac{1}{2^N}.$$

**Ejercicio 10.** Comprueba que para cada  $N$  la media de la variable  $X_N$  es cero y su varianza es  $N$ .

El comportamiento de esta expresión para  $m \ll N^1$ ,  $m, N \rightarrow \infty$  podemos verlo usando la fórmula de Stirling:

$$n! \sim (2\pi n)^{1/2} \left(\frac{n}{e}\right)^n \quad n \rightarrow \infty, \quad (2)$$

donde  $f(n) \sim g(n)$  para  $n \rightarrow \infty$  quiere decir que  $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 1$ . Usando esta aproximación podemos escribir:

$$w(m, N) \sim \sqrt{\frac{2}{\pi}} N^{N+1/2} \left(\frac{N-m}{N+m}\right)^{m/2} (N^2 - m^2)^{-\frac{N+1}{2}} \quad N \gg m. \quad (3)$$

---

<sup>1</sup>recordemos que  $m \ll N$  para  $N \rightarrow \infty$  es equivalente a  $\lim_{N \rightarrow \infty} \frac{m}{N} = 0$ .



o de forma equivalente:

$$w(m, N) \sim \sqrt{\frac{2}{\pi N}} \left( \frac{1 - m/N}{1 + m/N} \right)^{m/2} \left( 1 - \left( \frac{m}{N} \right)^2 \right)^{-\frac{N+1}{2}} \quad N \gg m. \quad (4)$$

Por otra parte, si  $m \ll N$  tenemos también que:

$$\left( \frac{1-x}{1+x} \right)^{m/2} = e^{\frac{m}{2} \log f(x)} = e^{\frac{m}{2} (-2x + O(x^2))} = e^{-mx + mO(x^2)}$$

de modo que

$$\left( \frac{1 - m/N}{1 + m/N} \right)^{m/2} = e^{-\frac{m^2}{N} (1 + O(\frac{m}{N}))}$$

y además

$$(1 - x^2)^{-\frac{N+1}{2}} = e^{-\frac{N+1}{2} \log(1-x^2)} = e^{-\frac{N+1}{2} (-x^2 + O(x^4))}$$

entonces

$$\left( 1 - \left( \frac{m}{N} \right)^2 \right)^{-\frac{N+1}{2}} = e^{\frac{m^2}{2N} (1 + O(\frac{m^2}{N^2}))}$$

Y así, reemplazando en (4) obtenemos la siguiente relación:

$$w(m, N) \sim \sqrt{\frac{2}{\pi N}} e^{-\frac{m^2}{N}} e^{\frac{m^2}{2N}} = \sqrt{\frac{2}{\pi N}} e^{-\frac{m^2}{2N}}.$$

El comportamiento asintótico para  $N \rightarrow \infty$ ,  $m \ll N$ , es finalmente:

$$w(m, N) = \frac{N!}{\left(\frac{N+m}{2}\right)! \left(\frac{N-m}{2}\right)!} \frac{1}{2^N} \sim \sqrt{\frac{2}{\pi N}} e^{-\frac{m^2}{2N}} \quad (5)$$

### 2.3. La función densidad de probabilidad

Si llamamos  $k := t/N$ ,  $h := x/m$  (tamaño de la malla), y teniendo en cuenta que en un intervalo de longitud  $\Delta x$  entran aproximadamente  $\frac{\Delta x}{2h}$  posiciones posibles (debido a que dependiendo de  $m$  y  $N$  tenemos algunas posiciones inaccesibles), la probabilidad de encontrar a la partícula en el intervalo  $[x - \frac{\Delta x}{2}, x + \frac{\Delta x}{2}]$  viene dada por:

$$\int_{x - \frac{\Delta x}{2}}^{x + \frac{\Delta x}{2}} u(y, t) dy := P\left(x - \frac{\Delta x}{2} \leq mh \leq x + \frac{\Delta x}{2}\right) \approx w(x/h, t/k) \frac{\Delta x}{2h} = \frac{\Delta x}{h} \sqrt{\frac{k}{2\pi t}} e^{-\frac{x^2}{2t} \frac{k}{h^2}} \quad (6)$$

Definiendo además el coeficiente de difusión  $D := \frac{h^2}{2k}$  (es la velocidad a la que se recorre la distancia  $h^2$  partida por dos) tenemos que la densidad de probabilidad viene dada (para  $h \rightarrow 0$ ,  $\Delta x \ll 1$ ) por la expresión

$$u(x, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}} \quad (7)$$

que corresponde a la solución fundamental de la ecuación del calor.

### 2.3.1. Pared reflejante

Supongamos que la partícula, al llegar a una posición  $r > 0$ , tiene probabilidad 1 de ir hacia la izquierda y probabilidad 0 de desplazarse a la derecha. La posición  $r$  hace “rebotar” el movimiento hacia la izquierda y debemos recalcular la probabilidad de algunos de los caminos anteriores. Consideremos la función  $w_{ref}(m, N; r) :=$  probabilidad de que la partícula llegue a la posición  $m$ , con  $m < r$ , al desplazarse  $N$  pasos.

Para fijar ideas tomemos un camino que rebote una sola vez en la posición  $r$ . La probabilidad de cada paso sigue siendo  $1/2$ , con excepción del paso dado en la posición  $r$ , que tiene probabilidad 1 de ir hacia la izquierda. En este caso es equivalente a multiplicar por dos la probabilidad que tiene el mismo camino de ocurrir en un desplazamiento libre, sin una pared reflejante. Es fácil convencerse de que la probabilidad de un camino que rebote  $n$  veces en la pared, es  $2^n$  por la probabilidad del mismo camino considerado como un proceso sin pared reflejante. Una forma gráfica de calcular estas probabilidades es considerar *caminos reflejados* al otro lado de la pared. Por ejemplo, si el camino rebota sólo una vez, podemos considerar otro camino que sale hacia la derecha en el punto de rebote y está reflejado al otro lado de la pared. Al realizar los  $N$  pasos el camino original llega a la posición  $m$  y el reflejado llegará a la posición  $2r - m$  (véase la figura)

Del mismo modo, cada camino libre que atraviese una sola vez la posición  $r$  (y que llegue a la posición  $m' = 2r - m$ ) corresponde a la reflexión de un camino admisible. Si sumamos la probabilidad del camino libre que llega a  $m$  y la del camino libre que llega a  $m'$  tenemos el mismo efecto que al multiplicar por dos la probabilidad del camino admisible (considerado como camino libre). Para generalizar esta idea, supongamos que el camino rebota dos veces en la pared. En este caso habrá que multiplicar por 4 a la probabilidad del camino libre, con lo cual habrá que construir 3 caminos más para seguir con la misma idea.

Dado el camino admisible, tenemos 3 caminos libres más para elegir, uno que llega a  $m$  y dos que llegan a  $m'$ . Consideremos el que llega hasta el punto 1, después de pasar la barrera se dirige al punto 2 y de ahí al punto  $m$ . Otro de ellos sigue el mismo camino admisible hasta el punto 2 y en lugar de rebotar en la pared se dirige al punto  $m'$ . Por último, el camino que a partir del punto 1 es una reflexión del camino admisible con respecto al eje que pasa por  $r$ . De este modo, podemos convencernos de que es válida la siguiente fórmula

$$w_{ref}(m, N; r) = w(m, N) + w(2r - m, N) \quad (m < r). \quad (8)$$

Una aclaración importante para  $m = r$ . El razonamiento que hemos hecho es válido solamente para posiciones  $m < r$ , ya que por cada rebote contra la pared debemos considerar el doble de caminos, pero si  $m = r$  esto no es verdad. En este caso debemos tener en cuenta que el último paso, si bien toca la pared, no debe ser multiplicado por dos. Véase el ejemplo de la figura

La probabilidad que buscamos es entonces la mitad que la obtenida con la fórmula (8) y tenemos entonces que

$$w_{ref}(r, N; r) = w(r, N).$$

¡Es decir que la probabilidad de llegar a la pared reflectante en tiempo  $N$  es exactamente la misma que la probabilidad de llegar por caminos libres!

Si tomamos  $N$  muy grande en (8) tenemos la fórmula asintótica:

$$w_{ref}(m, N; r) \sim \left(\frac{2}{\pi N}\right)^{\frac{1}{2}} \left(e^{-m^2/2N} + e^{-(2r-m)^2/2N}\right) \quad (m < r)$$

o en su versión continua (para la densidad de probabilidad):

$$u_{ref}(x, t) = \frac{1}{2\sqrt{\pi Dt}} \left(e^{-\frac{x^2}{4Dt}} + e^{-\frac{(2x_r-x)^2}{4Dt}}\right) \quad (x < x_r), \quad (9)$$

Donde  $x_r := rh$ , es la posición de la pared.

Nótese que la función definida en (9) satisface la ecuación:

$$\frac{\partial u_{ref}}{\partial x} \Big|_{x=x_r} = 0$$

### 2.3.2. Pared absorbente

En este caso calcularemos la probabilidad de encontrar a la partícula en la posición  $m$ , después de  $N$  pasos, pero considerando que una vez que atravesó una pared en la posición  $a$ , ya no regresa. Es decir, si el camino que seguimos atraviesa la posición  $a > 0$ , la probabilidad de ir hacia la izquierda es ahora nula. De los caminos libres que llegan a  $m$  tenemos que quitar los que en algún instante anterior han atravesado la barrera en  $a$ . Podemos hacer la misma construcción gráfica que para el caso anterior, esta vez teniendo en cuenta que *debemos descontar la probabilidad de los caminos que terminan en  $m'$* . Por cada camino prohibido (libre) que llega a  $m$  (esto es, que atraviesa a la pared en  $a$ ), podemos construir uno que llega a  $m'$  usando la reflexión con respecto al eje en  $a$  a partir de los puntos de contacto. Los caminos admisibles no pueden reflejarse ya que no hay puntos de contacto con el eje. Recíprocamente, por cada camino que sale de 0 y llega a  $m'$  podemos contar un camino prohibido que llega a  $m$ . De este modo podemos decir que la probabilidad que buscamos es

$$\begin{aligned} w_{abs}(m, N; a) &= w(m, N) - w(2a - m, N) \\ &\approx \left(\frac{2}{\pi N}\right)^{\frac{1}{2}} \left(e^{-m^2/2N} - e^{-(2a-m)^2/2N}\right) \quad m < a. \end{aligned}$$

y la función densidad es

$$u_{abs}(x, t) = \frac{1}{2\sqrt{\pi Dt}} \left(e^{-\frac{x^2}{4Dt}} - e^{-\frac{(2x_a-x)^2}{4Dt}}\right).$$

Con la propiedad

$$u_{abs}(x_a, t) = 0.$$

Para  $m = a$  el razonamiento anterior no es válido y solamente debemos quitar algunos de los caminos inadmisibles (no todos los caminos que llegan a  $m' = a$  son ahora inadmisibles). El problema está vinculado con el cálculo de la probabilidad de que la partícula sea absorbida por la pared después de  $N$  pasos. Debemos tener en cuenta que de todos los caminos que llegan a la posición  $a$ , debemos quitar todos los que han atravesado la barrera anteriormente. Una forma de contar los caminos inadmisibles es la siguiente. Para llegar a  $a$  en “tiempo”  $N$  hay dos posibilidades. O bien a tiempo  $N - 1$  estamos en  $a + 1$  o en  $a - 1$ . Algunos de los caminos que llegan a  $a - 1$  son admisibles y otros son inadmisibles. Los que llegan a  $a + 1$  son todos inadmisibles. Los inadmisibles que llegan a  $a - 1$  son *la misma cantidad* que los que llegan a  $a + 1$  para tiempo  $N - 1$ . Podemos ver esto con un argumento de reflexión en el eje  $a$ . Si

reflejamos cada camino inadmisibles que llega a  $a - 1$  (a partir del primer contacto con el eje  $a$ ) obtenemos otro que llega a  $a + 1$  al mismo tiempo. De este modo, la probabilidad que queremos será la de los caminos libres que llegan a  $a$  en tiempo  $N$  menos el doble de los caminos que llegan a  $a + 1$  en tiempo  $N - 1$ :

$$\begin{aligned} \text{número de caminos} &= \frac{N!}{\left(\frac{N+a}{2}\right)!\left(\frac{N-a}{2}\right)!} - 2 \frac{(N-1)!}{\left(\frac{(N-1)+(a+1)}{2}\right)!\left(\frac{(N-1)-(a+1)}{2}\right)!} \\ &= \frac{N!}{\left(\frac{N+a}{2}\right)!\left(\frac{N-a}{2}\right)!} \left(1 - 2 \frac{\binom{N-a}{2}}{N}\right) \\ &= \frac{a}{N} \frac{N!}{\left(\frac{N+a}{2}\right)!\left(\frac{N-a}{2}\right)!} \end{aligned}$$

La probabilidad de ser absorbido en la posición  $a$ , a tiempo  $N$  es entonces el número anterior dividido por  $2^N$

$$w_{abs}(a, N; a) = \frac{a}{N} \frac{N!}{2^N \left(\frac{N+a}{2}\right)!\left(\frac{N-a}{2}\right)!} = \frac{a}{N} w(a, N).$$

En el caso de  $N$  grande tenemos la fórmula asintótica

$$w_{abs}(a, N; a) \sim \frac{a}{N} \left(\frac{2}{\pi N}\right)^{\frac{1}{2}} e^{(-a^2/2N)},$$

que expresada en términos de  $x_a = ah$  y  $t = Nk$  viene dada por:

$$\begin{aligned} p(x_a, t) &= \frac{x_a k}{t} \left(\frac{2k}{\pi h^2 t}\right)^{\frac{1}{2}} e^{(-x_a^2/4Dt)} \\ &= \frac{x_a k}{t} \left(\frac{1}{\pi Dt}\right)^{\frac{1}{2}} e^{(-x_a^2/4Dt)} \end{aligned}$$

**Observación:** la función  $p$  no es la densidad de probabilidad, es simplemente una especie de interpolación de la función  $w_{abs}$ . La densidad de probabilidad  $q(x_a, t)$  respecto del tiempo la calculamos a continuación. La probabilidad de encontrar a la partícula en  $x_a$  en el intervalo de tiempo  $[t, t + \Delta t]$  es

$$\int_t^{t+\Delta t} q(x_a, t) dt \approx p(x_a, t) \frac{\Delta t}{2k} = \frac{x_a}{2t} \left(\frac{1}{\pi Dt}\right)^{\frac{1}{2}} e^{(-x_a^2/4Dt)} \Delta t$$

donde tuvimos que dividir por dos, ya que los tiempos posibles de llegada, fijado  $x_a$ , están definidos en intervalos de tamaño  $2k$  (dependiendo de la paridad de  $m, N$ ). Haciendo  $\Delta t \rightarrow 0$  obtenemos

$$q(x_a, t) = \frac{x_a}{2t} \left(\frac{1}{\pi Dt}\right)^{\frac{1}{2}} e^{(-x_a^2/4Dt)}$$

Destaquemos la siguiente relación

$$q(x_a, t) = -D \frac{\partial}{\partial x} (u_{abs}(x, t)) \Big|_{x=x_a}.$$

Si interpretamos la probabilidad como frecuencia a la que ocurre un suceso, hemos obtenido la fracción de partículas (de un número muy grande de ellas) que salen de  $x = 0$  y llegan por primera vez a la pantalla absorbente  $x_a$  por unidad de tiempo.

## 2.4. Una ecuación en diferencias para la densidad de probabilidad

Vamos a abordar el problema del cálculo de la densidad de probabilidad desde otro punto de vista, sin utilizar la solución explícita para las probabilidades obtenida en la sección anterior. En primer lugar observemos que la función  $w(m, N)$  satisface una ecuación en diferencias. La idea básica es la siguiente: la probabilidad de que la partícula llegue a  $x = mh$  en el “tiempo”  $t = (N + 1)k$  puede calcularse como la probabilidad de que llegue hasta  $(m - 1)h$  en el tiempo  $Nk$  por la probabilidad de desplazarse hacia la derecha en el siguiente paso temporal, más la probabilidad de que llegue a  $(m + 1)h$  en el instante  $Nk$  por la probabilidad de que se desplace hacia la izquierda en el siguiente paso. Como cada paso es una variable discreta independiente del movimiento anterior tenemos que:

$$w(m, N + 1) = \frac{1}{2}w(m - 1, N) + \frac{1}{2}w(m + 1, N), \quad (10)$$

donde un medio es la probabilidad de desplazarse a derecha o izquierda. Cuando  $m, N \rightarrow \infty$  tenemos que  $w(m, N) \rightarrow 0$  ya que si en el caso discreto tenemos cada vez más posiciones posibles, la probabilidad de que alguna de ellas ocurra va a tender a cero necesariamente. Por esto es mejor que consideremos una aproximación a la densidad de probabilidad  $u(x, t)$ :

$$u_{h,k}(x, t) := \frac{1}{2h}w(x/h, t/k). \quad (11)$$

La idea central que nos guía es encontrar una función de densidad límite para  $h, k \rightarrow 0$ . De momento, vamos a suponer que este límite existe y lo denotamos por  $u(x, t)$ :

$$u_{h,k}(x, t) = u(x, t) + \varepsilon_{h,k}(x, t) \quad \varepsilon_{h,k}(x, t) \rightarrow 0, \text{ para } h, k \rightarrow 0.$$

Si reemplazamos los valores correspondientes de  $u$  en la ecuación (10) tenemos que

$$u_{h,k}(x, t + k) = \frac{1}{2}u_{h,k}(x - h, t) + \frac{1}{2}u_{h,k}(x + h, t). \quad (12)$$

Como veremos más adelante, (12) es un esquema numérico que aproxima a la solución de una ecuación diferencial. En este caso tenemos el problema inverso al cálculo numérico: dada la ecuación en diferencias intentaremos determinar la ecuación diferencial de la función límite. Para hacer esto necesitamos saber con qué ecuación es consistente este esquema, y debemos reemplazar la solución continua  $u$  en el esquema dado. Aproximando las cantidades utilizando el desarrollo de Taylor:

$$\begin{aligned} u(x, t + k) &= u(x, t) + \frac{\partial u}{\partial t}(x, t)k + O(k^2) \\ \frac{1}{2}(u(x - h, t) + u(x + h, t)) &= u(x, t) + \frac{1}{2}\frac{\partial^2 u}{\partial x^2}(x, t)h^2 + O(h^4) \end{aligned}$$

(nótese el error de orden 4 debido a la cancelación de las potencias impares). Igualando las cantidades obtenidas tenemos que

$$\frac{\partial u}{\partial t}(x, t) + O(k) = \frac{h^2}{2k}\frac{\partial^2 u}{\partial x^2}(x, t) + O\left(\frac{h^4}{k}\right).$$

Si tomamos ahora límite para  $h, k \rightarrow 0$  de modo que el cociente  $D = \frac{h^2}{2k}$  permanezca constante, obtenemos la ecuación del calor unidimensional para la densidad de probabilidad  $u$ :

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2}.$$

Obsérvese que si  $h^2/k \rightarrow 0$  la ecuación es consistente con  $\partial u/\partial t = 0$ , que corresponde a un estado estacionario de la ecuación del calor.

