

Una presentació de les equacions diferencials estocàstiques*

Aureli Alabert & Rosario Delgado
Secció de Matemàtiques
Universitat Autònoma de Barcelona

Motivació i exemples

Durant el segle XVII Newton i Leibniz posaren les bases del que ara anomenem “càlcul diferencial”. Des de llavors les equacions diferencials han estat una eina matemàtica fonamental per modelar sistemes físics. Les lleis físiques que governen un sistema determinen les equacions corresponents, que després intentem resoldre, és a dir, de les quals intentem obtenir una expressió de l'estat del sistema a l'instant de temps t com a funció explícita de t .

De vegades el sistema físic que volem modelar és massa complex no tan sols per resoldre efectivament les equacions associades sinó fins i tot per arribar a formular un conjunt d'equacions diferencials que sigui prou representatiu de les seves característiques. Resulta que en moltes d'aquestes situacions “desesperades” és possible usar certa *informació parcial* a l'abast sobre les forces que hi interactuen i a canvi obtenir determinats *resultats parcials* del problema. Concretament, veurem com l'ús d'hipòtesis estadístiques en la formulació ens permet arribar a conclusions probabilístiques com a resultat. Això vol dir: si l'estat d'un sistema a l'instant t està representat per un vector n -dimensional $x(t)$, renunciem a assignar un valor a aquest vector, però fem ús del nostre coneixement del comportament estadístic dels elements que afecten $x(t)$ per tal de contestar preguntes com ara: “Quina és la probabilitat que $x(t)$ pertanyi a una certa regió A de $\mathbf{R}^n$?”, i altres de semblants.

*Aquest article és una versió escrita de la xerrada que, amb el mateix títol, realitzà el primer autor al *Seminari d'Equacions Diferencials, Càlcul Numèric i Aplicacions a l'Enginyeria*, a la Universitat Politècnica de Catalunya, el 31 de març de 1992. Agraïm al prof. Joan de Solà-Morales haver-nos donat aquesta oportunitat.

Exemple 1. Considereu una petita partícula (de diàmetre, posem, 1 micra) en suspensió dins un fluid. A través d'un microscopi hom pot observar que la partícula es mou de manera molt ràpida i irregular en totes direccions, i que aquest moviment no s'atura mai. Quan el fenomen va ser descrit per primera vegada al principi del segle passat, diverses teories (equivocades) sorgiren per explicar-ne les causes. Avui dia, amb l'existència dels àtoms fermament establerta, sabem que aquests moviments "caòtics" són deguts al bombardeig intensiu de la partícula per part de les molècules del fluid que l'envolta. Sota condicions normals de pressió i temperatura, la partícula suporta de l'ordre de 10^{21} col·lisions moleculars per segon, cadascuna amb la seva pròpia direcció i energia. Malgrat que cada col·lisió individual té un efecte inapreciable (les molècules són encara molt més petites que la partícula), la superposició d'un tal nombre de xocs sí que produeix un efecte observable.

És obvi que el càlcul diferencial clàssic pot ajudar ben poc a l'estudi d'aquesta situació. Què es pot fer, doncs, per resoldre (almenys parcialment) el problema?

Per simplicitat, suposem que cap altre camp de forces, com podria ser la gravetat, interactua amb el sistema que acabem de descriure, i normalitzem a 1 la massa de la partícula. Només hi ha aleshores dos orígens de forces a considerar: d'una banda, la fricció dinàmica sistemàtica soferta per la partícula en les seves trasllacions i, de l'altra, les col·lisions moleculars irregulars, negligibles individualment però amb efecte macroscòpic apreciable.

Si oblidem per un moment aquesta última font de moviment, i notem per $v(t)$ la velocitat a l'instant t , la llei de Newton ens porta immediatament a l'equació diferencial de primer ordre

$$\frac{dv(t)}{dt} = -\beta \cdot v(t), \quad v(0) = v_0, \quad (1)$$

on la constant β es determina per a partícules esfèriques mitjançant la primera llei de Stokes: $\beta = 6\pi r\eta$, sent r el radi de la partícula i η el coeficient dinàmic de viscositat del fluid (per a partícules no esfèriques cal canviar el factor 6π per un altre d'adequat).

Considerem la versió integral de l'equació (1):

$$v(t + \Delta t) - v(t) = -\beta \int_t^{t+\Delta t} v(s) ds, \quad v(0) = v_0. \quad (2)$$

Notem per $M(t)$ l'increment de moment net guanyat per la partícula degut a la segona font de forces esmentada durant l'interval temporal

$[0, t]$. Tenint en compte aquesta funció, (2) canvia clarament a

$$v(t + \Delta t) - v(t) = -\beta \int_t^{t+\Delta t} v(s) ds + M(t + \Delta t) - M(t), \quad v(0) = v_0. \quad (3)$$

$M(t)$ és una funció desconeguda, però si acceptem certes hipòtesis físiques sobre les forces que representa, aquestes es traduiran en propietats matemàtiques (probabilístiques) de M . Concretament, suposem que:

1. Les condicions de l'entorn (tals com pressió, temperatura, ...) són constants en el temps.
2. El bombardeig sofert per la partícula durant un interval de temps donat és independent del sofert en intervals de temps anteriors això és, no hi té relació física.
3. L'acceleració total fins a l'instant t canvia contínuament respecte de t .

Veurem més tard com es pot construir un model matemàtic que representi aquestes condicions. Deixant de banda la qüestió de si 1), 2) i 3) són prou adequades al sistema que hem descrit (això pot ser tema de molta discussió, especialment 2)), assenyalem només que el model que se'n deduirà produeix resultats que s'ajusten força bé a les observacions empíriques. $\square$

Aquest exemple és molt clàssic (vegeu, per exemple, Nelson [22]) i en continuarem la discussió a l'apartat 2, un cop introduïdes les eines matemàtiques necessàries per tractar-lo. Citem dos exemples més de situacions similars, sense especificar detalls.

Exemple 2. El conegut model de Verhulst per a l'evolució del volum d'individus d'una espècie ve donat per l'equació diferencial

$$\frac{dx(t)}{dt} = x(t)(a - bx(t)) \quad (4)$$

(l'equació logística), que es dedueix de suposar una taxa de natalitat constant a i una taxa de mortalitat proporcional a la població $bx(t)$. L'equació (4) es pot pensar com l'aproximació contínua del planteig discret

$$x(n+1) - x(n) = x(n)(a - bx(n)). \quad (5)$$

Típicament, les constants a i b són idealitzacions de quantitats sotmeses a fluctuacions imprevisibles *a priori*. Suposem, per exemple, que canviem a per una funció $a + \xi(n)$, on $\xi(n)$ representa una petita pertorbació de la taxa de natalitat, que és desconeguda i que és independent (i.e., no hi ha relació de causa-efecte) de les pertorbacions $\xi(m)$, $m \neq n$. Si notem per $M(n)$ l'acumulació d'aquestes pertorbacions des de l'instant 0 fins al $n - 1$, o sigui $M(n) = \sum_{k=0}^{n-1} \xi(k)$, obtindrem l'equació modificada

$$x(n+1) - x(n) = x(n)(a + M(n+1) - M(n) - bx(n)). \quad (6)$$

Construïm ara l'aproximació contínua de (6). L'anàleg dels increments $x(n+1) - x(n)$ en mode continu és $\frac{x(t+\Delta t) - x(t)}{\Delta t}$, i semblantment per a $M(n+1) - M(n)$. Escrivem, doncs,

$$x(t+\Delta t) - x(t) = x(t)(a - bx(t))\Delta t + x(t)(M(t+\Delta t) - M(t)). \quad (7)$$

Si suposem constant l'entorn en el sentit que les causes de les pertorbacions són invariables en el temps, mantenim la independència del valor de les pertorbacions que afecten el sistema en intervals disjunts, i finalment fem la hipòtesi raonable que la funció $M(t)$ ha de ser contínua en el temps, retrobem exactament les condicions 1), 2) i 3) proposades a l'exemple anterior com a propietats de la funció desconeguda M . En resulta doncs formalment la mateixa situació de l'exemple 1. $\square$

Exemple 3. Considerem ara el problema de portar un satèl·lit artificial a la seva òrbita geoestacionària, a partir de l'òrbita el·líptica de transferència, que és l'estadi intermedi entre la fase de llançament i el posicionament definitiu.

Per controlar aquesta delicada maniobra cal tenir una informació el més acurada possible de la posició del satèl·lit en cada instant. En una primera aproximació, tenint en compte només l'acció del camp gravitatori terrestre, les equacions del moviment es poden escriure en la forma

$$\frac{dx(t)}{dt} = f(x(t)) \quad (8)$$

amb $x(t) \in \mathbf{R}^6$ (tres paràmetres de posició, tres paràmetres de velocitat).

Altres elements que influeixen en la trajectòria del satèl·lit, com ara la no-esfericitat i no-homogeneïtat de la Terra, la influència d'altres cossos celestes, la pressió de radiació solar, etc., poden ser incorporats al model si es té un bon coneixement de com actuen. Tanmateix, hi intervindran

també pertorbacions la direcció i intensitat de les quals no són conegudes *a priori*.

Una heurística similar a la dels exemples anteriors ens porta a escriure

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

on M és una funció amb les característiques 1), 2), 3) llistades a l'exemple 1.

