

Aureli Alabert

Departament de Matemàtiques  
Universitat Autònoma de Barcelona

# INTRODUCCIÓN A LAS ECUACIONES DIFERENCIALES ESTOCÁSTICAS

Cali, Enero de 2004



## Índice

<i>Índice</i> .....	<i>i</i>
<b>1. Motivación y ejemplos</b> .....	1
<b>2. Conceptos básicos de probabilidad y procesos estocásticos</b> .....	5
<b>3. El proceso de Wiener</b> .....	9
<b>4. Los modelos de Einstein-Smoluchowski y de Ornstein-Uhlenbeck para el movimiento browniano</b> .....	15
4.1 El modelo de Einstein-Smoluchowski .....	15
4.2 El modelo de Ornstein-Uhlenbeck .....	17
<b>5. Integración estocástica (I)</b> .....	19
5.1 El problema .....	19
5.2 Integrales estocásticas elementales .....	22
<b>6. Integración estocástica (II)</b> .....	25
6.1 La integral estocástica de Itô .....	25
<b>7. Cálculo estocástico</b> .....	33
<b>8. Ecuaciones diferenciales estocásticas. Existencia y unicidad</b> .....	37
<b>9. Referencias</b> .....	41



## 1. Motivación y ejemplos

Desde que en el siglo XVII Newton y Leibniz pusieron las bases del que ahora llamamos “Cálculo Diferencial”, las ecuaciones diferenciales han sido una herramienta matemática fundamental para modelizar sistemas físicos. Las leyes físicas que gobiernan un sistema determinan las ecuaciones correspondientes, que después intentamos resolver, es decir, de las cuales intentamos obtener una expresión del estado del sistema en el instante de tiempo  $t$  como función explícita de  $t$ .

A veces el sistema físico que queremos modelizar es demasiado complejo no sólo para resolver efectivamente las ecuaciones asociadas sino incluso para llegar a formular un conjunto de ecuaciones diferenciales que sea suficientemente representativo de sus características. Resulta que en muchas de estas situaciones “desesperadas” es posible usar cierta *información parcial* al alcance sobre las fuerzas que interactúan y a cambio obtener determinados *resultados parciales* del problema. Concretamente, veremos cómo el uso de hipótesis estadísticas en la formulación nos permite llegar a conclusiones probabilísticas como resultado. Esto significa: si el estado de un sistema en el instante  $t$  está representado por un vector  $n$ -dimensional  $x(t)$ , renunciamos a asignar un valor a este vector, pero hacemos uso de nuestro conocimiento del comportamiento estadístico de los elementos que afectan  $x(t)$  para contestar preguntas tales como: “¿Cuál es la probabilidad de que  $x(t)$  se encuentre en una cierta región  $A$  de  $\mathbb{R}^n$ ?”, y otras parecidas.

### 1.1 El ejemplo fundamental: El movimiento browniano

Considérese una pequeña partícula (de diámetro, digamos, una micra) inmersa en un fluido. A través de un microscopio se puede observar que la partícula se mueve de manera muy rápida e irregular en todas direcciones, y que este movimiento no cesa nunca. Cuando el fenómeno fue descrito por primera vez a principios del siglo XIX, diversas teorías (equivocadas) surgieron para explicar sus causas. Actualmente, con la existencia de los átomos firmemente establecida, sabemos que estos movimientos “caóticos” son debidos al bombardeo intensivo de la partícula por parte de las moléculas del fluido que la rodea. Bajo condiciones normales de presión y temperatura, la partícula soporta del orden de  $10^{21}$  colisiones moleculares por segundo, cada una con su propia dirección y energía. A pesar de que cada colisión individual tienen un efecto inapreciable (les moléculas son todavía mucho más pequeñas que la partícula) la superposición de un número tan enorme de choques sí produce un efecto observable.

Es obvio que el cálculo diferencial clásico puede ayudar muy poco al estudio de esta situación. Qué puede hacerse entonces para resolver (al menos parcialmente) el problema?

Por simplicidad, supongamos que ningún otro campo de fuerzas, como podría ser la gravedad, interactúa con el sistema que acabamos de describir, y normalicemos a 1 la masa de la partícula. Sólo hay entonces dos orígenes de fuerzas a considerar: Por un lado, la fricción dinámica sistemática sufrida por la partícula en sus traslaciones, y por otro, las colisiones moleculares irregulares, negligibles individualmente pero con efecto macroscópico apreciable.

Si olvidamos por un momento esta última fuente de movimiento, y denotamos por  $v(t)$  la velocidad en el instante  $t$ , la segunda ley de Newton nos lleva inmediatamente a la ecuación diferencial de primer orden

$$\left. \begin{aligned} \frac{dv(t)}{dt} &= -\beta \cdot v(t) \\ v(0) &= v_0 \end{aligned} \right\}, \quad (1.1)$$

donde la constante  $\beta$  se determina para partículas esféricas mediante la ley de Stokes:  $\beta = 6\pi r\eta$ , siendo  $r$  el radio de la partícula y  $\eta$  el coeficiente dinámico de viscosidad del fluido (para partículas no esféricas hay que cambiar el factor  $6\pi$  por otro adecuado).

Consideremos la versión integral de la ecuación (1.1):

$$\left. \begin{aligned} v(t + \Delta t) - v(t) &= -\beta \int_t^{t+\Delta t} v(s) ds \\ v(0) &= v_0 \end{aligned} \right\}. \quad (1.2)$$

Denotemos por  $M(t)$  el incremento de momento neto ganado por la partícula debido a la segunda fuente de fuerzas citada durante el intervalo temporal  $[0, t]$ . Teniendo en cuenta esta función,

$$\left. \begin{aligned} v(t + \Delta t) - v(t) &= -\beta \int_t^{t+\Delta t} v(s) ds + M(t + \Delta t) - M(t) \\ v(0) &= v_0 \end{aligned} \right\}. \quad (1.3)$$

$M(t)$  es una función desconocida, pero si aceptamos ciertas hipótesis físicas sobre las fuerzas que representa, éstas se traducirán en propiedades matemáticas (probabilísticas) de  $M$ . Concretamente, supongamos:

- 1) Las condiciones del entorno (tales como presión, temperatura, ...) son constantes en el tiempo.
- 2) El bombardeo sufrido por la partícula durante un intervalo de tiempo dado es independiente de (esto es, no tiene relación física con) el sufrido en intervalos de tiempo anteriores.
- 3) La aceleración total hasta el instante  $t$  cambia continuamente respecto de  $t$ .
- 4) El medio es isótropo.

Veremos más tarde cómo puede construirse un modelo matemático que represente estas condiciones. Dejando de lado la cuestión de si 1), 2), 3) y 4) son lo bastante adecuadas al sistema que hemos descrito (esto puede ser tema de mucha discusión, especialmente 2)), señalemos sólo que el modelo que se deducirá produce resultados que se ajustan bastante bien a las observaciones empíricas.  $\square$

Este ejemplo es muy clásico (véase, por ejemplo, Nelson [22]) y continuaremos su discusión en capítulo sucesivos. Citemos dos ejemplos más de situaciones similares, sin especificar detalles.

## 1.2 Ejemplo: Verhulst

El conocido modelo de Verhulst para la evolución del volumen de individuos de una especie viene dado por la ecuación diferencial

$$\frac{dx(t)}{dt} = x(t)(a - bx(t)) \quad (1.4)$$

(la ecuación logística), que se deduce de suponer una tasa de natalidad constante  $a$  y una tasa de mortalidad proporcional a la población  $bx(t)$ . (1.4) se puede pensar como la aproximación continua del planteo discreto

$$x(n+1) - x(n) = x(n)(a - bx(n)) \quad (1.5)$$

Típicamente, las constantes  $a$  y  $b$  son idealizaciones de cantidades sometidas a fluctuaciones imprevisibles a priori. Supongamos, por ejemplo, que cambiamos  $a$  por una función  $a + \xi(n)$ , donde

$\xi(n)$  representa una pequeña perturbación de la tasa de natalidad, que es desconocida y que es independiente (i.e., no hay relación de causa-efecto) de las perturbaciones  $\xi(m)$ ,  $m \neq n$ . Si denotamos por  $M(n)$  la acumulación de estas perturbaciones des del instante 0 hasta el  $n - 1$ , o sea  $M(n) = \sum_{k=0}^{n-1} \xi(k)$ , obtendremos la ecuación modificada

$$x(n+1) - x(n) = x(n)(a + M(n+1) - M(n) - bx(n)) . \quad (1.6)$$

Construimos ahora la aproximación continua de (1.6). El análogo de los incrementos  $x(n+1) - x(n)$  en modo continuo es  $\frac{x(t+\Delta t) - x(t)}{\Delta t}$ , y análogamente para  $M(n+1) - M(n)$ . Escribiremos pues

$$x(t+\Delta t) - x(t) = x(t)(a - bx(t))\Delta t + x(t)(M(t+\Delta t) - M(t)) . \quad (1.7)$$

Si suponemos constante el entorno en el sentido de que las causas de las perturbaciones son invariables en el tiempo, mantenemos la independencia del valor de las perturbaciones que afectan al sistema en intervalos disjuntos, y finalmente hacemos la hipótesis razonable de que la función  $M(t)$  ha de ser continua en el tiempo, volvemos a encontrar exactamente las condiciones 1), 2) y 3) propuestas en el ejemplo anterior como propiedades de la función desconocida  $M$ . Resulta pues formalmente la misma situación del Ejemplo 1.1.  $\square$

### 1.3 Ejemplo: Filtraje

Consideremos ahora el problema de llevar un satélite artificial a su órbita geoestacionaria, a partir de la órbita elíptica de transferencia, que es el estadio intermedio entre la fase de lanzamiento y el posicionamiento definitivo.

Para controlar esta delicada maniobra hay que tener una información lo más precisa posible de la posición del satélite en cada instante. En primera aproximación, teniendo en cuenta sólo la acción del campo gravitatorio terrestre, las ecuaciones del movimiento pueden escribirse en la forma

$$\frac{dx(t)}{dt} = f(x(t)) \quad (1.8)$$

con  $x(t) \in \mathbb{R}^6$  (tres parámetros de posición, tres parámetros de velocidad).

Otros elementos que influyen en la trayectoria del satel.lit, como por ejemplo la no-esfericidad y no-homogeneidad de la Tierra, la influencia de otros cuerpos celestes, la presión de radiación solar, etc., pueden ser incorporados al modelo si se tiene un buen conocimiento de como actúan. Sin embargo, intervendrán también perturbaciones con dirección e intensidad no conocida a priori.

Una heurística similar a la de los ejemplos anteriores nos lleva a escribir

$$x(t+\Delta t) - x(t) = \int_t^{t+\Delta t} f(x(s)) ds + M(t+\Delta t) - M(t)$$

donde  $M$  es una función con las características 1), 2), 3) listadas en el Ejemplo 1.1.

Además, contamos con la información adicional aportada por las estaciones de seguimiento, que reciben una señal

$$y(t) = g(t, x(t)) ,$$

a su vez contaminada también por perturbaciones en la medición:

$$y(t+\Delta t) - y(t) = \int_t^{t+\Delta t} g(s, x(s)) ds + N(t+\Delta t) - N(t) ,$$

siendo  $N(t)$  desconocida e independiente (en el sentido de que las causas lo son) de  $M(t)$ . El problema, que se llama *filtraje*, consiste en obtener la mejor información posible sobre el valor de  $x(t)$ , teniendo en cuenta las observaciones  $y(t)$ .  $\square$

Nuestro propósito en los capítulos siguientes es presentar las herramientas matemáticas que permiten tratar los problemas de este tipo.





## 2. Conceptos básicos de probabilidad y procesos estocásticos

De costumbre, se dice que un experimento es aleatorio cuando su resultado no se puede conocer a priori, a pesar de que no sea la primera vez que se realiza. El experimento puede estar gobernado por claras leyes deterministas, pero renunciamos a efectuar los cálculos, o simplemente no sabemos cómo hacerlos. A menudo, sin embargo, es posible decir algo sobre cada posible resultado. Por ejemplo, al tirar repetidamente una moneda equilibrada, es de esperar que el cociente entre el número de cruces obtenidas y el número de lanzamientos (la llamada *frecuencia relativa*) se aproxime a  $1/2$  a la larga. Esto nos invita a enunciar: “ $1/2$  es la ‘probabilidad’ de obtener cruz en cada lanzamiento de una moneda equilibrada”.

Definir el concepto de probabilidad apelando a los límites de frecuencias relativas es bastante problemático. La formalización habitual es axiomática y el hecho de que las frecuencias relativas converjan en un cierto sentido a la probabilidad de un suceso se obtiene como teorema (leyes de los grandes números).

La formalización habitual es la siguiente:

Sea  $\Omega$  un conjunto y  $\mathcal{F} \subset \mathcal{P}(\Omega)$  una  $\sigma$ -álgebra de partes de  $\Omega$ . Una **probabilidad** es una medida positiva y finita sobre  $(\Omega, \mathcal{F})$  con masa total igual a 1. Ésto es decir, una probabilidad sobre  $(\Omega, \mathcal{F})$  es una aplicación  $P: \mathcal{F} \rightarrow [0, 1]$  tal que  $P(\emptyset) = 0$ ,  $P(\Omega) = 1$ , y para cada familia numerable  $\{A_i\}_{i \in I}$  de conjuntos de  $\mathcal{F}$  tal que  $A_i \cap A_j = \emptyset$ ,  $\forall i, j \in I$ , se tiene  $P(\cup_{i \in I} A_i) = \sum_{i \in I} P(A_i)$ . El espacio de medida  $(\Omega, \mathcal{F}, P)$  recibe el nombre de **espacio de probabilidad**. En este contexto, cada conjunto medible  $A \in \mathcal{F}$  se llama **suceso**. El número  $P(A)$  es la **probabilidad del suceso**  $A$ . Al modelizar un experimento real, cada  $\omega \in \Omega$  representa un posible resultado y  $P(A)$  es la “probabilidad” de que el resultado  $\omega$  pertenezca a  $A$ .

Sea  $(E, \mathcal{E})$  un otro espacio medible. Una **variable aleatoria  $E$ -valuada sobre  $(\Omega, \mathcal{F}, P)$**  es una aplicación medible  $X: (\Omega, \mathcal{F}, P) \rightarrow (E, \mathcal{E})$ . Cuando  $(E, \mathcal{E})$  es el conjunto de los números reales equipado con su  $\sigma$ -álgebra de Borel natural, y que denotaremos  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ , decimos simplemente que tenemos una **variable aleatoria (real)**. Las variables aleatorias representan el mecanismo mediante el cual se observa el experimento. Si observamos un conjunto  $B \in \mathcal{E}$ , deducimos que el resultado del experimento pertenece a  $X^{-1}(B) \in \mathcal{F}$ .

Dada una variable aleatoria real sobre  $(\Omega, \mathcal{F}, P)$ , podemos definir una nueva probabilidad  $\mu$  en  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  como  $\mu(B) := P(X^{-1}(B))$ ,  $\forall B \in \mathcal{B}(\mathbb{R})$ . Esta nueva probabilidad  $\mu$  asociada a  $P$  y  $X$  se llama la **ley de  $X$  bajo  $P$** . Si  $\mu$  es una probabilidad sobre  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ , la función  $F: \mathbb{R} \rightarrow \mathbb{R}$  definida por  $F(x) := \mu(-\infty, x]$  recibe el nombre de **función de distribución** de  $\mu$ .  $F$  es una función no-decreciente y continua por la derecha verificando  $\lim_{x \rightarrow -\infty} F(x) = 0$  y  $\lim_{x \rightarrow +\infty} F(x) = 1$ . Recíprocamente, toda función con estas propiedades determina unívocamente una probabilidad  $\mu$ . Si  $F$  es expresable como  $F(x) = \int_{-\infty}^x f(z) dz$ , para alguna función  $f: \mathbb{R} \rightarrow \mathbb{R}^+$ , entonces  $f$  es la **función de densidad** de  $\mu$  respecto a la medida de Lebesgue. En tal caso, tendremos que

$\forall B \in \mathcal{B}(\mathbb{R}), \mu(B) = \int_B f(z) dz$ . Estos conceptos se extienden fácilmente a variables aleatorias  $\mathbb{R}^n$ -valuadas. A menudo estas variables son llamadas **vectores aleatorios**, y su ley de probabilidad la **ley conjunta** de las componentes del vector.