Con esta forma de ver las cosas es relativamente sencillo imponer condiciones para tener una pared reflejante o absorbente. De acuerdo con la condición de reflexión la única posibilidad de llegar a la pared es desde la izquierda (en el tiempo anterior) con probabilidad  $\frac{1}{2}$ , de modo que:

$$w_{ref}(r, N; r) = \frac{1}{2}w_{ref}(r-1, N-1; r).$$

Por otra parte, la probabilidad de llegar en tiempo  $N+1$  a la posición  $r-1$  es:

$$w_{ref}(r-1, N+1; r) = \frac{1}{2}w_{ref}(r-2, N; r) + w_{ref}(r, N; r),$$

ya que tenemos probabilidad 1 de rebotar una vez que estamos en la posición  $r$ . Eliminando  $w_{ref}(r, N; r)$  de ambas ecuaciones obtenemos:

$$w_{ref}(r-1, N+1; r) = \frac{1}{2}w_{ref}(r-2, N; r) + \frac{1}{2}w_{ref}(r-1, N-1; r).$$

En términos de la densidad continua escribimos la ecuación anterior:

$$u_{ref}(x_r - h, t + k) = \frac{1}{2}u_{ref}(x_r - 2h, t) + \frac{1}{2}u_{ref}(x_r - h, t - k).$$

Obsérvese que usamos la densidad para puntos interiores, ya que la relación entre densidad y probabilidad discreta no es obvia para los puntos del borde (desde este punto de vista). Aproximando por Taylor obtenemos la identidad (todas las funciones están evaluadas en  $(x_r, t)$  y quitamos el *ref* del subíndice):

$$u - \frac{\partial u}{\partial x}h + \frac{\partial u}{\partial t}k + o(\sqrt{k^2 + h^2}) = \frac{1}{2}u - \frac{\partial u}{\partial x}h + \frac{1}{2}u - \frac{1}{2}\frac{\partial u}{\partial x}h - \frac{1}{2}\frac{\partial u}{\partial t}k + o(\sqrt{h^2 + k^2}),$$

y teniendo en cuenta que el límite es calculado para  $k = O(h^2)$ :

$$\frac{\partial u}{\partial x}h = O(h^2) \Rightarrow \frac{\partial u}{\partial x} = 0.$$

Esta es la misma condición, de tipo Neumann, que observamos para la solución explícita.

En el caso de la pared absorbente debemos observar que:

$$w_{abs}(a-1, N+1; a) = \frac{1}{2}w_{abs}(a-2, N; a),$$

ya que la probabilidad de que venga por la derecha (desde  $a$ ) es igual a cero con la pared absorbente. Usando la densidad continua:

$$u_{abs}(x_a - h, t + k) = \frac{1}{2}u_{abs}(x_a - 2h, t),$$

y aproximando por Taylor (funciones evaluadas en  $(x_a, t)$ ):

$$u + O(h+k) = \frac{1}{2}u + O(h+k),$$

por lo tanto

$$u = 0,$$

como obtuvimos anteriormente con la solución explícita para  $u_{abs}$ .

Resumiendo, hemos encontrado que, haciendo un camino aleatorio de paso espacial  $h$  y paso temporal  $k$ , la densidad de probabilidad de encontrar a una partícula para tiempo  $t = Nk$  en la posición  $x = mh$ , cuando  $m, N \rightarrow \infty$  con  $h^2/(2k) = D$ , es consistente con ecuación del calor

$$u_t = Du_{xx}. \quad (13)$$

Cuando hay una pared reflectante en  $x = x_r$  debemos añadir la condición de contorno

$$\frac{\partial u}{\partial x}(x_r, t) = 0$$

y cuando hay una pared absorbente tenemos la condición

$$u(x_a, t) = 0.$$

En ambos casos la condición inicial para  $u$  es una *delta de Dirac*. Es decir, para  $t = 0$  toda la probabilidad está concentrada en el origen, la densidad es cero para  $x \neq 0$  pero su integral en cualquier intervalo alrededor del origen es igual a 1. Esta es la solución fundamental de la ecuación del calor, podemos reconstruir a partir de ella cualquier condición inicial y también utilizar las ideas aplicadas para caminos discretos con paredes reflejantes y absorbentes.

## 2.5. La caminata aleatoria como límite de pasos independientes: El teorema de Laplace-De Moivre

Podemos obtener la distribución límite de la caminata aleatoria a través de un resultado clásico de suma de variables aleatorias discretas, conocido como el teorema de Laplace-De Moivre. La demostración es esencialmente la misma que la vista anteriormente para los “random-walks” (utilizando también el comportamiento asintótico de los factoriales).

(Laplace-De Moivre). Sean  $I_n$ ,  $n \in \mathbb{N}$ , variables aleatorias discretas independientes que toman valores 0 y 1 y tales que

$$\begin{cases} P(I_n = 1) &= p \\ P(I_n = 0) &= q \end{cases}$$

con  $p + q = 1$ . Si  $S_N := \sum_{n=1}^N I_n$ , entonces

$$\lim_{N \rightarrow \infty} P\left(a \leq \frac{S_N - Np}{\sqrt{Npq}} \leq b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx.$$

**Ejercicio 11.** Demuestra el teorema de Laplace-De Moivre siguiendo los siguientes pasos:

1) Verifica que la esperanza y la varianza de  $S_N$  vienen dadas por:

$$\begin{aligned} E[S_N] &= Np \\ V[S_N] &= Npq \end{aligned}$$

2) Identifica los posibles valores de la variable  $S_N$ , muestra que se satisface:

$$P(S_N = m) = p^m q^{(N-m)} \frac{N!}{(N-m)!m!} \quad (14)$$

3) Considera ahora los valores posibles de la variable  $\Xi_N = \frac{S_N - Np}{\sqrt{Npq}}$ , escribe

$$x = \frac{m - Np}{\sqrt{Npq}}.$$

Observa que si se deja fijo el valor de  $x$  y hacemos  $N \rightarrow \infty$ , entonces  $m \rightarrow \infty$ . Aplica la fórmula de Stirling (2) a (14) justificando su uso para  $x$  fijo,  $N \rightarrow \infty$ .

4) Escribe (14) en términos de  $x$  y  $N$ , y desarrollando los exponentes hasta orden 2 demuestra que:

$$P(S_N = m) = P(\Xi_N = x) = \frac{1}{\sqrt{2\pi Npq}} e^{-\frac{x^2}{2}} (1 + o(1))$$

Nota que:

$$1 - \frac{m}{N} = q \left(1 - \sqrt{\frac{p}{qN}} x\right), \quad \frac{m}{N} = p \left(1 + \sqrt{\frac{q}{pN}} x\right).$$

5) Suma las contribuciones del intervalo, teniendo en cuenta que la distancia entre los valores

$$x_m = \frac{m - Np}{\sqrt{Npq}}$$

es

$$\Delta x_m = x_{m+1} - x_m = \frac{1}{\sqrt{npq}}.$$

Si sumamos sobre los  $x_m$  que entran en el intervalo tenemos que

$$P(a \leq \Xi_N \leq b) = \frac{1}{\sqrt{2\pi}} \sum_{x_m \in [a,b]} e^{-\frac{x_m^2}{2}} \Delta x_m \rightarrow \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx \quad N \rightarrow \infty$$

Para obtener la distribución límite de la caminata aleatoria definimos ahora variables *independientes*  $B_n$  que pueden tomar los valores 1 y  $-1$  con probabilidad  $1/2$  respectivamente. La variable  $I_n = \frac{B_n+1}{2}$  satisface las condiciones del teorema, con  $p = q = 1/2$ . En este caso  $\sum_{n=1}^N I_n = \frac{1}{2} \sum_{n=1}^N B_n + \frac{N}{2}$  y el teorema nos dice que:

$$\lim_{N \rightarrow \infty} P\left(a \leq \frac{\sum_{n=1}^N B_n}{\sqrt{N}} \leq b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx.$$

Si cambiamos la escala de los pasos con una constante  $h$  y definimos  $k = t/N$  como el paso temporal, podemos escribir esto del siguiente modo

$$\lim_{N \rightarrow \infty} P\left(a\sqrt{t}\frac{h}{\sqrt{k}} \leq \sum_{n=1}^N hB_n \leq b\sqrt{t}\frac{h}{\sqrt{k}}\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx.$$

Llamando  $X_N$  a  $\sum_{n=1}^N hB_n$ , tenemos que

$$P\left(a\sqrt{t}\frac{h}{\sqrt{k}} \leq X_N \leq b\sqrt{t}\frac{h}{\sqrt{k}}\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx + \varepsilon_N,$$

donde  $\varepsilon_N \rightarrow 0$  para  $N \rightarrow \infty$  con  $a$  y  $b$  fijos. Obsérvese que si  $\frac{h}{\sqrt{k}} \rightarrow 0$  para  $N \rightarrow \infty$ , la densidad de  $X_N$  estará concentrada alrededor del origen. Eligiendo  $a, -b \gg 1$  tales que  $\frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx$  esté tan cerca del valor 1 como queramos y  $N$  suficientemente grande tal que  $-\varepsilon < [a\sqrt{t}\frac{h}{\sqrt{k}}, b\sqrt{t}\frac{h}{\sqrt{k}}] \subset (-\varepsilon, \varepsilon)$ , tenemos que:

$$P(-\varepsilon \leq X_N \leq \varepsilon) \geq P\left(a\sqrt{t}\frac{h}{\sqrt{k}} \leq X_N \leq b\sqrt{t}\frac{h}{\sqrt{k}}\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx + \varepsilon_N = 1 - \delta + \varepsilon_N,$$

donde  $\delta$  puede hacerse arbitrariamente pequeño eligiendo  $a$  y  $b$  convenientemente. Concluimos que

$$\lim_{N \rightarrow \infty} P(-\varepsilon \leq X_N \leq \varepsilon) = 1 \quad \text{si} \quad \frac{h}{\sqrt{k}} \rightarrow 0.$$



Obsérvese que esta es la escala que corresponde a la ley de los grandes números, donde tomaríamos  $h = 1/N = O(k)$ .

Si por el contrario tenemos que  $\frac{h}{\sqrt{k}} \rightarrow \infty$ , elegimos  $a \leq 0 \leq b$  de modo que  $\frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx < \delta$ , con  $\delta$  pequeño. En este caso, dados  $A, B$  arbitrarios (fijos), tendremos que para un  $N$  suficientemente grande  $[A, B] \subset (a\sqrt{t}\frac{h}{\sqrt{k}}, b\sqrt{t}\frac{h}{\sqrt{k}})$  y por lo tanto

$$P(A \leq X_N \leq B) \leq P\left(a\sqrt{t}\frac{h}{\sqrt{k}} \leq X_N \leq b\sqrt{t}\frac{h}{\sqrt{k}}\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx + \varepsilon_N = \delta + \varepsilon_N,$$

de modo que

$$\lim_{N \rightarrow \infty} P(A \leq X_N \leq B) \leq \delta$$

para  $\delta$  arbitrariamente pequeño. Podemos decir entonces que

$$\lim_{N \rightarrow \infty} P(A \leq X_N \leq B) = 0 \quad \text{si} \quad \frac{h}{\sqrt{k}} \rightarrow \infty.$$

Los casos considerados son límites que nos llevan a distribuciones de algún modo triviales. La primera, que concentra toda la medida en el origen, es una delta de Dirac. La segunda dispersa las posiciones de tal modo que la probabilidad de encontrarlas en un intervalo finito es nula. El caso que más nos interesa es el que viene a continuación.

Supongamos ahora que  $\lim_{N \rightarrow \infty} \frac{h}{\sqrt{k}} = \sqrt{2D} > 0$ . Si definimos  $A := a\sqrt{t}\frac{h}{\sqrt{k}} = a\sqrt{2Dt}$ ,  $B := b\sqrt{t}\frac{h}{\sqrt{k}} = b\sqrt{2Dt}$ , tenemos que

$$\begin{aligned} P(A \leq X_N \leq B) &= \frac{1}{\sqrt{2\pi}} \int_{B/\sqrt{2Dt}}^{A/\sqrt{2Dt}} e^{-\frac{x^2}{2}} dx + \varepsilon_N \\ &= \frac{1}{\sqrt{4\pi Dt}} \int_B^A e^{-\frac{x^2}{4Dt}} dx + \varepsilon_N \end{aligned}$$

de modo que

$$\lim_{N \rightarrow \infty} P(A \leq X_N \leq B) = \frac{1}{\sqrt{4\pi Dt}} \int_B^A e^{-\frac{x^2}{4Dt}} dx. \quad (15)$$

Este resultado es idéntico al que obtuvimos anteriormente y es el que corresponde al límite difusivo.

### 3. El proceso de Wiener.

#### 3.1. Definición y propiedades básicas del proceso de Wiener.

Hemos analizado uno de los procesos más simples en una dimensión: la caminata aleatoria (random walk). Hemos visto también que el límite de las distribuciones de probabilidad discretas puede llevarse al continuo por medio de un límite adecuado en el tamaño de los pasos espacial y temporal. Las escalas entre ambos pasos están vinculadas de una manera particular para que en el límite obtengamos una distribución de probabilidad no trivial. Concretamente, hemos visto que  $h$  debe ser del orden de  $\sqrt{k}$ , donde  $k$  es el paso temporal y  $h$  el espacial. Si elegimos  $h = \sqrt{k}$  ( $D = 1/2$ ), en el límite podemos generar un proceso continuo, o una familia de variables aleatorias reales  $W_t^2$ , indexadas por el tiempo  $t$ , que tiene las siguientes propiedades:

---

<sup>2</sup>Como es costumbre en estadística, utilizaremos letras mayúsculas para indicar a las variables aleatorias y minúsculas para indicar un valor particular de las mismas.

1.  $W_0 = 0$ .
2.  $W_t$  es una variable aleatoria con distribución normal,  $N(0, \sqrt{t})$ .
3. Si  $t > s$ ,  $W_t - W_s$  es independiente de  $W_s$  y tiene distribución  $N(0, \sqrt{t-s})$ . **Importante:** Esta propiedad no nos dice que  $W_t$  sea independiente de  $W_s$ , sino que el *incremento*  $W_t - W_s$  lo es. Las variables  $W_t$  y  $W_s$  tienen correlación no nula. Para fijar estas ideas es mejor tener en mente el mecanismo de construcción del proceso.

Las propiedades 1-3 pueden obtenerse a partir de los procesos discretos de caminata aleatoria que hemos definido en el apartado anterior. Escribamos nuevamente al proceso discreto como suma de variables aleatorias independientes, de tipo Bernouilli:

$$X_N = \sum_{n=1}^N \sqrt{k} B_n,$$

donde  $k := t/N$  es el “paso temporal” y  $B_n$ ,  $1 \leq n \leq N$  son variables aleatorias *independientes* que pueden tomar los valores 1 y  $-1$  con probabilidad  $1/2$  respectivamente. La media de  $X_N$  es cero para todo  $N$  y su varianza es la suma de las varianzas de cada incremento, es decir  $V[X_N] = NV[\sqrt{k}B_1] = Nk = t$ . Cuando  $N \rightarrow \infty$  (dejando a  $t$  constante, con  $k \rightarrow 0$ ), la variable  $W_t := \lim_{N \rightarrow \infty} X_N$  es una suma independiente de variables equidistribuidas y el teorema de Laplace-De Moivre (cf. (15)) nos dice que  $W_t$  tiene una distribución normal  $N(0, \sqrt{t})$ . Nótese que el mismo límite puede obtenerse sumando *variables normales* adecuadamente escaladas en cada paso en lugar de las  $B_n$  discretas. Supongamos que  $Z_n$  son variables normales de media 0 y varianza 1, y construimos el proceso (ahora con una distribución continua de posibles valores de la posición, pero con valores temporales discretos)

$$X_N = \sum_{n=1}^N \sqrt{k} Z_n \quad Z_n \stackrel{\text{dist}}{=} N(0, 1).$$

Es fácil mostrar que la suma de variables normales independientes sigue siendo normal, con media igual a la suma de las medias respectivas y varianza igual a la suma de las varianzas respectivas. En este caso,  $X_N$  tendrá una distribución normal de media cero y varianza  $t$  ¡para todo  $N$ ! La misma distribución límite puede obtenerse con la suma de variables independientes, idénticamente distribuidas y con varianza finita. Este resultado es conocido como *teorema central del límite*.

Para ver la propiedad número 3 del proceso tenemos que tener en cuenta que los incrementos a partir de un valor de tiempo  $s$  son todos independientes del valor del proceso a tiempo  $s$ , de modo que la variable  $W_t$  puede escribirse como:

$$W_t = W_s + \lim_{N \rightarrow \infty} \sum_{n=1}^N \sqrt{k} B_n, \quad k = (t-s)/N.$$

Dado que todas las pruebas  $B_n$  posteriores al tiempo  $s$  son independientes de la variable  $W_s$ , tenemos que  $W_t - W_s$  es independiente de  $W_s$  y que su distribución es normal con media cero.

Este proceso básico, que hemos estudiado desde distintos puntos de vista, tiene el nombre de *proceso de Wiener*, y es fundamental para la construcción de otros procesos en distintas disciplinas (física, química, biología, finanzas, etc.) El incremento de  $W_t$  en un intervalo  $[t, t+\Delta t]$  está dado por

$$\Delta W_t := W_{t+\Delta t} - W_t = \sqrt{\Delta t} Z_t, \quad Z_t \stackrel{\text{dist}}{=} N(0, 1).$$

Podemos escribir, para  $\Delta t \sim 0$ :

$$dW_t = \sqrt{dt}Z_t. \quad (16)$$

Una aclaración sobre el uso de subíndices en (16). El incremento infinitesimal del proceso  $W_t$  se obtiene realizando una muestra de una variable normal *independiente del valor actual de  $W_t$* , multiplicada por la raíz del incremento infinitesimal  $dt$ . El subíndice en  $Z_t$  se utiliza para indicar que la prueba aleatoria debe realizarse en tiempo  $t$  y que esta prueba es independiente de las que deban realizarse para calcular incrementos en otros valores de tiempo. Como variables aleatorias son idénticas (están idénticamente distribuidas) pero es importante comprender que la infinidad de pruebas de las variables  $Z_t$  que deben hacerse para construir el proceso  $W_t$  son todas independientes. Dicho de otro modo y salvando las diferencias, es como si tuviéramos que arrojar una moneda en cada instante de tiempo para calcular una muestra de la evolución del proceso. Como veremos a continuación, esto es equivalente a definir una medida de probabilidad en un espacio de caminos continuos.