A més, hom compta amb la informació addicional aportada per les estacions de seguiment, que reben un senyal

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

al seu torn contaminat també per pertorbacions en la mesura:

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

sent $N(t)$ desconeguda i independent (en el sentit que les causes ho són) de $M(t)$. El problema, que s'anomena *filtratge*, consisteix a obtenir la millor informació possible sobre el valor de $x(t)$, tenint en compte les observacions $y(t)$. $\square$

Podríem citar molts altres exemples que mostren que els problemes modelables de forma similar a les situacions presentades no es redueixen tan sols als camps de la física, la biologia o l'enginyeria.

El nostre propòsit en les seccions següents és presentar les eines matemàtiques que permeten tractar els problemes d'aquest tipus.

1 Conceptes bàsics de probabilitat i processos estocàstics

D'habitud, hom diu que un experiment és aleatori quan el seu resultat no es pot conèixer *a priori*, malgrat que no sigui el primer cop que es realitza. L'experiment pot estar governat per clares lleis deterministes, però renunciem a efectuar els càlculs, o simplement no sabem com fer-los. Sovint, però, és possible dir quelcom sobre cada possible resultat. Per exemple, en tirar repetidament una moneda equilibrada, és d'esperar que el quocient entre el nombre de creus obtingudes i el nombre de llançaments (l'anomenada *frequència relativa*) s'aproximi a 1/2 a la llarga. Això ens convida a

enunciar: “ $1/2$ és la ‘probabilitat’ d’obtenir creu en cada llançament d’una moneda equilibrada”. La frase és filosòficament molt interessant, però *certament no és una definició matemàtica*. La formalització habitual és com segueix:

Sigui Ω un conjunt i $\mathcal{F} \subset \mathcal{P}(\Omega)$ una σ -àlgebra de parts de Ω . Una **probabilitat** és una mesura positiva i finita sobre $(\Omega, \mathcal{F})$ amb massa total igual a 1. És a dir, una probabilitat sobre $(\Omega, \mathcal{F})$ és una aplicació $p: \mathcal{F} \rightarrow [0, 1]$ tal que $p(\emptyset) = 0$, $p(\Omega) = 1$, i per a cada família numerable $\{A_i\}_{i \in I}$ de conjunts de $\mathcal{F}$ tal que $A_i \cap A_j = \emptyset$, $\forall i, j \in I$, es té $p(\cup_{i \in I} A_i) = \sum_{i \in I} p(A_i)$. L’espai de mesura $(\Omega, \mathcal{F}, p)$ rep el nom d’**espai de probabilitat**. En aquest context, cada conjunt mesurable $A \in \mathcal{F}$ s’anomena **esdeveniment**. El nombre $p(A)$ és la **probabilitat de l’esdeveniment** A . En modelar un experiment real, cada $\omega \in \Omega$ representa un possible resultat i $p(A)$ és la “probabilitat” que el resultat ω pertanyi a A .

Sigui $(E, \mathcal{E})$ un altre espai mesurable. Una **variable aleatòria E -valuada sobre $(\Omega, \mathcal{F}, p)$** és una aplicació mesurable $X: (\Omega, \mathcal{F}, p) \rightarrow (E, \mathcal{E})$. Quan $(E, \mathcal{E})$ és el conjunt dels nombres reals equipat amb la seva σ -àlgebra de Borel natural, i que notarem $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$, diem simplement que tenim una **variable aleatòria (real)**. Les variables aleatòries representen el mecanisme mitjançant el qual s’observa l’experiment. Si observem un conjunt $B \in \mathcal{E}$, deduïm que el resultat de l’experiment pertany a $X^{-1}(B) \in \mathcal{F}$.

Donada una variable aleatòria real sobre $(\Omega, \mathcal{F}, p)$, podem definir una nova probabilitat μ en $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ com $\mu(B) := p(X^{-1}(B))$, $\forall B \in \mathcal{B}(\mathbf{R})$. Aquesta nova probabilitat μ associada a p i X s’anomena la **lleï de X sota p** . Si μ és una probabilitat sobre $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$, la funció $F: \mathbf{R} \rightarrow \mathbf{R}$ definida per $F(x) := \mu([-\infty, x])$ rep el nom de **funció de distribució** de μ . F és una funció no-decreixent i contínua per la dreta verificant $\lim_{x \rightarrow -\infty} F(x) = 0$ i $\lim_{x \rightarrow +\infty} F(x) = 1$. Recíprocament, tota funció amb aquestes propietats determina unívocament una probabilitat μ . Si F és expressable com $F(x) = \int_{-\infty}^x f(z) dz$, per alguna funció $f: \mathbf{R} \rightarrow \mathbf{R}^+$, aleshores f és la **funció de densitat** de μ respecte de la mesura de Lebesgue. En tal cas, tindrem que $\forall B \in \mathcal{B}(\mathbf{R})$, $\mu(B) = \int_B f(z) dz$. Aquests conceptes s’estenen fàcilment a variables aleatòries $\mathbf{R}^n$ -valuades. Sovint aquestes variables són anomenades **vectors aleatoris**, i la seva lleï de probabilitat la **lleï conjunta** de les components del vector.

Dos esdeveniments $A_1, A_2 \in \mathcal{F}$ s’anomenen **independents** si $p(A_1 \cap A_2) = p(A_1) \cdot p(A_2)$. Dues variables aleatòries reals X, Y definides sobre el mateix espai de probabilitat són **independents** si $\forall B_1, B_2 \in \mathcal{B}(\mathbf{R})$,

els esdeveniments $X^{-1}(B_1)$ i $Y^{-1}(B_2)$ són esdeveniments independents. Intuïtivament, això vol dir que l'observació d'un esdeveniment A a través de la variable aleatòria X no canvia la llei de probabilitat de Y . La definició és anàloga per a variables valuades en qualsevol espai i per a famílies de variables aleatòries.

Donada una variable aleatòria X , definim el q -èsim moment de X com $\int_{\Omega} X(\omega)^q p(dx)$, sempre que la integral existeixi i sigui finita. Arguments senzills de teoria de la mesura mostren que si μ és la llei de X sota p i F és la funció de distribució associada, el q -èsim moment de X es pot calcular com la integral de Stieljes $\int_{\mathbf{R}} x^q dF(x)$. Si, a més, F té una funció de densitat f , l'última integral coincideix amb la integral ordinària $\int_{\mathbf{R}} x^q \cdot f(x) dx$. El primer moment de X s'anomena **esperança o mitjana** de X . El notem per $E[X]$. Si per a dues variables aleatòries X, Y es té la igualtat $p(\{\omega : X(\omega) = Y(\omega)\}) = 1$, diem que $X = Y$ **quasi segur** (q. s.). Aquesta no és més que la coneguda equivalència *quasi per tot* de la teoria de la integració de Lebesgue. Mitjançant aquesta relació, hom defineix els habituals espais de Lebesgue $L^q(\Omega, \mathcal{F}, p)$ de (classes de) variables aleatòries que posseeixen q -èsim moment, $0 < q < \infty$. $L^\infty(\Omega, \mathcal{F}, p)$ serà l'espai de les variables aleatòries acotades, mentre que $L^0(\Omega, \mathcal{F}, p)$ representa l'espai vectorial de totes les variables aleatòries, sense cap condició d'integrabilitat. Recordem que en qualsevol espai de mesura finita es té $0 \leq r \leq q \leq \infty \Rightarrow L^q \subset L^r$.

Sigui T un conjunt ordenat. Un **procés estocàstic** indexat per T és una família de variables aleatòries (reals, si no s'especifica altra cosa) $\{X_t, t \in T\}$, totes elles definides sobre el mateix espai de probabilitat. Normalment, T és $[0, \infty[$ o $\mathbf{N} = \{0, 1, 2, \dots\}$, i s'interpreta com el temps, en mode continu o discret, de forma que un procés estocàstic és l'eina apropiada per modelar sistemes dinàmics quan l'estat del sistema en cada instant t està governat per una llei de probabilitat, la llei de X_t . Suposarem a partir d'ara $T = [0, \infty[$.

Per a cada ω fixat, la funció

$$\begin{array}{ccc} X(\omega) : T & \longrightarrow & \mathbf{R} \\ t & \longrightarrow & X_t(\omega) \end{array}$$

s'anomena una **trajectòria** del procés. Aquest punt de vista ens permet pensar en un procés estocàstic com una variable aleatòria valuada en algun

espai de funcions

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

Cada possible resultat aleatori w determina un element de l'espai funcional, que pot ser l'espai de totes les funcions $\mathbf{R}^T$ o un de més petit. Hom pot definir la llei d'un procés seguint la mateixa pauta que per a les variables aleatòries reals o $\mathbf{R}^n$ -valuades. Es pot demostrar, d'altra banda, que la llei del procés està completament determinada per les **lleis en dimensió finita**, p.e. les lleis de les projeccions de rang finit

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

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

Hom caracteritza diverses classes particulars de processos segons propietats de llurs trajectòries o bé propietats relacionades amb la llei del procés. Entre les primeres, citarem els següents:

Diem que un procés estocàstic és **continu** si, per a cada $\omega \in \Omega$, la trajectòria $X(\omega)$ és una funció contínua. Un procés és **càdlàg** si les seves trajectòries són contínues per la dreta amb límits per l'esquerra en cada punt (càdlàg és un acrònim del francès “continu à droite, avec des limites à gauche”). Similarment es defineix la propietat de ser **càglàd**.