Dos sucesos  $A_1, A_2 \in \mathcal{F}$  se llaman **independientes** si  $P(A_1 \cap A_2) = P(A_1) \cdot P(A_2)$ . Dos variables aleatorias reales  $X, Y$  definidas sobre el mismo espacio de probabilidad son **independientes** si  $\forall B_1, B_2 \in \mathcal{B}(\mathbb{R})$ , los sucesos  $X^{-1}(B_1)$  y  $Y^{-1}(B_2)$  son sucesos independientes. Intuitivamente, esto significa que la observación de un suceso  $A$  a través de la variable aleatoria  $X$  no cambia la ley de probabilidad de  $Y$ . La definición es análoga para variables valuadas en cualquier espacio y para familias de variables aleatorias.

Dada una variable aleatoria  $X$ , definimos el  **$p$ -ésimo momento de  $X$**  como  $\int_{\Omega} X(\omega)^p P(d\omega)$ , siempre que la integral exista y sea finita. Argumentos sencillos de teoría de la medida muestran que si  $\mu$  es la ley de  $X$  bajo  $P$  y  $F$  es la función de distribución asociada, el  $p$ -ésimo momento de  $X$  se puede calcular como la integral de Stieljes  $\int_{\mathbb{R}} x^p dF(x)$ . Si, además,  $F$  tiene una función de densidad  $f$ , la última integral coincide con la integral ordinaria  $\int_{\mathbb{R}} x^p f(x) dx$ . El primer momento de  $X$  se llama **esperanza** o **media** de  $X$ . Lo denotamos por  $E[X]$ . Si para dos variables aleatorias  $X, Y$  se tiene la igualdad  $P(\{\omega : X(\omega) = Y(\omega)\}) = 1$ , decimos que  $X = Y$  **casi seguro (c.s.)**. Esta no es más que la conocida equivalencia *casi por todo* de la teoría de la integración de Lebesgue. Mediante esta relación, se definen los habituales espacios de Lebesgue  $L^p(\Omega, \mathcal{F}, P)$  de (clases de) variables aleatorias que poseen  $p$ -ésimo momento, para  $0 < p < \infty$ .  $L^\infty(\Omega, \mathcal{F}, P)$  será el espacio de las variables aleatorias acotadas, mientras que  $L^0(\Omega, \mathcal{F}, p)$  representa el espacio vectorial de todas las variables aleatorias, sin ninguna condición de integrabilidad. En cualquier espacio de medida finita se tiene  $0 \leq q \leq p \leq \infty \Rightarrow L^p \subset L^q$ .

Sea  $T$  un conjunto ordenado. Un **proceso estocástico** indexado por  $T$  es una familia de variables aleatorias (reales, si no se especifica otra cosa)  $\{X_t, t \in T\}$ , todas ellas definidas sobre el mismo espacio de probabilidad. Normalmente,  $T$  es  $\mathbb{R}^+ = [0, \infty[$  o  $\mathbb{N} = \{0, 1, 2, \dots\}$ , y se interpreta como el tiempo, en modo continuo o discreto, de forma que un proceso estocástico es la herramienta apropiada para modelizar sistemas dinámicos cuando el estado del sistema en cada instante  $t$  está gobernado por una ley de probabilidad, la ley de  $X_t$ . Supondremos a partir de ahora  $T = [0, \infty[$ .

Para cada  $\omega$  fijado, la función

$$\begin{aligned} X(\omega): T &\longrightarrow \mathbb{R} \\ t &\longmapsto X_t(\omega) \end{aligned}$$

se llama una **trayectoria** del proceso. Este punto de vista nos permite pensar en un proceso estocástico como una variable aleatoria valuada en algún espacio  $E$  de funciones:

$$\begin{aligned} X: \Omega &\longrightarrow E \\ \omega &\longmapsto X(\omega) \end{aligned}$$

Cada posible resultado aleatorio  $w$  determina un elemento del espacio funcional, que puede ser el espacio de todas las funciones  $\mathbb{R}^T$  o uno más pequeño. Se puede definir la ley de un proceso siguiendo la misma pauta que para las variables aleatorias reales o  $\mathbb{R}^n$ -valuadas. Se puede demostrar, por otro lado, que la ley del proceso está completamente determinada por las **leyes en dimensión finita**, i.e. las leyes de las proyecciones de rango finito

$$\begin{aligned} X: \Omega &\longrightarrow \mathbb{R}^n \\ \omega &\longmapsto (X_{t_1}(\omega), \dots, X_{t_n}(\omega)) \end{aligned}$$

$n \in \mathbb{N}, t_1, \dots, t_n \in T$ .

Se caracterizan diversas clases particulares de procesos según propiedades de sus trayectorias o bien propiedades relacionadas con la ley del proceso. Por ejemplo, decimos que un proceso estocástico es **continuo** si, para cada  $\omega \in \Omega$ , la trayectoria  $X(\omega)$  es una función continua.

La continuidad es una propiedad de las trayectorias. Son propiedades relacionadas con la ley las siguientes: Un proceso tiene **incrementos estacionarios** si la ley de la variable aleatoria  $X_t - X_s$  depende sólo de la diferencia  $t - s$ , para todos  $t$  y  $s$  en  $T$ . Decimos que un proceso estocástico tiene **incrementos independientes (del pasado)** si para todos  $t, s \in T$ ,  $s < t$ , la variable  $X_t - X_s$  es independiente de la familia  $\{X_r, r \in T, r \leq s\}$ .

Por su importancia en aplicaciones, la **ley gaussiana** (o **normal**) merece especial atención. Una variable aleatoria real  $X$  es gaussiana si su ley de probabilidad tiene función de densidad

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} \quad (2.1)$$

para ciertos  $m \in \mathbb{R}$  y  $\sigma^2 > 0$ . Lo escribimos  $X \sim N(m, \sigma^2)$ . Se comprueba fácilmente que las cantidades  $E[X]$  y  $\text{Var}[X] := E[(X - E[X])^2]$ , llamadas **media** y **varianza** de  $X$  valen  $m$  y  $\sigma^2$  respectivamente. Por tanto, la media y la varianza de una variable aleatoria gaussiana determinan completamente su ley. Es conveniente considerar también el caso degenerado  $\sigma^2 = 0$  como gaussiano. Si hacemos tender  $\sigma^2$  a cero en (2.1), observamos que  $f(x)$  converge a una  $\delta$  de Dirac, de forma que para  $\sigma^2 = 0$  no hay una verdadera densidad. La ley representada por esta “densidad generalizada” concentra toda su masa en el punto  $m$ .

El análogo de las variables gaussianas en  $\mathbb{R}^n$  son los vectores aleatorios con densidad

$$f(x) = ((2\pi)^n \det D)^{-1/2} \exp\left\{-\frac{1}{2}(x-m)D^{-1}(x-m)\right\}, \quad x \in \mathbb{R}^n, \quad (2.2)$$

donde  $m \in \mathbb{R}^n$  es el **vector de medias** y  $D$  es la **matriz de covarianzas**, que ha de ser definida-positiva (o sólo semidefinida-positiva en los casos degenerados en que la ley se concentra en una variedad afín de  $\mathbb{R}^n$ ). Cualquier proyección de un vector aleatorio gaussiano es de nuevo gaussiano.

La ley gaussiana  $n$ -dimensional juega un papel central en el estudio de las probabilidades en  $\mathbb{R}^n$ . El análogo en dimensión infinita a los vectores aleatorios gaussianos son los **procesos estocásticos gaussianos**. Decimos que  $\{X_t, t \in \mathbb{R}^+\}$  es gaussiano si sus leyes en dimensión finita son gaussianas.

Los siguientes teoremas, que relacionan algunos de los conceptos introducidos, nos serán útiles después. (Véase Yeh [27, pàg. 202 y 255] para una demostración detallada).

## 2.1 Teorema

Sea  $X \equiv \{X_t, t \in \mathbb{R}^+\}$  un proceso real tal que

**(H.1)** es continuo y con incrementos independientes del pasado.

Entonces, existen una función continua  $m: \mathbb{R}^+ \rightarrow \mathbb{R}$  y una función continua no-decreciente  $v: \mathbb{R}^+ \rightarrow \mathbb{R}$  tales que  $\forall s, t \in \mathbb{R}^+$ ,  $s < t$ , la variable aleatoria  $X_t - X_s$  es gaussiana con

$$E[X_t - X_s] = m(t) - m(s), \quad \text{Var}[X_t - X_s] = v(t) - v(s). \quad (2.3)$$

Si, además,

**(H.2)**  $X$  tiene incrementos estacionarios

entonces  $m$  y  $v$  satisfacen

$$\begin{aligned} m(t) - m(s) &= \mu \cdot (t - s) \\ v(t) - v(s) &= \sigma^2 \cdot (t - s) \end{aligned}$$

para ciertas constantes  $\mu \in \mathbb{R}$  y  $\sigma^2 \in \mathbb{R}^+$ .

$\mu$  y  $\sigma^2$  se llaman **coeficiente de deriva** y **coeficiente de difusión** respectivamente.  $\square$

El recíproco del Teorema es cierto, en el sentido de que dadas las funciones  $m$  continua y  $v$  continua no-decreciente, existe un proceso continuo con incrementos independientes  $X$  verificando las igualdades (2.3).

## 2.2 Teorema

Bajo la hipótesis **(H.1)**, si  $X_0$  es una variable gaussiana, entonces  $\{X_t, t \in \mathbb{R}^+\}$  es un proceso gaussiano.

En particular, un proceso continuo con incrementos independientes tal que  $X_0 \equiv x \in \mathbb{R}$ , ha de ser forzosamente gaussiano.  $\square$

## 2.3 Definición

Un proceso estocástico real  $W \equiv \{W_t, t \in \mathbb{R}^+\}$  es un **proceso de Wiener** (o **movimiento browniano**) si

- i) es continuo,
- ii) tiene incrementos independientes del pasado,
- iii)  $\forall t, s, s < t, W_t - W_s$  es una variable aleatoria gaussiana con esperanza cero y varianza  $\sigma^2 \cdot (t - s)$ , para algún  $\sigma^2 \in \mathbb{R}^+$ .

Un proceso de Wiener **comienza en**  $x \in \mathbb{R}$  si  $W_0 = x$  c.s.

Si, además,  $\sigma^2 = 1$ ,  $W$  se llama **proceso de Wiener estándar comenzando en**  $x \in \mathbb{R}$ .  $\square$

Obsérvese que un proceso de Wiener estándar es un proceso gaussiano con incrementos estacionarios, y por otro lado, si  $X \equiv \{X_t, t \in \mathbb{R}^+\}$  es un proceso continuo con incrementos estacionarios e independientes del pasado, con  $X_0 = 0$  c.s. y coeficiente de deriva  $\mu = 0$ , entonces  $X = \sigma W$ , donde  $W$  es un proceso de Wiener estándar comenzando en 0.

## 2.4 Ejemplo (el movimiento browniano, 2)

Usando los conceptos introducidos a el capítulo anterior, podemos ahora dar una versión matemática precisa de las hipótesis físicas sobre la función  $M(t)$  del Ejemplo 1.1. Imaginaremos, para simplificar, que el problema es unidimensional (la partícula se mueve sobre una recta). De hecho, será evidente en seguida que el problema real se puede estudiar coordenada por coordenada.

Primeramente, nos hemos referido a  $M(t)$  como una “función desconocida”. Por tanto, se puede pensar de manera natural como un proceso estocástico  $\{M_t(\omega), t \in \mathbb{R}^+\}$  definido en algún espacio de probabilidad  $(\Omega, \mathcal{F}, P)$ . El hecho de que las causas de las fuerzas resumidas en  $M_t$  son invariantes en el tiempo (condición 1) se modeliza imponiendo que la ley de  $M_t - M_s$  dependa sólo de la diferencia  $t - s$ . En otras palabras, el proceso  $\{M_t, t \in \mathbb{R}^+\}$  ha de tener *incrementos estacionarios*. La incorrelación física entre las fuerzas que actúan en un intervalo dado y las que han actuado previamente (condición 2) significa que el proceso ha de tener *incrementos independientes del pasado*. Finalmente, la condición 3 de continuidad simplemente se traduce en que  $\{M_t, t \in \mathbb{R}^+\}$  es un *proceso continuo*. Podemos además imponer arbitrariamente  $M_0 = 0$  en el origen del tiempo. Según los Teoremas 2.1 y 2.2, en esta situación el proceso ha de ser gaussiano con  $E[M_t] = \mu t$  y  $\text{Var}[M_t] = \sigma^2 t$ . La isotropía del medio en que se mueve la partícula (no hay ninguna dirección privilegiada para las fuerzas provenientes de los choques moleculares) la representamos poniendo coeficiente de deriva  $\mu = 0$ .

Ahora sólo hay que usar las observaciones que siguen a la definición de proceso de Wiener para concluir que  $M_t = \sigma W_t$ , siendo  $W$  un proceso de Wiener estándar comenzando en 0. Sobre el valor concreto de  $\sigma$  volveremos más adelante.

La ecuación (1.3) puede escribirse pues

$$\left. \begin{aligned} v(t + \Delta t) - v(t) &= -\beta \int_t^{t+\Delta t} v(s) ds + \sigma (W(t + \Delta t) - W(t)) \\ v(0) &= v_0 \end{aligned} \right\}. \quad (2.4)$$

### 3. El proceso de Wiener

Hay ciertas propiedades de las trayectorias del proceso de Wiener definido en el Capítulo 2 que es importante remarcar. Estas són:

- Las trayectorias  $\{W_t(\omega), t \in \mathbb{R}^+\}$  del proceso de Wiener son no-derivables en ningún punto, casi seguro. Esto es consecuencia esencialmente de la propiedad de incrementos independientes.
- Las trayectorias del proceso de Wiener tienen variación total infinita en todo intervalo  $[a, b]$ ,  $a < b$ , casi seguro. Es una consecuencia inmediata de la anterior.
- La variación cuadrática de las trayectorias del proceso de Wiener es no nula y finita. El sentido preciso de esta afirmación se verá cuando especifiquemos qué entendemos por variación cuadrática en este contexto. Intuitivamente, dado que  $E[|W_t - W_s|^2] = t - s$ , la variación  $|W_t - W_s|$  será posiblemente del orden de  $\sqrt{t - s}$ , lo que hace que, dada una partición  $t_0 < \dots < t_n$  de un intervalo  $[a, b]$ , sea lógico esperar que

$$\sum_{k=1}^n |W_{t_k} - W_{t_{k-1}}|^2 \approx b - a .$$

En este capítulo se enuncian y demuestran rigurosamente estas propiedades, que por otra parte pueden encontrarse en multitud de libros.

#### 3.1 Proposición

*Casi seguro, las trayectorias de un proceso de Wiener  $W$  son no-diferenciables en ningún punto.*

*Demostración:* Supongamos para simplificar que estamos hablando de un proceso de Wiener estándar. (Si no, algunas de las cantidades que aparecerán vendrán afectadas por una constante, pero los argumentos no cambian.) Para  $k, n \in \mathbb{N}$ , sean

$$X_{nk} := \max \{ |W_{k2^{-n}} - W_{(k-1)2^{-n}}|, |W_{(k+1)2^{-n}} - W_{k2^{-n}}|, |W_{(k+2)2^{-n}} - W_{(k+1)2^{-n}}| \}$$

y

$$Y_n := \min_{\frac{k}{n} \leq 2^{-n}} X_{nk} .$$

Las tres variables dentro de los valores absolutos son independientes y tienen la misma ley que  $2^{-n/2}W_1$ , de donde

$$P\{X_{nk} \leq \varepsilon\} = (P\{|W_1| \leq \varepsilon 2^{n/2}\})^3 \leq (2\varepsilon 2^{n/2})^3 ,$$

y obtenemos

$$P\{Y_n \leq \varepsilon\} \leq \sum_{k=0}^{n2^n} P\{X_{nk} \leq \varepsilon\} \leq n2^n (2^{n/2+1}\varepsilon)^3 .$$

Denotemos

$$A := \{\omega \in \Omega : W(\omega) \text{ tiene derivada en algún punto}\} .$$

Si  $W(\omega)$  es derivable en el punto  $t$  y  $D$  es su derivada, entonces existirá  $\delta = \delta(t, \omega)$  tal que  $|s - t| \leq \delta \Rightarrow |W_s(\omega) - W_t(\omega)| \leq (|D| + 1)|s - t|$ .