### 3.2. La medida de Wiener

En el apartado anterior hemos definido el proceso de Wiener como el límite de una suma de variables aleatorias de un modo parecido al que utilizamos para integrar una función continua: Sumamos una gran cantidad de variables de varianza pequeña. Conocemos la distribución límite de estas variables, pero no sabemos si les corresponde algún espacio de probabilidad, con una  $\sigma$ -álgebra de conjuntos medibles bien definida. Veamos en primer lugar que el mismo proceso de construcción, y la escala utilizada para añadir las pruebas independientes, hacen que una muestra del proceso de Wiener sea una función continua (en un sentido probabilístico). Calculemos la siguiente probabilidad:

$$P(|W_{t+\Delta t} - W_t| > \varepsilon) = P(|\sqrt{\Delta t}Z_t| > \varepsilon) = \frac{1}{\sqrt{2\pi\Delta t}} \int_{|x|>\varepsilon} e^{-\frac{x^2}{2\Delta t}} dx$$

Por otra parte, para  $\varepsilon/\sqrt{\Delta t} > 1$ :

$$\int_{|x|>\varepsilon} e^{-\frac{x^2}{2\Delta t}} dx = 2 \int_{\varepsilon}^{\infty} e^{-\frac{x^2}{2\Delta t}} dx \leq 2 \int_{\varepsilon}^{\infty} e^{-\frac{x}{2\sqrt{\Delta t}}} dx = \frac{e^{-\frac{\varepsilon}{2\sqrt{\Delta t}}}}{\sqrt{\Delta t}}$$

De modo que:

$$P(|W_{t+\Delta t} - W_t| > \varepsilon) = o((\Delta t)^n) \quad \forall n \geq 0, \Delta t \rightarrow 0, \quad (17)$$

donde la última igualdad se obtiene teniendo en cuenta que  $\lim_{u \rightarrow \infty} u^n e^{-u} = 0$  para todo  $n > 0$ . No debemos confundir (17) con la noción de regularidad de una función en el sentido usual. Si bien esta relación nos dice que la probabilidad de dar un salto es de un orden arbitrariamente pequeño respecto de  $\Delta t$ , esto no implica regularidad alguna. Es decir, puede haber una muestra límite (que tendríamos que definir con precisión) que sea discontinua, ya que en principio *cualquier gráfica* se podría obtener al sumar la infinidad de variables independientes. La noción de la regularidad de estas muestras no tiene sentido hasta que no definamos un ambiente natural en el que deben vivir los procesos límite y le asignemos una medida de probabilidad. En ese caso podríamos intentar demostrar una afirmación del tipo: “el conjunto de funciones discontinuas tiene medida cero”, o “el conjunto de funciones diferenciables tiene medida cero”.

**Construcción de la medida de Wiener.** El proceso de construcción de la medida se basa en el hecho de que la familia de variables  $W_t$  está incluida de forma natural dentro del conjunto de funciones reales definidas en la semirrecta real positiva. Es decir, una *muestra* del proceso límite es como asignar un valor real a cada  $t > 0$ , y en ese sentido nuestro espacio muestral debería ser el conjunto enorme de las funciones reales definidas en  $[0, +\infty)$ . Sin embargo, teniendo

en cuenta (17) es posible restringirse al conjunto  $\Omega$  de funciones continuas  $\omega : [0, +\infty) \mapsto \mathbb{R}$ , tales que  $\omega(0) = 0^3$ . De hecho, N. Wiener asumió la continuidad como un postulado más en su trabajo original.

Imaginemos ahora que, dada una muestra  $\omega \in \Omega$ , la variable  $W_t : \omega \mapsto \mathbb{R}$  nos devuelve el valor de la función continua  $\omega$  en  $t$ :

$$W_t(\omega) = \omega(t).$$

La distribución de probabilidad de  $W_t$  nos permite calcular la probabilidad de encontrar a la variable en un intervalo real determinado. Esta *distribución marginal* nos permitiría definir la medida para unos subconjuntos de  $\Omega$  muy particulares:

$$P(a \leq W_t \leq b) = \text{medida de los } \omega \text{ que pasan por el intervalo } [a, b] \text{ en } t.$$

El conjunto

$$C_{[a,b];t} := \{\omega \in \Omega : a \leq \omega(t) \leq b\}$$

se llama *conjunto cilíndrico* o *ventana* y su medida de probabilidad la definimos entonces como:

$$P(C_{[a,b];t}) := \int_a^b \phi(x, t) dx, \quad \phi(x, t) := \frac{e^{-\frac{x^2}{2t}}}{\sqrt{2\pi t}}, \quad (18)$$

esto es, usando la distribución de la variable  $W_t$ , que es normal de varianza  $t$  y media cero.

Para definir la probabilidad de la intersección de dos ventanas  $C_{[a_1, b_1]; t_1}$  y  $C_{[a_2, b_2]; t_2}$  debemos tener en cuenta la distribución conjunta de las variables  $W_{t_1}$  y  $W_{t_2}$  ( $t_1 < t_2$ ) caracterizada por la independencia del incremento  $\Delta W_{t_1} := W_{t_2} - W_{t_1}$  respecto de  $W_{t_1}$ . La probabilidad que queremos medir es la de las funciones continuas que pasan en  $t_1$  por un intervalo  $[a_1, b_1]$  y en  $t_2$  por un intervalo  $[a_2, b_2]$ . Para ello notemos que:

$$\begin{aligned} P(a_1 \leq W_{t_1} \leq b_1, a_2 \leq W_{t_2} \leq b_2) &\stackrel{(1)}{=} \int_{a_1}^{b_1} \int_{a_2}^{b_2} p(x_1, t_1; x_2, t_2) dx_1 dx_2 \\ &= \int_{a_1}^{b_1} \int_{a_2}^{b_2} p(x_2, t_2 | x_1, t_1) p(x_1, t_1) dx_1 dx_2 = \int_{a_1}^{b_1} p(x_1, t_1) \left( \int_{a_2}^{b_2} p(x_2, t_2 | x_1, t_1) dx_2 \right) dx_1 \end{aligned}$$

donde en la igualdad (1) usamos la densidad de la distribución conjunta

$$p(x_1, t_1; x_2, t_2) dx_1 dx_2 := P(x_1 \leq W_{t_1} \leq x_1 + dx_1, x_2 \leq W_{t_2} \leq x_2 + dx_2).$$

La independencia de los incrementos nos dice que

$$\begin{aligned} p(x_2, t_2 | x_1, t_1) dx_2 &= P(x_2 \leq W_{t_2} \leq x_2 + dx_2 | W_{t_1} = x_1) \\ &= P(x_2 - x_1 \leq \Delta W_{t_2 - t_1} \leq x_2 - x_1 + dx_2) \\ &= \phi(x_2 - x_1, t_2 - t_1), \end{aligned} \quad (19)$$

donde  $\Delta W_{t_2 - t_1} := W_{t_2} - W_{t_1}$ . Finalmente:

$$P(C_{[a_1, b_1]; t_1} \cap C_{[a_2, b_2]; t_2}) := \int_{a_1}^{b_1} \phi(x_1, t_1) \left( \int_{a_2}^{b_2} \phi(x_2 - x_1, t_2 - t_1) dx_2 \right) dx_1.$$

Habiendo definido la probabilidad de la intersección de eventos elementales, tenemos entonces definida la de la unión:

$$P(C_{[a_1, b_1]; t_1} \cup C_{[a_2, b_2]; t_2}) = P(C_{[a_1, b_1]; t_1}) + P(C_{[a_2, b_2]; t_2}) - P(C_{[a_1, b_1]; t_1} \cap C_{[a_2, b_2]; t_2}).$$

---

<sup>3</sup>La demostración de esta afirmación nos llevaría un poco lejos a cuestiones técnicas de teoría de la medida. Véase [2].

Esto nos permite definir una medida en el álgebra generada por los conjuntos cilíndricos  $\mathcal{F}_0$  (uniones e intersecciones finitas de estos conjuntos). A partir de aquí podemos definir la sigma álgebra  $\mathcal{F}$  como la mínima que contiene a  $\mathcal{F}_0$  y la medida se extiende en forma continua a todos ellos. No entraremos en los detalles técnicos para mostrar que la medida está bien definida, véase [2] o [5].

La definición de esta medida, debida a Norbert Wiener, nos permite integrar funcionales definidos en  $\Omega$  (si probamos previamente que el funcional es medible...). Podemos interpretar, dado un funcional  $F(\omega)$ , a su integral

$$\int_{\Omega} F(\omega) dP_{\omega} \underset{\text{notación}}{=} \int_{\Omega} F(\omega) dW$$

como el valor esperado del funcional  $F$  cuando  $\omega$  recorre las funciones continuas en un intervalo dado.

EJEMPLO (Chorin-Hald) Dado  $T > 0$ , consideremos el funcional

$$F(\omega) := \omega(T)^2.$$

En este caso

$$\int_{\Omega} F(\omega) dW = \int_{-\infty}^{+\infty} x^2 P(\omega : x \leq \omega(T) < x + dx) = \int_{-\infty}^{+\infty} \phi(x, T) x^2 dx = T,$$

que es la varianza del proceso de Wiener en tiempo  $T$ .

## 4. Interpretación estocástica de esquemas numéricos para EDP'S.

### 4.1. La ecuación del calor: solución numérica.

En la sección 4.1 identificamos una ecuación de difusión como la ecuación consistente con el esquema en diferencias para las funciones de densidad  $u_{h,k}$  (11). En el apéndice A mostramos que esta ecuación tiene una solución explícita bien definida con una condición inicial de tipo delta de Dirac. En este apartado veremos que los valores de  $u_{h,k}$  pueden interpretarse como valores en los nodos de un *esquema numérico convergente* para la ecuación del calor. Esto cierra el método iniciado con la ecuación en diferencias para  $w(m, N)$  y nos aporta una nueva interpretación estocástica de la solución de la ecuación en derivadas parciales.

Consideramos como punto de partida a la ecuación del calor (13), con un coeficiente de difusión  $D > 0$ , y un mallado espacial de paso  $h$ , y temporal de paso  $k$ . Aproximamos los términos de la ecuación diferencial utilizando el desarrollo de Taylor para  $t > 0$ :

$$\frac{\partial u}{\partial t}(x, t) = \frac{u(x, t+k) - u(x, t)}{k} + O(k)$$

$$\frac{\partial^2 u}{\partial x^2}(x, t) = \frac{u(x+h, t) - 2u(x, t) + u(x-h, t)}{h^2} + O(h^2)$$

El *error de truncatura local* viene dado por:

$$\begin{aligned} \tau_m^n &= \left( \frac{u_m^{n+1} - u_m^n}{k} - D \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{h^2} \right) - \left( \frac{\partial u}{\partial t}(mh, nk) - D \frac{\partial^2 u}{\partial x^2}(mh, nk) \right) \\ &= O(k + h^2), \end{aligned}$$

donde  $u_m^n = u(mh, nk)$  (esto es válido para cualquier función suave sobre la malla, no necesariamente tiene que ser la solución de la ecuación del calor!). La propiedad  $\tau_m^n \rightarrow 0$  para  $h, k \rightarrow 0$  se denomina *consistencia* en análisis numérico.

Definamos entonces  $U_m^n$  como la solución del siguiente problema discreto ( $m \in \mathbb{Z}$ ,  $n \in \mathbb{N}_0$ ):

$$\frac{U_m^{n+1} - U_m^n}{k} = D \frac{U_{m+1}^n - 2U_m^n + U_{m-1}^n}{h^2}, \quad U_m^0 = u_0(mh),$$

o puesto en forma explícita:

$$U_m^{n+1} = U_m^n(1 - 2\lambda) + \lambda(U_{m+1}^n + U_{m-1}^n); \quad U_m^0 = u_0(mh), \quad (20)$$

donde  $\lambda := \frac{Dk}{h^2}$ . A partir del dato inicial  $U_m^0$ , podemos calcular los niveles temporales sucesivos por medio de la fórmula explícita (20). Nótese que para  $\lambda = 1/2$  nos queda exactamente el mismo esquema que el deducido para la función de probabilidad discreta  $w(m, N)$ . Independientemente de esto, veamos la convergencia del esquema numérico discreto. Sea  $e_m^n = u_m^n - U_m^n$ . La ecuación en diferencias para  $e$  viene dada por:

$$e_m^{n+1} = e_m^n(1 - 2\lambda) + \lambda(e_{m+1}^n + e_{m-1}^n) + k\tau_m^n, \quad e_m^0 = 0.$$

Intentamos comparar el error del paso temporal  $n + 1$  con el del paso  $n$ -ésimo. Si tomamos el error máximo para cada nivel de tiempo obtenemos:

$$e^n := \max_m(|e_m^n|),$$

y definiendo

$$\tau^n = \max_m|\tau_{m,n}|; \quad \tau = \max_n \tau^n.$$

Si  $1 - 2\lambda \geq 0$  (ó  $\lambda \leq \frac{1}{2}$ ) tenemos la siguiente acotación:

$$e^{n+1} \leq e^n + k\tau \leq e^{n-1} + 2k\tau \leq \dots \leq e^0 + (n + 1)k\tau.$$

Teniendo en cuenta que  $e^0 = 0$  y que nos interesa estudiar la convergencia del esquema para un instante de tiempo finito (ie.  $nk = t$ ,  $n \rightarrow \infty$ ,  $h \rightarrow 0$ ), tenemos que el esquema es convergente para  $\lambda \leq 1/2$ . En el caso límite  $\lambda = 1/2$ , como mencionamos anteriormente, el esquema discreto coincide con el de las probabilidades para un camino aleatorio unidimensional y podemos reinterpretar el resultado en términos estadísticos.

Realicemos algunas iteraciones “hacia atrás” del esquema discreto (20) para  $\lambda = 1/2$ :

$$\begin{aligned} U_m^n &= \frac{1}{2}(U_{m-1}^{n-1} + U_{m+1}^{n-1}) \\ &= \frac{1}{4}(U_{m-2}^{n-2} + 2U_m^{n-2} + U_{m+2}^{n-2}) \end{aligned}$$

Ya en estas dos iteraciones podemos observar que  $\frac{1}{2^j}$  multiplica a todos los términos, donde  $j$  es el número de iteraciones realizadas. Los coeficientes que acompañan a cada uno de los valores  $U_m^{n-2}$  corresponden al número de caminos que pueden realizarse dando dos pasos temporales, para llegar a  $U_m^n$  (véase la figura).

esqnumerico.eps

Desde la posición  $(m + 1, n - 2)$  no puede accederse a  $(m, n)$  en sólo dos pasos y por lo tanto le corresponde un cero en el coeficiente. A medida que seguimos iterando, el número de caminos partido por  $2^j$  no es otra cosa que la probabilidad de llegar al punto  $(i, n - j)$  desde la posición  $(m, n)$  cuando realizamos una caminata aleatoria con probabilidad  $1/2$  de ir hacia cualquiera de los costados. De este modo podemos escribir

$$U_m^n = \sum_i w(i - m, n) U_i^0. \quad (21)$$

Esta ecuación nos permite interpretar a  $U_m^n$  como el *valor esperado* de los valores iniciales  $U_i^0$  sobre el espacio de probabilidad de los caminos que unen el punto  $(m, n)$  con la recta del nivel inicial. Utilizando la fórmula asintótica para  $w(m, n)$  (véase (5)),  $U_i^0 = u(ih, 0)$ ,  $t = nk$ ,  $x = mh$ , y el hecho de que solamente la mitad de los puntos aportan al promedio, podemos escribir

$$U_{m,n} = \frac{1}{2} \sum_i \sqrt{\frac{2}{\pi t}} \frac{\sqrt{k}}{h} e^{-\frac{k(ih-x)^2}{2h^2t}} u(ih, 0)h.$$

En este caso  $\frac{k}{h^2} = \frac{1}{2D}$  y obtenemos

$$U_{x/h,t/k} = \frac{1}{2\sqrt{\pi Dt}} \sum_i e^{-\frac{(ih-x)^2}{4Dt}} u(ih, 0)h,$$

y utilizando el resultado de convergencia numérica para  $h, k \rightarrow 0$ :

$$u(x, t) = \frac{1}{2\sqrt{\pi Dt}} \int_{-\infty}^{+\infty} e^{-\frac{(y-x)^2}{4Dt}} u(y, 0)dy,$$

donde recuperamos la expresión calculada mediante la transformada de Fourier en A, pero con una nueva interpretación: La función  $\frac{e^{-\frac{(y-x)^2}{4Dt}}}{2\sqrt{\pi Dt}} dy$ , es una medida sobre la recta, para cada  $x, t$  fijos y la solución de la ecuación del calor es el *valor esperado* del dato inicial con respecto a esta densidad de probabilidad para cada  $x, t$ . Nótese que ahora estamos *mirando el problema hacia atrás*. Es decir, asignamos una densidad de probabilidad en el nivel de tiempo inicial recorriendo hacia atrás los caminos que salen de  $(x, t)$  y llegan al nivel inicial.

Por otra parte, teniendo en cuenta que la variable aleatoria es el proceso de Wiener a tiempo  $t$ , podemos interpretar el resultado anterior como el siguiente valor esperado:

$$u(x, t) = E [u_0(x + W_t)],$$

donde  $E$  es el valor esperado de la variable aleatoria  $W_t$  definida en el conjunto de caminos continuos que salen del origen y la integración se realiza respecto de la medida de Wiener.

## 4.2. La ecuación del calor con una fuente

Consideremos el siguiente problema de valor inicial (caso en que  $D = 1/2$ )

$$\frac{\partial u}{\partial t} = \frac{1}{2} \frac{\partial^2 u}{\partial x^2} + \eta(x), \quad u(x, 0) = \psi(x) \quad (22)$$

Aproximemos la ecuación por el esquema explícito en diferencias:

$$U_m^{n+1} = \lambda U_{m+1}^n + (1 - 2\lambda)U_m^n + \lambda U_{m-1}^n + k\eta_m \quad \lambda := \frac{1}{2} \frac{k}{h^2}.$$

Mostremos la convergencia del esquema, sea  $e_m^n := U_m^n - u_m^n$

$$e_m^{n+1} = \lambda e_{m+1}^n + (1 - 2\lambda) e_m^n + \lambda e_{m-1}^n - (u_m^{n+1} - u_m^n) + \lambda (u_{m+1}^n - 2u_m^n + u_{m-1}^n) + k\eta_m$$

Donde, al ser válida la consistencia del método

$$u_m^{n+1} - u_m^n = \frac{\partial u}{\partial t} k + o(k)$$

$$\lambda (u_{m+1}^n - 2u_m^n + u_{m-1}^n) = \frac{k}{2} \frac{\partial^2 u}{\partial x^2} + k o_h(1) \quad (o_h \rightarrow 0, h \rightarrow 0)$$

De este modo

$$\begin{aligned} - (u_m^{n+1} - u_m^n) + \lambda (u_{m+1}^n - 2u_m^n + u_{m-1}^n) &= \left( -\frac{\partial u}{\partial t} + \frac{1}{2} \frac{\partial^2 u}{\partial x^2} \right) k + o(k) + k o_h(1) \\ &= -k\eta_m + o(k) + k o_h(1) \end{aligned}$$

Obtenemos entonces:

$$e_m^{n+1} = \lambda e_{m+1}^n + (1 - 2\lambda) e_m^n + \lambda e_{m-1}^n + o(k) + k o_h(1)$$

Si  $\lambda \leq 1/2$  obtenemos la acotación, definiendo  $e^n = \max_m |e_m^n|$ :

$$e^{n+1} \leq e^n + o(k) \leq \dots \leq e^0 + (n+1)(o(k) + k o_h(1))$$

De modo que el esquema es convergente para  $h, k \rightarrow 0$  con  $\lambda \leq 1/2$ , es decir  $k \leq h^2$ .