Definicions relacionades amb la llei són: un procés és **estacionari** (o **temporalment homogeni**) si les lleis en dimensió finita són invariants per trasllacions temporals (p.e., la llei de $(X_{t_1}, \dots, X_{t_n})$ coincideix amb la de $(X_{t_1+s}, \dots, X_{t_n+s})$ per a tot $s > 0$). Un procés té **increments estacionaris** si la llei de la variable aleatòria $X_t - X_s$ depèn només de la diferència $t - s$, per a tots t i s en T . Diem que un procés estocàstic té **increments independents (del passat)** si per a tots $t, s \in T$, $s < t$, la variable $X_t - X_s$ és independent de la família $\{X_r, r \in T, r \leq s\}$.

Per la seva importància en aplicacions, l'anomenada **llei gaussiana** (o **normal**) mereix especial atenció. Una variable aleatòria real X és gaussiana si la seva llei de probabilitat té funció de densitat

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp \left\{ \frac{-(x-m)^2}{2\sigma^2} \right\} \quad (9)$$

per a certs $m \in \mathbf{R}$ i $\sigma^2 > 0$. Ho escrivim $X \sim \mathcal{N}(m, \sigma^2)$. Es comprova fàcilment que les quantitats $E[X]$ i $\text{Var}[X] := E[(X - E[X])^2]$, anomenades **mitjana** i **variància** de X valen m i σ^2 respectivament. Per tant, la

mitjana i la variància d'una variable aleatòria gaussiana determinen completament la llei. És convenient considerar també el cas degenerat $\sigma^2 = 0$ com a gaussià. Si fem tendir σ^2 a zero en (9) observem que $f(x)$ convergeix a una “funció” δ de Dirac (que no és cap funció, és clar), de forma que per a $\sigma^2 = 0$ no hi ha una vertadera densitat. La llei representada per aquesta “densitat generalitzada” concentra tota la massa en el punt $m \in \mathbf{R}$.

L'anàleg de les variables gaussianes a $\mathbf{R}^n$ són els vectors aleatoris amb densitat

$$f(\vec{x}) = ((2\pi)^n \det D)^{-1/2} \cdot \exp \left\{ -\frac{1}{2} \cdot (D^{-1}(\vec{x} - \vec{m}), \vec{x} - \vec{m}) \right\} \quad (10)$$

on $\vec{m} \in \mathbf{R}^n$ és el **vector de mitjanes** i D és la **matriu de covariàncies**, que ha de ser definida-positiva (o només semidefinida-positiva en els casos degenerats en què la llei es concentra en una varietat afí de $\mathbf{R}^n$). Qualsevol projecció d'un vector aleatori gaussià és de nou gaussià.

La llei gaussiana n -dimensional té un paper central en l'estudi de les probabilitats a $\mathbf{R}^n$. L'anàleg en dimensió infinita als vectors aleatoris gaussianes són els **processos estocàstics gaussianes**. Diem que $\{X_t, t \in T\}$ és gaussià si les lleis en dimensió finita són gaussianes.

Farem ús a la secció que segueix dels teoremes següents, que relacionen alguns dels conceptes introduïts. (Vegeu Yeh [27, pàg. 202 i 255]).

Teorema 1.1. Sigui $X \equiv \{X_t, t \in \mathbf{R}^+\}$ un procés real tal que

(H.1) és continu i amb increments independents del passat.

llavors, existeixen una funció contínua $m: \mathbf{R}^+ \rightarrow \mathbf{R}$ i una funció contínua no-decreixent $v: \mathbf{R}^+ \rightarrow \mathbf{R}$ tals que $\forall s, t \in \mathbf{R}^+, s < t$, la variable aleatòria $X_t - X_s$ és gaussiana amb

$$E[X_t - X_s] = m(t) - m(s), \quad \text{Var}[X_t - X_s] = v(t) - v(s). \quad (11)$$

Si, a més,

(H.2) X té increments estacionaris

llavors m i v satisfan

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

per a certes constants $\mu \in \mathbf{R}$ i $\sigma^2 \in \mathbf{R}^+$.

μ i σ^2 s'anomenen **coeficient de deriva** i **coeficient de difusió**, respectivament. $\square$

El recíproc del teorema és cert en el sentit que, donades les funcions m contínua i v contínua no-decreixent, existeix un procés continu amb increments independents X que verifiquen les igualtats (11).

Teorema 1.2. Sota la hipòtesi (H.1), si X_0 és una variable gaussiana, llavors $\{X_t, t \in \mathbf{R}^+\}$ és un procés gaussià.

En particular, un procés continu amb increments independents tal que $X_0 \equiv x \in \mathbf{R}$, ha de ser forçosament gaussià. $\square$

Definició. Un procés estocàstic real $W \equiv \{W_t, t \in \mathbf{R}^+\}$ és un **procés de Wiener** (o **moviment brownià**) si

- i) és continu,
- ii) té increments independents del passat,
- iii) $\forall t, s, s < t, W_t - W_s$ és una variable aleatòria gaussiana amb esperança zero i variància $\sigma^2 \cdot (t - s)$, per a algun $\sigma^2 \in \mathbf{R}^+$.

Un procés de Wiener **comença en** $x \in \mathbf{R}$ si $W_0 = x$ q. s.

Si, a més, $\sigma^2 = 1$, W s'anomena **procés de Wiener estàndard començant en** $x \in \mathbf{R}$. $\square$

Observeu que un procés de Wiener estàndard és un procés gaussià amb increments estacionaris i, d'altra banda, si $X \equiv \{X_t, t \in \mathbf{R}^+\}$ és un procés continu amb increments estacionaris i independents del passat, amb $X_0 = 0$ q. s. i coeficient de deriva $\mu = 0$, llavors $X = \sigma W$, on W és un procés de Wiener estàndard començant en 0. Aquest fet és la base de la discussió del proper apartat.

És interessant fer aquí una petita digressió històrica. Els moviments erràtics de partícules petites en un fluid presentat a l'exemple 1 foren estudiats per primer cop pel botànic anglès Robert Brown al voltant de 1827 ([5], [6]) (sembla haver-hi també un treball anterior del neerlandès Jan Ingenhousz, cap al 1785). Des d'aleshores, aquest fenomen ha estat conegut amb el nom de *moviment brownià*. La primera descripció matemàtica del moviment brownià fou presentada per A. Einstein [10] el 1905 (vegeu Nelson [22] per a un breu resum de la història del tema), qui proposà el procés de Wiener com a model matemàtic per a les trajectòries de les partícules. Per cert, és força curiós el fet que Einstein no coneixia l'existència del treball de Brown; ell deduí l'existència de moviments

observables de les petites partícules a través de raonaments purament teòrics, com a conseqüència de la teoria atòmica. Avui dia, tant moviment brownià com procés de Wiener són noms igualment utilitzats per designar la classe de processos que acabem de definir.

Més tard, però, el model d'Einstein va ser millorat per Ornstein i Uhlenbeck (vegeu Uhlenbeck-Ornstein [26], Chandrasekar [7], Wang-Uhlenbeck [28], i especialment Doob [9]), els quals usaren el procés de Wiener en el seu model però no directament com a versió matemàtica del moviment brownià. Per evitar confusions, doncs, usarem sempre el nom de procés de Wiener en referir-nos a l'objecte matemàtic.

Hi ha certes propietats de les trajectòries del procés de Wiener que es demostren sense gaire dificultat a partir de la definició i que necessitarem més tard. Clourem l'apartat amb elles.

Teorema 1.3. Les trajectòries $\{W_t(\omega), t \in \mathbf{R}^+\}$ del procés de Wiener són no-derivables en cap punt (q. s.). $\square$

Corol·lari. Les trajectòries del procés de Wiener tenen variació total infinita en tot interval $[a, b]$, $a < b$, (q. s.). $\square$

Aquesta propietat, que per a la partícula browniana implicaria recórrer camins de longitud infinita en temps finit, estimulà l'abandó del model d'Einstein.

El resultat següent dóna una informació més precisa de la magnitud de les oscil·lacions de les trajectòries de W . Intuïtivament, com que $E[|W_t - W_s|^2] = t - s$, la variació $|W_t - W_s|$ serà possiblement de l'ordre de $\sqrt{t - s}$, cosa que, donada una partició $t_0 < \dots < t_n$ d'un interval $[t, t + \Delta t]$, sigui lògic esperar que

$$\sum_{k=1}^n |W_{t_k} - W_{t_{k-1}}|^2 \approx \Delta t,$$

això és, que tinguem una *variació quadràtica* no nula i finita. Recordem el concepte de variació d'ordre k d'una funció.