Existe un  $n_0 = n_0(t, \omega)$  tal que

$$n \geq n_0 \Rightarrow \begin{cases} n > t , \\ n > 2(|D| + 1) , \\ 2^{-n} < \delta/2 . \end{cases}$$

Para cada  $n \geq n_0$ , escogemos  $k$  tal que  $k2^{-n} \leq t < \frac{k+1}{2^n}$ . Entonces  $|t - i2^{-n}| < \delta$  para  $i = k - 1, k, k + 1, k + 2$ . Por tanto,

$$X_{nk} \leq (|D| + 1)2 \cdot 2^{-n} \leq n2^{-n} ,$$

y en consecuencia, como  $k \leq 2^nt < 2^n n$ ,

$$Y_n \leq n2^{-n} .$$

Hemos probado que  $\omega \in A \Rightarrow \omega \in A_n := \{Y_n \leq n2^{-n}\}$ , para todo  $n$  a partir de un cierto  $n_0$ . Es decir,

$$A \subset \varliminf_n A_n .$$

Finalmente,

$$\begin{aligned} P(A) &\leq P(\varliminf_n A_n) = P(\bigcup_n \bigcap_{k \geq n} A_k) = \lim_{n \rightarrow \infty} P(\bigcap_{k \geq n} A_k) \\ &\leq \lim_{n \rightarrow \infty} P(A_n) \leq \lim_{n \rightarrow \infty} n2^n (2 \cdot 2^{n/2} n2^{-n})^3 = 0 . \quad \square \end{aligned}$$

### 3.2 Definiciones

Dados una función  $f: [a, b] \rightarrow \mathbb{R}$  y una partición  $\pi = \{a = t_0 < t_1 < \dots < t_n = b\}$  de  $[a, b]$ , se llama **variación de  $f$  en  $[a, b]$  respecto la partición  $\pi$**  a la cantidad

$$V_{[a,b],\pi}(f) := \sum_{t_i, t_{i+1} \in \pi} |f(t_{i+1}) - f(t_i)| .$$

Se llama **variación de  $f$  en  $[a, b]$**  a la cantidad

$$V_{[a,b]}(f) := \sup_{\pi} \sum_{t_i, t_{i+1} \in \pi} |f(t_{i+1}) - f(t_i)| ,$$

donde  $\pi$  varía en el conjunto de todas las particiones de  $[a, b]$ .

Una función tal que  $V_{[a,b]}(f) < \infty$  se dice que es de **variación acotada** (o **finita**) en  $[a, b]$ .  $\square$

Toda función real de variable real con variación acotada en un intervalo  $[a, b]$  es derivable en  $[a, b]$  excepto en un conjunto de medida de Lebesgue cero. Por tanto, como consecuencia de la Proposición 3.1, las trayectorias del proceso de Wiener son de variación no-acotada en todo intervalo no degenerado  $[a, b]$ , casi seguro. Veremos que esto se deduce también del hecho de que las trayectorias tienen casi seguro variación cuadrática no nula (véase la Proposición 3.7).

La definición de variación cuadrática que se usa en Análisis es la que parece natural: La variación cuadrática de  $f$  en un intervalo  $[a, b]$  es

$$V_{[a,b]}^2(f) := \sup_{\pi} \sum_{t_i, t_{i+1} \in \pi} |f(t_{i+1}) - f(t_i)|^2,$$

y en general se puede definir la variación de orden  $k$  de manera similar. No obstante, cuando hablamos de variación cuadrática en la teoría de procesos estocásticos no nos referimos exactamente a esto, sino al concepto que definimos seguidamente.

Denotaremos por  $\|\pi\| := \max_{t_i, t_{i+1} \in \pi} |t_{i+1} - t_i|$  la **norma** de una partición  $\pi$ . Una sucesión  $\{\pi_n\}_n$  de particiones de  $[a, b]$  decimos que es **refinante** si  $\pi_n \subset \pi_{n+1}$ ,  $\forall n$ .

### 3.3 Definición

Sea  $\{\pi_n\}_n$  una sucesión refinante de particiones de  $[a, b]$  tal que  $\lim_{n \rightarrow \infty} \|\pi_n\| = 0$ .

Se llama **variación cuadrática** de  $f: [a, b] \rightarrow \mathbb{R}$  al límite, si existe y no depende de la sucesión  $\{\pi_n\}_n$  escogida,

$$V_{[a,b]}^2(f) := \lim_{n \rightarrow \infty} V_{[a,b], \pi_n}^2(f) := \lim_{n \rightarrow \infty} \sum_{t_i, t_{i+1} \in \pi_n} |f(t_{i+1}) - f(t_i)|^2. \quad \square$$

Para funciones continuas, se tiene que

$$V_{[a,b]}(f) := \lim_{n \rightarrow \infty} \sum_{t_i, t_{i+1} \in \pi_n} |f(t_{i+1}) - f(t_i)|,$$

si  $\{\pi_n\}_n$  es cualquier sucesión refinante con  $\lim_{n \rightarrow \infty} \|\pi_n\| = 0$ . Pero esto no es cierto para la variación cuadrática.

### 3.4 Proposición

*Las trayectorias de un proceso de Wiener estándar  $W$  tienen casi seguro variación cuadrática igual a la constante  $b - a$  en todo intervalo  $[a, b]$ .*

*Idea de la demostración:* Este resultado se puede demostrar de la manera siguiente:

- 1) Tomar una sucesión refinante  $\{\pi_n\}_n$  de  $[a, b]$  con  $\lim_{n \rightarrow \infty} \|\pi_n\| = 0$  y ver que

$$L^2(\Omega)\text{-} \lim_{n \rightarrow \infty} V_{[a,b], \pi_n}^2(W) = b - a.$$

- 2) Demostrar que

$$\text{c.s.-} \lim_{n \rightarrow \infty} V_{[a,b], \pi_n}^2(W)$$

existe.

- 3) De 1) y 2) se obtiene que el límite casi seguro de  $V_{[a,b], \pi_n}^2(W)$  ha de ser  $b - a$ .  $\square$

La manera eficiente de demostrar la parte 2) de la proposición anterior requiere profundizar previamente en la teoría de martingalas, y no lo haremos aquí. Nos limitaremos a demostrar la parte 1), con una pequeña ampliación: Como es habitual cuando se tiene convergencia en  $L^2$  (o simplemente en probabilidad), si podemos asegurar que la convergencia es suficientemente rápida, se obtiene también convergencia casi segura, vía el fundamental Lema de Borel–Cantelli, que puede encontrarse en cualquier referencia básica de probabilidades.

### 3.5 Proposición

*Sea  $W$  un proceso de Wiener estándar. Sea  $\{\pi_n\}_n$  una sucesión refinante de particiones de  $[a, b]$  tal que  $\lim_{n \rightarrow \infty} \|\pi_n\| = 0$ .*

Entonces

$$L^2\text{-}\lim_{n \rightarrow \infty} V_{[a,b],\pi_n}^2(W) = b - a .$$

Si, además, se cumple  $\sum_n \|\pi_n\| < \infty$ , entonces

$$\text{c.s.-}\lim_{n \rightarrow \infty} V_{[a,b],\pi_n}^2(W) = b - a .$$

*Demostración:* Simplificaremos la notación poniendo

$$\begin{aligned} \Delta_i &:= t_{i+1} - t_i \\ \dot{W}(\Delta_i) &:= W_{t_{i+1}} - W_{t_i} \\ V_n^2 &:= V_{[a,b],\pi_n}^2(W) = \sum_{t_i, t_{i+1} \in \pi} \dot{W}(\Delta_i)^2 . \end{aligned}$$

Usando la independencia de incrementos del proceso de Wiener y las propiedades de la ley normal, tenemos que

$$\begin{aligned} \mathbb{E}[(V_n^2 - (b-a))^2] &= \sum_{i,j} \mathbb{E}[\dot{W}(\Delta_i)^2 \dot{W}(\Delta_j)^2] - 2(b-a) \sum_i \mathbb{E}[\dot{W}(\Delta_i)^2] + (b-a)^2 \\ &= \sum_{i \neq j} \Delta_i \Delta_j + 3 \sum_i \Delta_i^2 - 2(b-a) \sum_i \Delta_i + (b-a)^2 \\ &= 3 \sum_i \Delta_i^2 + \sum_{i \neq j} \Delta_i \Delta_j - \left( \sum_i \Delta_i \right)^2 + \left( \sum_i \Delta_i - (b-a) \right)^2 \\ &= 3 \sum_i \Delta_i^2 - \sum_i \Delta_i^2 \\ &\leq 2\|\pi_n\| \sum_i \Delta_i \\ &= 2\|\pi_n\|(b-a) \xrightarrow{n \rightarrow \infty} 0 . \end{aligned}$$

Supongamos ahora que  $\|\pi\| \xrightarrow{n \rightarrow \infty} 0$ . Denotemos  $X_n = V_n^2 - (b-a)$ . Acabamos de ver que  $\mathbb{E}[X_n^2] \leq 2\|\pi_n\|(b-a)$ , para todo  $n$ . Por la desigualdad de Chebishef, si ponemos  $A_n := \{|X_n| > \varepsilon\}$ ,

$$P(A_n) \leq \frac{\mathbb{E}[X_n^2]}{\varepsilon^2} \leq \frac{2\|\pi_n\|(b-a)}{\varepsilon^2}, \Rightarrow \sum_{n=1}^{\infty} P(A_n) < \infty .$$

Por el Lema de Borel-Cantelli,  $P\left(\bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m^c\right) = 1$ . Es decir, para todo  $\omega$  en este conjunto de probabilidad 1, existe  $n$  tal que si  $m \geq n$ , se tiene  $|X_m(\omega)| \leq \varepsilon$ , y obtenemos la convergencia casi segura.  $\square$

### 3.6 Observación

Si el proceso de Wiener tiene varianza  $\sigma^2$ , entonces la variación cuadrática en  $[a, b]$  es  $\sigma^2(b-a)$ .  $\square$

Ya hemos razonado que las trayectorias del proceso de Wiener han de ser casi seguro de variación no-acotada. Veamos como se puede deducir esto de la Proposición 3.5.

### 3.7 Proposición

*Las trayectorias de un proceso de Wiener estándar  $W$  tienen casi seguro variación no-acotada en todo intervalo no degenerado  $[a, b]$ .*



*Demostración:* En el conjunto  $\{V_{[a,b]}(W) < \infty\}$ , tendremos

$$\begin{aligned} b - a &= \lim_{n \rightarrow \infty} \sum_{t_i, t_{i+1} \in \pi_n} |W_{t_{i+1}} - W_{t_i}|^2 \\ &\leq \lim_{n \rightarrow \infty} \sup_{t_i, t_{i+1} \in \pi_n} |W_{t_{i+1}} - W_{t_i}| \sum_{t_i, t_{i+1} \in \pi_n} |W_{t_{i+1}} - W_{t_i}| \\ &\leq V_{[a,b]}(W) \lim_{n \rightarrow \infty} \sup_{t_i, t_{i+1} \in \pi_n} |W_{t_{i+1}} - W_{t_i}| \\ &= 0, \quad \text{c.s.}, \end{aligned}$$

puesto que las trayectorias de  $W$  son uniformemente continuas en  $[a, b]$ , y la desigualdad es absurda. Por tanto, el conjunto  $\{V_{[a,b]}(W) < \infty\}$  tiene probabilidad cero.

Además, el conjunto de probabilidad cero donde la variación sí es acotada, no depende del intervalo  $[a, b]$  escogido. En efecto, esto es obviamente cierto si nos restringimos a intervalos con extremos racionales. Pero cualquier intervalo abierto es unión numerable de intervalos de este tipo.  $\square$

### 3.8 Ejemplo (movimiento browniano, 3)

Sabemos que las trayectorias del proceso de Wiener son no-derivables en ningún punto. Por tanto, esta claro que en la ecuación (2.4) no podemos dividir por  $\Delta t$  y tomar  $\lim_{\Delta t \rightarrow 0}$  para llegar a una ecuación diferencial ordinaria, parametrizada por  $\omega$ . Hemos de quedarnos con la forma integral

$$v_t = v_0 - \beta \int_0^t v_s ds + \sigma W_t. \quad (3.1)$$

La ecuación (3.1), a pesar de todo, se puede resolver derivándola formalmente y aplicando técnicas elementales de ecuaciones diferenciales ordinarias. La solución que se obtiene,

$$v_t = e^{-\beta t} \left( v_0 + \sigma \int_0^t \dot{W}_s e^{\beta s} ds \right),$$

puede expresarse, después de integrar por partes, como

$$v_t = e^{-\beta t} \left( v_0 + \sigma W_t e^{\beta t} - \sigma \beta \int_0^t W_s e^{\beta s} ds \right), \quad (3.2)$$

que vuelve a tener sentido riguroso y se puede comprobar directamente que satisface (3.1). A partir de la velocidad encontramos, simplemente integrando, la función que nos da la posición de la partícula browniana en cada instante:

$$x_t = x_0 + \int_0^t \left[ e^{-\beta s} \left( v_0 + \sigma W_s e^{\beta s} - \sigma \beta \int_0^s W_r e^{\beta r} dr \right) \right] ds. \quad (3.3)$$

Los procesos  $\{x_t, t \in \mathbb{R}^+\}$  y  $\{v_t, t \in \mathbb{R}^+\}$  reciben el nombre de **proceso de posición y proceso de velocidad de Ornstein-Uhlenbeck**, respectivamente. Las fórmulas (3.2) y (3.3) dan explícitamente las trayectorias de  $x$  y  $v$  en función de las de  $W$ , y a partir de ellas podemos calcular la ley de estos procesos, que es en realidad la información útil que buscamos.  $x$  y  $v$  resultan ser procesos gaussianos con

$$\begin{aligned} \mathbb{E}[v_t] &= v_0 e^{-\beta t} \\ \text{Cov}[v_t, v_s] &= \frac{\sigma^2}{2\beta} (e^{2\beta(t \wedge s)} - 1) e^{-\beta(t+s)} \\ \mathbb{E}[x_t] &= x_0 + \frac{v_0}{\beta} (1 - e^{-\beta t}) \\ \text{Cov}[x_t, x_s] &= \frac{\sigma^2}{\beta^2} (s \wedge t) + \frac{\sigma^2}{2\beta^3} \left( -e^{-\beta(t+s)} - e^{-\beta|t-s|} + 2(e^{-\beta s} + e^{-\beta t} - 1) \right) \end{aligned}$$

donde  $\text{Cov}[X, Y] := \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$  se llama **covarianza** de las variables  $X$  e  $Y$ . La ley de un proceso gaussiano queda determinada por las medias y las covarianzas entre las variables del proceso. Los cálculos anteriores están detallados, por ejemplo, en Arnold [1].



## 4. Los modelos de Einstein-Smoluchowski y de Ornstein-Uhlenbeck para el movimiento browniano

Los movimientos erráticos de partículas pequeñas en un fluido del Ejemplo 1.1 fueron estudiados por primera vez por el botánico inglés Robert Brown alrededor de 1827 ([5], [6]). (Parece ser que hay también un trabajo anterior del neerlandés Jan Ingenhousz, hacia el 1785.) Desde entonces, este fenómeno ha sido conocido con el nombre de **movimiento browniano**. La primera descripción matemática del movimiento browniano fue presentada por Albert Einstein [10] en 1905 (véase Nelson [22] para un breve resumen de la historia del tema), quien propuso el proceso de Wiener como modelo matemático para las trayectorias de las partículas. Por cierto, Einstein dice que no conocía el trabajo de Brown; su deducción de la existencia de movimientos observables de las pequeñas partículas es puramente teórica, y se obtiene como consecuencia de la teoría atómica. Actualmente, “movimiento browniano” y “proceso de Wiener” se usan frecuentemente como sinónimos.