Elegimos ahora  $\lambda = 1/2$  y escribimos algunos pasos del método numérico

$$\begin{aligned} U_m^n &= \frac{1}{2} U_{m+1}^{n-1} + \frac{1}{2} U_{m-1}^{n-1} + k\eta_m \\ &= \frac{1}{4} (U_{m+2}^{n-2} + 2U_m^{n-2} + U_{m-2}^{n-2}) + \frac{1}{2} k\eta_{m+1} + \frac{1}{2} k\eta_{m-1} + k\eta_m \\ &= E [U_{m+X_2}^{n-2}] + kE [\eta(x_m + hX_1)] + kE [\eta(x_m + hX_0)] \end{aligned}$$

donde

$$X_n = \sum_{i=1}^n B_i, \quad X_0 = 0 \tag{23}$$

es una caminata aleatoria con pruebas independientes dadas por:

$$B_i = \begin{cases} 1 & \text{con probabilidad } 1/2 \\ -1 & \text{con probabilidad } 1/2 \end{cases}$$

y  $E[\cdot]$  representa el valor esperado de la variable aleatoria correspondiente. Si seguimos este esquema hasta el nivel de tiempo inicial, podemos escribir

$$U_m^n = E [\psi(x_m + hX_n)] + \sum_{i=1}^n kE [\eta(x_m + hX_i)]$$

tomando límite para  $h, k \rightarrow 0$  con  $h^2/k = 1$ , podemos escribir

$$\begin{aligned} u(x, t) &= \int dW \psi(x + W_t(\omega)) + \int_0^t ds \int dW \eta(x + W_s(\omega)) \\ &= E \left[ \psi(x + W_t) + \int_0^t \eta(x + W_s) ds \right] \end{aligned} \tag{24}$$

donde, como antes, la integral se realiza en el espacio de caminos continuos definidos en el intervalo  $[0, t]$  respecto de la medida de Wiener. Esta representación es otro caso simple de la fórmula de Feynman-Kac.



## Verificación

La ecuación con fuente puede resolverse formalmente sumando una solución particular con valor inicial cero y una solución de la ecuación homogénea ( $\eta \equiv 0$ ). La solución particular viene dada por la expresión:

$$\begin{aligned} u_p(x, t) &= \int_0^t e^{\frac{(t-s)}{2} \frac{\partial^2}{\partial x^2}} \eta(x) ds \\ &= \int_0^t \left( \int_{-\infty}^{+\infty} \phi(x - x', t - s) \eta(x) dx' \right) ds \\ &\stackrel{h=x'-x, \tau=t-s}{=} \int_0^t \left( \int_{-\infty}^{+\infty} \phi(h, \tau) \eta(x + h) dh \right) d\tau \\ &= \int_0^t E[\eta(x + W_\tau)] d\tau \end{aligned}$$

Tenemos entonces la misma expresión que antes.

## 4.3. La ecuación del calor con potencial

Consideremos ahora la ecuación

$$\frac{\partial u}{\partial t} = \frac{1}{2} \frac{\partial^2 u}{\partial x^2} + V(x)u \quad u(x, 0) = \phi(x) \quad (25)$$

donde  $V$  representa un *potencial* que asumiremos acotado, esto es  $|V(x)| \leq M$ .

Proponemos la discretización numérica

$$\frac{U_m^{n+1} - U_m^n}{k} = \frac{1}{2} \frac{U_{m+1}^n - 2U_m^n + U_{m-1}^n}{h^2} + \frac{1}{2} (U_{m+1}^n V_{m+1} + U_{m-1}^n V_{m-1}) \quad (26)$$

El esquema del lado derecho está elegido especialmente para conseguir una interpretación estocástica más directa. Reescribimos la relación en la forma

$$U_m^{n+1} = \left( \lambda + \frac{k}{2} V_{m+1} \right) U_{m+1}^n + (1 - 2\lambda) U_m^n + \left( \lambda + \frac{k}{2} V_{m-1} \right) U_{m-1}^n \quad (27)$$

Veamos que el término nuevo en (26) no destruye la consistencia del método numérico:

$$\begin{aligned} \frac{1}{2} (u_{m+1}^n V_{m+1} + u_{m-1}^n V_{m-1}) &= \frac{1}{2} \left( u_m^n + h \frac{\partial u}{\partial x} + o(h) \right) \left( V_m + \frac{\partial V}{\partial x} h + o(h) \right) + \\ &\quad \frac{1}{2} \left( u_m^n - h \frac{\partial u}{\partial x} + o(h) \right) \left( V_m - \frac{\partial V}{\partial x} h + o(h) \right) \\ &= u_m^n V_m + o(h) \end{aligned}$$

La convergencia del esquema puede mostrarse siguiendo pasos análogos a los de la ecuación del calor con una fuente.

$$e_m^{n+1} = \left( \lambda + \frac{k}{2} V_{m+1} \right) e_{m+1}^n + (1 - 2\lambda) e_m^n + \left( \lambda + \frac{k}{2} V_{m-1} \right) e_{m-1}^n + o(k) + k o_h(1)$$

Podemos acotar asumiendo que  $\lambda \leq 1/2$  y que  $kM \leq 2\lambda$ , así garantizamos que los paréntesis sean positivos:

$$\begin{aligned}
e^{n+1} &\leq (1 + kM) e^n + o(k) + ko_h(1) \\
&\leq (1 + kM)^2 e^{n-1} + (1 + kM) (o(k) + ko_h(1)) + (o(k) + ko_h(1)) \\
&\dots \leq (1 + kM)^{n+1} e^0 + \frac{1 - (1 + kM)^{n+1}}{1 - (1 + kM)} (o(k) + ko_h(1)) \\
&= \frac{(1 + kM)^{n+1} - 1}{M} (o_k(1) + o_h(1)) \rightarrow \frac{e^{Mt} - 1}{M} (o_k(1) + o_h(1)) .
\end{aligned}$$

Donde vemos que el algoritmo es convergente para  $h, k \rightarrow 0$ .

Elegimos ahora  $\lambda = 1/2$  y obtenemos el esquema

$$U_m^n = \frac{1}{2} U_{m+1}^{n-1} (1 + kV_{m+1}) + \frac{1}{2} U_{m-1}^{n-1} (1 + kV_{m-1})$$

Escribamos

$$e^{kV_j} = 1 + kV_j + o(k)$$

de modo que el algoritmo nos queda

$$U_m^n = \frac{1}{2} U_{m+1}^{n-1} e^{kV_{m+1}} + \frac{1}{2} U_{m-1}^{n-1} e^{kV_{m-1}} + o(k)$$

donde  $o(k)$  está uniformemente acotado con  $|V| \leq M$ .

Una iteración más del algoritmo nos da la expresión:

$$\begin{aligned}
U_m^n &= \frac{1}{4} (U_{m+2}^{n-2} e^{kV_{m+2}} + U_m^{n-2} e^{kV_m}) e^{kV_{m+1}} + \frac{1}{4} (U_m^{n-2} e^{kV_m} + U_{m-2}^{n-2} e^{kV_{m-2}}) e^{kV_{m-1}} + o(k) \\
&= \frac{1}{4} U_{m+2}^{n-2} e^{kV_{m+1} + kV_{m+2}} + \frac{1}{4} U_m^{n-2} (e^{kV_{m+1} + kV_m} + e^{kV_{m-1} + kV_m}) + \frac{1}{4} U_{m-2}^{n-2} e^{kV_{m-1} + kV_{m-2}} + 2o(k)
\end{aligned}$$

Esto lo escribimos ahora en una notación más sugerente:

$$\begin{aligned}
U_m^n &= \sum_{\text{caminos } \omega} \frac{1}{2^2} U_{m+X_2(\omega)}^{n-2} e^{\sum_{i=1}^2 kV_{m+X_i(\omega)}} + 2o(k) \\
&= \sum_{\text{caminos } \omega} \frac{1}{2^n} U_{m+X_n(\omega)}^0 e^{\sum_{i=1}^n kV_{m+X_i(\omega)}} + no(k) \\
&\xrightarrow{k \rightarrow 0} \int_{\Omega} \psi(x_m + W_t(\omega)) e^{\int_0^t V(x_m + W_s(\omega)) ds} dW_{\omega}
\end{aligned}$$

donde  $X_i$ , es la caminata aleatoria discreta definida en (23) y denotamos con  $\omega$  a los caminos discretos de esta variable.

Utilizando el resultado de convergencia obtenemos la fórmula de Feynman-Kac para la solución de la ecuación del calor con potencial:

$$u(x, t) = \int_{\Omega} \psi(x + W_t(\omega)) e^{\int_0^t V(x + W_s(\omega)) ds} dW_{\omega} \quad (28)$$

## Observación

Vamos a interpretar el lado derecho de la fórmula (28). Suponiendo la existencia de una solución  $u$  de (25), construimos el siguiente proceso estocástico basado en el proceso de Wiener

$$\xi_s = u(x + W_s, t - s) e^{\int_0^s V(x+W_\tau) d\tau}$$

Para  $s = 0$  la variable  $\xi_0$  es determinista (esto es, toma un único valor con probabilidad 1) que viene dado por

$$\xi_0 = u(x, t).$$

Si dejamos evolucionar a  $\xi_s$  hasta  $s = t$  tenemos que

$$\xi_t = u(x + W_t, 0) e^{\int_0^t V(x+W_\tau) d\tau} = \psi(x + W_t) e^{\int_0^t V(x+W_\tau) d\tau}$$

Tomando ahora valor esperado sobre los caminos continuos obtenemos la expresión

$$E[\xi_t] = E[\psi(x + W_t) e^{\int_0^t V(x+W_\tau) d\tau}]$$

Si queremos que esta fórmula represente a la solución de (25), esto debe coincidir con  $u(x, t) = \xi_0$ . Es decir que

$$E[\xi_t] = \xi_0 \quad \forall t > 0.$$

Una forma de verificar esto es calcular la siguiente derivada

$$\frac{d}{ds} E[u(x + W_s, t - s) e^{\int_0^s V(x+W_\tau) d\tau}]$$

y comprobar que es cero para todo valor de  $s$ . En ese caso, la solución coincidirá con la fórmula de Feynman-Kac para la ecuación con potencial.

## 5. Introducción

Los conocimientos básicos de Teoría de las Probabilidades necesarios para seguir el curso pueden encontrarse en el capítulo 2 del libro [Evans]. Dicho texto también proporciona un enfoque complementario sobre algunos de los temas expuestos en estas notas.

En la sección primera recapitulemos brevemente algunos de los conceptos vistos en la parte primera del curso. A continuación, en la sección segunda describiremos con detalle la construcción de la medida de Wiener y del movimiento Browniano.

## 6. Construcción de una caminata aleatoria infinitesimal

Hemos considerado el movimiento de una partícula que describe una *caminata aleatoria* en la recta real  $\mathbb{R}$ . La posición  $X_n$  de la partícula tras  $n \in \mathbb{N}$  pasos temporales viene regida por las siguientes leyes:

1. En el instante inicial la partícula se encuentra en el origen:  $X_0 = 0$ .
2. A cada paso temporal, la partícula se mueve una distancia  $h > 0$ , fija, a la derecha o a la izquierda con probabilidad  $1/2$ . En otras palabras, para  $n \in \mathbb{N}$ ,

$$X_n = X_{n-1} + hB_n, \tag{29}$$

donde  $B_n$  es una variable aleatoria tal que

$$P(B_n = 1) = P(B_n = -1) = 1/2. \tag{30}$$

3. El incremento de la posición de la partícula en un paso de tiempo  $n$  es independiente de los incrementos correspondientes a los pasos  $1, \dots, n-1$ . Más precisamente, las variables aleatorias  $B_n$ ,  $n \in \mathbb{N}$ , son independientes dos a dos.

El problema principal al que hemos dedicado el primer capítulo del curso fue el de construir una ley de propagación  $W_t$ , definida para instantes de tiempo  $t \in [0, \infty)$ , que fuera una versión *infinitesimal*, de la caminata aleatoria. Con ello queremos decir:

- (i) para cada instante  $t > 0$  la partícula se desplaza hacia la derecha o la izquierda de forma aleatoria.
- (ii) al igual que ocurre con la caminata aleatoria, el incremento de la posición  $W_{t+h} - W_t$  a partir de un instante  $t$  es independiente del recorrido previo de la partícula  $W_s$  para  $s \leq t$ .

Dicho problema fue contemplado originalmente por Albert Einstein (1905) en su intento formular una teoría cuantitativa que describiera las observaciones realizadas, entre otros, por el botánico Robert Brown (alrededor de 1827) sobre el movimiento, aparentemente caótico, que describen pequeñas partículas de polen en suspensión sobre un líquido.<sup>4</sup> Debido a consideraciones físicas – las partículas no saltan instantáneamente de una posición a otra – requeriremos también que el proceso  $W_t$  satisfaga:

- (iii) los caminos  $(W_t)_{t \geq 0}$  son continuos.

Para construir  $W_t$  hicimos lo siguiente. Fijado  $t \in [0, \infty)$  dividimos el intervalo  $[0, t]$  en  $N \in \mathbb{N}$  partes iguales; escribiremos

$$k := t/N,$$

y, para  $n = 0, \dots, N$ ,

$$W_{nk}^k := X_n,$$

con la intención de definir  $W_t$  como el límite, cuando  $k \rightarrow 0^+$ , de los  $W_t^k = W_{Nk}^k$ .

En media, la posición  $W_t^k$  de la partícula es cero:

$$E [W_t^k] = 0.$$

No obstante, su desplazamiento cuadrático medio es:

$$E [(W_t^k)^2] = h^2 N = \frac{h^2}{k} t. \tag{31}$$

Así pues, si hacemos tender  $k$  a cero sin más, obtenemos que dicho desplazamiento converge a infinito. Si pretendemos obtener un límite no trivial, hemos de reescalar también el parámetro espacial  $h$ . Si hiciéramos tender  $h$  a cero de modo que  $h^2/k \rightarrow 0$  la identidad (31) sugiere que el proceso límite obtenido es determinista, en el sentido de que la partícula permanece indefinidamente en su posición inicial  $W_0 = 0$ .

Si elegimos  $h$  de modo que  $h^2/k \rightarrow 2D > 0$  entonces el teorema de Laplace-de Moivre garantiza que:

$$\lim_{\substack{h, k \rightarrow 0^+ \\ h^2/k \rightarrow 2D}} P(A \leq W_t^k \leq B) = \frac{1}{\sqrt{4\pi Dt}} \int_A^B e^{-\frac{x^2}{4Dt}} dx. \tag{32}$$

Por tanto, la ley de probabilidad de un eventual proceso límite  $W_t$  ha de ser una normal de media cero y desviación típica  $\sqrt{2Dt}$ . Lo hecho hasta ahora *no prueba* aún la existencia del

---

<sup>4</sup>Un comentario más detallado de los aspectos históricos del descubrimiento del movimiento Browniano puede encontrarse en el libro [Nelson67].

proceso límite  $W_t$ . Simplemente permite afirmar que, de existir  $W_t$ , éste ha de estar distribuido según una ley normal  $\mathcal{N}(0; \sqrt{2Dt})$ .

Para obtener la ley de  $W_t$  nos hemos servido de una caminata aleatoria cuyos incrementos están distribuidos según una ley muy particular (30). Es natural preguntarse si el resultado final (32) depende de esta circunstancia. Dicho de otro modo: tomemos incrementos  $B_n$  independientes e idénticamente distribuidos – no necesariamente según (30) – y formemos la correspondiente caminata aleatoria (29); ¿puede el límite de los nuevos  $W_t^k$  seguir una ley que no sea normal? La respuesta a esta pregunta es negativa, y se conoce como *Teorema Central del Límite*.

**Teorema 6.1** (Teorema Central del Límite). Sean  $B_n$ ,  $n \in \mathbb{N}$ , variables aleatorias independientes, idénticamente distribuidas, con media cero  $E[B_n] = 0$  y varianza finita  $E[B_n^2] = \sigma^2$ . Entonces, para todos los reales  $a < b$  se tiene

$$\lim_{n \rightarrow \infty} P\left(a \leq \frac{B_1 + \dots + B_n}{\sqrt{n}} \leq b\right) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_a^b e^{-\frac{x^2}{2\sigma^2}} dx.$$

Vemos por tanto que la distribución normal aparece de forma natural como la única distribución posible para el proceso límite  $W_t$ ; dicho fenómeno se conoce como *universalidad* y se suele decir que el  $W_t$  es un *objeto universal*. Vimos también en el capítulo 3 de la primera parte, procediendo de forma análoga, que para  $t > s \geq 0$  necesariamente el incremento  $W_t - W_s$  ha de ser independiente de  $W_\tau$  si  $\tau \leq s$  y ha de tener una distribución  $\mathcal{N}(0; \sqrt{t-s})$ .

## 7. El movimiento Browniano y la medida de Wiener

Lo hecho hasta ahora nos conduce a introducir la siguiente definición.

**Definición 7.1.** Una familia de variables aleatorias reales  $W_t$ ,  $t \in [0, \infty)$ , definidas sobre un espacio de probabilidad<sup>5</sup>  $(\Omega, \mathbb{P})$  es un *proceso de Wiener*<sup>6</sup> o *movimiento Browniano*<sup>7</sup> estándar si cumple las siguientes condiciones:

B1.  $W_0 = 0$ .

B2. Dados  $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$  los incrementos

$$W_{t_n} - W_{t_{n-1}}, \quad W_{t_{n-1}} - W_{t_{n-2}}, \quad \dots \quad W_{t_2} - W_{t_1}$$

son independientes.

B3. Para todos  $t \geq 0$  y  $h > 0$ , el incremento  $W_{t+h} - W_t$  tiene una distribución normal  $\mathcal{N}(0; \sqrt{h})$ .

B4. Para casi todo  $\omega \in \Omega$  se tiene que  $t \mapsto W_t(\omega)$  es una aplicación continua.

Esta definición es un poco distinta de la que dimos en el capítulo 3 de la primera parte de las notas. Recuérdese la notación:

$$\phi(x, t) := \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}, \quad \text{para } t > 0.$$

Escribiremos también:

$$\phi(x, 0) := \delta_0(x),$$

<sup>5</sup>Salvo en los casos en que pueda haber ambigüedad, no haremos referencia explícita a la sigma álgebra de  $\Omega$  sobre que la que definimos la probabilidad  $\mathbb{P}$ .

<sup>6</sup>Esta es la terminología preferida por los físicos.

<sup>7</sup>Es el término más utilizado en la literatura matemática.

siendo  $\delta_0$  la medida de Dirac centrada en el origen  $x = 0$ . La condición B1 expresa que la distribución de  $W_0$  es precisamente  $\delta_0$ .<sup>8</sup>

La hipótesis B2 puede parecer un poco extraña en un primer contacto. Claramente, implica que  $W_{t+h} - W_t$  es independiente de  $W_s$  para todos los  $s \leq t$ . El hecho de introducir esta condición más fuerte es que determina de forma única la distribución conjunta de las variables

$$W_{t_1}, W_{t_2}, \dots, W_{t_n} \quad (33)$$

para cualquier elección de  $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$ . Dicha ley necesariamente es:

$$\begin{aligned} & \mathbb{P}(W_{t_1} \in A_1, W_{t_2} \in A_2, \dots, W_{t_n} \in A_n) \\ &= \int_{A_1} \int_{A_2} \dots \int_{A_n} \phi(x_1, t_1) \phi(x_2 - x_1, t_2 - t_1) \dots \phi(x_n - x_{n-1}, t_n - t_{n-1}) dx_1 \dots dx_n, \end{aligned} \quad (34)$$

para  $A_1, \dots, A_n \subset \mathbb{R}$  Borelianos.<sup>9</sup> En otras palabras, para toda función  $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$  medible, se tiene

$$\mathbb{E}[\varphi(W_{t_1}, \dots, W_{t_n})] = \int_{\mathbb{R}^d} \varphi(x_1, \dots, x_n) \phi(x_1, t_1) \dots \phi(x_n - x_{n-1}, t_n - t_{n-1}) dx_1 \dots dx_n, \quad (35)$$

siendo  $\mathbb{E}$  la esperanza tomada respecto a  $\mathbb{P}$ .