Sigui $k \geq 0$. Sigui π una partició d'un interval $[a, b]$. La **variació d'ordre k d'una funció f en $[a, b]$ relativa a la partició π** és

$$\pi^{(k)}(f) := \sum_{t_i, t_{i+1} \in \pi} |f(t_{i+1}) - f(t_i)|^k.$$

Es defineix la **variació d'ordre k** de f en $[a, b]$ com el suprem dels valors $\pi^{(k)}(f)$ quan π recorre el conjunt $\mathcal{P}$ de totes les possibles particions

de $[a, b]$:

$$V^{(k)}(f) := \sup_{\pi \in \mathcal{P}} \pi^{(k)}(f) \quad (\in [0, +\infty]).$$

Si f és contínua, llavors per a tota successió refinant $\{\pi_n\}_{n \in \mathbf{N}}$ de particions de $[a, b]$ tal que $\lim_{n \rightarrow \infty} \|\pi_n\| = 0$, es té $V^{(k)}(f) = \lim_{n \rightarrow \infty} \pi_n^{(k)}(f)$ (vegeu Favard [11, pàg. 229]), la qual cosa permet demostrar fàcilment que si per a una funció contínua f i un cert k_0 es verifica $0 < V^{(k_0)}(f) < +\infty$, llavors $V^{(k)}(f) = +\infty$, $\forall k < k_0$, i $V^{(k)}(f) = 0$, $\forall k > k_0$.

El següent resultat precisa que per a les trajectòries del procés de Wiener el valor crític k_0 és precisament 2 i estableix el valor de $V^{(2)}(W(\omega))$. La demostració és més tècnica que la dels teoremes anteriors (vegi's Doob [8, pàg. 395]).

Teorema 1.4. Sigui $W \equiv \{W_t, t \in \mathbf{R}^+\}$ un procés de Wiener.

Donats $s, t \in \mathbf{R}^+$, $s < t$, i $\omega \in \Omega$, la variació quadràtica de $W(\omega)$ en $[s, t]$ és $V^{(2)}(W(\omega)) = t - s$ (q. s.). $\square$

2 Un model matemàtic per a problemes amb informació parcial

Usant els conceptes introduïts a l'apartat anterior, podem ara donar una versió matemàtica precisa de les hipòtesis físiques sobre la funció $M(t)$ de l'exemple 1. Imaginarem, per simplificar, que el problema és unidimensional (la partícula es mou sobre una recta). De fet, serà evident de seguida que el problema real es pot estudiar coordinada per coordinada.

Primer de tot, ens hem referit a $M(t)$ com una “funció desconeguda”. Per tant, es pot pensar de manera natural com un procés estocàstic $\{M_t(\omega), t \in \mathbf{R}^+\}$ definit en algun espai de probabilitat $(\Omega, \mathcal{F}, p)$. El fet que les causes de les forces resumides en M_t són invariants en el temps (condició 1) es modela imposant que la llei de $M_t - M_s$ depengui només de la diferència $t - s$. En altres paraules, el procés $\{M_t, t \in \mathbf{R}^+\}$ ha de tenir *increments estacionaris*. La incorrelació física entre les forces que actuen en un interval donat i les que han actuat prèviament (condició 2) significa que el procés ha de tenir *increments independents del passat*. Finalment, la condició 3 de continuïtat simplement es tradueix en el fet que $\{M_t, t \in \mathbf{R}^+\}$ és un *procés continu*. Podem, a més, imposar arbitràriament $M_0 = 0$ quan comencem a comptar el temps. Segons els teoremes 1.1 i 1.2, en aquesta situació el procés ha de ser gaussià amb $E[M_t] = \mu t$ i $\text{Var}[M_t] = \sigma^2 t$. La isotropia del medi en què es mou la partícula (no hi ha

cap direcció privilegiada per a les forces provinents dels xocs moleculars) la representem posant coeficient de deriva $\mu = 0$.

Ara només cal usar les observacions que segueixen a la definició de procés de Wiener de l'apartat anterior per concloure que $M_t = \sigma W_t$, sent W un procés de Wiener estàndard començant en 0. Sobre el valor concret de σ hi tornarem a l'apartat 5.

Sabem que les trajectòries del procés de Wiener (i, per tant, també les de M_t) són no-derivables en cap punt. Per tant, és clar que a l'equació (3) no podem dividir per Δt i prendre $\lim_{\Delta t \rightarrow 0}$ per arribar a una equació diferencial ordinària, parametritzada per ω . Hem de quedar-nos amb la forma integral

$$v_t = v_0 - \beta \int_0^t v_s ds + \sigma W_t. \quad (12)$$

L'equació (12), malgrat tot, es pot resoldre derivant-la formalment i aplicant tècniques elementals d'equacions diferencials ordinàries. La solució que s'obté,

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

pot expressar-se, després d'integrar per parts, com

$$v_t = \exp\{-\beta t\} \left(v_0 + \sigma W_t \exp\{\beta t\} - \sigma \beta \int_0^t W_s \exp\{\beta s\} ds \right), \quad (13)$$

que torna a tenir sentit rigorós i es pot comprovar directament que satisfà (12). A partir de la velocitat trobem aleshores la funció que ens dóna la posició de la partícula browniana a cada instant:

$$x_t = x_0 + \int_0^t \left[\exp\{-\beta s\} \left(v_0 + \sigma W_s \exp\{\beta s\} - \sigma \beta \int_0^s W_r \exp\{\beta r\} dr \right) \right] ds. \quad (14)$$

Els processos $\{x_t, t \in \mathbf{R}^+\}$ i $\{v_t, t \in \mathbf{R}^+\}$ reben el nom de procés de posició i procés de velocitat d'Ornstein-Uhlenbeck. No obstant això, l'equació (12) fou ja plantejada per Langevin [17] el 1908, i es coneix amb el seu nom. Les fórmules (13) i (14) donen explícitament les trajectòries de x i v en funció de les de W , i a partir d'elles podem calcular la llei d'aquests processos, que és en realitat la informació útil que busquem. x

i v resulten ser processos gaussians amb

$$E[v_t] = v_0 e^{-\beta t}$$

$$\text{Cov}[v_t, v_s] = \frac{\sigma^2}{2\beta} (e^{2\beta(t \wedge s)} - 1) e^{-\beta(t+s)}$$

$$E[x_t] = x_0 + \frac{v_0}{\beta} (1 - e^{-\beta t})$$

$$\text{Cov}[x_t, x_s] = \frac{\sigma^2}{\beta^2} (s \wedge t) + \frac{\sigma^2}{2\beta^3} (-e^{-\beta(t+s)} - e^{-\beta|t-s|} + 2(e^{-\beta s} + e^{-\beta t} - 1))$$

on $\text{Cov}[X, Y] := E[(X - E[X])(Y - E[Y])]$ s'anomena **covariància de X i Y** . La llei d'un procés gaussià queda determinada per les mitjanes i les covariàncies entre les variables del procés. Els càlculs anteriors estan detallats, per exemple, en Arnold [1].

Hem pogut resoldre l'equació (12) sense usar res més que els trucs habituals del càlcul, però es tracta de fet d'un cas molt particular. La hipòtesi d'isotropia del medi en l'espai i en el temps ens ha portat a una expressió molt senzilla del procés $\{M_t, t \in \mathbf{R}^+\}$ en termes del procés de Wiener. El punt clau és que, encara que no tingui sentit parlar del diferencial dW_t per a cap trajectòria, en canvi sí que es pot donar un sentit evident a l'expressió $\int_0^t dW_s$. Només cal definir aquest símbol com $W_t - W_0$, que no és més que el valor que s'obté mitjançant sumes de Riemann. En general, si f_t és una funció (o un procés) de variació fitada, i tenint en compte que W és continu, hom pot definir

$$\int_0^t f_s dW_s$$

mitjançant sumes de Riemann per a cada valor del paràmetre ω i el resultat coincideix amb el que s'obtingria integrant formalment per parts:

$$f_t W_t - f_0 W_0 - \int_0^t W_s df_s,$$

on l'última integral té sentit com a integral de Riemann-Stieltjes clàssica.

En situacions més generals pot passar molt bé que la influència d'una pertorbació externa a un sistema depengui del temps i de l'estat del sistema a cada instant. Si la magnitud de la pertorbació durant un interval de temps $[s, t]$ ve donada per l'increment $M(t) - M(s)$ d'una funció M , freqüentment es pot modelar el sistema resultant mitjançant una equació diferencial del tipus

$$dx(t) = f(t, x(t)) dt + g(t, x(t)) dM(t), \quad (15)$$

on $dx(t) = f(t, x(t))dt$ és l'equació corresponent al sistema sense pertorbar i g és una funció que representa la “sensitivitat” del sistema a la pertorbació M .

Si M és coneguda i prou regular, la situació no té més complicacions que les habituals. Però si M representa una pertorbació de magnitud desconeguda governada per certes lleis de probabilitat (i, per tant, estem tractant un procés, i no pas una sola funció), la situació esdevé diferent. La solució de (15) serà aleshores també un procés estocàstic i l'interès girarà cap a l'estudi de la llei del procés solució, ja que el valor puntual $x(t)$ ha deixat de ser abastable. Actualment, aquesta situació apareix en força aplicacions de la física i d'altres camps diversos.

La dificultat fonamental en aquests casos, quan M és un procés de Wiener, o qualsevol altre amb trajectòries igualment irregulars, rau en el fet que l'expressió (15) no té sentit, ni tan sols si s'escriu en forma integral

$$x_t = x_0 + \int_0^t f(s, x_s) ds + \int_0^t g(s, x_s) dW_s, \quad (16)$$

llevat que g sigui tan regular que suavitzzi les irregularitats que x_t hereta de W_t , com era el cas del nostre exemple ($g \equiv \sigma$).

Cal, doncs, donar un sentit al símbol

$$\int_0^t g(s, x_s) dW_s \quad (17)$$

de forma que es pugui interpretar com l'aportació total de la pertorbació aleatòria a l'estat del sistema en l'interval $[0, t]$.