Más tarde el modelo de Einstein fue mejorado por Ornstein y Uhlenbeck (véase Uhlenbeck-Ornstein [26], Chandrasekar [7], Wang-Uhlenbeck [28], y especialmente Doob [9]) los cuales hicieron intervenir el proceso de Wiener en su modelo pero no directamente como versión matemática del movimiento browniano. Para evitar confusiones, parece mejor usar el nombre de proceso de Wiener para el objeto matemático, y el de movimiento browniano para el fenómeno físico.

### 4.1 El modelo de Einstein-Smoluchowski

La primera explicación mínimamente satisfactoria del movimiento browniano fue dada por Einstein (1905) y también independientemente por Marian von Smoluchowski (1906), quien trabajó más en el tema tanto desde el punto de vista teórico como experimental. La modelización del movimiento browniano como proceso de Wiener se conoce como el **modelo de Einstein-Smoluchowski**.

El razonamiento que hizo Einstein es como sigue (no hay mucho rigor matemático, pero las ideas son buenas):

Se supone que cada partícula individual ejecuta un movimiento que es independiente de los movimientos de las demás partículas. Se supone también que los movimientos de una determinada partícula en intervalos disjuntos de tiempo son procesos físicos independientes, siempre que estos intervalos “no se tomen demasiado pequeños”.

Introducimos un intervalo temporal de longitud  $\tau$ , muy pequeño comparado con los intervalos temporales observables, pero todavía lo bastante grande para que en dos intervalos sucesivos, los movimientos ejecutados puedan pensarse como independientes uno del otro.

Sea  $N$  el número total de partículas en suspensión en un líquido. En un intervalo temporal  $\tau$ , las coordenadas de las partículas individuales se incrementarán en una cantidad  $\Delta$ , donde, para cada partícula,  $\Delta$  tendrá un valor diferente (positivo o negativo).  $\Delta$  seguirá una cierta “ley de frecuencias”; la cantidad  $dN$  de partículas que experimentaran un desplazamiento que está entre  $\Delta$  y  $\Delta + d\Delta$  será expresable como

$$dN = N\phi(\Delta) d\Delta ,$$

donde

$$\int_{-\infty}^{+\infty} \phi(\Delta) d\Delta = 1 ,$$

y  $\phi$  es diferente de cero sólo para valores muy pequeños de  $\Delta$ , y satisface

$$\phi(\Delta) = \phi(-\Delta) .$$

Nos restringiremos al caso en que la concentración (cantidad de partículas por unidad de volumen) depende sólo de la coordenada  $x$  y el tiempo  $t$ . Denotémosla por  $f(x, t)$ .

Calcularemos la distribución de las partículas en el instante  $t + \tau$  a partir de la distribución en el instante  $t$ . La cantidad de partículas que en el instante  $t + \tau$  se encontrarán entre los puntos  $x$  y  $x + dx$  será

$$f(x, t + \tau) dx = dx \int_{-\infty}^{+\infty} f(x + \Delta, t) \phi(\Delta) d\Delta . \quad (4.1)$$

Como  $\tau$  es muy pequeño, podemos poner

$$f(x, t + \tau) = f(x, t) + \tau \frac{\partial f}{\partial t}(x, t) .$$

Por otra parte, desarrollemos  $f(x + \Delta, t)$  en potencias de  $\Delta$ :

$$f(x + \Delta, t) = f(x, t) + \Delta \frac{\partial f(x, t)}{\partial x} + \frac{\Delta^2}{2!} \frac{\partial^2 f(x, t)}{\partial x^2} + \dots$$

Podemos usar esta serie para hacer la integral de (4.1), porque sólo nos interesan valores pequeños de  $\Delta$ . Obtenemos

$$f + \tau \frac{\partial f}{\partial t} = f \int_{-\infty}^{\infty} \phi(\Delta) d\Delta + \frac{\partial f}{\partial x} \int_{-\infty}^{\infty} \Delta \phi(\Delta) d\Delta + \frac{\partial^2 f}{\partial x^2} \int_{-\infty}^{\infty} \frac{\Delta^2}{2} \phi(\Delta) d\Delta + \dots \quad (4.2)$$

Por la simetría de  $\phi$ , los términos segundo, cuarto, etc, de la derecha se anulan. Entre los otros, cada término es muy pequeño comparado con el anterior. Nos quedaremos sólo con el primero y el tercero. Introducimos la constante

$$D := \frac{1}{\tau} \int_{-\infty}^{+\infty} \frac{\Delta^2}{2} \phi(\Delta) d\Delta$$

y nos queda

$$\frac{\partial f}{\partial t} = D \frac{\partial^2 f}{\partial x^2} , \quad (4.3)$$

la *ecuación de la difusión*, ya conocida en tiempos de Einstein. La difusión es un fenómeno físico que consiste en el hecho de que un conjunto de partículas disueltas en un fluido tienden a desplazarse de las zonas de más concentración a las zonas de menos concentración, como por ejemplo una gota de tinta en un cubo de agua, que se va esparciendo y perdiendo color, hasta desaparecer.  $D$  es una constante dependiente del medio que se llama **coeficiente de difusión**.

La solución es, suponiendo que hacemos difusión de  $N$  partículas a partir del origen, y despreciando la interacción entre ellas,

$$f(x, t) = \frac{N}{\sqrt{4\pi Dt}} \exp\{-x^2/4Dt\} . \quad (4.4)$$

En  $\mathbb{R}^3$ , obtendríamos la ecuación

$$\frac{df}{dt} = D \sum_{i=1}^3 \frac{\partial^2 f}{\partial x_i^2}, \quad (4.5)$$

y el resultado sería

$$f(t, x) = \frac{N}{(4\pi Dt)^{3/2}} \exp\left\{-\frac{|x|^2}{4Dt}\right\}. \quad (4.6)$$

Para una sola partícula ( $N = 1$ ), (4.6) puede interpretarse como la densidad de probabilidad de encontrar la partícula en una región del espacio determinada al cabo de un tiempo  $t$ . Es decir, la probabilidad de encontrar la partícula en la región  $A$  viene dada por

$$\int_A (4\pi Dt)^{-3/2} \exp\{-|x|^2/4Dt\} dx. \quad (4.7)$$

Esta densidad esencialmente caracteriza al proceso de Wiener (el proceso de Wiener en  $\mathbb{R}^3$  tiene por coordenadas procesos de Wiener reales independientes). Más precisamente, es la densidad de  $X_t = \sqrt{2D}W_t$ , donde  $W$  es un proceso de Wiener estándar (tridimensional). Esto justifica el nombre de coeficiente de difusión para la constante  $\sigma^2$  en el Teorema 2.1. (En realidad, sale  $\sigma^2 = 2D$ . Si se quiere que el 2 no aparezca, hay que cambiar  $D$  por  $D/2$  en las ecuaciones (4.3) y (4.5).

Los diversos artículos de Einstein sobre el movimiento browniano están recogidos en el libro de Furth [12]. Los artículos clásicos posteriores de Uhlenbeck y Ornstein, Chandrasekar, Wang y Uhlenbeck, y Doob, se encuentran reunidos también en forma de libro (Wax [29]).

El proceso de Wiener fue estudiado por primera vez con rigor por Norbert Wiener, en una serie de trabajos entre 1923 y 1934, y de ahí su nombre. Wiener se benefició de los recientes avances en teoría de la medida llevados a cabo por Émile Borel, Henri Lebesgue y J. P. Daniell, para hacer una construcción del proceso a partir del espacio de trayectorias. Einstein no disponía todavía de estas herramientas y su construcción se limita a la ley; no a las trayectorias.

## 4.2 El modelo de Ornstein-Uhlenbeck

Modelizar el movimiento browniano como un proceso de Wiener (modelo de Einstein-Smoluchowski) tiene el inconveniente de que la propiedad de variación total infinita de las trayectorias implica que la partícula browniana recorre caminos de longitud infinita en tiempo finito. Esto estimuló el abandono de este modelo.

En cambio, el modelo que estamos desarrollando (conocido como modelo de Ornstein-Uhlenbeck) da lugar a velocidades finitas. Son las aceleraciones las que pasan a tener valores infinitos.

Hemos empezado formulando unas ciertas hipótesis físicas de tipo microscópico (de comportamiento mecánico molecular) con la idea que a partir de ellas se puedan predecir de forma teórica los comportamientos macroscópicos observables. Esto nos ha llevado a la ecuación “física” (1.3). Posteriormente, hemos visto las herramientas que posibilitan formular un modelo con pleno sentido matemático, (2.4) o (3.1). Un último paso, puramente notacional en este momento, es expresar la ecuación (3.1) en “forma diferencial”

$$dv_t = -\beta v_t dt + \sigma dW_t, \quad (4.8)$$

o

$$\dot{v}_t = -\beta v_t + \sigma \dot{W}_t. \quad (4.9)$$

A pesar de que  $dW_t$  no existe, podemos tomar (4.8) y (4.9) simplemente como símbolos que abrevian la formulación integral (3.1). En todo caso, constituyen un primer ejemplo (muy sencillo!) de lo que llamamos una *ecuación diferencial estocástica*.

Vamos a completar la formulación del modelo de Ornstein-Uhlenbeck para el movimiento browniano, determinando el único elemento de la ecuación que las hipótesis iniciales no nos han permitido determinar: el coeficiente  $\sigma$ .

De la fórmula para  $\text{Cov}[v_t, v_s]$  del Capítulo 3, deducimos que

$$\text{Var}[v_t] = \frac{\sigma^2}{2\beta}(1 - e^{-2\beta t}) ,$$

y por tanto, haciendo  $t \rightarrow \infty$ , obtenemos una varianza límite de  $\sigma^2/2\beta$ , sea cual sea la velocidad inicial  $v_0$ . Análogamente, se observa que  $\lim_{t \rightarrow \infty} E[v_t] = 0$ . La ley de equipartición de la energía en mecánica estadística afirma que la energía cinética media de un sistema de partículas (las moléculas del fluido) en estado de equilibrio termodinámico ha de valer  $\frac{1}{2}kT$  por grado de libertad del sistema, siendo  $k$  la constante de Boltzmann y  $T$  la temperatura absoluta. La partícula sumergida en el fluido ha de compartir esta misma energía media. La esperanza de la energía cinética de la partícula se puede calcular pues como

$$E[\frac{1}{2}v^2] = \frac{\sigma^2}{4\beta}$$

(donde  $v$  representa la velocidad límite y seguimos haciendo el convenio de que la masa de la partícula vale 1), y obtenemos por tanto la igualdad

$$\frac{\sigma^2}{4\beta} = \frac{kT}{2} . \quad (4.10)$$

Por otro lado, hay una relación sencilla entre el coeficiente de difusión  $D$  y el de fricción  $\beta$ , demostrada por Einstein:

$$D = \frac{kT}{\beta} . \quad (4.11)$$

Deducimos de (4.10) y (4.11) que, en el caso de dimensión 1,  $\sigma^2$  se expresa en términos de los coeficientes de fricción y de difusión como

$$\sigma^2 = 2\beta^2 D .$$

Nuestra ecuación diferencial estocástica (4.8) para la velocidad de la partícula browniana es pues

$$dv_t = -\beta v_t dt + \beta\sqrt{2D} dW_t ,$$

y queda completo el modelo. La nueva hipótesis que hemos introducido para poder determinar  $\sigma$  es que el sistema de moléculas verifica la ecuación térmica de los gases ideales

$$pV = nkT ,$$

que es de donde se deduce la ley de equipartición. Aquí  $p$  es la presión,  $V$  el volumen y  $n$  el número total de moléculas.

## 5. Integración estocástica (I)

### 5.1 El problema

Hemos podido resolver la ecuación (3.1) sin usar nada más que los trucos habituales del cálculo, pero se trata de hecho de un caso muy particular. El punto clave es que, aunque no tenga sentido hablar del diferencial  $dW_t$  para ninguna trayectoria, en cambio sí se puede dar un sentido evidente a la expresión  $\int_0^t dW_s$ . Tan solo hay que definir este símbolo como  $W_t - W_0$ , que no es más que el valor que se obtiene mediante sumas de Riemann. En general, si  $X_t$  es una función (o un proceso) de variación acotada, y teniendo en cuenta que  $W$  es un proceso continuo, se puede definir

$$\int_0^t X_s dW_s$$

mediante sumas de Riemann para cada valor del parámetro  $\omega$ . El resultado coincide con el que se obtendría integrando formalmente por partes:

$$X_t W_t - X_0 W_0 - \int_0^t W_s dX_s ,$$

donde la última integral tiene sentido como integral de Riemann–Stieljes clásica. (Recuérdese que, más en general, para toda función de variación acotada  $\alpha$ , se puede definir la integral de Lebesgue respecto  $d\alpha$

$$\int_a^b f(t) d\alpha(t)$$

de cualquier función  $f$  que sea Lebesgue-integrable en  $[a, b]$  respecto la variación total de  $\alpha$ .)

#### 5.1 Ejemplo (Verhulst, 2)

Retomando el Ejemplo 1.2 del Capítulo 1 (modelo de Verhulst con tasa de natalidad perturbada), apliquemos a la ecuación discretizada (1.7) el mismo razonamiento que hemos usado en el movimiento browniano. Obtenemos, en forma integral, (tal vez con alguna constante  $\sigma$  multiplicando a  $W$ )

$$x_t = x_0 + \int_0^t x_s(a - bx_s) ds + \int_0^t x_s dW_s ,$$

i en forma diferencial,

$$dx_t = x_t(a - bx_t) dt + x_t dW_t , \tag{5.1}$$

con la condición inicial  $x_0$ .  $\square$

Como se ha visto en los ejemplos, hay situaciones en las cuales no funciona la integración por partes para definir la integral respecto de un proceso estocástico con trayectorias irregulares. En efecto, obsérvese en (5.1) si el proceso solución  $x_t$  fuera de variación acotada, también lo sería integral respecto  $dt$  y en consecuencia la integral respecto  $dW_t$  también debería serlo; por otra parte, integrando por partes esta integral el resultado es claramente de variación no acotada. Por tanto no es en absoluto obvio el sentido que hay que dar al símbolo

$$\int_0^t x_t dW_t$$

en estos casos. La ecuación (5.1) es una “genuina” ecuación diferencial estocástica, en el sentido que no es una pura aleatorización de una ecuación diferencial ordinaria.

En general, si la influencia de una perturbación externa a un sistema depende del tiempo y del estado del sistema en cada instante, frecuentemente se puede modelizar el sistema resultante mediante una ecuación diferencial del tipo

$$dx(t) = f(t, x(t)) dt + g(t, x(t)) dM(t) , \quad (5.2)$$

donde  $dx(t) = f(t, x(t))dt$  es la ecuación correspondiente al sistema sin perturbar y  $g$  es una función que representa la “sensitividad” del sistema a la perturbación  $M$ .

Cuando  $M$  es aleatoria e irregular, es frecuente que pueda modelizarse mediante un proceso de Wiener, y se obtiene una ecuación que no sólo no tiene sentido clásico en forma diferencial, sino que tampoco lo tiene en forma integral

$$x_t = x_0 + \int_0^t f(s, x_s) ds + \int_0^t g(s, x_s) dW_s , \quad (5.3)$$

a menos que  $g$  sea tan regular que suavice las irregularidades que  $x_t$  hereda de  $W_t$ , como sucedía en nuestro primer ejemplo ( $g$  constante).

Hay que dar un sentido pues al símbolo

$$\int_0^t g(s, x_s) dW_s \quad (5.4)$$

de forma que se pueda interpretar como la aportación total de la perturbación aleatoria al estado del sistema en el intervalo  $[0, t]$ .

El resultado de integrar un proceso en un intervalo fijado  $[0, t]$  será naturalmente una variable aleatoria. Por tanto lo que buscamos es un operador con dominio un cierto espacio de procesos, a valores en  $L^0(\Omega, \mathcal{F}, P)$  y que para merecer el nombre de integral, debería ser

- 1) lineal,
- 2) continuo en algún sentido (lo que equivale a tener algún teorema de convergencia).

Queremos, además, que

- 3) si  $X_s(\omega)$  es de variación acotada para todo  $\omega$ , se tenga

$$\left[ \int_0^t X_s dW_s \right](\omega) = \int_0^t X_s(\omega) dW_s(\omega) ,$$

extendiendo de este modo el concepto de integral de Stieljes usual.