- Es fácil demostrar que si un proceso verifica que la ley conjunta de (33) viene dada por (34) entonces automáticamente se tiene B2 y B3.
- Consideremos la siguiente condición: B3'. *Para todos  $t \geq 0$  y  $h > 0$ , los incrementos  $W_{t+h} - W_t$  están idénticamente distribuidos.* Puede demostrarse, utilizando el Teorema Central del Límite que B2, B3' y B4 implican B3. Los procesos que verifican B2 y B3' se conocen como *procesos de Lévy*.

A continuación describimos cómo construir un espacio de probabilidad  $(\Omega, \mathbb{P})$  dotado de un movimiento Browniano  $W_t$ .

## 7.1. Sobre el espacio muestral

Supongamos por un momento que hemos construido un movimiento Browniano  $W_t^\Omega$  en un espacio de probabilidad  $(\Omega, \mathbb{P}_\Omega)$ ; esto es,  $W_t^\Omega$  satisface B1, ..., B4. Cada  $\omega \in \Omega$  define a través del movimiento Browniano un camino

$$[0, \infty) \ni t \mapsto W_t^\Omega(\omega).$$

Dicho de otro modo, un movimiento Browniano induce una aplicación:

$$\begin{aligned} W^\Omega : \Omega &\rightarrow \mathbb{R}^{[0, \infty)} \\ \omega &\mapsto (W_t(\omega))_{t \in [0, \infty)}; \end{aligned}$$

es conviene recordar que el producto cartesiano  $\mathbb{R}^{[0, \infty)}$  es por definición el conjunto de *todas* las funciones de  $[0, \infty)$  en  $\mathbb{R}$ . La aplicación  $W^\Omega$  es Borel medible, puesto que cada una de sus componentes  $W_t^\Omega$  lo es. La probabilidad  $\mathbb{P}_\Omega$  induce, a través de la aplicación  $W^\Omega$ , una medida

<sup>8</sup>Para una función medible  $\varphi$  de  $\mathbb{R}$  se tiene  $\int_{\mathbb{R}} \varphi(x) d\delta_0(x) = \varphi(0)$ .

<sup>9</sup>Si  $t_1 = 0$  hay que interpretar en (34) la integral respecto a la medida  $\phi(x_1, 0) dx_1$  como la integral respecto a  $\delta_0(x_1)$ .

de probabilidad  $\mathbb{P}$  en los Borelianos de  $\mathbb{R}^{[0,\infty)}$  del siguiente modo: para  $E \subset \mathbb{R}^{[0,\infty)}$  Boreliano, ponemos

$$\begin{aligned}\mathbb{P}(E) &:= \mathbb{P}_\Omega(W^\Omega \in E) \\ &= \mathbb{P}_\Omega(\{\omega \in \Omega : W^\Omega(\omega) \in E\}).\end{aligned}$$

La medida  $\mathbb{P}$  se suele llamar *distribución* del proceso  $W_t^\Omega$ . De hecho,  $\mathbb{P}$  es la distribución conjunta de *todas* las variables aleatorias  $W_t$  para  $t \in [0, \infty)$ . Si  $F : \mathbb{R}^{[0,\infty)} \rightarrow \mathbb{R}$  es medible entonces se tiene

$$\int_\Omega F(W^\Omega) d\mathbb{P}_\Omega = \int_{\mathbb{R}^{[0,\infty)}} F(\omega) d\mathbb{P}(\omega). \quad (36)$$

Es sencillo demostrar el siguiente resultado.

**Proposición 7.2.** Todos los movimientos Brownianos tienen la misma distribución  $\mathbb{P}$ . Más aún, si  $W_t : \mathbb{R}^{[0,\infty)} \rightarrow \mathbb{R}$  está definida por  $W_t(\omega) := \omega(t)$  entonces  $W_t$  es un movimiento Browniano definido sobre el espacio de probabilidad  $(\mathbb{R}^{[0,\infty)}, \mathbb{P})$ .

Vemos pues que existe una representación *canónica* del movimiento Browniano. El problema de su construcción es equivalente al de encontrar una medida de probabilidad sobre  $\mathbb{R}^{[0,\infty)}$  para la cual el proceso  $W_t(\omega) := \omega(t)$  es un movimiento Browniano. Dicha medida existe y se conoce como la medida de Wiener. Describiremos a continuación un modo de construirla. En lo sucesivo utilizaremos el símbolo  $\mathbb{P}$  exclusivamente para referirnos a la medida de Wiener y escribiremos, para  $\omega \in \mathbb{R}^{[0,\infty)}$ ,  $W_t(\omega) := \omega(t)$ .

De los axiomas B1, ..., B4 de la definición del movimiento Browniano inferimos las siguientes propiedades para la medida de Wiener  $\mathbb{P}$ .

1. Si  $E \subset \mathbb{R}^{[0,\infty)}$  es una intersección de ventanas elementales, esto es, para ciertos  $0 \leq t_1 \leq \dots \leq t_n$ :

$$\begin{aligned}E &= C_{[a_1, b_1]; t_1} \cap \dots \cap C_{[a_n, b_n]; t_n} \\ &= \{\omega \in \mathbb{R}^{[0,\infty)} : a_j \leq \omega(t_j) \leq b_j \text{ para } j = 1, \dots, n\}\end{aligned}$$

entonces, utilizando (34), deducimos:

$$\mathbb{P}(E) = \int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} \phi(x_1, t_1) \dots \phi(x_n - x_{n-1}, t_n - t_{n-1}) dx_1 \dots dx_n. \quad (37)$$

2. La fórmula anterior se puede generalizar del siguiente modo. Si  $F$  es una función medible de la forma  $F(\omega) = F(\omega(t_1), \dots, \omega(t_n))$  entonces, debido a (35) y (36):

$$\begin{aligned}&\int_{\mathbb{R}^{[0,\infty)}} F(\omega(t_1), \dots, \omega(t_n)) d\mathbb{P}(\omega) \\ &= \int_{\mathbb{R}^n} F(x_1, \dots, x_n) \phi(x_1, t_1) \dots \phi(x_n - x_{n-1}, t_n - t_{n-1}) dx_1 \dots dx_n\end{aligned} \quad (38)$$

3. Una consecuencia inmediata de (37) es que

$$\mathbb{P}(\{\omega \in \mathbb{R}^{[0,\infty)} : \omega(0) \neq 0\}) = 0.$$

4. La propiedad B4 de la definición de  $W_t$  es equivalente a afirmar:

$$\mathbb{P}(C_0([0, \infty); \mathbb{R})) = 1, \quad (39)$$

donde  $C_0([0, \infty); \mathbb{R})$  denota el conjunto de los caminos continuos  $\omega$  de  $[0, \infty)$  en  $\mathbb{R}$  tales que  $\omega(0) = 0$ .

La construcción que daremos de la medida  $\mathbb{P}$  se basa en probar que la definición a través de (38) permite extender de forma única la medida  $\mathbb{P}$  a todos los Borelianos de  $\mathbb{R}^{[0,\infty)}$ . El instrumento que permite esta extensión es el Teorema de Riesz-Markov.

## 7.2. El teorema de Riesz-Markov

El Teorema de Representación de Riesz-Markov permite identificar ciertas medidas de Borel<sup>10</sup> en un espacio topológico  $X$  con funcionales lineales definidos sobre  $C_c(X)$ , el espacio de las funciones continuas, a valores reales y con soporte compacto, definidas sobre  $X$ .

Supondremos que  $X$  es *separable* y *localmente compacto*.

**Teorema 7.3** (Riesz-Markov). Sea  $L$  una aplicación lineal de  $C_c(X)$  en  $\mathbb{R}$ . Si  $L$  es positiva, esto es,  $Lf \geq 0$  para  $f \geq 0$ , entonces existe una única medida  $\mu$  definida sobre los borelianos de  $X$  que representa a  $L$ :

$$Lf = \int_X f(x) d\mu(x), \quad \text{para toda } f \in C_c(X). \quad (40)$$

La medida  $\mu$  verifica además que  $\mu(K)$  es finita para todo compacto  $K \subset X$ .

La demostración y varias de las aplicaciones de este resultado pueden leerse en [Rudin]. Hagamos algunas observaciones antes de proseguir.

1. Si  $L$  es positiva entonces es automáticamente continua. Basta observar que, para  $f \in C_c(X)$  con soporte en un compacto  $K \subset X$  se tiene:  $|Lf| \leq L|f| \leq C_K \max_{x \in X} |f(x)|$  con  $C_K := L\varphi$  para  $\varphi \in C_c(X)$ , con  $\varphi \geq 0$  e idénticamente igual a uno sobre  $K$ .
2. La medida  $\mu$  que verifica (40) se define del siguiente modo. En primer lugar, para un conjunto abierto  $V \subset X$  ponemos:

$$\mu(V) := \sup \{L\varphi : \varphi \in C_c(X), \quad 0 \leq \varphi \leq 1, \quad \text{supp } \varphi \subset V\}.$$

Una vez hecho esto, definimos para  $E \subset X$  arbitrario:

$$\mu(E) := \inf \{\mu(V) : V \subset X \text{ abierto y } E \subset V\}. \quad (41)$$

3. La expresión (41) no define una medida sobre todos los subconjuntos de  $X$ . Sí es posible probar que  $\mu$  es una medida sobre los borelianos de  $X$ . De hecho,  $\mu$  puede definirse sobre una sigma álgebra  $\mathcal{A}_\mu$  un poco mayor, que se forma añadiendo a los borelianos todos los conjuntos  $E \subset X$  con la propiedad de que  $E$  está contenido en un boreliano  $B$  con  $\mu(B) = 0$ .
4. Si  $X$  no es demasiado patológico<sup>11</sup> entonces es posible demostrar que  $\mu$  además verifica:

$$\mu(E) = \sup \{\mu(K) : K \text{ compacto y } K \subset E\}. \quad (42)$$

5. El Teorema de Riesz proporciona un modo de definir la integral de Lebesgue en  $\mathbb{R}^d$  a partir de la de Riemann. Consideramos la aplicación lineal sobre  $C_c(\mathbb{R}^d)$ :

$$Lf := \int_{\mathbb{R}^d} f$$

donde la integral se interpreta en el sentido de Riemann. Claramente,  $L$  es positiva por lo que el Teorema de Riesz garantiza la existencia de una medida  $m$  que satisface:

$$\int_{\mathbb{R}^d} f = \int_{\mathbb{R}^d} f(x) dm(x),$$

siendo la segunda integral la definida mediante la teoría de la medida. Es fácil probar que  $m$  es precisamente la medida de Lebesgue sobre  $\mathbb{R}^d$ .

<sup>10</sup>Se llama *medida de Borel* sobre un espacio topológico  $X$  a una medida definida sobre la sigma álgebra generada por los conjuntos abiertos de  $X$ . Los elementos de dicha sigma álgebra se denominan *Borelianos*.

<sup>11</sup>“No demasiado patológico” quiere decir aquí que todos los abiertos de  $X$  son unión numerable de conjuntos compactos.



6. Otra medida importante es la que resulta de evaluar en un punto. Tomemos  $x_0 \in X$  y definamos

$$L_{x_0} : C_c(X) \rightarrow \mathbb{R} \\ \varphi \mapsto \varphi(x_0) .$$

De nuevo  $L_{x_0}$  es lineal y positivo. Por tanto existe una medida  $\delta_{x_0}$  con la propiedad:

$$\int_X \varphi(x) d\delta_{x_0}(x) = \varphi(x_0) .$$

Dicha medida se conoce como la *masa de Dirac* centrada en  $x_0$ . Es inmediato comprobar que  $\delta_{x_0}(E) = 0$  si  $x_0 \notin E$  y que  $\delta(\{x_0\}) = 1$ .

En el caso de  $X = \mathbb{R}^d$  (o un abierto de  $\mathbb{R}^d$ ) el teorema de representación de Riesz puede refinarse notablemente.

**Teorema 7.4** (Schwartz). Sea  $T \in \mathcal{D}'(\mathbb{R}^d)$  una distribución positiva, esto es, que verifica  $T(\varphi) \geq 0$  para toda  $\varphi \in C_c^\infty(\mathbb{R}^d)$  tal que  $\varphi \geq 0$ . Entonces existe una medida  $\mu$ , con las mismas propiedades que en el Teorema 7.3 que verifica:

$$T(\varphi) = \int_{\mathbb{R}^d} \varphi(x) d\mu(x) .$$

La demostración puede encontrarse, por ejemplo, en el altamente aconsejable [LiebLoss]. Vemos pues que, por el simple hecho de ser positiva,  $T$  puede extenderse de  $C_c^\infty(\mathbb{R}^d)$  a todas las funciones Borel medibles de soporte compacto.

### 7.3. La construcción de la medida de Wiener

La construcción que daremos es debida a E. Nelson [Nelson64]. Probaremos en esta sección el siguiente resultado.

**Proposición 7.5.** Existe una medida  $\mathbb{P}$  sobre los Borelianos de  $\mathbb{R}^{[0,\infty)}$  que satisface (37), (38). Como consecuencia de esto, el proceso en  $(\mathbb{R}^{[0,\infty)}, \mathbb{P})$  definido por  $W_t(\omega) := \omega(t)$  cumple la propiedades B1, B2 y B3.

La idea que seguiremos es la de definir un funcional  $\mathbb{L}$  que se comporte como (38) sobre funciones que dependen de un número finito de coordenadas. Más concretamente, si  $F : \mathbb{R}^{[0,\infty)} \rightarrow \mathbb{R}$  es de la forma

$$F(\omega) = F(\omega(t_1), \dots, \omega(t_n)), \quad \text{para ciertos } 0 \leq t_1 \leq \dots \leq t_n, \quad (43)$$

entonces definimos, basándonos en (38),

$$\mathbb{L}(F) := \int_{\mathbb{R}} \dots \int_{\mathbb{R}} F(x_1, \dots, x_n) \phi(x_1, t_1) \dots \phi(x_n - x_{n-1}, t_n - t_{n-1}) dx_1 \dots dx_n. \quad (44)$$

Se nos presentan dos grandes dificultades para aplicar el Teorema 7.3 en este contexto.

- Si  $X$  es “infinito dimensional” entonces no es localmente compacto – por ejemplo,  $C([0, \infty); \mathbb{R})$ ,  $\mathbb{R}^{\mathbb{N}}$  o un espacio de Hilbert de dimensión infinita no son localmente compactos. Más precisamente, es fácil probar que un producto de espacios localmente compactos es localmente compacto si y sólo si todos los factores, salvo eventualmente un número finito, son compactos.
- Lo que es más importante, una función continua de la forma (43) no pertenece a  $C_c(\mathbb{R}^{[0,\infty)})$ .

Un modo de rodear estas dificultades es el siguiente. Denotemos por  $\dot{\mathbb{R}}$  la compactificación de  $\mathbb{R}$  obtenida añadiendo un punto al infinito; es decir,  $\dot{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$  y los entornos abiertos de  $\infty$  se definen como los complementarios de los compactos de  $\mathbb{R}$ .

El espacio producto

$$\Omega := \dot{\mathbb{R}}^{[0, \infty)}$$

es automáticamente compacto<sup>12</sup> y contiene  $\mathbb{R}^{[0, \infty)}$ ; más aún,  $C_c(\Omega) = C(\Omega)$ . Es en este marco donde aplicaremos el teorema de Riesz. Obsérvese que  $\Omega$  consta simplemente de los caminos de  $\mathbb{R}$  que pueden eventualmente pasar por el infinito y que  $\mathcal{C} := C(\Omega)$  se puede identificar con las funciones de  $C(\mathbb{R}^{[0, \infty)})$  cuyas coordenadas tienen límites finitos e idénticos en  $\pm\infty$ .

Denotemos por  $\mathcal{C}_{\text{fin}}$  al conjunto de las funciones  $F$  continuas de  $\Omega$  en  $\mathbb{R}$  de la forma (43). En otras palabras,  $\mathcal{C}_{\text{fin}}$  es subconjunto de  $\mathcal{C} := C(\Omega)$  constituido por las funciones que dependen únicamente de un número finito de variables. Definimos sobre  $\mathcal{C}_{\text{fin}}$  un funcional  $\mathbb{L}$  mediante la formula (44). El funcional  $\mathbb{L}$  está bien definido, puesto que:

1. La integral que lo define es siempre finita. Más aún,

$$|\mathbb{L}(F)| \leq \sup_{\omega \in \mathbb{R}^{[0, \infty)}} |F(\omega)|. \quad (45)$$

2. La definición (44) es consistente. Si  $F$  no depende de la coordenada  $t_n$  claramente se tiene:

$$\begin{aligned} & \int_{\mathbb{R}} \dots \int_{\mathbb{R}} F(x_1, \dots, x_{n-1}) \phi(x_1, t_1) \dots \phi(x_n - x_{n-1}, t_n - t_{n-1}) dx_1 \dots dx_n \\ &= \int_{\mathbb{R}} \dots \int_{\mathbb{R}} F(x_1, \dots, x_{n-1}) \phi(x_1, t_1) \dots \phi(x_{n-1} - x_{n-2}, t_{n-1} - t_{n-2}) dx_1 \dots dx_n. \end{aligned}$$

Es evidente que  $\mathbb{L}$  es positivo sobre  $\mathcal{C}_{\text{fin}}$ . El teorema de Stone-Weierstrass<sup>13</sup> – véase [Dugundji] – implica que  $\mathcal{C}_{\text{fin}}$  es denso en  $\mathcal{C}$ ; ello, junto con la estimación (45) permite asegurar que  $\mathbb{L}$  se puede extender a un funcional lineal positivo sobre  $\mathcal{C}$ .

Estamos en condiciones de aplicar el teorema de Riesz-Markov y así obtener la existencia de una medida que representa a  $\mathbb{L}$ . Denotamos por  $\mathbb{P}$  a la restricción de dicha medida a  $\mathbb{R}^{[0, \infty)}$ ; claramente  $\mathbb{P}$  satisface (37). Queda por comprobar que también verifica (39).

## 7.4. La continuidad de las trayectorias del movimiento Browniano

Concluiremos la demostración de la existencia de la medida de Wiener probando:

**Teorema 7.7** (Wiener). Las trayectorias del movimiento Browniano son continuas con probabilidad uno:

$$\mathbb{P}(C_0([0, \infty); \mathbb{R})) = 1.$$

<sup>12</sup>Un producto – eventualmente no numerable – de espacios compactos es compacto. Este resultado se conoce como el teorema de Tychonov. Véase [Dugundji] para una demostración.

<sup>13</sup>La extensión de M. Stone del resultado clásico de aproximación de funciones continuas por polinomios de K. Weierstrass reza así:

**Teorema 7.6.** Sea  $X$  un espacio topológico y  $D \subset C(X)$  un conjunto que contiene una función constante no nula y que separa puntos – esto es, dados  $x, y \in X$  con  $x \neq y$  existe  $f \in D$  tal que  $f(x) \neq f(y)$ . Entonces el álgebra – respecto a la suma y el producto – generada por  $D$  es densa en  $C(X)$ .

Sea

$$\rho(\varepsilon, \delta) := \sup_{0 < t \leq \delta} \int_{|x| > \varepsilon} \phi(x, t) dt;$$

claramente

$$\rho(\varepsilon, \delta) = o(\delta) \quad \text{cuando } \delta \rightarrow 0.$$

Vamos a acotar la probabilidad de que un camino tenga una variación de tamaño mayor que  $\varepsilon$  en un intervalo de longitud  $\delta$  en función de  $\rho$ . En todo lo que sigue, supondremos implícitamente que todos los caminos considerados verifican  $\omega(0) = 0$ .