La definició de (17) donarà sentit a l'equació (16), que al seu torn s'escriu
habituàlment (per conveni) com

$$dx_t = f(t, x_t) dt + g(t, x_t) dW_t. \quad (18)$$

En tal situació diem que tenim una **equació diferencial estocàstica**. (Aquesta terminologia és deguda a Bernštein [3]).

El nostre problema consisteix, doncs, a definir la integral $\int_0^t f_s dW_s$, on $\{W_t, t \in \mathbf{R}^+\}$ és un procés de Wiener i $\{f_t, t \in \mathbf{R}^+\}$ és un procés amb trajectòries no necessàriament de variació fitada.

El resultat d'integrar un procés en un interval fixat $[0, t]$ serà naturalment una variable aleatòria. Per tant, el que busquem és un operador amb domini d'un cert espai de processos, a valors en $L^0(\Omega, \mathcal{F}, p)$ i que per tal de merèixer el nom d'integral, haurà de ser

1. lineal,
2. continu en algun sentit (la qual cosa equival a tenir algun teorema de convergència).

Volem, a més, que

3. si $f_s(\omega)$ és de variació fitada per a tot ω , es tingui

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

estenenent d'aquesta manera el concepte d'integral de Stieljes usual.

Un objectiu modest, a la vista de la situació creada per l'equació (16), és intentar que l'operador actuï almenys sobre processos continus. Les condicions 2) i 3) suggereixen aleshores definir la integral d'un procés continu $f_t(\omega)$ respecte al procés de Wiener com el límit de les sumes

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

si el límit de la dreta existeix i és independent de la successió de particions refinant $\{\pi_n\}_{n \in \mathbb{N}}$ amb $\|\pi_n\| \rightarrow 0$ i els punts $\xi_i \in [t_i, t_{i+1}]$ escollits. I això per a cada ω fixat o, més feblement, prenent el límit en la topologia natural de $L^0(\Omega, \mathcal{F}, p)$, p.e. la topologia de la convergència en mesura. Això correspondria a definir un operador continu entre l'espai de processos continus amb la topologia induïda per la convergència uniforme (en (t, ω)) i l'espai de variables aleatòries $L^0(\Omega, \mathcal{F}, p)$ amb la convergència en mesura.

Malauradament, però, això és impossible. En efecte:

Teorema 2.1. Sigui $\alpha(t)$ una funció contínua en $[0, 1]$ (càdlàg és suficient).

Sigui $\{\pi_n\}_{n \in \mathbb{N}}$ una successió refinant de particions de $[0, 1]$ tal que la successió de normes tendeix a zero.

Llavors,

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

convergeix a un límit finit quan $n \rightarrow \infty$ per a tota funció contínua f si i només si $\alpha(t)$ és de variació fitada. $\square$

Reproduïm aquí una demostració simple d'aquest teorema, suggerida per Meyer [20] i que conté la clau "per sortir d'aquesta situació". La demostració usa el conegut teorema de Banach–Steinhauss, que recordem:

Lema (*teorema de Banach–Steinhaus*)

Siguin B_1 un espai de Banach i B_2 un espai normat. Sigui $\{T_i\}_{i \in I}$ una família d'operadors lineals acotats $T_i: B_1 \rightarrow B_2$.

Si per a cada $x \in B_1$, $\sup_{i \in I} \|T_i x\| < \infty$, aleshores $\sup_{i \in I} \|T_i\| < \infty$. $\square$

Demostració del teorema

$\Leftarrow$) És conegut. Convergeix a la integral de Riemann–Stieljes $\int_0^1 f(t) d\alpha(t)$.

$\Rightarrow$) Siguin B_1 l'espai de funcions contínues en $[0, 1]$ amb la norma del suprem, $B_2 = \mathbf{R}$ amb el valor absolut.

Per a cada $n \in \mathbf{N}$, sigui $T_n: B_1 \rightarrow B_2$ l'operador definit per

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

Per hipòtesi, $\lim_{n \rightarrow \infty} T_n f$ existeix, i per tant $\sup_n |T_n f| < \infty$. Gràcies al lema, $\sup_n \|T_n\| < \infty$.

Vegem ara que la variació de α ,

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

és més petita que $\sup_n \|T_n\|$, amb la qual cosa haurem acabat. Efectivament, per a cada $n \in \mathbf{N}$ fixat, sigui f una funció contínua tal que $f(t_i) = \text{signe}\{\alpha(t_{i+1}) - \alpha(t_i)\}$ i $\|f\|_\infty = 1$. Llavors:

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

$$\|T_n\| \geq \sum_{t_i, t_{i+1} \in \pi_n} |\alpha(t_{i+1}) - \alpha(t_i)| \Rightarrow$$

$$\sup_n \|T_n\| \geq \sup_n \sum_{t_i, t_{i+1} \in \pi_n} |\alpha(t_{i+1}) - \alpha(t_i)| = V^{(1)}(\alpha).$$

(Recordem la caracterització de les variacions d'una funció contínua per a successions de particions que hem comentat a l'apartat 1 i observeu que el $\sup_n$ és substituïble per $\lim_n$.) $\square$

Aquest teorema demostra la impossibilitat d'estendre la definició de la integral de Riemann–Stieljes més enllà d'integradors de variació fitada, si

es vol integrar, com a mínim, totes les funcions contínues. Es pot pensar que, en ser l'operador que busquem valuat en $L^0(\Omega, \mathcal{F}, p)$ en lloc de $\mathbf{R}$, potser sí que es pot obtenir la convergència en mesura de les sumes

$$\sum_{t_i, t_{i+1} \in \pi_n} f_{t_i}(\omega) \cdot (\alpha_{t_{i+1}}(\omega) - \alpha_{t_i}(\omega)) \quad (19)$$

en lloc de la convergència quasi segura en ω . Però la situació no és millor, ja que en el conjunt $A = \{\omega \in \Omega : \alpha_t(\omega) \text{ no és de variació fitada}\}$ podríem construir una successió parcial de (19) convergent quasi segur, i repetir la demostració del teorema per a aquesta parcial. Per tant, si $p(A) > 0$, arribem a la mateixa dificultat que en el cas determinista, en què ω no hi és.

3 Integració estocàstica i càlcul estocàstic

Malgrat tot el que acabem d'exposar, resulta que sí que pot definir-se la integral de processos continus (i més generals fins i tot) respecte al procés de Wiener com a límit de sumes de Riemann, i això és el que farem en aquest apartat.

El punt de partida és adonar-se que el paràmetre aleatori ω ha fet un paper d'espectador passiu en tota la discussió anterior: simplement hem intentat definir la integral respecte al procés de Wiener trajectòria per trajectòria, per a cada ω fixat, i hem vist que no era possible. En segon lloc, és evident que caldrà imposar alguna restricció a l'integrand, però no estem disposats a demanar més regularitat a les seves trajectòries. Hauran de ser restriccions d'una altra mena.

Amb aquestes idees en ment, repassem un moment la demostració de l'últim teorema. Per veure la necessitat que l'integrand sigui de variació fitada hem usat funcions amb valors prefixats sobre els punts de cada partició:

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

Observem que, per assignar aquests valors, cal "mirar endavant" la funció α ; és a dir, el valor de f en t_i depèn del valor de α en t_{i+1} . Si prohibim aquesta dependència, la demostració queda invalidada i, de fet, com veurem, el resultat és fals.

El nostre objectiu és, doncs, posar aquesta restricció en termes matemàtics, i observar després que no impedeix modelar problemes reals amb el formalisme de (18). Cal introduir primer uns quants conceptes addicionals.

Donat un espai de probabilitat $(\Omega, \mathcal{F}, p)$ i un conjunt ordenat T , qualsevol família creixent $\{\mathcal{F}_t, t \in T\}$ de sub- σ -àlgebres de $\mathcal{F}$ s'anomena una **filtració**, i l'estructura $(\Omega, \mathcal{F}, p, \{\mathcal{F}_t, t \in T\})$ és un **espai de probabilitat filtrat**. Una filtració pren nota de l'evolució en el temps de la informació disponible a cada instant. Expliquem aquesta frase. Una variable aleatòria X definida en $(\Omega, \mathcal{F}, p)$ induïx una sub- σ -àlgebra de $\mathcal{F}$ (la σ -àlgebra **inicial** o **generada** per X), que és la més petita que fa mesurable l'aplicació X , i que es nota habitualment per $\sigma\langle X \rangle$. Aquesta σ -àlgebra està formada pels únics conjunts que són observables a través de X . Dit d'altra manera, dels únics esdeveniments que podem decidir si s'han verificat o no (p.e., si ω hi pertany o no) amb la sola informació del resultat concret $X(\omega)$. En aquest sentit $\sigma\langle X \rangle$ és la informació subministrada per X . Ara, si $X \equiv \{X_t, t \in T\}$ és un procés estocàstic, la informació total subministrada pel procés fins a l'instant t és la més petita σ -àlgebra que fa mesurables totes les aplicacions de la família $\{X_s, s \leq t\}$. Ho escrivim $\sigma\langle \{X_s, s \leq t\} \rangle$. Els conjunts d'aquesta σ -àlgebra són els observables a través del procés un cop aquest ha arribat a l'instant t . $\{\sigma\langle \{X_s, s \leq t\} \rangle, t \in T\}$ és la **filtració generada** pel procés X .