Un objetivo modesto es intentar que el operador actúe al menos sobre procesos continuos. Las condiciones 2) y 3) sugieren entonces definir la integral de un proceso continuo  $f_t(\omega)$  respecto el proceso de Wiener como el límite de las sumas

$$\left[ \int_0^t X_s dW_s \right](\omega) := \lim_{n \rightarrow \infty} \sum_{t_i, t_{i+1} \in \pi_n} X_{\xi_i}(\omega) \cdot (W_{t_{i+1}}(\omega) - W_{t_i}(\omega)) ,$$



si el límite de la derecha existe y es independiente de la sucesión de particiones refinante  $\{\pi_n\}_{n \in \mathbb{N}}$  con  $\|\pi_n\| \rightarrow 0$  y los puntos  $\xi_i \in [t_i, t_{i+1}]$  escogidos. Y esto para cada  $\omega$  fijado o, más débilmente, tomando el límite en la topología natural de  $L^0(\Omega, \mathcal{F}, P)$ , i.e. la topología de la convergencia en medida. Esto correspondería a definir un operador continuo entre el espacio de procesos continuos con la topología inducida por la convergencia uniforme (en  $(t, \omega)$ ) y el espacio de variables aleatorias  $L^0(\Omega, \mathcal{F}, p)$  con la convergencia en medida.

Pero desgraciadamente esto es imposible. En efecto:

## 5.2 Teorema

Sea  $\alpha(t)$  una función continua en  $[0, 1]$ .

Sea  $\{\pi_n\}_{n \in \mathbb{N}}$  una sucesión refinante de particiones de  $[0, 1]$  tal que la sucesión de normas tiende a cero.

Entonces,

$$\sum_{t_i, t_{i+1} \in \pi_n} f(t_i) \cdot (\alpha(t_{i+1}) - \alpha(t_i))$$

converge a un límite finito cuando  $n \rightarrow \infty$  para toda función continua  $f$  si y sólo si  $\alpha(t)$  es de variación acotada.  $\square$

Reproducimos aquí una demostración simple de este Teorema, sugerida por Meyer [20] y que contiene la clave para evitar estas dificultades. La demostración usa el conocido Teorema de Banach–Steinhaus, que recordamos:

## 5.3 Lema (Teorema de Banach–Steinhaus).

Sean  $B_1$  un espacio de Banach y  $B_2$  un espacio normado. Sea  $\{T_i\}_{i \in I}$  una familia de operadores lineales acotados  $T_i: B_1 \rightarrow B_2$ .

Si para cada  $x \in B_1$ ,  $\sup_{i \in I} \|T_i x\| < \infty$ , entonces  $\sup_{i \in I} \|T_i\| < \infty$ .

*Demostración del Teorema:*

$\Leftarrow$ ) Es conocido. Converge a la integral de Riemann–Stieljes  $\int_0^1 f(t) d\alpha(t)$ .

$\Rightarrow$ ) Sean  $B_1$  el espacio de funciones continuas en  $[0, 1]$  con la norma del supremo,  $B_2 = \mathbb{R}$  con el valor absoluto.

Para a cada  $n \in \mathbb{N}$ , sea  $T_n: B_1 \rightarrow B_2$  el operador definido por

$$T_n f := \sum_{t_i, t_{i+1} \in \pi_n} f(t_i) \cdot (\alpha(t_{i+1}) - \alpha(t_i)) .$$

Por hipótesis,  $\lim_{n \rightarrow \infty} T_n f$  existe, y por tanto  $\sup_n \|T_n f\| < \infty$ . Gracias al Lema,  $\sup_n \|T_n\| < \infty$ .

Vamos a ver que la variación de  $\alpha$ ,

$$V^{(1)}(\alpha) := \sup_{\pi \in \mathcal{P}} \sum_{t_i, t_{i+1} \in \pi} |\alpha(t_{i+1}) - \alpha(t_i)| ,$$

es más pequeña que  $\sup_n \|T_n\|$ , con lo que habremos acabado. Efectivamente, para cada  $n \in \mathbb{N}$  fijado, sea  $f$  una función continua tal que  $f(t_i) = \text{signo}\{\alpha(t_{i+1}) - \alpha(t_i)\}$  y  $\|f\|_\infty = 1$ . Entonces:

$$\begin{aligned} T_n f &= \sum_{t_i, t_{i+1} \in \pi_n} f(t_i) (\alpha(t_{i+1}) - \alpha(t_i)) = \sum_{t_i, t_{i+1} \in \pi_n} |\alpha(t_{i+1}) - \alpha(t_i)| \Rightarrow \\ \|T_n\| &\geq \sum_{t_i, t_{i+1} \in \pi_n} |\alpha(t_{i+1}) - \alpha(t_i)| \Rightarrow \\ \sup_n \|T_n\| &\geq \sup_n \sum_{t_i, t_{i+1} \in \pi_n} |\alpha(t_{i+1}) - \alpha(t_i)| = V^{(1)}(\alpha) , \end{aligned}$$

la última igualdad gracias a que el  $\sup_n$  es sustituible por  $\lim_n$ , y la caracterización de la variación de una función en términos de sucesiones de particiones mencionada en el Capítulo 3.  $\square$

Se puede pensar que, al ser el operador que buscamos valuado en  $L^0(\Omega, \mathcal{F}, p)$  en vez de  $\mathbb{R}$ , tal vez sí que se puede obtener la convergencia en medida de las sumas

$$\sum_{t_i, t_{i+1} \in \pi_n} f_{t_i}(\omega) \cdot (\alpha_{t_{i+1}}(\omega) - \alpha_{t_i}(\omega)) \quad (5.5)$$

en vez de la convergencia casi segura en  $\omega$ . Pero la situación no es mejor, ya que en el conjunto  $A = \{\omega \in \Omega : \alpha_t(\omega) \text{ no es de variación acotada}\}$  podríamos construir una sucesión parcial de (5.5) convergente casi seguro, y repetir la demostración del Teorema para esta parcial. Por tanto, si  $P(A) > 0$ , llegamos a la misma dificultad que en el caso determinista, en que  $\omega$  no està.

## 5.2 Integrales estocásticas elementales

Llamamos **integrales estocásticas** a las expresiones del tipo

$$\int_a^b X_t dZ_t, \quad (5.6)$$

en que  $Z$  es un proceso estocástico, y  $X$  es una función determinista o, tal vez, otro proceso estocástico. Hay tres tipos de integrales estocásticas que podemos llamar “elementales”:

Hemos visto que si  $X$  y  $Z$  tienen trayectorias de variación acotada, podemos definir de manera natural la variable aleatoria (5.6) como

$$\left[ \int_a^b X_t dZ_t \right](\omega) := \int_a^b X_t(\omega) dZ_t(\omega). \quad (5.7)$$

y en general podemos usar esta definición siempre que tenga sentido el término de la derecha con probabilidad 1.

Si  $Z$  no tiene trayectorias de variación acotada pero  $X$  sí, entonces podemos usar la integración por partes (estamos suponiendo que ahora  $X$  y  $Z$  son continuos para simplificar):

$$\int_a^b X_t dZ_t := X_t Z_t - X_0 Z_0 - \int_a^b Z_t dX_t,$$

donde la última integral se define por  $\omega$  como en (5.7).

El tercer tipo de integral que podríamos todavía llamar “elemental”, pero que empieza a mostrar algunos elementos característicos de la integración estocástica no trivial, es la que puede construirse cuando el integrando es determinista, y el integrador es de cuadrado integrable ( $Z_t \in L^2(\Omega), \forall t$ ) y con incrementos independientes. Esta integral tiene interés por sí misma aparte de las ecuaciones diferenciales estocásticas, pues se usa en la teoría de procesos estacionarios y en el estudio de series temporales.

Queremos definir la integral de  $f$  respecto  $W$  en  $[a, b]$  para toda función  $f$  del espacio  $L^2([a, b], dt)$ . Sea  $\mathcal{S} \subset L^2([a, b], dt)$  el conjunto de todas las funciones  $f$  de la forma

$$f(t) := \sum_{i=1}^n c_i \mathbf{1}_{[t_{i-1}, t_i]}(t),$$

$a = t_0 < t_1 < \dots < t_n = b$  (funciones “escalonadas”).

Para  $f \in \mathcal{S}$  definimos

$$\int_a^b f(t) dW_t := \sum_{i=1}^n c_i \cdot (W_{t_i} - W_{t_{i-1}}).$$

Esta definición es independiente de la representación particular de la función  $f$ .

#### 5.4 Proposición

La aplicación

$$\begin{aligned} I: \mathcal{S} \subset L^2([a, b], dt) &\longrightarrow L^2(\Omega, \mathcal{F}, P) \\ f &\longmapsto \int_a^b f(t) dW_t \end{aligned}$$

tiene las propiedades siguientes:

- 1)  $I(c_1 f_1 + c_2 f_2) = c_1 I(f_1) + c_2 I(f_2)$ .
- 2)  $E[I(f)] = 0$ .
- 3)  $E[I(f)I(g)] = \int_a^b f(t)g(t) dt$ , y, en particular, preserva la norma:  $\|I(f)\|_{L^2(\Omega)} = \|f\|_{L^2([a, b], dt)}$ .

*Demostración:*

- (1) Inmediato.
- (2) Sale del hecho de que  $W$  es un proceso centrado ( $E[W_t] = 0, \forall t$ ).
- (3) Representamos  $f$  y  $g$  respecto la misma partición:

$$\begin{aligned} f(t) &= \sum_{i=1}^n c_i \mathbf{1}_{]t_{i-1}, t_i]} , \\ g(t) &= \sum_{i=1}^n d_i \mathbf{1}_{]t_{i-1}, t_i]} . \end{aligned}$$

Entonces,

$$\begin{aligned} E[I(f)I(g)] &= E\left[\sum_{i,j=1}^n c_i d_j (W_{t_i} - W_{t_{i-1}})(W_{t_j} - W_{t_{j-1}})\right] \\ &= \sum_{i=1}^n c_i d_i E[(W_{t_i} - W_{t_{i-1}})^2] = \sum_{i=1}^n c_i d_i (t_i - t_{i-1}) \\ &= \int_a^b f(t)g(t) dt . \quad \square \end{aligned}$$

#### 5.5 Proposición

El conjunto  $\mathcal{S}$  es un subespacio denso de  $L^2([a, b], dt)$ .  $\square$

Sea  $f \in L^2([a, b], dt)$ , y sea  $\{f_n\}_n \subset \mathcal{S}$  una sucesión convergente a  $f$  en  $L^2([a, b], dt)$ . Esta sucesión será de Cauchy en  $L^2([a, b], dt)$  y, por la propiedad de isometría (Proposición 5.4, 3), su imagen  $\{I(f_n)\}_n$  será de Cauchy en  $L^2(\Omega)$ , y por tanto convergente.

#### 5.6 Definición

En la situación anterior, se define la integral estocástica de  $f$  respecto  $W$  en  $[a, b]$  como

$$\int_a^b f(t) dW_t := \lim_{n \rightarrow \infty} \int_a^b f_n(t) dW_t . \quad \square$$

Para que la definición tenga sentido, tiene que ser independiente de la sucesión  $\{f_n\}_n$  escogida. Se puede comprobar que lo es.

Hemos definido pues un operador

$$\begin{aligned} I: L^2([a, b], dt) &\longrightarrow L^2(\Omega) \\ f &\longmapsto \int_a^b f(t) dW_t \end{aligned}$$

### 5.7 Proposición

El operador integral estocástica  $I$  tiene las propiedades 1), 2) y 3) de la Proposición 5.4.

*Demostración:* Sólo hay que ver que las citadas propiedades se conserven por paso al límite en  $L^2([a, b], dt)$ : La linealidad es obvia; la convergencia en  $L^2$  implica la convergencia de las esperanzas; por último, en todo espacio de Hilbert, si  $x_n \rightarrow x$  y  $y_n \rightarrow y$ , entonces  $\langle x_n, y_n \rangle \rightarrow \langle x, y \rangle$ .

□

Se puede considerar la integral estocástica como un proceso: Dada  $f \in \bigcap_{t \in \mathbb{R}^+} L^2([0, t])$ , podemos considerar el **proceso integral estocástica**

$$\left\{ \int_0^t f(s) dW_s, t \in \mathbb{R}^+ \right\}.$$

### 5.8 Proposición

El proceso integral estocástica tiene incrementos independientes. □

Esta construcción de la integral estocástica es válida usando cualquier proceso  $Z$  con incrementos independientes y continuo por la derecha en  $L^2$  (es decir, tal que  $E[|Z_t - Z_s|^2] \rightarrow 0$  cuando  $t \searrow s$ ) como integrador. En tal caso, existe una función de distribución (i.e., una función creciente continua por la derecha)  $F$  tal que

$$F(b) - F(a) = E[|Z_b - Z_a|^2], \quad 0 \leq a \leq b. \quad (5.8)$$

Las funciones  $f$  que podemos usar como integrandos son las del espacio  $L^2([a, b], dF)$ , y la integral estocástica es una isometría

$$\begin{aligned} I: L^2([a, b], dF) &\longrightarrow L^2(\Omega) \\ f &\longmapsto \int_a^b f(t) dZ_t \end{aligned}$$

No tiene dificultad repetir todo lo anterior en esta situación más general. Además, se tiene el siguiente resultado de tipo Radon-Nikodým:

### 5.9 Proposición

Sea  $F$  la función de distribución asociada al proceso con incrementos independientes  $Z$ . Sean  $f \in L^2([0, t], dF)$  y  $Y_t := \int_0^t f(t) dZ_t$ .

Entonces, el proceso con incrementos independientes  $Y$  tiene función de distribución asociada  $G(t) = \int_0^t f^2(t) dF(t)$ , y si  $g \in L^2([0, t], dG)$  se tiene

$$\int_0^t g(s) dY_s = \int_0^t g(s) f(s) dZ_s, \quad c.s.$$

## 6. Integración estocástica (II)

### 6.1 La integral estocástica de Itô

Para construir la integral estocástica con integrandos aleatorios, la construcción anterior no es suficiente. La demostración del Teorema 5.2 nos dará una pista de cómo proceder ahora. En principio, este teorema parece que nos previene de poder hacer una definición de la integral mediante sumas de Riemann; no obstante, eso es precisamente lo que vamos a hacer.

Está claro que tendremos que imponer alguna restricción al integrando, pero no queremos pedir más regularidad a sus trayectorias, porque no obtendríamos nada útil de cara a las ecuaciones diferenciales estocásticas. Buscaremos restricciones de otro tipo.

Repasemos un momento la demostración del último Teorema. Para ver la necesidad de que el integrador sea de variación acotada hemos usado funciones que tomaban unos valores determinados sobre los puntos de cada partición:

$$f(t_i) := \text{signo}\{\alpha(t_{i+1}) - \alpha(t_i)\}$$

Observemos que, para asignar estos valores,  $f$  “mira hacia adelante” la función  $\alpha$ ; es decir, el valor de  $f$  en  $t_i$  depende del valor de  $\alpha$  en  $t_{i+1}$ . Si prohibimos esta dependencia, la demostración queda invalidada.

Nuestro objetivo es pues poner esta restricción en términos matemáticos, y observar después que no impide modelizar problemas reales con el formalismo de (5.4). Hay que introducir primero unos cuantos conceptos adicionales.