**Lema 7.8.** Sean  $\varepsilon, \delta > 0$  y  $0 \leq t_1 \leq \dots \leq t_n$  tales que  $t_n - t_1 < \delta$ . Consideremos el conjunto

$$E := \{\omega \in \Omega : |\omega(t_i) - \omega(t_j)| > 2\varepsilon \text{ para ciertos } i, j = 1, \dots, n\}.$$

Entonces

$$\mathbb{P}(E) \leq 2\rho\left(\frac{\varepsilon}{2}, \delta\right).$$

**Demostración.** Basta probar que

$$\mathbb{P}(A) \leq 2\rho\left(\frac{\varepsilon}{2}, \delta\right),$$

siendo

$$A := \{\omega \in \Omega : |\omega(t_1) - \omega(t_j)| > \varepsilon \text{ para algún } j = 2, \dots, n\},$$

puesto que  $E \subset A$ .

Vamos a descomponer  $A$  en distintas partes y estimar cada una de ellas. Sean

$$B := \{\omega \in \Omega : |\omega(t_1) - \omega(t_n)| > \varepsilon/2\},$$

$$C_j := \{\omega \in \Omega : |\omega(t_j) - \omega(t_n)| > \varepsilon/2\},$$

$$D_j := \{\omega \in \Omega : |\omega(t_1) - \omega(t_j)| > \varepsilon \text{ y } |\omega(t_1) - \omega(t_k)| \leq \varepsilon \text{ para } k \leq j\}.$$

Entonces se tiene:

$$A \subset B \cup C$$

siendo

$$C := \bigcup_{j=2}^n (C_j \cap D_j).$$

Esto es debido a que, dado  $\omega \in A$  existe un índice menor  $j_0$  para el cual  $\omega \in D_{j_0}$ . Si  $\omega \notin B$  entonces necesariamente  $\omega \in C_{j_0}$ . Así,

$$\begin{aligned} \mathbb{P}(A) &\leq \mathbb{P}(B) + \mathbb{P}(C) \\ &\leq \rho\left(\frac{\varepsilon}{2}, \delta\right) + \mathbb{P}(C). \end{aligned}$$

Ahora bien, visto que los incrementos  $\omega(t_n) - \omega(t_j)$  son independientes, tenemos

$$\begin{aligned} \mathbb{P}(C_j \cap D_j) &= \mathbb{P}(C_j) \mathbb{P}(D_j) \\ &\leq \rho\left(\frac{\varepsilon}{2}, \delta\right) \mathbb{P}(D_j). \end{aligned}$$

Finalmente, como los  $D_j$  son disjuntos, se concluye la demostración. ★

A partir de aquí la prueba del teorema es técnica. Comencemos extendiendo el resultado a un intervalo continuo de tiempos.

**Lema 7.9.** Sea ahora, para  $0 \leq a < b$  con  $b - a \leq \delta$  y

$$E(a, b, \varepsilon) := \{\omega \in \Omega : |\omega(t) - \omega(s)| > 2\varepsilon \text{ para ciertos } t, s \in [a, b]\}.$$

Nuevamente,

$$\mathbb{P}(E(a, b, \varepsilon)) \leq 2\rho\left(\frac{\varepsilon}{2}, \delta\right).$$

**Demostración.** Sea  $S \subset [a, b]$  finito y escribamos

$$E(a, b, \varepsilon, S) := \{\omega \in \Omega : |\omega(t) - \omega(s)| > 2\varepsilon \text{ para ciertos } t, s \in S\}. \quad (46)$$

El lema anterior garantiza que

$$\mathbb{P}(E(a, b, \varepsilon, S)) \leq 2\rho\left(\frac{\varepsilon}{2}, \delta\right).$$

Claramente,  $E(a, b, \varepsilon)$  es la unión de los abiertos  $E(a, b, \varepsilon, S)$  cuando  $S$  recorre los subconjuntos finitos de  $[a, b]$ . Más aún,

$$\mathbb{P}(E(a, b, \varepsilon)) = \sup_S \mathbb{P}(E(a, b, \varepsilon, S)),$$

puesto que, dado  $r > 0$  existe  $K \subset E(a, b, \varepsilon)$  compacto – esto es una consecuencia del teorema de Riesz – tal que

$$\mathbb{P}(K) > \mathbb{P}(E(a, b, \varepsilon)) - r.$$

Necesariamente,  $K$  está contenido en la unión de un número finito de conjuntos (46) y dicha unión es también un conjunto de la forma  $E(a, b, \varepsilon, S)$ . Por tanto

$$\mathbb{P}(E(a, b, \varepsilon, S)) > \mathbb{P}(E(a, b, \varepsilon)) - r,$$

lo cual concluye la demostración. ★

**Lema 7.10.** Sea  $k \in \mathbb{N}$  y supongamos que  $\delta^{-1} \in \mathbb{N}$ . Sea

$$F(k, \varepsilon, \delta) := \{\omega \in \Omega : |\omega(t) - \omega(s)| > 4\varepsilon \text{ para ciertos } t, s \in [0, k], |t - s| < \delta\}.$$

Entonces

$$\mathbb{P}(F(k, \varepsilon, \delta)) \leq 2\frac{k}{\delta}\rho\left(\frac{\varepsilon}{2}, \delta\right).$$

**Demostración.** Podemos escribir  $[0, k]$  como unión de los subintervalos

$$[0, \delta], [\delta, 2\delta], \dots, [(k-1)\delta, k\delta].$$

Si  $\omega \in F(k, \varepsilon, \delta)$  entonces  $|\omega(t) - \omega(s)| > 4\varepsilon$  para  $t, s$  en el mismo subintervalo o en subintervalos consecutivos. En cualquier caso  $\omega$  pertenece a  $E(a, b, \varepsilon)$  siendo  $[a, b]$  uno de los dos subintervalos. ★

Para concluir la prueba del Teorema basta observar que:

$$C_0([0, \infty); \mathbb{R}) = \bigcap_{k \in \mathbb{N}} \bigcap_{n \in \mathbb{N}} \bigcup_{m \in \mathbb{N}} F\left(k, \frac{1}{n}, \frac{1}{m}\right)^c.$$

En particular, como cada uno de los  $F(k, 1/n, 1/m)^c$  es cerrado,  $C_0([0, \infty); \mathbb{R})$  es un Boreliano. Por otra parte,

$$C_0([0, \infty); \mathbb{R})^c = \bigcup_{k \in \mathbb{N}} \bigcup_{n \in \mathbb{N}} \bigcap_{m \in \mathbb{N}} F\left(k, \frac{1}{n}, \frac{1}{m}\right),$$

y claramente,

$$\mathbb{P}\left(\bigcap_{m \in \mathbb{N}} F\left(k, \frac{1}{n}, \frac{1}{m}\right)\right) = \lim_{m \rightarrow \infty} \mathbb{P}\left(F\left(k, \frac{1}{n}, \frac{1}{m}\right)\right) = 0,$$

con lo que

$$\mathbb{P}(C_0([0, \infty); \mathbb{R})) = 1.$$

★

La construcción privilegiada por muchos textos se basa en probar que una medida que satisface (37) se puede extender a una medida de probabilidad sobre la sigma álgebra  $\mathcal{B}_0$  generada por las ventanas elementales (se utiliza el conocido como Teorema de extensión de Kolmogorov). Este enfoque tiene algunos inconvenientes técnicos. Por ejemplo, como resultado de ello, el conjunto  $C_0([0, \infty); \mathbb{R})$  no es  $\mathcal{B}_0$ -medible – aunque puede probarse que existe un conjunto que difiere de él en un conjunto de medida nula que sí lo es. Ello es debido a que  $\mathcal{B}_0$  es estrictamente menor que la sigma álgebra de los Borelianos. Para darse cuenta de ello basta observar que cualquier union o intersección numerable de ventanas elementales sólo depende de una cantidad numerable de componentes  $t_j$ , mientras que un conjunto del tipo  $\{\omega : \omega(t) > a \text{ para algún } t \in \mathbb{R}\}$  es un abierto – y por tanto un Boreliano – de  $\mathbb{R}^{[0, \infty)}$ . En particular, los análogos de los lemas 7.9 y 7.10, que en nuestro caso son triviales, son mucho más difíciles de demostrar si uno trabaja con  $\mathcal{B}_0$ .

La construcción original de N. Wiener se basa en tomar en  $\mathbb{R}^{\mathbb{N}}$  (equipado de una medida Gaussiana) una sucesión  $G_n$ ,  $n = 0, 1, 2, \dots$  de variables  $\mathcal{N}(0; 1)$  independientes y definir para  $\omega \in \mathbb{R}^{\mathbb{N}}$  el camino  $W_t(\omega)$  mediante una serie de Fourier con coeficientes aleatorios:

$$W_t(\omega) := G_0(\omega)t + \sqrt{2} \sum_{n=1}^{\infty} \sum_{k=2^{n-1}}^{2^n-1} G_k(\omega \pi k t) \frac{\text{sen} \pi k t}{\pi k}.$$

Wiener demostró que la serie converge para casi todo  $\omega$  a una función continua en  $t$ . En el capítulo 3 de [Evans] puede encontrarse una demostración basada en ideas similares, aunque técnicamente mucho más sencilla, debida a P. Lévy.

Es posible demostrar que las trayectorias del movimiento Browniano son más regulares de lo que hemos probado. Sea  $C_0^\alpha([0, \infty); \mathbb{R})$  el espacio de los caminos Hölder de exponente  $\alpha > 0$  que arrancan del origen, esto es,  $\omega \in C_0^\alpha([0, \infty); \mathbb{R})$  si  $\omega(0) = 0$  y, dado  $k > 0$ , existe  $K$  tal que, para todos  $t, s \in [0, k]$

$$|\omega(t) - \omega(s)| \leq K |t - s|^\alpha.$$

**Teorema 7.11.** Si  $0 < \alpha < 1/2$  entonces  $\mathbb{P}(C_0^\alpha([0, \infty); \mathbb{R})) = 1$ . Si  $\alpha \geq 1/2$  entonces  $\mathbb{P}(C_0^\alpha([0, \infty); \mathbb{R})) = 0$ . En particular, con probabilidad uno las trayectorias del movimiento Browniano no son diferenciables.

**Ejercicio 7.12.** Sea  $S \subset \mathbb{R}$  un Boreliano de medida cero. Pruébese que el conjunto

$$\Omega_S := \{\omega \in \Omega : \omega(t) \in S \text{ para un conjunto de tiempos } t \text{ de medida positiva}\}$$

verifica

$$\mathbb{P}(\Omega_S) = 0.$$

Advertencia: ante todo hay que probar que  $\Omega_S$  es medible.

## 8. Fórmula de Feynman-Kac y ecuaciones en derivadas parciales

### 8.1. Introducción

Vimos en la primera parte del curso que la medida de Wiener conducía a una representación muy sugerente para las soluciones de la ecuación del calor. Concretamente, si  $u$  es solución de

$$\begin{cases} \partial_t u(x, t) - \frac{1}{2} \Delta u(x, t) = 0, & (x, t) \in \mathbb{R}^d \times [0, \infty), \\ u(x, 0) = \varphi(x), \end{cases} \quad (47)$$

con, pongamos,  $\varphi \in L^2(\mathbb{R}^d)$ , entonces  $u$  puede representarse mediante la *fórmula de Feynman-Kac*:

$$u(x, t) = \int_{\Omega} \varphi(x + \omega(t)) d\mathbb{P}, \quad (48)$$

donde  $\Omega := C_0([0, \infty); \mathbb{R})$  y  $\mathbb{P}$  denota la medida de Wiener.<sup>14</sup> Esta representación proporciona un marco común para la resolución de ecuaciones en derivadas parciales de caracteres, en principio, muy diversos.

Ilustremos esto con la ecuación:

$$\begin{cases} \partial_t u(x, t) - v \cdot \nabla_x u(x, t) = 0, & (x, t) \in \mathbb{R}^d \times [0, \infty), \\ u(x, 0) = \varphi(x), \end{cases} \quad (49)$$

para  $v \in \mathbb{R}^d$ . Claramente,

$$u(x, t) = \varphi(x - tv),$$

y podemos escribir,

$$u(x, t) = \int_{\Omega} \varphi(x + \omega(t)) d\mathbb{M}_v,$$

siendo  $\mathbb{M}_v$  la medida de probabilidad sobre  $\Omega$  que satisface:

$$\mathbb{M}_v \left( \left\{ (-tv)_{t \in [0, \infty)} \right\} \right) = 1.$$

Pese a que (47) es parabólica y (49) es hiperbólica, ambas ecuaciones admiten una solución en términos de una integral del dato inicial en el espacio de caminos  $\Omega$ .

Llevados por esta analogía, podemos decir que *las trayectorias del movimiento Browniano son las curvas características de la ecuación del calor*. De hecho, esta afirmación es más que una mera analogía. En la sección siguiente mostraremos una aplicación de la fórmula de Feynman-Kac (en su versión con potencial, que describimos a continuación) al estudio del comportamiento de los autovalores de un operador de Schrödinger.

### 8.2. La fórmula de Feynman-Kac para $-\frac{1}{2}\Delta + V$

Nuestro objetivo será derivar una fórmula análoga a (48) para las soluciones de la ecuación del calor con potencial:

$$\begin{cases} \partial_t u(x, t) - \frac{1}{2} \Delta u(t, x) + V(x) u(t, x) = 0, & (x, t) \in \mathbb{R}^d \times [0, \infty), \\ u(x, 0) = \varphi(x). \end{cases} \quad (50)$$

Concretamente, probaremos el resultado siguiente:

---

<sup>14</sup>Conviene observar que la demostración de esta fórmula es elemental una vez que se conoce la fórmula explícita para la solución fundamental para la ecuación del calor en  $\mathbb{R}^d$ .

**Teorema 8.1** (Fórmula de Feynman-Kac). Sea  $V$  una función continua y acotada de  $\mathbb{R}^d$  en  $\mathbb{R}$ , y sea  $\varphi \in C(\mathbb{R}^d)$  una función que se anula en el infinito.<sup>15</sup> Entonces, para todo  $x \in \mathbb{R}^d$  se tiene que la solución  $u$  de (50) viene dada por:

$$u(x, t) = \int_{\Omega} \varphi(x + \omega(t)) e^{-\int_0^t V(x + \omega(s)) ds} d\mathbb{P}. \quad (51)$$

Antes que nada, conviene saber que la fórmula (51) es válida bajo hipótesis mucho más débiles sobre  $V$  y  $\varphi$ . Concretamente, basta suponer, por ejemplo,  $\varphi \in L^2(\mathbb{R}^d)$  y  $V \in L^1_{\text{loc}}(\mathbb{R}^d)$  esencialmente acotado inferiormente [Nelson64, TayPDEII].

La prueba que presentamos es debida a Nelson [Nelson64] y está basada en la fórmula de Lie para las soluciones de una ecuación diferencial asociadas a la suma de dos campos de vectores (que en el marco infinito-dimensional se conoce como fórmula de Trotter). Comencemos recordando el resultado clásico de Lie:

**Teorema 8.2** (Fórmula de Lie). Sean  $A$  y  $B$  matrices cuadradas  $d \times d$ . Entonces:

$$e^{t(A+B)} = \lim_{n \rightarrow \infty} (e^{t/nA} e^{t/nB})^n. \quad (52)$$

Este resultado expresa que, para calcular las soluciones de la ecuación diferencial:

$$\dot{x} = (A + B)x$$

basta saber resolver alternativamente las ecuaciones

$$\dot{y} = Ay, \quad \dot{z} = Bz,$$

en intervalos de tiempo cortos  $t/n$ . Como es bien sabido, si  $A$  y  $B$  conmutan  $AB = BA$  entonces la fórmula anterior es una identidad:

$$e^{t(A+B)} = e^{tA} e^{tB} = e^{tB} e^{tA}.$$

**Ejercicio 8.3.** Demuestra la fórmula de Lie (52).

La fórmula de Lie es válida en un marco más general que el que presentamos aquí (en concreto, es cierta en el contexto de espacios de Banach y generadores de grupos de contracciones, véase [TayPDEII], Apéndice 11.A). Nos limitamos a enunciar el resultado parcial que nos será útil; en cuanto a su demostración, remitimos al lector a las referencias antes citadas.

**Proposición 8.4.** Supongamos que  $V$  y  $\varphi$  satisfacen las hipótesis del Teorema 8.1. Entonces la solución  $u(x, t)$  de (50) viene dada por:

$$u(x, t) = \lim_{n \rightarrow \infty} (e^{t/n\Delta/2} e^{-t/nV})^n \varphi(x). \quad (53)$$

De hecho, la convergencia en (53) es uniforme sobre  $\mathbb{R}^d$ .

El aspecto notable de la identidad (53) es que sólo involucra resolver la ecuación del calor sin potencial  $e^{t\Delta/2}$  y multiplicar por la función  $e^{-tV}$ ; ambos pasos son explícitos.

Comencemos usando la fórmula (48) para deducir:

$$\begin{aligned} e^{t/n\Delta/2} e^{-t/nV} \varphi(x) &= \int_{\Omega} e^{-t/nV(x + \omega(t/n))} \varphi(x + \omega(t/n)) d\mathbb{P} \\ &= \int_{\mathbb{R}^d} e^{-t/nV(x + x_1)} \varphi(x + x_1) \phi(x_1, t/n) dx_1; \end{aligned}$$

<sup>15</sup>Con esto queremos decir que dado  $\varepsilon > 0$  existe  $K \subset \mathbb{R}^d$  compacto tal que  $|\varphi(x)| < \varepsilon$  para  $x \in \mathbb{R}^d \setminus K$ .

ahora, repitiendo el cálculo anterior y realizando dos sencillos cambios de variable obtenemos:

$$\begin{aligned}
& (e^{t/n\Delta/2}e^{-t/nV})^2 \varphi(x) \\
&= \int_{\Omega} e^{-t/nV(x+\omega(t/n))} \int_{\mathbb{R}^d} e^{-t/nV(x+x_1)} \varphi(x+\omega(t/n)+x_1) \phi(x_1, t/n) dx_1 d\mathbb{P} \\
&= \int_{\mathbb{R}^{2d}} e^{-t/nV(x+x_2)} e^{-t/nV(x+x_1+x_2)} \varphi(x+x_1+x_2) \phi(x_1, t/n) \phi(x_2, t/n) dx_1 dx_2 \\
&= \int_{\mathbb{R}^{2d}} e^{-t/nV(x+x_2)} e^{-t/nV(x+x_1)} \varphi(x+x_2) \phi(x_1, t/n) \phi(x_2-x_1, t/n) dx_1 dx_2,
\end{aligned}$$

y, reproduciendo el proceso anterior, obtenemos en general:

$$\begin{aligned}
(e^{t/n\Delta/2}e^{-t/nV}\varphi)^n(x) &= \int_{\mathbb{R}^{nd}} e^{-t/n(V(x+x_n)+\dots+V(x+x_1))} \varphi(x+x_n) \\
&\quad \phi(x_1, t/n) \phi(x_2-x_1, t/n) \dots \phi(x_n-x_{n-1}, t/n) dx_1 \dots dx_n.
\end{aligned}$$

Esta expresión no es más que:

$$(e^{t/n\Delta/2}e^{-t/nV}\varphi)^n(x) = \int_{\Omega} e^{-t/n\sum_{j=1}^n V(x+\omega(j/nt))} \varphi(x+\omega(t)) d\mathbb{P}, \quad (54)$$

esto es, la esperanza de la variable aleatoria:

$$e^{-R_n(t,\omega)} \varphi(x+\omega(t)), \quad \text{donde } R_n(t,\omega) := \frac{t}{n} \sum_{j=1}^n V(x+\omega(j/nt)).$$

Ahora bien, si  $\omega(t)$  es continuo entonces, para todo  $t \in \mathbb{R}$ ,

$$\lim_{n \rightarrow \infty} R_n(t,\omega) = \int_0^t V(x+\omega(s)) ds,$$

con lo que para  $\mathbb{P}$ -casi todo  $\omega$  se tiene:

$$\lim_{n \rightarrow \infty} e^{-R_n(t,\omega)} \varphi(x+\omega(t)) = e^{-\int_0^t V(x+\omega(s)) ds} \varphi(x+\omega(t)).$$

Como  $V$  es acotada, tenemos que  $e^{-R_n(t,\omega)} \varphi(x+\omega(t))$  están uniformemente acotadas (con respecto a  $\omega$ ). Así, podemos aplicar el teorema de la convergencia dominada y concluir, utilizando (54) y (53) la fórmula (51).