Donat un espai de probabilitat filtrat $(\Omega, \mathcal{F}, p, \{\mathcal{F}_t, t \in T\})$, un procés estocàstic X sobre $(\Omega, \mathcal{F}, p)$ parametritzat per T es diu que és **adaptat a** $\{\mathcal{F}_t, t \in T\}$ si X_t és $\mathcal{F}_t$ -mesurable (mesurable respecte a $\mathcal{F}_t$) per a cada $t \in T$. Si $\{\mathcal{F}_t, t \in T\}$ és la filtració generada per un altre procés Y diem que X és **adaptat a** Y .

Sigui $W \equiv \{W_t, t \in \mathbf{R}^+\}$ un procés de Wiener. Una filtració $\{\mathcal{F}_t, t \in \mathbf{R}^+\}$ es diu que **no-anticipa** W si se satisfan les condicions següents:

1. W és adaptat a $\{\mathcal{F}_t, t \in \mathbf{R}^+\}$.
2. $\forall s \in \mathbf{R}, \mathcal{F}_s$ i $\sigma\langle \{W_t - W_s, t \geq s\} \rangle$ són σ -àlgebres independents (això és, tot esdeveniment d'una d'elles és independent de tot esdeveniment de l'altra).

Tals filtracions efectivament existeixen. Prengueu per exemple la induïda pel mateix W , o sigui $\{\sigma\langle \{W_s, s \leq t\} \rangle, t \in \mathbf{R}^+\}$.

Ara tenim a la nostra disposició els ingredients necessaris per modelar la condició “no val mirar endavant” de què parlàvem. Sigui X un procés adaptat a una filtració no-anticipativa. És clar en les definicions que les funcions f usades a la demostració del teorema 2.1 no poden ser trajectòries de X , perquè *depenen del futur de l'integrador*, i ara això està prohibit. Com veurem de seguida, la condició de no-anticipació de

l'integrand permet desenvolupar una teoria satisfactòria de la integració respecte al procés de Wiener.

Sigui $T = [0, \tau]$, amb $\tau \in \mathbf{R}^+ \cup \{+\infty\}$ fixat. Sigui $L^2(T \times \Omega)$ el conjunt de processos estocàstics reals tals que considerats com a aplicacions

$$\begin{aligned} X : T \times \Omega &\longrightarrow \mathbf{R} \\ (t, \omega) &\longrightarrow X_t(\omega) \end{aligned}$$

són mesurables respecte a la σ -àlgebra producte $\mathcal{B}(T) \otimes \mathcal{F}$ (on $\mathcal{B}(T)$ denota la σ -àlgebra de Borel en T) i satisfan

$$\mathbb{E} \left[\int_T |X_t(\omega)|^2 dt \right] < \infty.$$

Fixada una filtració no-anticipativa $\{\mathcal{F}_t, t \in T\}$, considerem el subespai vectorial $L_a^2(T \times \Omega) \subset L^2(T \times \Omega)$ dels processos adaptats a $\{\mathcal{F}_t, t \in T\}$. $L_a^2(T \times \Omega)$ és un subespai vectorial complet i, per tant, hereta l'estructura hilbertiana de $L^2(T \times \Omega)$.

Sigui $\mathcal{S}$ el conjunt dels processos X de la forma

$$X_t(\omega) := c_0(\omega) \cdot 1_{\{0\}} + \sum_{k=0}^{n-1} c_k(\omega) \cdot 1_{[t_k, t_{k+1}]}(t),$$

per alguna partició $\{0 = t_0 < t_1 < \dots < t_n = \tau\}$ de T , i variables aleatòries $c_k(\omega)$ $\mathcal{F}_{t_k}$ -mesurables i de quadrat integrable. $\mathcal{S}$ és clarament un subespai vectorial de $L_a^2(T \times \Omega)$.

Definició. Per processos $X \in \mathcal{S}$, definim la **integral estocàstica de X respecte al procés de Wiener** “à la Riemann”:

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

□

No és difícil demostrar la proposició següent.

Proposició 3.1. Per a tot $X \in \mathcal{S}$,

1. la variable aleatòria $\int_T X_t dW_t$ pertany a $L^2(\Omega, \mathcal{F}, p)$.

$$2. \mathbb{E} \left[\int_T X_t dW_t \right] = 0.$$

$$3. \mathbb{E} \left[\left| \int_T X_t dW_t \right|^2 \right] = \mathbb{E} \left[\int_T |X_t|^2 dt \right]$$

$$(\text{és a dir, } \left\| \int_T X_t dW_t \right\|_{L^2(\Omega)} = \|X\|_{L_a^2(T \times \Omega)}). \quad \square$$

D'altra banda, tenim també:

Proposició 3.2. $\mathcal{S}$ és un subespai dens de $L_a^2(T \times \Omega)$. $\square$

Per tant, acabem de definir una aplicació lineal que preserva la norma entre un subespai dens de $L_a^2(T \times \Omega)$ i $L^2(\Omega)$. En conseqüència, podem estendre-la a un operador lineal isomètric

$$I: L_a^2(T \times \Omega) \rightarrow L^2(\Omega).$$

L'extensió és única i ve donada per

$$I(X) := L^2(\Omega) - \lim_{n \rightarrow \infty} \int_T X_t^{(n)} dW_t,$$

on $\{X^{(n)}\}_{n \in \mathbb{N}}$ és qualsevol successió de processos de $\mathcal{S}$ convergint a X en $L_a^2(T \times \Omega)$.

$I(X)$ és la **integral estocàstica de X respecte a W** que volíem definir. Naturalment, la notació més habitual és $\int_T X_t dW_t$. Tots els apartats de la proposició 3.1 es verifiquen per a cada procés de $L_a^2(T \times \Omega)$. Això es demostra per arguments estàndards de pas al límit.

Així doncs, hem pogut definir la integral estocàstica sense haver d'imposar condicions dràstiques de regularitat a l'integrand, al preu, això sí, de certs requeriments de mesurabilitat respecte de la variable ω .

És evident que si $X \in L_a^2(T \times \Omega)$, llavors també $X \in L_a^2(U \times \Omega)$ per a tot interval $U \subset T$. Per tant, és natural demanar-se per les propietats del procés $\{I_t(X) := \int_{[0,t]} X_s dW_s, t \in T\}$. En primer lloc, tenim la proposició següent. La demostració és fàcil per a processos de $\mathcal{S}$. El resultat general se segueix per densitat.

Proposició 3.3. El procés $\{I_t(X), t \in \mathbf{R}^+\}$ és adaptat a $\{\mathcal{F}_t, t \in \mathbf{R}^+\}$. $\square$

Aquest procés és, doncs, un element de $L_a^2(T \times \Omega)$. Com a element d'aquest espai, $[I_t(X)](\omega)$ està definit llevat d'un conjunt de mesura zero respecte $dt \otimes dp$ i, per tant, no té sentit parlar de les seves propietats trajectorials. No obstant això, es pot afirmar el següent:

Proposició 3.4. Existeix un representant de la classe $\{I_t(X), t \in T\} \in L_a^2(T \times \Omega)$ que és un procés continu. $\square$

A la vista d'aquest resultat, definirem el **procés integral estocàstic de X en T** com el representant continu de $\{I_t(X), t \in T\}$ (determinat q. s.). Sempre ens referim a aquest representant quan escrivim $\{\int_0^t X_s dW_s, t \in T\}$.

En el càlcul diferencial i integral clàssic, els teoremes forts com ara els de canvi de variable, integració per parts, teorema fonamental del càlcul, etc. ens permeten calcular integrals sense embrutar-nos les mans amb els complicats passos al límit que es desprenen de les definicions. És raonable esperar que resultats anàlegs puguin ser demostrats per al càlcul associat a les integrals estocàstiques, al qual anomenem *càlcul estocàstic*. Això és així, efectivament, però les fórmules que en resulten no són formalment iguals a les corresponents del càlcul clàssic. Començarem enunciant la fórmula de canvi de variable, de la qual poden deduir-se les altres.

Teorema 3.5 (*canvi de variable*)

Sigui $f \in L_a^2(T \times \Omega)$.

Sigui $\{X_t, t \in T\}$ el procés definit per $X_t = \int_0^t f_s dW_s$.

Sigui $\varphi: \mathbf{R} \rightarrow \mathbf{R}$ una funció dues vegades derivable amb continuïtat.

Aleshores:

$$\varphi(X_t) = \varphi(X_0) + \int_0^t f_s \cdot \varphi'(X_s) dW_s + \frac{1}{2} \int_0^t f_s^2 \cdot \varphi''(X_s) ds. \quad (20)$$

$\square$

Si comparem aquesta fórmula amb l'anàloga del càlcul integral habitual observem que hi apareix el terme addicional $\frac{1}{2} \int_0^t f_s^2 \cdot \varphi''(X_s) ds$.

Exemple. Siguin $X_t = W_t$ i $\varphi(x) = x^n$. Obtindrem:

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

resultat diferent al que s'obté amb el càlcul diferencial tradicional.

Una conseqüència immediata del teorema 3.5 és que el conjunt de processos que són integrals estocàstiques no és tancat per transformacions tan regulars com ara $x \rightarrow x^2$ de l'exemple anterior. Una dW_s -integral es converteix en un procés suma d'una dW_s -integral i una ds -integral. (Una ds -integral és clarament de variació fitada mentre que la dW_s -integral té variació quadràtica no nul·la; d'aquí que cap d'elles es pot convertir en l'altra.)