Dado un espacio de probabilidad  $(\Omega, \mathcal{F}, P)$ , cualquier familia creciente  $\{\mathcal{F}_t, t \in \mathbb{R}^+\}$  de sub- $\sigma$ -álgebras de  $\mathcal{F}$  se llama una **filtración**, y la estructura  $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t, t \in \mathbb{R}^+\})$  es un **espacio de probabilidad filtrado**. Una filtración *toma nota de la evolución en el tiempo de la información disponible en cada instante*. Expliquemos esta frase. Una variable aleatoria  $X$  definida en  $(\Omega, \mathcal{F}, P)$  induce una sub- $\sigma$ -álgebra de  $\mathcal{F}$ , llamada  $\sigma$ -álgebra **generada** por  $X$ , que es la más pequeña que hace medible la aplicación  $X$ , y que se denota habitualmente por  $\sigma\langle X \rangle$ . Esta  $\sigma$ -álgebra está formada por los únicos conjuntos que son *observables* a través de  $X$ ; dicho de otro modo, de los únicos sucesos que podemos decidir si se han verificado o no (i.e., si  $\omega$  pertenece a él o no) con la sola información del resultado concreto  $X(\omega)$ . En este sentido  $\sigma\langle X \rangle$  es la información suministrada por  $X$ . Ahora, si  $X \equiv \{X_t, t \in \mathbb{R}^+\}$  es un proceso estocástico, la información total suministrada por el proceso hasta en el instante  $t$  es la más pequeña  $\sigma$ -álgebra que hace medibles todas las aplicaciones de la familia  $\{X_s, s \leq t\}$ . Lo escribimos  $\sigma\langle X_s, s \leq t \rangle$ . Los conjuntos

de esta  $\sigma$ -álgebra son los observables a través del proceso una vez éste ha llegado al instante  $t$ .  $\{\sigma\langle X_s, s \leq t \rangle, t \in \mathbb{R}^+\}$  es la **filtración generada** por el proceso  $X$ .

Dado un espacio de probabilidad filtrado  $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t, t \in \mathbb{R}^+\})$ , un proceso estocástico  $X$  sobre  $(\Omega, \mathcal{F}, P)$  se dice que es **adaptado a**  $\{\mathcal{F}_t, t \in \mathbb{R}^+\}$  si  $X_t$  es  $\mathcal{F}_t$ -medible (medible respecto  $\mathcal{F}_t$ ) para cada  $t \in \mathbb{R}^+$ . Si  $\{\mathcal{F}_t, t \in \mathbb{R}^+\}$  es la filtración generada por un otro proceso  $Y$  decimos que  $X$  es **adaptado a**  $Y$ .

Sea  $W \equiv \{W_t, t \in \mathbb{R}^+\}$  un proceso de Wiener. Una filtración  $\{\mathcal{F}_t, t \in \mathbb{R}^+\}$  se dice que **no-anticipa**  $W$  si se satisfacen las condiciones siguientes:

- 1)  $W$  es adaptado a  $\{\mathcal{F}_t, t \in \mathbb{R}^+\}$ .
- 2)  $\forall s \in \mathbb{R}, \mathcal{F}_s$  y  $\sigma\langle W_t - W_s, t \geq s \rangle$  son  $\sigma$ -álgebras independientes (esto es, todo suceso de una de ellas es independiente de todo suceso de la otra).

Tales filtraciones efectivamente existen. Tómese por ejemplo la inducida por el propio  $W$ , o sea  $\{\sigma\langle W_s, s \leq t \rangle, t \in \mathbb{R}^+\}$ , gracias a la propiedad de incrementos independientes de  $W$ .

Un proceso estocástico  $X$  se dice que es **medible** si, como aplicación  $X: \mathbb{R}^+ \times \Omega \rightarrow \mathbb{R}$ , es medible respecto la  $\sigma$ -álgebra producto de los borelianos de  $\mathbb{R}^+$  con  $\mathcal{F}$ .

Diremos que un proceso estocástico  $X$  es **no-anticipativo** respecto un proceso de Wiener  $W$  si:

- 1)  $X$  es un proceso medible.
- 2)  $X$  es adaptado a alguna filtración no-anticipativa de  $W$ .

A veces los llamaremos simplemente procesos “medibles y adaptados”.

Ahora tenemos a la nuestra disposición los ingredientes necesarios para modelizar la condición “no vale mirar hacia adelante” de la que hablábamos. Como veremos en seguida, la condición de no-anticipación del integrando permite desarrollar una teoría satisfactoria de la integración respecto al proceso de Wiener.

### 6.1 Notación

Denotaremos por  $\mathcal{P}[a, b]$  la colección de procesos no-anticipativos  $X$  tales que  $P\{\omega : X(\omega) \in L^2([a, b])\} = 1$ .

Normalmente, identificaremos en  $\mathcal{P}[a, b]$  dos procesos que sean *indistinguibles* (es decir, con las mismas trayectorias excepto un conjunto de probabilidad cero). De esta manera estamos haciendo un paso al cociente, pero seguiremos usando la notación  $\mathcal{P}[a, b]$  para este conjunto cociente.  $\square$

### 6.2 Proposición

*El conjunto  $\mathcal{P}[a, b]$  es un espacio vectorial con las operaciones usuales de suma y producto por escalares.*

*Demostración:* Si  $X^1, \dots, X^n$  son procesos no-anticipativos y  $H: \mathbb{R}^n \rightarrow \mathbb{R}$  es una función medible, está claro que  $H(X^1, \dots, X^n)$  es un proceso no-anticipativo. En particular, es cierto si  $H$  es una función lineal.  $\square$

Vamos a introducir una convergencia a el espacio vectorial  $\mathcal{P}[a, b]$ , y veremos que le dota de estructura de espacio métrico.

### 6.3 Definición

Sea  $\{X^n\}_n \subset \mathcal{P}[a, b]$ .

Diremos que  $\{X^n\}_n$  converge a  $X \in \mathcal{P}[a, b]$  si la sucesión de variables aleatorias

$$\left\{ \omega \mapsto \int_a^b |X_t^n(\omega) - X_t(\omega)|^2 dt \right\}_n$$

converge a cero en probabilidad. Dicho de otra manera, si

$$\|X^n - X\|_{L^2([a,b])} \xrightarrow{P} 0. \quad \square$$

## 6.4 Proposición

*La aplicación*

$$d: \mathcal{P}[a, b] \times \mathcal{P}[a, b] \longrightarrow \mathbb{R}^+$$

$$(X, Y) \longmapsto \mathbb{E} \left[ \frac{\|X - Y\|_{L^2([a,b])}}{1 + \|X - Y\|_{L^2([a,b])}} \right]$$

es una distancia y metriza la convergencia de la Definición 6.3.

*Demostración:* Sabemos que  $d(X, Y) := \|X - Y\|$  es una distancia en  $L^2([a, b])$ . Por otra parte, siempre que  $d$  es una distancia en un espacio métrico, se tiene que  $d'(x, y) = \frac{d(x, y)}{1 + d(x, y)}$  es también una distancia en el mismo espacio (equivalente al anterior). Esto nos da la desigualdad triangular

$$\frac{\|X - Y\|_{L^2([a,b])}}{1 + \|X - Y\|_{L^2([a,b])}} \leq \frac{\|X - Z\|_{L^2([a,b])}}{1 + \|X - Z\|_{L^2([a,b])}} + \frac{\|Z - Y\|_{L^2([a,b])}}{1 + \|Z - Y\|_{L^2([a,b])}}.$$

Tomando esperanzas, resulta la desigualdad triangular de la aplicación  $d$ , que es la única dificultad por comprobar que es una distancia.

Veamos ahora que metriza la convergencia que hemos definido en  $\mathcal{P}[a, b]$ : Supongamos primer que

$$\forall \varepsilon > 0, P\{\|X^n - X\|_{L^2([a,b])} > \varepsilon\} \xrightarrow{n \rightarrow \infty} 0.$$

Entonces,

$$\begin{aligned} & \mathbb{E} \left[ \frac{\|X^n - X\|_{L^2([a,b])}}{1 + \|X^n - X\|_{L^2([a,b])}} \right] \\ &= \int_{\{\|X^n - X\|_{L^2([a,b])} \geq \varepsilon\}} \frac{\|X^n - X\|_{L^2([a,b])}}{1 + \|X^n - X\|_{L^2([a,b])}} dP + \int_{\{\|X^n - X\|_{L^2([a,b])} < \varepsilon\}} \frac{\|X^n - X\|_{L^2([a,b])}}{1 + \|X^n - X\|_{L^2([a,b])}} dP \\ &\leq P\{\|X^n - X\|_{L^2([a,b])} \geq \varepsilon\} + \varepsilon P\{\|X^n - X\|_{L^2([a,b])} < \varepsilon\} \\ &\leq P\{\|X^n - X\|_{L^2([a,b])} \geq \varepsilon\} + \varepsilon \xrightarrow{n \rightarrow \infty} \varepsilon. \end{aligned}$$

Siendo esto válido para todo  $\varepsilon > 0$ , resulta que  $d(X^n, X) \xrightarrow{n \rightarrow \infty} 0$ , como queríamos ver.

Recíprocamente, suponemos que  $d(X^n, X) \xrightarrow{n \rightarrow \infty} 0$ . Fijemos  $\varepsilon > 0$ .

$$\begin{aligned} P\{\|X^n - X\|_{L^2([a,b])} \geq \varepsilon\} &= P\left\{ \frac{\|X^n - X\|_{L^2([a,b])}}{1 + \|X^n - X\|_{L^2([a,b])}} \geq \frac{\varepsilon}{1 + \varepsilon} \right\} \\ &\leq \frac{1 + \varepsilon}{\varepsilon} \mathbb{E} \left[ \frac{\|X^n - X\|_{L^2([a,b])}}{1 + \|X^n - X\|_{L^2([a,b])}} \right] \xrightarrow{n \rightarrow \infty} 0. \quad \square \end{aligned}$$

### 6.5 Notación

Denotaremos por  $L_a^2([a, b] \times \Omega)$  el conjunto de procesos de  $L^2([a, b] \times \Omega)$  adaptados a la filtración no-anticipativa  $\{\mathcal{F}_t, t \in \mathbb{R}^+\}$ :

$$L_a^2([a, b] \times \Omega) := \left\{ X : X \text{ es no-anticipativo y } \mathbb{E} \left[ \int_a^b X_t^2 dt \right] < \infty \right\}.$$

Está claro que el conjunto  $L_a^2([a, b] \times \Omega)$  es un subespacio de el espacio de Hilbert  $L^2([a, b] \times \Omega)$  y que está contenido en  $\mathcal{P}[a, b]$ .  $\square$

### 6.6 Definición

Un proceso estocástico no-anticipativo  $X$  diremos que es **simple** (o **escalonado**) en  $[a, b]$  si existe una partición  $\{a = t_0 < t_1 < \dots < t_n = b\}$  tal que

$$X_t(\omega) = X_{t_k}(\omega), \quad \forall \omega \in \Omega, \quad \forall t \in [t_k, t_{k+1}[.$$

Equivalentemente, si  $X$  se puede escribir en  $[a, b]$  como

$$X_t(\omega) = \sum_{k=0}^{n-1} c_k(\omega) \mathbf{1}_{[t_k, t_{k+1}[}(t),$$

con  $c_k$  variables aleatorias  $\mathcal{F}_{t_k}$ -medibles.

Denotaremos por  $\mathcal{S}[a, b]$  el conjunto de procesos simples, que es claramente un subespacio vectorial de  $\mathcal{P}[a, b]$ .  $\square$

### 6.7 Definición

Sea  $X \in \mathcal{S}[a, b]$ .

Definimos la integral de  $X$  respecto el proceso de Wiener  $W$  como la variable aleatoria sobre  $(\Omega, \mathcal{F}, P)$

$$\left[ \int_a^b X_t dW_t \right](\omega) := \sum_{k=0}^{n-1} c_k(\omega) [W_{t_{k+1}} - W_{t_k}](\omega). \quad \square$$

### 6.8 Proposición

*La aplicación entre espacios vectoriales*

$$\begin{array}{ccc} \mathcal{S}[a, b] \subset \mathcal{P}[a, b] & \longrightarrow & L^0(\Omega) \\ X & \longmapsto & \int_a^b X_t dW_t \end{array}$$

*está bien definida y es lineal.*

### 6.9 Proposición

Sea  $X \in \mathcal{S}[a, b]$ . Entonces:

1) Si  $[a', b'] \subset [a, b]$ , entonces  $X \in \mathcal{S}[a', b']$ ,  $X \cdot \mathbf{1}_{[a', b']} \in \mathcal{S}[a, b]$ , y

$$\int_{a'}^{b'} X_t dW_t = \int_a^b X_t \cdot \mathbf{1}_{[a', b']}(t) dW_t.$$

2) Si  $c \in ]a, b[$ , entonces  $X \in \mathcal{S}[a, c]$ ,  $X \in \mathcal{S}[c, b]$ , y

$$\int_a^b X_t dW_t = \int_a^c X_t dW_t + \int_c^b X_t dW_t.$$

3)  $\left[ \mathbb{E}[X_t] < \infty, \forall t \in [a, b] \right] \implies \mathbb{E} \left[ \int_a^b X_t dW_t \right] = 0.$

4)  $\left[ \mathbb{E}[X_t^2] < \infty, \forall t \in [a, b] \right] \implies \mathbb{E} \left[ \left| \int_a^b X_t dW_t \right|^2 \right] = \int_a^b \mathbb{E}[X_t^2] dt.$



### 6.10 Proposición

Sea  $X \in \mathcal{S}[a, b]$ .

Entonces,  $\forall \varepsilon > 0, \forall N > 0$ ,

$$P\left\{\left|\int_a^b X_t dW_t\right| > N\right\} \leq P\left\{\int_a^b X_t^2 dW_t > \varepsilon\right\} + \frac{\varepsilon}{N^2}.$$

### 6.11 Proposición

$\mathcal{S}[a, b]$  es denso en  $\mathcal{P}[a, b]$ .

*Demostración:* Se demuestra siguiendo estos tres pasos:

- 1)  $\mathcal{S}$  es denso en los procesos continuos.
- 2) Los procesos continuos son densos en los acotados.
- 3) Los procesos acotados son densos en  $\mathcal{P}[a, b]$ .  $\square$

### 6.12 Proposición

La aplicación entre espacios métricos

$$\begin{array}{ccc} \mathcal{S}[a, b] \subset \mathcal{P}[a, b] & \longrightarrow & L^0(\Omega) \\ X & \longmapsto & \int_a^b X_t dW_t \end{array}$$

es continua y transforma sucesiones de Cauchy en sucesiones de Cauchy.

### 6.13 Lema (Teorema de extensión de funciones continuas.)

Sean  $E$  un espacio métrico y  $F$  un espacio métrico completo.

Sean  $A \subset \bar{A} = E$ , y  $f: A \rightarrow F$  una función continua.

Entonces: Existe una extensión continua de  $f$  a  $E$  si y sólo si  $f$  transforma sucesiones de Cauchy en sucesiones de Cauchy.

En caso afirmativo, la extensión es única y viene dada por

$$\begin{array}{ccc} \bar{f}: E & \longrightarrow & F \\ x & \longmapsto & \bar{f}(x) := \lim_{n \rightarrow \infty} f(x_n) \end{array}$$

donde  $\{x_n\}_n \subset A$  es una sucesión cualquiera que converge a  $x$ .

### 6.14 Definición

Sea  $X \in \mathcal{P}[a, b]$ .

Sea  $\{X^n\}_n \subset \mathcal{S}[a, b]$  una sucesión cualquiera que converja a  $X$  en el espacio métrico  $\mathcal{P}[a, b]$ .

Entonces definimos la integral estocástica de  $X$  respecto el proceso de Wiener  $W$  como la variable aleatoria sobre  $(\Omega, \mathcal{F}, P)$

$$\int_a^b X_t dW_t := P\text{-}\lim_{n \rightarrow \infty} \int_a^b X_t^n dW_t.$$

### 6.15 Proposición

La aplicación entre espacios vectoriales

$$\begin{array}{ccc} \mathcal{P}[a, b] & \longrightarrow & L^0(\Omega) \\ X & \longmapsto & \int_a^b X_t dW_t \end{array}$$

obtenida en la Definición 6.14 es lineal.

### 6.16 Proposición

Sea  $X \in \mathcal{P}[a, b]$ .

Entonces,  $\forall \varepsilon > 0, \forall N > 0$ ,

$$P\left\{\left|\int_a^b X_t dW_t\right| > N\right\} \leq P\left\{\int_a^b X_t^2 dW_t > \varepsilon\right\} + \frac{\varepsilon}{N^2}.$$

### 6.17 Proposición

Sea  $X \in L_a^2([a, b] \times \Omega)$ .