### 8.3. La fórmula de Feynman para la ecuación de Schrödinger

La idea de interpretar las soluciones de una ecuación en derivadas parciales en términos de una integral sobre un espacio de caminos es debida originalmente a Feynman en el contexto de sus investigaciones sobre la mecánica cuántica. Feynman se interesó a interpretar las soluciones de la ecuación de Schrödinger en términos de trayectorias clásicas (caminos). Tomemos la ecuación, para  $h > 0$ :

$$\begin{cases} ih\partial_t u(x,t) - \frac{h^2}{2}\Delta u(t,x) + V(x)u(t,x) = 0, & (x,t) \in \mathbb{R}^d \times \mathbb{R}, \\ u(x,0) = \varphi(x) \in L^2(\mathbb{R}^d). \end{cases} \quad (55)$$

Esta ecuación es muy similar a la ecuación del calor, excepto por la presencia de la unidad imaginaria  $i$  frente a la derivada temporal (la constante  $h$  es irrelevante desde un punto de



vista puramente matemático). Su presencia cambia radicalmente el carácter de la ecuación; supongamos por un momento  $V = 0$  y  $h = 1$ , entonces la solución de (55) se calcula (utilizando, por ejemplo la transformada de Fourier) y resulta en:

$$u(x, t) = e^{it\Delta} \varphi(x) = \frac{1}{(2\pi it)^{d/2}} \int_{\mathbb{R}^d} e^{i|x-y|^2/2t} \varphi(y) dy.$$

Para  $h > 0$ , la solución de (55) es simplemente  $u(x, t) = e^{iht\Delta} \varphi(x)$ . En particular, se comprueba fácilmente que, a diferencia de la ecuación del calor, la norma  $L^2(\mathbb{R}^d)$  de las soluciones de (55) se conserva:  $\|u(t, \cdot)\|_{L^2(\mathbb{R}^d)} = \|\varphi\|_{L^2(\mathbb{R}^d)}$ .

Al igual que sucede con la ecuación del calor, es posible demostrar la siguiente fórmula de Trotter (suponemos, hasta que digamos lo contrario, que  $h = 1$ ):

**Proposición 8.5.** Supongamos que  $V$  continua y acotada. Para todo  $\varphi \in L^2(\mathbb{R}^d)$ , la solución  $u(x, t)$  de (55) viene dada por:

$$u(x, t) = \lim_{n \rightarrow \infty} (e^{it/n\Delta/2} e^{-it/nV})^n \varphi(x). \quad (56)$$

Calculemos dicho límite. Comenzamos con

$$e^{it/n\Delta/2} e^{-it/nV} \varphi(x) = \frac{1}{(2\pi it)^{d/2}} \int_{\mathbb{R}^d} e^{i\frac{t}{n} \left( \frac{|x-y|^2}{2(t/n)^2} - V(y) \right)} \varphi(y) dy,$$

el cuadrado satisface:

$$\begin{aligned} & (e^{it/n\Delta/2} e^{-it/nV})^2 \varphi(x) \\ &= \frac{1}{(2\pi it)^{2d/2}} \int_{\mathbb{R}^{2d}} e^{i\frac{t}{n} \left( \frac{|x-x_2|^2}{2(t/n)^2} - V(x_2) \right)} e^{i\frac{t}{n} \left( \frac{|x_2-x_1|^2}{2(t/n)^2} - V(x_1) \right)} \varphi(x_1) dx_1 dx_2 \\ &= \frac{1}{(2\pi it)^{2d/2}} \int_{\mathbb{R}^{2d}} e^{i\frac{t}{n} \left( \frac{|x-x_1|^2}{2(t/n)^2} - V(x_1) \right)} e^{i\frac{t}{n} \left( \frac{|x_2-x_1|^2}{2(t/n)^2} - V(x_2) \right)} \varphi(x_2) dx_1 dx_2, \end{aligned}$$

y, en general, escribiendo  $x_0 = x$ :

$$(e^{it/n\Delta/2} e^{-it/nV})^n \varphi(x) = \frac{1}{(2\pi it)^{nd/2}} \int_{\mathbb{R}^{nd}} e^{i\frac{t}{n} \sum_{j=1}^n \left( \frac{|x_j-x_{j-1}|^2}{2(t/n)^2} - V(x_j) \right)} \varphi(x_n) dx_1 \dots dx_n.$$

Si escribimos

$$S_n(t, x_0, \dots, x_n) := \frac{t}{n} \sum_{j=1}^n \left( \frac{|x_j - x_{j-1}|^2}{2(t/n)^2} - V(x_j) \right) \quad (57)$$

entonces si  $\omega(t)$  es un camino definido en  $[0, t]$  tal que  $\omega(0) = x_0 = x$  y  $\omega(jt/n) = x_j$  para  $j = 1, \dots, n$  obtenemos que, formalmente, (57) es la suma de Riemann correspondiente a la acción clásica:

$$S(t, \omega) := \int_0^t \left( \frac{1}{2} |\dot{\omega}(s)|^2 - V(\omega(s)) \right) ds.$$

Este cálculo motivó la escritura de solución  $u(x, t)$  de (55) como:

$$\begin{aligned} u(x, t) &= \lim_{n \rightarrow \infty} \frac{1}{(2\pi it)^{nd/2}} \int_{\mathbb{R}^{nd}} e^{iS_n(t, x, \dots, x_n)} \varphi(x_n) dx_1 \dots dx_n \\ &= \int_{\mathbb{R}^d} \int_{C_{x,y}^1([0,t])} e^{iS(t, \omega)} \varphi(y) dC(\omega) dy \end{aligned}$$

siendo  $C_{x,y}^1([0,t])$  el conjunto de los caminos diferenciables que satisfacen  $\omega(0) = x$  y  $\omega(t) = y$ . Este cálculo es meramente formal. De hecho, es posible demostrar que no existe ninguna medida sobre el espacio de caminos para la que la fórmula anterior es válida (véase [Cameron]). No obstante, el proceso de paso al límite define una integral a la Riemann. Más detalles sobre esta construcción pueden encontrarse en [Nelson64].

Si tomamos  $h > 0$  obtenemos:

$$u(x, t) = \int_{\mathbb{R}^d} \int_{C_{x,y}^1([0,t])} e^{iS(t,\omega)/h} \varphi(y) dC(\omega) dy.$$

**Ejercicio 8.6.** Aplica formalmente el principio de la fase estacionaria para obtener el límite, cuando  $h \rightarrow 0$ , de la integral

$$\int_{C_{x,y}^1([0,t])} e^{iS(t,\omega)/h} dC(\omega).$$

## 9. Aplicación al estudio de los autovalores de un operador elíptico

### 9.1. Los autovalores del operador de Schrödinger $-\frac{1}{2}\Delta + V(x)$

Sea  $V \in L_{\text{loc}}^1(\mathbb{R}^d; \mathbb{R})$  acotado inferiormente,

$$C_V = \text{essinf}_{x \in \mathbb{R}^d} V(x) > -\infty, \quad (58)$$

y tal que

$$\lim_{R \rightarrow \infty} \text{essinf}_{|x| \geq R} V(x) = +\infty. \quad (59)$$

Bajo estas condiciones existe una base ortonormal de autofunciones de  $-\frac{1}{2}\Delta + V$ .

**Proposición 9.1.** Supongamos que  $V$  satisface (58) y (59). Entonces el operador  $A := -\frac{1}{2}\Delta_x + V(x)$  es autoadjunto en  $L^2(\mathbb{R}^d)$  con

$$D(A) := \left\{ u \in H^1(\mathbb{R}^d) : \int_{\mathbb{R}^d} V(x) |u(x)|^2 dx < \infty \right\}.$$

Además, existe una base ortonormal  $(\psi_n)_{n \in \mathbb{N}}$  de  $L^2(\mathbb{R}^d)$  formada por autofunciones de  $A$ :

$$-\frac{1}{2}\Delta_x \psi_n + V \psi_n = \lambda_n \psi_n.$$

Los autovalores correspondientes son reales, mayores que  $C_V$ , y verifican  $\lambda_n \rightarrow \infty$  cuando  $n \rightarrow \infty$ .

**Demostración.** La demostración es estándar. Comencemos observando que podemos, sin pérdida de generalidad suponer que  $C_V \geq 0$  – de lo contrario, basta aplicar el resultado obtenido en ese caso a  $V + C_V$ . Sea  $f \in L^2(\mathbb{R}^d)$ ; claramente  $u \in D(A)$  es solución de

$$-\frac{1}{2}\Delta u + Vu + u = f$$

si y sólo si,

$$\begin{aligned} Q_A(u, \varphi) &:= \int_{\mathbb{R}^d} \left( \frac{1}{2} \nabla_x u(x) \cdot \nabla_x \varphi(x) + u(x) \varphi(x) + V(x) u(x) \varphi(x) \right) dx = \\ &= \int_{\mathbb{R}^d} f(x) \varphi(x) dx =: L_f(\varphi) \end{aligned}$$

para toda  $\varphi \in D(A)$ .  $Q_A(\cdot, \cdot)$  no es más que el producto escalar canónico en  $D(A)$  (que convierte a dicho espacio en un espacio de Hilbert); mientras que  $L_f$  define una aplicación continua de  $D(A)$  en  $\mathbb{R}$ . Podemos pues aplicar el teorema de Riesz y asegurar que, dada  $f \in L^2(\mathbb{R}^d)$  existe un único  $u \in D(A)$  que verifica  $Q_A(u, \varphi) = L_f(\varphi)$  para toda  $\varphi \in D(A)$ . Tenemos pues que:

$$(A + 1)^{-1} : L^2(\mathbb{R}^d) \rightarrow D(A) \hookrightarrow L^2(\mathbb{R}^d)$$

está bien definida y es continua. Probaremos más adelante que la inyección de  $D(A)$  en  $L^2(\mathbb{R}^d)$  es compacta, de lo que se sigue que  $(A + 1)^{-1}$  es un operador compacto de  $L^2(\mathbb{R}^d)$  en si mismo. Teniendo en cuenta que también es autoadjunto, concluimos – véase, por ejemplo, [Davies] o cualquier otro libro de análisis funcional – que existe una base ortonormal de  $L^2(\mathbb{R}^d)$  formada por funciones  $(\psi_n)_{n \in \mathbb{N}}$  que satisfacen:

$$(A + 1)^{-1} \psi_n = \mu_n \psi_n,$$

siendo  $\mu_n \in \mathbb{R}$  y  $\lim_{n \rightarrow \infty} \mu_n = 0$ . Es inmediato comprobar que

$$A\psi_n = \left( \frac{1}{\mu_n} - 1 \right) \psi_n.$$

con lo  $\lambda_n := \frac{1}{\mu_n} - 1$  son autovalores de  $A$ . Para concluir que su límite es  $+\infty$  basta observar que, si  $\psi$  es una autofunción de  $A$  de norma uno y con autovalor  $\lambda$  entonces se tiene:

$$C_V \leq \int_{\mathbb{R}^d} \frac{1}{2} |\nabla_x \psi|^2 + V(x) |\psi(x)|^2 dx = \lambda.$$

Habremos completado la demostración del resultado una vez que hayamos probado la compacidad de la inyección de  $D(A)$  en  $L^2(\mathbb{R}^d)$ .

**Lema 9.2.** Si  $(u_n)$  es una sucesión acotada en  $D(A)$  entonces existe una subsucesión que converge fuertemente en  $L^2(\mathbb{R}^d)$ .

**Demostración.** Sea  $\rho \in C_c(\mathbb{R}^d)$  tal que  $\text{supp} \rho \subset B(0; 2)$ ,  $\rho|_{B(0;1)} \equiv 1$  y  $0 \leq \rho \leq 1$ . Escribamos, para  $N \in \mathbb{N}$ :

$$u_n^N(x) := \rho\left(\frac{x}{N}\right) u_n(x).$$

Fijado  $N$ , la sucesión  $(u_n^N)$  está acotada en  $H_0^1(B(0; 2))$ . Podemos por tanto aplicar el teorema de Rellich (véase [LiebLoss]) y concluir la existencia de una subsucesión convergente en  $L^2(\mathbb{R}^d)$ . De hecho, aplicando el argumento diagonal de Cantor, podemos garantizar la existencia de una subsucesión  $(u_{n'}^N)$  que converja fuerte en  $L^2(\mathbb{R}^d)$  para todo  $N \in \mathbb{N}$ . Ahora bien, dado  $N > 0$  se tiene, utilizando que  $\int_{\mathbb{R}^d} V(x) |u_n(x)|^2 dx < C$  para todo  $n \in \mathbb{N}$ :

$$\begin{aligned} \int_{\mathbb{R}^d} |u_{n'}(x) - u_{n'}^N(x)|^2 dx &\leq \int_{\mathbb{R}^d} \left(1 - \rho\left(\frac{x}{N}\right)\right)^2 |u_n(x)|^2 dx \\ &\leq \int_{|x|>N} |u_{n'}(x)|^2 dx \\ &\leq \frac{1}{\text{ess inf}_{|x|>N} V(x)} \int_{\mathbb{R}^d} V(x) |u_{n'}(x)|^2 dx \\ &\leq C_N, \end{aligned}$$

siendo  $\lim_{N \rightarrow \infty} C_N = 0$ . De aquí es fácil – utilizando el argumento de  $\varepsilon/3$  – concluir que  $(u_{n'})$  es de Cauchy en  $L^2(\mathbb{R}^d)$ . ★

## 9.2. Aplicación de Feynman-Kac al estudio del comportamiento asintótico de los autovalores

La solución de la ecuación del calor con potencial:

$$\begin{cases} \partial_t u(x, t) - \frac{1}{2} \Delta u(x, t) + V(x) u(x, t) = 0, & (x, t) \in \mathbb{R}^d \times [0, \infty), \\ u(x, 0) = \varphi(x), \end{cases} \quad (60)$$

puede expresarse también como una integral en el espacio de los caminos:

$$u(x, t) := \int_{\Omega} e^{-\int_0^t V(x+\omega(\tau)) d\tau} \varphi(x + \omega(t)) d\mathbb{P}.$$

De hecho, la fórmula anterior es válida si  $V$  satisface las hipótesis (58) y (59). Para más información sobre las hipótesis que hay que requerir a  $V$  puede consultarse [Simon]. Aplicaremos dicha fórmula a la obtención de resultados finos sobre el comportamiento asintótico de los autovalores del operador  $A$ .

Ante todo, escribamos:

$$H(x, p) := \frac{1}{2} |p|^2 + V(x).$$

En el enunciado del siguiente resultado requeriremos que nuestro potencial satisfaga tres hipótesis adicionales:

$$V \in C(\mathbb{R}^d; \mathbb{R}), \quad e^{-tV} \in L^1(\mathbb{R}^d), \quad t \in [0, \infty), \quad (61)$$

$$\lim_{\delta \rightarrow 0^+} \liminf_{t \rightarrow 0^+} \frac{\int_{\mathbb{R}^d} e^{-tV_\delta(x)} dx}{\int_{\mathbb{R}^d} e^{-tV(x)} dx} = 1, \quad \text{siendo } V_\delta(x) := \max_{y \in B(x; \delta)} V(y), \quad (62)$$

$$\lim_{\lambda \rightarrow \infty} \frac{\text{vol}(\{(x, p) \in \mathbb{R}^d \times \mathbb{R}^d : H(x, p) \leq \lambda\})}{(2\pi)^d \lambda^\gamma} = C, \quad \gamma \geq 0. \quad (63)$$

La condición (62) es meramente técnica; simplifica enormemente la prueba. La condición (63) cuantifica la tasa de crecimiento de  $V$  en el infinito. Se satisface para una gran clase de funciones.

**Teorema 9.3.** Supongamos que  $V$  satisface las condiciones (58), (59), (61), (62), (63). Denotemos por  $N(\lambda)$  al número de autovalores de  $A$  menores que  $\lambda$ . Entonces:

$$N(\lambda) \sim \frac{1}{(2\pi)^d} \text{vol}(\{(x, p) \in \mathbb{R}^d \times \mathbb{R}^d : H(x, p) \leq \lambda\})$$

cuando  $\lambda \rightarrow \infty$ .

La demostración del teorema se hará en varios pasos. Sea

$$Q(x, y, t) := \sum_{n \in \mathbb{N}} e^{-t\lambda_n} \psi_n(x) \psi_n(y);$$

entonces para toda  $\varphi \in L^2(\mathbb{R}^d)$  la función:

$$u(x, t) := \int_{\mathbb{R}^d} Q(x, y, t) \varphi(y) dy$$

es la única solución de (60). La expresión

$$\int_{\mathbb{R}^d} Q(x, x, t) dx = \sum_{n \in \mathbb{N}} e^{-t\lambda_n}$$

es la *traza* del núcleo del calor. Si denotamos por  $\mu_c$  a la medida:

$$\mu_c(\lambda) := \sum_{n \in \mathbb{N}} \delta_{\lambda_n}(\lambda)$$

entonces tenemos que:

$$\sum_{n \in \mathbb{N}} e^{-t\lambda_n} = \int_0^\infty e^{-t\lambda} d\mu_c(\lambda)$$

y

$$N(\lambda) := \mu_c([0, \lambda]). \quad (64)$$

Nuestro objetivo será probar el siguiente resultado:

**Proposición 9.4.** Cuando  $t \rightarrow 0^+$ , se tiene:

$$\sum_{n \in \mathbb{N}} e^{-t\lambda_n} = \int_0^\infty e^{-t\lambda} d\mu_c(\lambda) \sim \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d \times \mathbb{R}^d} e^{-tH(x,p)} dx dp. \quad (65)$$

Esta última integral se puede escribir como:

$$\frac{1}{(2\pi)^d} \int_{\mathbb{R}^d \times \mathbb{R}^d} e^{-tH(x,p)} dx dp = \frac{1}{(2\pi)^d} \int_0^\infty e^{-tE} d\mu_H(E)$$

siendo, para  $I \subset \mathbb{R}$  medible:

$$\mu_H(I) := \text{vol}(\{(x,p) \in \mathbb{R}^d \times \mathbb{R}^d : H(x,p) \in I\}).$$

Obsérvese que por hipótesis tenemos que

$$\lim_{E \rightarrow \infty} \frac{\mu_H((-\infty, E])}{(2\pi)^d E^\gamma} = C. \quad (66)$$

Los siguientes resultados permiten relacionar (66) con el comportamiento asintótico de (64) para  $\lambda \rightarrow \infty$  a través de la propiedad (65).