La raó intuïtiva per la qual apareix el terme addicional és la següent: usant la fórmula de Taylor per a φ en el punt W_t , avaluada en $W_{t+\Delta t}$, i dividint per Δt , es té:

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

Si W_t fos una funció derivable, el terme quadràtic i tots els d'ordre superior convergrien clarament a zero quan $\Delta t \rightarrow 0$ i obtindríem la coneguda regla de la cadena per a la composició $\varphi \circ W$. En el nostre cas, *definint* $\frac{dW_t}{dt} = \lim_{\Delta t \rightarrow 0} \frac{W_{t+\Delta t} - W_t}{\Delta t}$ (això és precisament el que estem fent quan posem $W_t - W_s := \int_s^t dW_r$), i observant que $(W_{t+\Delta t} - W_t)^2$ és de l'ordre de Δt quan $\Delta t \rightarrow 0$ (variació quadràtica de W), obtenim justament (20) per al cas $X_t = W_t$. El mateix succeeix quan X_t és un procés integral estocàstic.

Definició. Siguin $f \in L_a^2(T \times \Omega)$ i $g \in L_a^1(T \times \Omega)$ (el significat de $L_a^1(T \times \Omega)$ és el natural).

Sigui X_0 una variable aleatòria $\mathcal{F}_0$ -mesurable.

Un procés $\{X_t, t \in T\}$ de la forma

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

s'anomena **procés d'Itô**. $\square$

És immediat a la definició que un procés d'Itô és sempre continu i adaptat. La classe dels processos d'Itô és tancada per canvi de variable. Formularem una versió més general del teorema 3.5, deixant que φ actui sobre n processos d'Itô al mateix temps (en lloc d'una sola integral estocàstica) i que depengui també de t . Aquesta versió es coneix com a *fórmula d'Itô* i és l'eina bàsica del càlcul estocàstic.

Teorema 3.6 (*fórmula d'Itô*)

Siguin $\{X^{(i)}\}_{i=1}^n$, n processos d'Itô amb representació

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

per a t en un cert interval $T = [0, \tau]$. Notem $\vec{X}_t = (X_t^{(1)}, \dots, X_t^{(n)})$.

Sigui

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

una funció tal que $\frac{\partial \varphi}{\partial t}$, $\frac{\partial \varphi}{\partial x_i}$, $\frac{\partial^2 \varphi}{\partial x_i \partial x_j}$ existeixen i són contínues.

Aleshores, $Y_t := \varphi(t, \vec{X}_t)$ és de nou un procés d'Itô en T , i es pot escriure com

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

(sumant d'1 a n sobre els índexs repetits). $\square$

Definició. Si $X_t = X_0 + \int_0^t f_s ds + \int_0^t g_s dW_s$ és un procés de Itô, diem que té **diferencial estocàstica**

$$dX_t = f_t ds + g_t dW_s$$

$\square$

Per exemple, hem vist que $dW_t^2 = 2W_t dW_t + dt$. La notació de diferencials estocàstiques permet una economia en la escriptura i (vegeu l'apartat 2) i dóna el significat precís a una equació diferencial estocàstica. La fórmula d'Itô posada en forma diferencial,

$$dY_t = \left[\frac{\partial \varphi}{\partial t}(t, \vec{X}_t) + \frac{\partial \varphi}{\partial x_i}(t, \vec{X}_t) f_t^{(i)} + \frac{1}{2} \frac{\partial^2 \varphi}{\partial x_i \partial x_j}(t, \vec{X}_t) \cdot g_t^{(i)} g_t^{(j)} \right] dt \quad (21)$$

$$+ \frac{\partial \varphi}{\partial x_i}(t, \vec{X}_t) g_t^{(i)} dW_t \quad (22)$$

suggereix el nom *regla de la cadena*, amb el qual també se la coneix. Si fem el conveni $(dW_t)^2 = dt$, $dW_t dt = dt dW_t = 0$, $dt dt = 0$ i $dX_1 dX_2$ és

la multiplicació dels respectius binomis, podem escriure (21) en la forma més compacta

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

Aplicant la regla de la cadena a les diferencials estocàstiques

$$dX_t^{(1)} = f_t^{(1)} dW_t + g_t^{(1)} dt$$

$$dX_t^{(2)} = f_t^{(2)} dW_t + g_t^{(2)} dt$$

amb $\varphi(t, x_1, x_2) = x_1 \cdot x_2$ obtenim la fórmula d'integració per parts

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

Per exemple, podem calcular altre cop $\int_0^t W_s dW_s$ integrant per parts:

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

Aplicant el teorema 3.5 a $X_t = \int_0^t dW_s = W_t$, obtenim, per a φ de classe C^2 ,

$$\varphi(W_t) = \varphi(W_s) + \int_s^t \varphi'(W_r) dW_r + \frac{1}{2} \int_s^t \varphi''(W_r) dr,$$

un teorema fonamental del càlcul, o

$$\int_s^t \varphi'(W_r) dW_r = \varphi(W_t) - \varphi(W_s) - \frac{1}{2} \int_s^t \varphi''(W_r) dr,$$

en forma de *regla de Barrow*.

Veiem ara com es resol una equació diferencial estocàstica molt senzilla. A diferència de la de l'exemple 1, que era resoluble sense abandonar el marc del càlcul clàssic, aquesta involucra una integral estocàstica i, per tant, hem de fer ús de les regles del càlcul estocàstic. Volem trobar un procés x_t que sigui solució del problema de Cauchy

$$dx_t = x_t dW_t, \quad x_0 = 1.$$

És a dir, un procés que verifiqui $x_t = 1 + \int_0^t x_s dW_s$.

Si W_t fos una funció diferenciable, usualment escriuríem $\frac{dx_t}{x_t} = \dot{W}_t dt$, seguit de $d(\log x_t) = \dot{W}_t dt$, i la solució es troba de seguida integrant i prenent exponencials.

Seguint la mateixa línia de pensament, calculem $d(\log x_t)$ usant la regla de la cadena:

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

Integrant i usant $x_0 = 1$, obtenim

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

i prenent exponencials

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

Per analogia amb el càlcul clàssic, anomenem a x_t la **funció exponencial estocàstica**, que difereix de la usual en el factor $\exp\{-\frac{1}{2}t\}$. Acabem de presentar una petita mostra del càlcul estocàstic que es dedueix de la definició de la integral estocàstica. La nostra intenció ha estat simplement mostrar que és possible desenvolupar un concepte d'integral i unes regles de càlcul ben adaptats a l'estudi de sistemes amb incertesa. Per acabar de justificar aquesta afirmació només cal mirar si situacions com les dels exemples vistos a l'apartat 0 poden ser modelades amb el formalisme de (18). Òbviament, això serà cert sempre que tinguem la seguretat *a priori* que les integrals que hi ha implícites a l'equació diferencial estocàstica tenen sentit. La qüestió crítica és, sens dubte, si l'integrand de la integral estocàstica és o no adaptat a una filtració no-anticipativa del procés de Wiener.

Suposem que combinem l'equació

$$dx_t = f(t, x_t) dt + g(t, x_t) dW_t$$

amb una condició inicial $x_0 \in \mathbf{R}$ (és a dir, una variable aleatòria constant) i que l'evolució de x_t no té altra causa d'estocasticitat que la induïda per l'evolució del procés $W \equiv \{W_t, t \in \mathbf{R}^+\}$. En tal cas, és clar que el procés solució és adaptat a la filtració generada per W i el formalisme funciona.

Més generalment, suposem que la condició inicial és una variable aleatòria x_0 independent de totes les variables del procés W . Llavors la solució serà adaptada a la filtració $\{\mathcal{F}_t, t \in \mathbf{R}\}$ definida per

$$\mathcal{F}_t = \sigma\langle x_0 \rangle \vee \sigma\langle \{W_s, s \leq t\} \rangle,$$

on $\mathcal{F} \vee \mathcal{G}$ denota la σ -àlgebra generada per (mínima σ -àlgebra que conté a) $\mathcal{F}$ i $\mathcal{G}$. La σ -àlgebra $\mathcal{F}_s$ és independent de la generada per les variables $W_t - W_s$, i per tant la filtració és no-anticipativa i l'equació diferencial tindrà sentit.

4 Equacions diferencials estocàstiques de primer ordre

Una forma força general d'equació diferencial estocàstica de primer ordre en dimensió 1 és

$$dx_t = f(t, x_t) dt + g(t, x_t) dW_t \quad , \quad t \in T = [0, \tau], \quad (23)$$

on x_t és el procés incògnita, W_t és el procés de Wiener, i $f(t, x)$ i $g(t, x)$ són funcions reals definides i mesurables en $T \times \mathbf{R}$, que anomenem coeficients de l'equació.

En aquest apartat precisarem el concepte de solució de (23) i enunciem dos resultats clàssics (d'existència i unicitat, i de dependència de condicions inicials) per a aquest tipus d'equacions. No es tracta dels millors enunciats possibles; només pretenem illustrar algunes analogies i diferències amb els resultats homòlegs de la teoria d'equacions diferencials ordinàries.

Comencem amb la definició de solució de (23), la qual cosa equival a precisar exactament què entenem per l'expressió (23).