Entonces, existe  $\{X^n\}_n \in \mathcal{S}[a, b] \cap L_a^2([a, b] \times \Omega)$  tal que

$$1) \|X_t^n - X_t\|_{L^2([a, b] \times \Omega)} \xrightarrow{n \rightarrow \infty} 0.$$

$$2) \left\| \int_a^b X_t^n dW_t - \int_a^b X_t dW_t \right\|_{L^2(\Omega)} \xrightarrow{n \rightarrow \infty} 0.$$

### 6.18 Observación

La Proposición 6.17 sugiere que si nos restringimos a procesos de  $L_a^2$ , podemos definir la integral estocástica como un operador continuo entre espacios de Hilbert

$$L_a^2([a, b] \times \Omega) \longrightarrow L^2(\Omega).$$

Efectivamente, podemos repetir todo el que hemos hecho en esta situación más agradable, pero menos general.

### 6.19 Proposición

Sea  $X \in \mathcal{P}[a, b]$ . Entonces:

1) Si  $[a', b'] \subset [a, b]$ , entonces  $X \in \mathcal{S}[a', b']$ ,  $X \cdot \mathbf{1}_{[a', b']} \in \mathcal{S}[a, b]$ , y

$$\int_{a'}^{b'} X_t dW_t = \int_a^b X_t \cdot \mathbf{1}_{[a', b']}(t) dW_t.$$

2) Si  $c \in ]a, b[$ , entonces  $X \in \mathcal{S}[a, c]$ ,  $X \in \mathcal{S}[c, b]$ , y

$$\int_a^b X_t dW_t = \int_a^c X_t dW_t + \int_c^b X_t dW_t.$$

Si, además,  $X \in L_a^2([a, b] \times \Omega)$ , entonces

$$3) \mathbb{E} \left[ \int_a^b X_t dW_t \right] = 0.$$

$$4) \mathbb{E} \left[ \left| \int_a^b X_t dW_t \right|^2 \right] = \int_a^b \mathbb{E}[X_t^2] dt.$$

### 6.20 Proposición (Aproximación por sumas de Riemann.)

Sea  $X \in \mathcal{P}[a, b]$  un proceso continuo c.s.

Entonces, para cualquier sucesión de particiones  $\pi_n = \{a = t_0^{(n)} < t_1^{(n)} < \dots < t_{m_n}^{(n)} = b\}$  de  $[a, b]$  tal que  $\|\pi_n\| \rightarrow 0$ , se cumple

$$\sum_{k=0}^{m_n-1} X_{t_k^{(n)}} [W_{t_{k+1}^{(n)}} - W_{t_k^{(n)}}] \xrightarrow[n \rightarrow \infty]{P} \int_a^b X_t dW_t.$$

### 6.21 Observación

Si, en la situación de la Proposición 6.20, consideramos

$$S_n = \sum_{k=0}^{m_n-1} W_{u_k^{(n)}} [W_{t_{k+1}^{(n)}} - W_{t_k^{(n)}}] ,$$

donde cada  $u_k^{(n)}$  es un punto arbitrario de  $[t_k^{(n)}, t_{k+1}^{(n)}]$ , entonces la existencia y el valor concreto de  $P\text{-}\lim_{n \rightarrow \infty} S_n$  depende de la elección de los  $u_k^{(n)}$ .

Por ejemplo, si ponemos

$$u_k^{(n)} = \alpha t_{k+1}^{(n)} + (1 - \alpha)t_k^{(n)} , \quad \text{para un cierto } \alpha \in [0, 1], \forall n, \forall k,$$

se obtiene

$$S_n \xrightarrow[n \rightarrow \infty]{L^2} \frac{1}{2}(W_b^2 - W_a^2) + \left(\alpha - \frac{1}{2}\right)(b - a) .$$

En particular, usando la Proposición 6.20, esto nos dice que

$$\int_a^b W_t dW_t = \frac{1}{2}(W_b^2 - W_a^2) - \frac{1}{2}(b - a) ,$$

y concluimos que la integral de Itô no sigue las reglas del cálculo newtoniano.

Se puede definir una integral estocástica que respete las reglas del cálculo newtoniano? El ejemplo que hemos visto sugiere que podríamos hacerlo, al menos para procesos continuos, definiendo

$$\int_a^b X_t \circ dW_t := P\text{-}\lim_{n \rightarrow \infty} \sum_{k=0}^{m_n-1} X_{\frac{1}{2}(t_k^{(n)} + t_{k+1}^{(n)})} [W_{t_{k+1}^{(n)}} - W_{t_k^{(n)}}] .$$

Esto se puede hacer, en efecto, y obtendríamos, para  $X_t = W_t$

$$\int_a^b W_t \circ dW_t = \frac{1}{2}(W_b^2 - W_a^2) .$$

Pero esta integral (llamada **integral de Stratonovich**) tiene otros inconvenientes.



## 7. Cálculo estocástico

En el cálculo diferencial e integral clásico, los teoremas fuertes como por ejemplo los de Cambio de Variable, Integración por Partes, Teorema Fundamental del Cálculo, etc. nos permiten calcular integrales sin recurrir a los complicados pasos al límite que se desprenden de las definiciones. Es razonable esperar que resultados análogos puedan ser demostrados para el cálculo asociado a las integrales estocásticas, al que llamamos *Cálculo Estocástico*. Esto es así, efectivamente, pero las fórmulas que resulten no son formalmente iguales a las correspondientes del cálculo clásico. Empezaremos enunciando la fórmula de Cambio de Variable, de la cual pueden deducirse las demás.

Denotemos  $\mathcal{P}^2 := \cap_t \mathcal{P}[0, t]$ , con  $t \in \mathbb{R}^+$  o  $t \in [0, T]$ .

### 7.1 Teorema (Cambio de Variable).

Sea  $g \in \mathcal{P}^2$ .

Sea  $X_t$  el proceso definido por  $X_t := \int_0^t g_s dW_s$ .

Sea  $\varphi: \mathbb{R} \rightarrow \mathbb{R}$  una función dos veces derivable con continuidad.

Entonces:

$$\varphi(X_t) = \varphi(X_0) + \int_0^t \varphi'(X_s) \cdot g_s dW_s + \frac{1}{2} \int_0^t \varphi''(X_s) \cdot g_s^2 ds . \quad \square \quad (7.1)$$

Si comparamos esta fórmula con la análoga del cálculo integral habitual observamos que aparece un término adicional.

### 7.2 Ejemplo

Sean  $X_t = W_t$  y  $\varphi(x) = x^n$ . Obtendremos:

$$\begin{aligned} W_t^n &= \int_0^t nW_s^{n-1} dW_s + \frac{1}{2} \int_0^t n(n-1)W_s^{n-2} ds \Rightarrow \\ \int_0^t W_s^{n-1} dW_s &= \frac{W_t^n}{n} - \frac{n-1}{2} \int_0^t W_s^{n-2} ds . \\ \text{En particular: } \int_0^t W_s dW_s &= \frac{W_t^2}{2} - \frac{t}{2} , \end{aligned}$$

resultado diferente al que se obtiene con el cálculo diferencial tradicional.

Una consecuencia inmediata del Teorema 7.1 es que el conjunto de procesos que son integrales estocásticas no es cerrado por transformaciones tan regulares como  $x \mapsto x^2$  del ejemplo anterior.

Una  $dW_s$ -integral se convierte en un proceso suma de una  $dW_s$ -integral y una  $ds$ -integral. (Una  $ds$ -integral es claramente de variación acotada mientras que la  $dW_s$ -integral tiene variación cuadrática no nula; de ahí que ninguna de ellas se puede convertir en la otra).

La razón intuitiva por la que aparece el término adicional es la siguiente: Usando la fórmula de Taylor para  $\varphi$  en el punto  $W_t$ , evaluada en  $W_{t+\Delta t}$ , y dividiendo por  $\Delta t$ , se tiene:

$$\frac{\varphi(W_{t+\Delta t}) - \varphi(W_t)}{\Delta t} = \varphi'(W_t) \cdot \frac{W_{t+\Delta t} - W_t}{\Delta t} + \frac{1}{2} \varphi''(W_t) \cdot \frac{(W_{t+\Delta t} - W_t)^2}{\Delta t} + \dots$$

Si  $W_t$  fuese una función derivable, el término cuadrático y todos los de orden superior convergerían claramente a cero cuando  $\Delta t \rightarrow 0$  y obtendríamos la conocida regla de la cadena para la composición  $\varphi \circ W$ . En el nuestro caso, definiendo  $\frac{dW_t}{dt} = \lim_{\Delta t \rightarrow 0} \frac{W_{t+\Delta t} - W_t}{\Delta t}$  (esto es precisamente lo que estamos haciendo cuando ponemos  $W_t - W_s := \int_s^t dW_r$ ), y observando que  $(W_{t+\Delta t} - W_t)^2$  es del orden de  $\Delta t$  cuando  $\Delta t \rightarrow 0$  (variación cuadrática de  $W$ ), obtenemos justamente (7.1) para el caso  $X_t = W_t$ . Lo mismo sucede cuando  $X_t$  es un proceso integral estocástica.

Denotemos por  $\mathcal{P}^1$  el conjunto de procesos no-anticipativos con trayectorias integrables en el intervalo que se esté considerando.

### 7.3 Definición

Sean  $g \in \mathcal{P}^2$  y  $f \in \mathcal{P}^1$ .

Sea  $X_0$  una variable aleatoria  $\mathcal{F}_0$ -medible.

Un proceso  $X$  de la forma

$$X_t = X_0 + \int_0^t f_s ds + \int_0^t g_s dW_s$$

se llama **proceso de Itô**.  $\square$

Es inmediato de la definición que un proceso de Itô es siempre continuo y adaptado. La clase de los procesos de Itô es cerrada por Cambio de Variable. Formularemos una versión más general del Teorema 7.1, dejando que  $\varphi$  actúe sobre  $n$  procesos de Itô al mismo tiempo (en lugar de una sola integral estocástica) y que dependa también de  $t$ . Esta versión se conoce como a **Fórmula de Itô** y es la herramienta básica del Cálculo Estocástico.

### 7.4 Teorema (Fórmula de Itô).

Sean  $\{X^{(i)}\}_{i=1}^n$ ,  $n$  procesos de Itô con representación

$$X_t^{(i)} = X_0^{(i)} + \int_0^t f_s^{(i)} ds + \int_0^t g_s^{(i)} dW_s, \quad i = 1, \dots, n.$$

Denotemos  $\vec{X}_t = (X_t^{(1)}, \dots, X_t^{(n)})$ .

Sea

$$\begin{aligned} \varphi: \mathbb{R} \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (t, \vec{x}) &\longmapsto \varphi(t, \vec{x}) \end{aligned}$$

una función tal que  $\frac{\partial \varphi}{\partial t}$ ,  $\frac{\partial \varphi}{\partial x_i}$ ,  $\frac{\partial^2 \varphi}{\partial x_i \partial x_j}$  existen y son continuas.

Entonces,  $Y_t := \varphi(t, \vec{X}_t)$  es de nuevo un proceso de Itô, y se puede escribir como

$$\begin{aligned} Y_t = Y_0 + \int_0^t &\left[ \frac{\partial \varphi}{\partial s}(s, \vec{X}_s) + \frac{\partial \varphi}{\partial x_i}(s, \vec{X}_s) f_s^{(i)} + \frac{1}{2} \frac{\partial^2 \varphi}{\partial x_i \partial x_j}(s, \vec{X}_s) g_s^{(i)} g_s^{(j)} \right] ds \\ &+ \int_0^t \frac{\partial \varphi}{\partial x_i}(s, \vec{X}_s) g_s^{(i)} dW_s \end{aligned}$$

(sumando de 1 a  $n$  sobre los índices repetidos).  $\square$

## 7.5 Definición

Si  $X_t = X_0 + \int_0^t f_s ds + \int_0^t g_s dW_s$  es un proceso de Itô, decimos que tiene **diferencial estocástica**

$$dX_t = f_t dt + g_t dW_t . \quad \square$$

Por ejemplo, hemos visto que  $dW_t^2 = 2W_t dW_t + dt$ . La notación de diferenciales estocásticas permite una economía en la escritura. La Fórmula de Itô puesta en forma diferencial,

$$dY_t = \left[ \frac{\partial \varphi}{\partial t}(t, \vec{X}_t) + \frac{\partial \varphi}{\partial x_i}(t, \vec{X}_t) f_t^{(i)} + \frac{1}{2} \frac{\partial^2 \varphi}{\partial x_i \partial x_j}(t, \vec{X}_t) \cdot g_t^{(i)} g_t^{(j)} \right] dt + \frac{\partial \varphi}{\partial x_i}(t, \vec{X}_t) g_t^{(i)} dW_t \quad (7.2)$$

sugiere el nombre **Regla de la cadena**, con el que también se la conoce. Si hacemos el convenio  $dW_t \cdot dW_t = dt$ ,  $dW_t \cdot dt = dt \cdot dW_t = 0$ ,  $dt \cdot dt = 0$ , y  $dX_1 dX_2$  es la multiplicación de los respectivos binomios, podemos escribir (7.2) en la forma más compacta

$$dY_t = \frac{\partial \varphi}{\partial t}(t, \vec{X}_t) dt + \frac{\partial \varphi}{\partial x_i}(t, \vec{X}_t) dX_t^{(i)} + \frac{1}{2} \frac{\partial^2 \varphi}{\partial x_i \partial x_j}(t, \vec{X}_t) dX_t^{(i)} dX_t^{(j)} .$$

Aplicando la regla de la cadena a las diferenciales estocásticas

$$\begin{aligned} dX_t^{(1)} &= f_t^{(1)} dt + g_t^{(1)} dW_t \\ dX_t^{(2)} &= f_t^{(2)} dt + g_t^{(2)} dW_t \end{aligned}$$

con  $\varphi(t, x_1, x_2) = x_1 \cdot x_2$  obtenemos la **Fórmula de Integración por Partes**

$$d(X_t^{(1)} X_t^{(2)}) = X_t^{(1)} dX_t^{(2)} + X_t^{(2)} dX_t^{(1)} + dX_t^{(1)} dX_t^{(2)} .$$

Por ejemplo, podemos calcular otra vez  $\int_0^t W_s dW_s$  integrando por partes:

$$\begin{aligned} \int_0^t W_s dW_s &= W_t^2 - \int_0^t W_s dW_s - \int_0^t ds \Rightarrow \\ \int_0^t W_s dW_s &= \frac{W_t^2}{2} - \frac{t}{2} . \end{aligned}$$

Aplicando el Teorema 7.1 a  $X_t = \int_0^t dW_s = W_t$ , obtenemos, para  $\varphi$  de clase  $C^2$ ,

$$\varphi(W_t) = \varphi(W_0) + \int_0^t \varphi'(W_s) dW_s + \frac{1}{2} \int_0^t \varphi''(W_s) ds ,$$

que sería un **Teorema Fundamental del Cálculo**, o bien

$$\int_0^t \varphi'(W_s) dW_s = \varphi(W_t) - \varphi(W_0) - \frac{1}{2} \int_0^t \varphi''(W_s) ds ,$$

en forma de **Regla de Barrow**.

## 7.6 Ejemplo

Veamos ahora como se resuelve una ecuación diferencial estocástica muy sencilla usando el cálculo estocástico. Se trata de la ecuación del modelo de Verhulst sin mortalidad:

$$dX_t = aX_t dt + \sigma X_t dW_t$$

con condición inicial  $X_0$  dada.