**Teorema 9.5** (Teorema Abeliano). Sea  $\mu$  una medida de Borel en  $[0, \infty)$  que satisface, para ciertos  $\gamma, C \geq 0$ ,

$$\lim_{E \rightarrow \infty} E^{-\gamma} \mu((-\infty, E]) = C.$$

Entonces,

$$\lim_{t \rightarrow 0^+} t^\gamma \int_0^\infty e^{-tE} d\mu(E) = C\Gamma(\gamma + 1).$$

De ese resultado se deduce inmediatamente que,

$$\lim_{t \rightarrow 0^+} \frac{t^\gamma}{(2\pi)^d} \int_{\mathbb{R}^d \times \mathbb{R}^d} e^{-tH(x,p)} dx dp = C\Gamma(\gamma + 1). \quad (67)$$

Ahora bien, el siguiente cuasi recíproco es cierto:

**Teorema 9.6** (Teorema Tauberiano). Sea  $\mu$  una medida de Borel en  $[0, \infty)$  que satisface  $\int_0^\infty e^{-t\lambda} d\mu(\lambda) < \infty$  para todo  $t > 0$  y, para ciertos  $\gamma, D \geq 0$ ,

$$\lim_{t \rightarrow 0^+} t^\gamma \int_0^\infty e^{-tE} d\mu(E) = D.$$

Entonces,

$$\lim_{\lambda \rightarrow \infty} \lambda^{-\gamma} \mu((-\infty, \lambda]) = \frac{D}{\Gamma(\gamma + 1)}.$$

Así, si se cumple (65) tenemos por (67):

$$\lim_{t \rightarrow 0^+} t^\gamma \int_0^\infty e^{-t\lambda} d\mu_c(\lambda) = C\Gamma(\gamma + 1),$$

y el teorema Tauberiano permite concluir que:

$$\lim_{\lambda \rightarrow \infty} \lambda^{-\gamma} \mu_c((-\infty, \lambda]) = C,$$

o en otras palabras, para  $\lambda \rightarrow \infty$ :

$$N(\lambda) \sim \frac{\text{vol}(\{(x, p) \in \mathbb{R}^d \times \mathbb{R}^d : H(x, p) \leq \lambda\})}{(2\pi)^d}.$$

El resto de esta sección estará dedicado a demostrar la Proposición 9.4.

### 9.2.1. Aplicación de la fórmula de Feynman-Kac

Definamos

$$f_\varepsilon(x) := \frac{\gamma_d}{\varepsilon^d} \mathbf{1}_{B(0;1)}\left(\frac{x}{\varepsilon}\right)$$

de modo que  $\int_{\mathbb{R}^d} f_\varepsilon(x) dx = 1$ . Claramente,

$$\int_{\mathbb{R}^d} Q(x, y, t) f_\varepsilon(y - x) dy \rightarrow Q(x, x, t) = \sum_{n \in \mathbb{N}} e^{-t\lambda_n} |\psi_n(x)|^2, \quad \text{cuando } \varepsilon \rightarrow 0^+.$$

Sea

$$C_\varepsilon^t := \{\omega \in \Omega : |\omega(t)| < \varepsilon\};$$

utilizando la fórmula de Feynman-Kac obtenemos:

$$\begin{aligned} Q(x, x, t) &= \lim_{\varepsilon \rightarrow 0^+} \int_{\mathbb{R}^d} Q(x, y, t) f_\varepsilon(y - x) dy \\ &= \lim_{\varepsilon \rightarrow 0^+} \int_{\Omega} e^{-\int_0^t V(x+\omega(\tau)) d\tau} f_\varepsilon(\omega(t)) d\mathbb{P}. \end{aligned}$$

Ahora bien,

$$\begin{aligned} \int_{\Omega} e^{-\int_0^t V(x+\omega(\tau)) d\tau} f_\varepsilon(\omega(t)) d\mathbb{P} &= \frac{\gamma_d}{\varepsilon^d} \int_{C_\varepsilon^t} e^{-\int_0^t V(x+\omega(\tau)) d\tau} d\mathbb{P} \\ &= \frac{K_\varepsilon}{(2\pi t)^{d/2} \mathbb{P}(C_\varepsilon^t)} \int_{C_\varepsilon^t} e^{-\int_0^t V(x+\omega(\tau)) d\tau} d\mathbb{P}, \end{aligned}$$

donde hemos escrito:

$$K_\varepsilon := \frac{\gamma_d}{\varepsilon^d} (2\pi t)^{d/2} \mathbb{P}(C_\varepsilon^t) = \frac{\gamma_d}{\varepsilon^d} \int_{|x| < \varepsilon} e^{-|x|^2/2t} dx;$$

claramente,  $K_\varepsilon \rightarrow 1$ , cuando  $\varepsilon \rightarrow 0^+$ .

### 9.2.2. El puente Browniano

Vamos a examinar con detalle el límite:

$$\lim_{\varepsilon \rightarrow 0^+} \frac{1}{\mathbb{P}(C_\varepsilon^t)} \int_{C_\varepsilon^t} \varphi(\omega(\tau)) d\mathbb{P} = \lim_{\varepsilon \rightarrow 0^+} \mathbb{E}[\varphi(W_\tau) | C_\varepsilon^t],$$

para  $\tau \in [0, t]$  y  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$  medible. Formalmente, esperamos tener

$$\lim_{\varepsilon \rightarrow 0^+} \frac{1}{\mathbb{P}(C_\varepsilon^t)} \int_{C_\varepsilon^t} \varphi(\omega(\tau)) d\mathbb{P} = \mathbb{E}[\varphi(W_\tau) | W_t = 0];$$

y de hecho obtenemos:

$$\begin{aligned} & \frac{1}{\mathbb{P}(C_\varepsilon^t)} \int_{C_\varepsilon^t} \varphi(\omega(\tau)) d\mathbb{P} \\ &= \frac{t^{d/2}}{(2\pi\tau(t-\tau))^{d/2}} D_\varepsilon^{-1} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi(y) \mathbf{1}_{B(0;1)}(x) e^{-\frac{t|y|^2}{2\tau(t-\tau)}} e^{-\varepsilon \frac{|x|^2 - 2x \cdot y}{2(t-\tau)}} dx dy, \end{aligned}$$

con

$$D_\varepsilon = \int_{|x| < 1} e^{-\varepsilon \frac{|x|^2}{2t}} dx.$$

A la vista de esto, se tiene:

$$\lim_{\varepsilon \rightarrow 0^+} \frac{1}{\mathbb{P}(C_\varepsilon^t)} \int_{C_\varepsilon^t} \varphi(\omega(\tau)) d\mathbb{P} = \frac{t^{d/2}}{(2\pi\tau(t-\tau))^{d/2}} \int_{\mathbb{R}^d} \varphi(y) e^{-\frac{t|y|^2}{2\tau(t-\tau)}} dy.$$

Claramente, el mismo razonamiento da, para  $0 \leq \tau_1 \leq \dots \leq \tau_n \leq t$ :

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0^+} \frac{1}{\mathbb{P}(C_\varepsilon^t)} \int_{C_\varepsilon^t} \varphi(\omega(\tau_1), \dots, \omega(\tau_n)) d\mathbb{P} \\ &= \frac{t^{d/2}}{(t-\tau_n)^{d/2}} \int_{\mathbb{R}^{nd}} \varphi(x_1, \dots, x_n) e^{-\frac{|x_n|^2}{2(t-\tau_n)}} \phi_{\tau_1, \dots, \tau_n}(x_1, \dots, x_n) dx_1 \dots dx_n. \end{aligned}$$

Esta última expresión es precisamente la distribución conjunta en  $\tau_1 \dots \tau_n$  del proceso:

$$W_\tau^{0,t} := W_\tau - \frac{\tau}{t} W_t, \quad \tau \in [0, t],$$

conocido como *puente Browniano*. Denotemos por  $\mathbb{P}^{0,t}$  a su distribución. Sirviéndonos del argumento basado en el teorema de Riesz-Markov que utilizamos para construir la medida de Wiener, podemos demostrar que, para toda función medible y acotada  $\varphi : \mathbb{R}^{[0,\infty)} \rightarrow \mathbb{R}$  se tiene:

$$\lim_{\varepsilon \rightarrow 0^+} \frac{1}{\mathbb{P}(C_\varepsilon^t)} \int_{C_\varepsilon^t} \varphi(\omega) d\mathbb{P} = \int_{\Omega} \varphi(\omega) d\mathbb{P}^{0,t}.$$

En particular,

$$Q(x, x, t) = \frac{1}{(2\pi t)^{d/2}} \int_{\Omega} e^{-\int_0^t V(x+\omega(\tau)) d\tau} d\mathbb{P}^{0,t}.$$

### 9.2.3. Estimaciones de $\text{tr}(e^{-t(\Delta/2-V)})$

Comencemos estimando superiormente utilizando la desigualdad de Jensen:

$$\begin{aligned}
\sum_{n \in \mathbb{N}} e^{-t\lambda_n} &= \frac{1}{(2\pi t)^{d/2}} \int_{\mathbb{R}^d} \int_{\Omega} e^{-\int_0^t V(x+\omega(\tau))d\tau} d\mathbb{P}^{0,t} dx \\
&\leq \frac{1}{(2\pi t)^{d/2}} \int_{\mathbb{R}^d} \int_{\Omega} \frac{1}{t} \int_0^t e^{-tV(x+\omega(\tau))} d\tau d\mathbb{P}^{0,t} dx \\
&= \frac{1}{(2\pi t)^{d/2}} \int_{\mathbb{R}^d} \int_0^t \frac{1}{t} \frac{t^{d/2}}{(2\pi\tau(t-\tau))^{d/2}} \int_{\mathbb{R}^d} e^{-tV(x+y)} e^{-\frac{t|y|^2}{2\tau(t-\tau)}} dy d\tau dx \\
&= \frac{1}{(2\pi t)^{d/2}} \int_{\mathbb{R}^d} e^{-tV(x)} dx.
\end{aligned}$$

Por otra parte, si escribimos:

$$\begin{aligned}
\Delta_\delta^t &:= \{\omega \in \Omega : |\omega(\tau)| \leq \delta \text{ para } \tau \in (0, t) \text{ y } \omega(0) = \omega(t) = 0\}, \\
V_\delta(x) &:= \max_{y \in B(x; \delta)} V(y),
\end{aligned}$$

tenemos, para todo  $\delta > 0$ ,

$$\begin{aligned}
\sum_{n \in \mathbb{N}} e^{-t\lambda_n} &\geq \frac{1}{(2\pi t)^{d/2}} \int_{\mathbb{R}^d} \int_{\Delta_\delta^t} e^{-\int_0^t V(x+\omega(\tau))d\tau} d\mathbb{P}^{0,t} dx \\
&\geq \frac{1}{(2\pi t)^{d/2}} \int_{\mathbb{R}^d} \int_{\Delta_\delta^t} e^{-tV_\delta(x)} d\mathbb{P}^{0,t} dx \\
&= \frac{\mathbb{P}^{0,t}(\Delta_\delta^t)}{(2\pi t)^{d/2}} \int_{\mathbb{R}^d} e^{-tV_\delta(x)} dx.
\end{aligned}$$

Ahora bien, razonando como en las demostraciones de los lemas 7.9 y 7.10 obtenemos:

$$\begin{aligned}
\mathbb{P}^{0,t}(\Delta_\delta^t) &\geq 1 - \sup_{0 \leq \tau \leq t} \frac{t^{d/2}}{(2\pi\tau(t-\tau))^{d/2}} \int_{|y| > \delta} e^{-\frac{t|y|^2}{2\tau(t-\tau)}} dy \\
&= 1 - \sup_{0 \leq \tau \leq t} \int_{|y| > \frac{t\delta}{\tau(t-\tau)}} e^{-\frac{|y|^2}{2}} \frac{dy}{(2\pi)^{d/2}} \\
&= 1 - \int_{|y| > \frac{4\delta}{t}} e^{-\frac{|y|^2}{2}} \frac{dy}{(2\pi)^{d/2}} \\
&= \int_{|y| \leq \frac{4\delta}{t}} e^{-\frac{|y|^2}{2}} \frac{dy}{(2\pi)^{d/2}}.
\end{aligned}$$

Ahora bien,  $\lim_{t \rightarrow 0^+} \mathbb{P}^{0,t}(\Delta_\delta^t) = 1$  y, por hipótesis también tenemos:

$$\lim_{\delta \rightarrow 0^+} \liminf_{t \rightarrow 0^+} \frac{\int_{\mathbb{R}^d} e^{-tV_\delta(x)} dx}{\int_{\mathbb{R}^d} e^{-tV(x)} dx} = 1.$$

Por tanto,

$$\begin{aligned}
1 &\leq \liminf_{t \rightarrow 0^+} \frac{(2\pi t)^{d/2} \sum_{n \in \mathbb{N}} e^{-t\lambda_n}}{\int_{\mathbb{R}^d} e^{-tV(x)} dx} \\
&\leq \limsup_{t \rightarrow 0^+} \frac{(2\pi t)^{d/2} \sum_{n \in \mathbb{N}} e^{-t\lambda_n}}{\int_{\mathbb{R}^d} e^{-tV(x)} dx} \leq 1
\end{aligned}$$

con lo que el resultado queda probado.



# Apéndice

## A. La ecuación del calor: solución analítica.

Para hallar una expresión para la función de densidad límite debemos resolver la ecuación del calor en una dimensión espacial. Repasemos brevemente la existencia de soluciones de esta ecuación por medio de la transformada de Fourier. Definimos

$$\hat{f}(s) := \int_{-\infty}^{+\infty} e^{-isx} f(x) dx \quad (68)$$

y asumimos la validez de la fórmula de inversión

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{isx} \hat{f}(s) ds. \quad (69)$$

Podemos ver fácilmente que:

$$\begin{aligned} \widehat{f''}(s) &= \int_{-\infty}^{+\infty} e^{-isx} f''(x) dx \\ &= -s^2 \int_{-\infty}^{+\infty} e^{-isx} f(x) dx \\ &= -s^2 \hat{f}(s). \end{aligned}$$

Si transformamos la ecuación (13) con respecto a la variable  $x$  a ambos lados obtenemos:

$$\frac{\partial \hat{u}}{\partial t}(s, t) = -Ds^2 \hat{u}(s, t),$$

tomamos a  $s$  como un parámetro de la ecuación e integramos respecto de  $t$ :

$$\log \left| \frac{\hat{u}}{\hat{u}_0} \right| = -Ds^2 t \Rightarrow \hat{u}(s, t) = \hat{u}_0(s) e^{-Ds^2 t}.$$

Podemos recuperar la función  $u$  formalmente a través de la fórmula de inversión (69):

$$u(x, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{isx - Ds^2 t} \hat{u}_0(s) ds.$$

Escribiendo la transformada de  $u_0$  según (68) obtenemos:

$$u(x, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{isx - Ds^2 t} \left( \int_{-\infty}^{+\infty} e^{-isq} u_0(q) dq \right) ds.$$

E intercambiando el orden de las integrales:

$$\begin{aligned} u(x, t) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} u_0(q) \left( \int_{-\infty}^{+\infty} e^{is(x-q) - Ds^2 t} ds \right) dq \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-\frac{(x-q)^2}{4Dt}} u_0(q) \left( \int_{-\infty}^{+\infty} e^{\left( is\sqrt{Dt} + \frac{(x-q)}{2\sqrt{Dt}} \right)^2} ds \right) dq \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-\frac{(x-q)^2}{4Dt}} u_0(q) \left( \int_{-\infty - i\frac{x-q}{2\sqrt{Dt}}}^{\infty - i\frac{x-q}{2\sqrt{Dt}}} e^{-p^2} \frac{dp}{\sqrt{Dt}} \right) dq \\ &= \frac{1}{2\sqrt{\pi Dt}} \int_{-\infty}^{+\infty} e^{-\frac{(x-q)^2}{4Dt}} u_0(q) dq. \end{aligned} \quad (70)$$

Cuando  $u_0$  es la función delta de Dirac centrada en el origen obtenemos la solución fundamental

$$u(x, t) = \frac{1}{2\sqrt{\pi Dt}} e^{-\frac{x^2}{4Dt}}, \quad (71)$$

idéntica a la expresión obtenida para la densidad de probabilidad (7). De esta manera recuperamos a la misma función a través de la solución de una ecuación en derivadas parciales. En (70) podemos ver que la solución para una condición inicial arbitraria se obtiene a partir de la solución fundamental a través de la convolución, esto es equivalente a distribuir en forma continua infinidad de fuentes a lo largo de la recta.

## Referencias

- [1] Alexandre Chorin, Ole Hald: *Stochastic Tools in Mathematics and Science. Surveys and Tutorials in the Applied Mathematical Sciences* 1, Springer 2006.
- [2] Leo Breiman: *Probability*, Addison-Wesley Series in Statistics, 1968.
- [3] S. Chandrasekar: *Stochastic Problems in Physics and Astronomy*, en *Reviews of Modern Physics*, Vol 15, N. 1, 1943.
- [4] D.R. Cox & H.D. Miller: *The theory of stochastic processes*.
- [5] J.L. Doob: *Stochastic Processes*, John Wiley and Sons 1953.
- [6] C.C. Lin, L.A. Segel: *Mathematics applied to deterministic problems in the Natural Sciences*. Classics in Applied Mathematics 1, SIAM.
- [7] C.W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*. Springer-Verlag.
- [8] H.P. McKean: *Application of Brownian motion to the equation of Kolmogorov-Petrovskii-Piskunov*. Communications on pure and applied mathematics, Vol. XXVI-II, 323-331 (1975).
- [Cameron] Cameron, R. H. A family of integrals serving to connect the Wiener and Feynman integrals. *J. Math. and Phys.* **39** (1960/1961) 126–140.
- [Davies] Davies, E. B. *Spectral theory and differential operators*. Cambridge Studies in Advanced Mathematics, **42**. Cambridge University Press, Cambridge, 1995.
- [Dugundji] Dugundji, J. *Topology*. Allyn and Bacon, Inc., Boston, Mass. 1966.
- [Evans] Evans, L. C. *An introduction to Stochastic Differential Equations*. Disponible en la URL: <http://math.berkeley.edu/~evans/>.
- [LiebLoss] Lieb, E. H.; Loss, M. *Analysis*. Second edition. Graduate Studies in Mathematics, **14**. American Mathematical Society, Providence, RI, 2001.
- [Nelson64] Nelson, E. Feynman integrals and the Schrödinger equation. *J. Mathematical Phys.* **5** (1964) 332–343.
- [Nelson67] Nelson, E. *Dynamical theories of Brownian motion*. Princeton University Press, Princeton, N.J. 1967. Disponible en la URL: <http://www.math.princeton.edu/~nelson/books.html>.

- [Rudin] Rudin, W. *Real and complex analysis*. Third edition. McGraw-Hill Book Co., New York, 1987.
- [Simon] Simon, B. *Functional integration and quantum physics*. Pure and Applied Mathematics, **86**. Academic Press, Inc., New York-London, 1979.
- [TayPDEII] Taylor, M. E. *Partial differential equations. II. Qualitative studies of linear equations*. Applied Mathematical Sciences, **116**. Springer-Verlag, New York, 1996.