Si  $W$  fuera derivable, escribiríamos

$$\frac{dX_t}{X_t} = (a + \sigma \dot{W}_t) dt ,$$

y la solución se encuentra en seguida integrando y tomando exponenciales. Siguiendo la misma idea, calculamos  $d(\log X_t)$  usando la fórmula de Itô:

$$\begin{aligned} d(\log X_t) &= \frac{1}{X_t} dX_t + \frac{1}{2} \cdot \frac{-1}{X_t^2} (dX_t)^2 \\ &= \frac{1}{X_t} (aX_t dt + \sigma X_t dW_t) - \frac{1}{2} \cdot \frac{-1}{X_t^2} \sigma^2 X_t^2 dt \\ &= (a - \sigma^2/2) dt + \sigma dW_t , \end{aligned}$$

de donde

$$\begin{aligned} \log X_t &= \log X_0 + \int_0^t (a - \sigma^2/2) ds + \int_0^t \sigma dW_s \\ &= \log X_0 + (a - \sigma^2/2)t + \sigma W_t , \end{aligned}$$

y finalmente

$$X_t = X_0 \cdot \exp\{(a - \sigma^2/2)t + \sigma W_t\} .$$

### 7.7 Ejemplo

Análogamente al ejemplo anterior, se resuelve la ecuación diferencial

$$\left. \begin{aligned} dX_t &= X_t dW_t \\ X_0 &= 1 \end{aligned} \right\} .$$

$$d(\log X_t) = \frac{1}{X_t} dX_t + \frac{1}{2} \cdot \frac{-1}{X_t^2} (dX_t)^2 dt = dW_t - \frac{1}{2} dt .$$

Integrando y usando  $X_0 = 1$ , obtenemos

$$\log X_t = W_t - \frac{1}{2}t ,$$

y tomando exponenciales

$$X_t = \exp\{W_t - \frac{1}{2}t\} .$$

Por analogía con el cálculo clásico, llamamos a  $X_t$  la **función exponencial estocástica**, que difiere de la usual en el factor  $\exp\{-\frac{1}{2}t\}$ .



## 8. Ecuaciones diferenciales estocásticas. Existencia y unicidad

Hemos desarrollado un concepto de integral y unas reglas de cálculo que pretenden ser bien adaptados al estudio de sistemas con incertidumbre. Para acabar de justificar esta afirmación sólo hay que comprobar si situaciones como las de los ejemplos vistos pueden ser modelizadas con el formalismo de las ecuaciones diferenciales estocásticas del tipo de (4.8) o (5.1). Esto será cierto siempre que tengamos la seguridad que las integrales que hay implícitas en la ecuación diferencial estocástica tienen sentido. La cuestión crítica es sin duda si el integrando de la integral estocástica es o no adaptado a una filtración no-anticipativa del proceso de Wiener.

Una forma bastante general de ecuación diferencial estocástica de primer orden en dimensión 1 es

$$dX_t = f(t, X_t) dt + g(t, X_t) dW_t, \quad t \in \mathbb{R}^+, \quad (8.1)$$

donde  $X_t$  es el proceso incógnita,  $W_t$  es el proceso de Wiener, y  $f(t, x)$  y  $g(t, x)$  son funciones reales definidas y medibles en  $\mathbb{R}^+ \times \mathbb{R}$ , que llamamos **coeficientes de la ecuación**.

Supongamos que combinamos esta ecuación con una condición inicial  $x_0 \in \mathbb{R}$  (es decir, una variable aleatoria constante) y que la evolución de  $X_t$  no tiene otra causa de aleatoriedad que la inducida por la evolución del proceso  $W \equiv \{W_t, t \in \mathbb{R}^+\}$ . En tal caso, no tiene sentido pensar que la solución hasta el instante  $t$  vaya a depender de lo que el azar deparará después del instante  $t$ . Por lo tanto, es razonable buscar soluciones sin esta dependencia, y el formalismo funciona.

Más generalmente, supongamos que la condición inicial es una variable aleatoria  $X_0$  independiente de todas las variables del proceso  $W$ . Entonces esperaremos soluciones que hasta el instante  $t$  dependan de la condición inicial y del proceso  $W$  hasta  $t$ , pero no más allá.

En este capítulo precisaremos el concepto de solución de (8.1) y enunciaremos dos resultados clásicos (de existencia y unicidad, y de dependencia de condiciones iniciales) para este tipo de ecuaciones. No se trata de los mejores enunciados posibles; sólo pretenden ilustrar algunas analogías y diferencias con los resultados homólogos de la teoría de ecuaciones diferenciales ordinarias.

Comencemos con la definición de solución de (8.1), lo que equivale a precisar exactamente qué entendemos por la expresión (8.1).

### 8.1 Definición

Un proceso  $\{X_t, t \in \mathbb{R}^+T\}$  es solución de (8.1) con condición inicial  $X_0 = \xi$  si y sólo si:

- 1)  $\{X_t, t \in \mathbb{R}^+\}$  es adaptado a alguna filtración  $\{\mathcal{F}_t, t \in T\}$  no-anticipativa de  $W$ .
- 2)  $\bar{g}_t := g(t, X_t) \in \mathcal{P}^2$  i  $\bar{f}_t := f(t, X_t) \in \mathcal{P}^1$ .
- 3)  $X_t$  tiene diferencial estocástica

$$dx_t = \bar{f}_t dt + \bar{g}_t dW_t .$$

Equivalentemente,  $X_t$  satisface

$$X_t = \xi + \int_0^t f(s, X_s) ds + \int_0^t g(s, X_s) dW_s, \quad t \in \mathbb{R}^+. \quad \square \quad (8.2)$$

Las condiciones 1) y 2) de la definición garantizan que las expresiones de 3) tienen sentido. En particular, obsérvese que

$$\left. \begin{aligned} dX_t &= f(t, X_t) dt + g(t, X_t) dW_t \\ X_0 &= \xi \end{aligned} \right\} \quad (8.3)$$

sólo podrá tener solución si  $\xi$  es una variable aleatoria independiente de todo el proceso  $W$ .

Para el problema de Cauchy (8.3) con  $g \equiv 0$  (ecuación diferencial ordinaria), el resultado clásico de existencia y unicidad de solución requiere que  $f$  cumpla alguna condición de Lipschitz. Esto permite demostrar mediante iteraciones de Picard la existencia y unicidad. Es natural intentar el mismo tipo de planteo en el caso  $g \neq 0$ .

## 8.2 Teorema

Supongamos que:

- 1) existe  $K > 0$  tal que

$$|f(t, x) - f(t, y)| + |g(t, x) - g(t, y)| \leq K \cdot |x - y|, \quad \forall t \in [0, T], \quad \forall x, y \in \mathbb{R}.$$

- 2) existe  $L > 0$  tal que

$$|f(t, x)|^2 + |g(t, x)|^2 \leq L \cdot (1 + |x|^2), \quad \forall t \in [0, T], \quad \forall x, y \in \mathbb{R}.$$

- 3)  $\xi$  es una variable aleatoria independiente de  $\{W_t, t \in T\}$ .

Entonces:

- a) Existe una solución de (8.3) que es un proceso continuo.  
 b) Si  $X^1$  y  $X^2$  son procesos continuos solución de (8.3), entonces  $\sup_{t \in [0, T]} |X_t^1 - X_t^2| = 0$ , c.s.  
 □

Sólo se asegura la unicidad de solución continua debido a la ambigüedad en la definición del proceso integral estocástica. De hecho, la exigencia de continuidad puede incorporarse a la definición de solución, y entonces se puede afirmar que dos soluciones sólo difieren en un conjunto de trayectorias de probabilidad nula.

La dependencia de la solución respecto de la condición inicial es similar a la que se encuentra en el caso determinista. En particular, se tiene el siguiente resultado de diferenciabilidad de primer orden respecto a una condición inicial constante.

## 8.3 Teorema

Supongamos:

- 1) Las funciones  $f(t, x)$  y  $g(t, x)$  satisfacen las hipótesis 1) y 2) del Teorema 8.1.  
 2) Las derivadas parciales  $\frac{\partial}{\partial x} f(t, x)$  y  $\frac{\partial}{\partial x} g(t, x)$  existen y son continuas y acotadas.

Entonces, la solución  $\{X_t(c), t \in \mathbb{R}^+\}$  del problema de Cauchy

$$\left. \begin{aligned} dx_t &= f(t, x_t) dt + g(t, x_t) dW_t \\ x_0 &= c \end{aligned} \right\},$$

con  $\xi \in \mathbb{R}$ , es derivable en  $L^2(\Omega)$  respecto  $c$  (esto es, existe  $L^2(\Omega)$ - $\lim_{h \rightarrow 0} \frac{X_t(c+h) - X_t(c)}{h}$ ) y el proceso derivada  $\{y_t(c) := \frac{d}{dc}x_t(c), t \in \mathbb{R}^+\}$  satisface la ecuación variacional

$$\left. \begin{aligned} dy_t(c) &= \frac{\partial}{\partial x} f(t, x_t(c)) \cdot y_t(c) dt + \frac{\partial}{\partial x} g(t, x_t(c)) \cdot y_t(c) dW_t \\ y_0(c) &= 1 \end{aligned} \right\}. \quad \square \quad (8.4)$$

Nótese que (8.4) es una ecuación lineal (en el sentido que los coeficientes son lineales respecto el proceso incógnita) en  $Y_t(c)$ , pero con coeficientes aleatorios. Es por tanto, en realidad, una ecuación de primer orden de un tipo más general que (8.1).

I.V. Girsanov encontró, en 1962, un ejemplo muy sencillo de ecuación diferencial estocástica que tiene un comportamiento diferente que su análoga determinista por lo que respecta a la unicidad de soluciones. La mayoría de libros sobre ecuaciones diferenciales estocásticas la mencionan:

La ecuación diferencial ordinaria

$$\left. \begin{aligned} dx_t &= |x_t|^\alpha dt \\ x_0 &= 0 \end{aligned} \right\}, \quad t \in [0, T], \quad (8.5)$$

con  $0 < \alpha < 1$ , tiene una infinidad de soluciones. En efecto, para todo  $s \in [0, T]$ , la función

$$x_t = \begin{cases} 0, & \text{si } 0 \leq t \leq s \\ [(1-\alpha)(t-s)]^{(1-\alpha)^{-1}}, & \text{si } t \geq s \end{cases}$$

es solución de (8.5).

En cambio, la ecuación estocástica

$$\left. \begin{aligned} dX_t &= |X_t|^\alpha dW_t \\ X_0 &= 0 \end{aligned} \right\}, \quad t \in [0, T], \quad (8.6)$$

resulta tener infinitas soluciones si  $0 < \alpha < 1/2$ , pero sólo una única solución (la trivial) para  $\alpha \geq 1/2$ , aunque no cumple las condiciones del Teorema 8.1.

Más detalles sobre los resultados apuntados en este capítulo y en general sobre la teoría de las ecuaciones diferenciales estocásticas se pueden encontrar en Gihhman-Skorohod [14], Arnold [1] o Gard [13].

Uno puede preguntarse si podemos aproximar la solución de una ecuación diferencial estocástica por soluciones de ecuaciones diferenciales con una perturbación aleatoria “más regular”, que permita ser tratada con los medios del cálculo clásico, y obtener la solución de la ecuación estocástica por paso al límite.

Supongamos que  $\{W^{(n)}\}_n$  es una sucesión de procesos estocásticos, con trayectorias de clase  $C^1$ , tal que para todo  $\omega$  (con probabilidad 1), converge al proceso de Wiener

$$W_t^{(n)}(\omega) \longrightarrow W_t(\omega)$$

uniformemente en  $t$  sobre intervalos acotados. Para cada  $\omega$ , sea  $X_t^{(n)}(\omega)$  la solución de la correspondiente ecuación determinista

$$dX_t = f(t, X_t) dt + g(t, X_t) dW_t^{(n)}.$$

Se demuestra que la sucesión de procesos  $\{X^{(n)}\}_n$  converge también uniformemente sobre intervalos acotados a un cierto proceso  $X$ . Este proceso resulta ser la solución de una ecuación diferencial estocástica con los mismos coeficientes, *siempre y cuando la integral estocástica se interprete en el sentido de Stratonovich* (véase el Capítulo 6):

$$X_t = X_0 + \int_0^t f(s, X_s) ds + \int_0^t g(s, X_s) \circ dW_t . \quad (8.7)$$

Esto es un argumento a favor de la interpretación de Stratonovich al formular una ecuación diferencial estocástica. Sin embargo, la característica de “no mirar hacia adelante” de la integral de Itô puede hacerla adecuada en muchos casos. Además, desde el punto de vista probabilístico, la integral de Itô es más fácil de tratar.

Afortunadamente, hay una relación relativamente simple entre las soluciones de una ecuación de Stratonovich y una ecuación de Itô, por lo que el desarrollo de la teoría para uno de los dos tipos produce también muchos resultados para el otro. En efecto, bajo hipótesis apropiadas, se tiene que la solución de la ecuación de Stratonovich (8.7) coincide con la de la ecuación de Itô modificada

$$X_t = X_0 + \int_0^t [f(s, X_s) + \frac{1}{2}g'(s, X_s)g(s, X_s)] ds + \int_0^t g(s, X_s) dW_t ,$$

donde  $g'$  denota la derivada de  $g$  respecto al segundo argumento.

Finalmente, hablaremos de un concepto de solución que no tiene análogo en el caso determinista. En primer lugar, notemos que en el teorema de existencia y unicidad 8.2 no se especifica la filtración no-anticipativa respecto la que existe una solución. En realidad, bajo las hipótesis dadas, se puede demostrar que tal filtración puede tomarse como la generada por la variable condición inicial y el proceso  $W$ . Cuando esto sucede, hablamos de una **solución fuerte** de la ecuación.

En cambio, cuando sólo los datos  $X_0$ ,  $f$ ,  $g$  vienen dados, la pregunta más natural, desde el punto de vista de la modelización, es si existen un espacio de probabilidad, un proceso de Wiener  $W$ , una filtración no-anticipativa, y un proceso  $X$  tales que se cumple (8.2). En tal caso se habla de **solución débil**.

El concepto de unicidad del Teorema 8.2 se llama unicidad fuerte o trayectorial; mientras que la unicidad en el sentido de las soluciones débiles significa que dos soluciones cualesquiera tienen la misma ley.

Evidentemente, toda solución fuerte es una solución débil. El recíproco no es cierto, y de hecho el siguiente ejemplo muestra una ecuación sin solución fuerte, pero con una única solución débil:

$$\left. \begin{aligned} dX_t &= \text{signo}(X_t) dW_t \\ x_0 &= 0 \end{aligned} \right\} ,$$

donde

$$\text{signo}(x) = \begin{cases} +1, & \text{si } x \geq 0 \\ -1, & \text{si } x < 0 \end{cases} .$$

## 9. Referencias

- [1] L. Arnold: *Stochastic Differential Equations: Theory and Applications*, John Wiley & Sons, 1974.
- [5] R. Brown: *A brief Account of Microscopical Observations made in the Months of June, July and August, 1827, on the Particles contained in the Pollen of Plants; and on the general Existence of active Molecules in Organic and Inorganic Bodies*, Philosophical Magazine N.S., **4** (1828).
- [6] R. Brown: *Additional Remarks on Active Molecules*, Philosophical Magazine N.S., **6** (1829).
- [7] S. Chandrasekar: *Stochastic Problems in Physics and Astronomy*, Reviews of Modern Physics, **15** (1943).
- [9] J.L. Doob: *The Brownian Movement and Stochastic Equations*, Annals of Mathematics, **43** (1942).
- [10] A. Einstein: *On the Movement of Small Particles Suspended in a Stationary Liquid Demanded by the Molecular–Kinetic Theory of Heat*, Annalen der Physik, **17** (1905).
- [12] R. Fürth (ed.): *Albert Einstein: Investigations on the Theory of the Brownian Movement*, Dover, 1956.
- [13] T.C. Gard: *Introduction to Stochastic Differential Equations*, Marcel Dekker, 1988.
- [14] I.I. Gihman, A.V. Skorohod: *Stochastic Differential Equations*, Springer–Verlag, 1972.
- [20] P.A. Meyer: *Geometrie Différentielle Stochastique*, Colloque en l’Honneur de Laurent Schwartz, Asterisque **131**, 1985.
- [22] E. Nelson: *Dynamical Theories of Brownian Motion*, Princeton University Press, 1967.
- [24] B. Øksendal: *Stochastic Differential Equations*, fifth edition, Springer–Verlag, 1998.
- [26] G.E. Uhlenbeck, L.S. Ornstein: *On the Theory of Brownian Motion*, Physical Review, **36** (1930).
- [27] J. Yeh: *Stochastic Processes and the Wiener Integral*, Marcel Dekker, 1973.
- [28] M.C. Wang, G.E. Uhlenbeck: *On the Theory of Brownian Motion II*, Reviews of Modern Physics, **17** (1945).
- [29] N. Wax (ed.): *Selected Papers on Noise and Stochastic Processes*, Dover, 1954.