Definició. Un procés $\{x_t, t \in T\}$ és solució de (4.1) si i només si:

1. $\{x_t, t \in T\}$ és adaptat a alguna filtració $\{\mathcal{F}_t, t \in T\}$ no anticipativa de W .
2. $\bar{g}_t := g(t, x_t) \in L^2(T \times \Omega)$ i $\bar{f}_t := f(t, x_t) \in L^1(T \times \Omega)$.
3. x_t té diferencial estocàstica

$$dx_t = \bar{f}_t dt + \bar{g}_t dW_t, \quad t \in T.$$

Equivalentment, x_t satisfà

$$x_t = x_0 + \int_0^t f(s, x_s) ds + \int_0^t g(s, x_s) dW_s, \quad t \in T. \quad \square$$

Les condicions 1) i 2) de la definició, conjuntament, garanteixen que les expressions de 3) tenen sentit. En particular, observeu que el problema de Cauchy

$$dx_t = f(t, x_t) dt + g(t, x_t) dW_t, \quad x_0 = \xi \quad (24)$$

només podrà tenir solució si ξ és una variable aleatòria independent de $\{W_t, t \in T\}$.

Per al problema de Cauchy (24) amb $g \equiv 0$ (equació diferencial ordinària), el resultat clàssic d'existència i unicitat de solució requereix que f compleixi alguna condició de Lipschitz. És natural intentar el mateix tipus de planteig en el cas $g \not\equiv 0$.

Teorema 4.1. Supposem que:

1. existeix $K > 0$ tal que

$$|f(t, x) - f(t, y)| + |g(t, x) - g(t, y)| \leq K \cdot |x - y|, \quad \forall t \in T, \quad \forall x, y \in \mathbf{R}$$

2. existeix $L > 0$ tal que

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

3. ξ és una variable aleatòria independent de $\{W_t, t \in T\}$, i tal que $\xi \in L^2(\Omega)$.

Llavors:

- a) Existeix una solució de (24) que és un procés continu.
- b) Si $\{x_t^1, t \in T\}$ i $\{x_t^2, t \in T\}$ són processos continus solució de (24), llavors $\sup_{t \in T} |x_t^1 - x_t^2| = 0$, q.s. $\square$

En aquest teorema trobem, a part de l'esperada condició de Lipschitz (1), una hipòtesi de restricció en el creixement dels coeficients (2) i l'exigència que la variable inicial ξ sigui de quadrat integrable i que no tingui res a veure amb el procés de Wiener que governa l'equació (3). Aquestes hipòtesis permeten demostrar que qualsevol procés candidat a ser solució verifica les condicions 1) i 2) de la definició.

Pel que fa a la tesi del teorema, només s'assegura la unicitat de solució contínua a causa de l'ambigüitat en la definició del procés integral estocàstic (vegeu els comentaris a les proposicions 3.3 i 3.4). De fet, l'exigència

de continuïtat pot incorporar-se a la definició de solució, i llavors es pot afirmar que dues solucions només difereixen en un conjunt de trajectòries de probabilitat nulla.

La dependència de la solució respecte de la condició inicial és similar a la que es troba en el cas determinista. En particular, es té el següent resultat de diferenciabilitat de primer ordre respecte a una condició inicial constant.

Teorema 4.2. Suposem:

1. Les funcions $f(t, x)$ i $g(t, x)$ satisfan les hipòtesis 1) i 2) del teorema 4.1.
2. Les derivades parcials $\frac{\partial}{\partial x}f(t, x)$ i $\frac{\partial}{\partial x}g(t, x)$ existeixen i són contínues i acotades.

Llavors, la solució $\{x_t(c), t \in T\}$ del problema de Cauchy

$$dx_t = f(t, x_t) dt + g(t, x_t) dW_t, \quad x_0 = \xi,$$

amb $c \in \mathbf{R}$, és derivable en $L^2(\Omega)$ respecte a c (això és, existeix el límit $L^2(\Omega)\text{-}\lim_{h \rightarrow 0} \frac{x_t(c+h) - x_t(c)}{h}$) i el procés derivada $\{y_t(c) := \frac{d}{dc}x_t(c), t \in T\}$ satisfà l'equació variacional

$$dy_t(c) = \frac{\partial}{\partial x}f(t, x_t(c)) \cdot y_t(c) dt + \frac{\partial}{\partial x}g(t, x_t(c)) \cdot y_t(c) dW_t, \quad y_0(c) = 1. \quad (25)$$

□

Noteu que (25) és una equació lineal (en el sentit que els coeficients són lineals respecte al procés incògnita) en $Y_t(c)$, però amb coeficients aleatoris. És, per tant, en realitat, una equació de primer ordre d'un tipus més general que (23).

I.V. Girsanov donà, el 1962, un exemple ben senzill d'equació diferencial estocàstica que té un comportament diferent que la seva anàloga determinista pel que fa a la unicitat de solucions. La majoria de llibres sobre equacions diferencials estocàstiques en fan esment.

L'equació diferencial ordinària

$$dx_t = |x_t|^\alpha dt, \quad t \in T = [0, \tau], \quad x_0 = 0. \quad (26)$$

amb $0 < \alpha < 1$, té una infinitat de solucions. En efecte, per a tot $s \in T$, la funció

$$x_t = \begin{cases} 0, & \text{si } 0 \leq t \leq s \\ [(1 - \alpha)(t - s)]^{(1-\alpha)^{-1}}, & \text{si } t \geq s \end{cases} \quad (27)$$

és solució de (26).

En canvi, l'equació estocàstica

$$dx_t = |x_t|^\alpha dW_t, \quad t \in T = [0, \tau], \quad x_0 = 0 \quad (28)$$

resulta tenir infinites solucions si $0 < \alpha < 1/2$, però només una única solució (la trivial) per a $\alpha \geq 1/2$, tot i no complir les condicions del teorema 4.1.

Es poden trobar més detalls sobre els resultats apuntats en aquest apartat i en general sobre la teoria de les equacions diferencials estocàstiques en el clàssic Gihman-Skorohod [14], o en Arnold [1] o Gard [13].

5 Comentaris finals

La matematització del concepte de probabilitat que hem presentat a l'apartat 1 és deguda a A. N. Kolmogorov (1933). Amb ella queda en certa manera resolt el problema de definir què és la probabilitat: simplement no es defineix. Tan sols es proposen unes certes propietats com a axiomes. La comoditat del planteig axiomàtic ha fet minvar força l'interès que hi hagué en un cert temps per discernir la natura del concepte de probabilitat com a “quelcom físic” i fer una definició matemàtica constructiva a partir de la idea intuïtiva d'atzar. La crítica principal que va rebre la idea de Kolmogorov és que condemna la teoria de la probabilitat a ser un subconjunt de la teoria de funcions de conjunt additives, i queda així desprovista d'entitat pròpia independent. Sobre aquest tema és força interessant el llibre divulgatiu de von Mises [21]. Per a una exposició més detallada de la teoria axiomàtica recomanem el llibre bàsic d'E. Bonnet [4] i el curs de D. Nualart – M. Sanz [23], també la referència estàndard R.B. Ash [2], que enquadra perfectament la teoria de la probabilitat com a subdisciplina de l'anàlisi matemàtica. $\square$

La teoria dels processos estocàstics és extensíssima i es ramifica en moltes direccions. La majoria de textos que tracten, amb més o menys profunditat però amb vocació introductòria, la integració o les equacions

diferencials estocàstiques, ofereixen com a preliminars la part de la teoria de processos que és rellevant per als propòsits de l'autor. $\square$

Entre els processos estocàstics a temps continu, el procés de Wiener és possiblement el més estudiat, per la seva importància teòrica i aplicada. Hom pot trobar arreu moltes variants en la definició, però gairebé totes són essencialment equivalents. El procés de Wiener fou estudiat per primer cop amb rigor per N. Wiener, en una sèrie de treballs entre 1923 i 1934, i d'aquí el seu nom. Wiener es beneficià dels recents avenços en teoria de la mesura portats a terme per Borel, Lebesgue i Daniell, per tal de fer una construcció del procés a partir de l'espai de trajectòries. Einstein, encara sense aquestes eines, deduí en els seus treballs sobre moviment brownià, a partir de certes hipòtesis físiques (més aviat vagament formulades), que la llei de probabilitat que segueix a l'instant t la posició de la partícula browniana que parteix en $t = 0$ de l'origen de coordenades té una densitat $f(t, x)$ que ha de satisfer l'equació

$$\frac{\partial f}{\partial t} = \frac{D}{2} \Delta f \quad (29)$$

on Δ és l'operador de Laplace. A aquesta conclusió havia arribat també abans (1900) L. Bachelier, encara que el seu treball va passar més aviat desapercbut a l'època. (29) no és més que l'equació de la difusió, ja coneguda en temps d'Einstein. La difusió és un fenomen físic que consisteix en el fet que un conjunt de partícules dissoltes en un fluid tendeixen a desplaçar-se de les zones de més concentració a les zones de menys concentració, com ara una gota de tinta en un cubell d'aigua, que es va escampant i va perdent color, fins a desaparèixer. En la interpretació original, $f(t, x)$ és la concentració de partícules. (Vegeu Jou [15, cap. 3 i 4], per a una exposició planera dels fenòmens de sedimentació i difusió, i la seva relació amb el moviment brownià.) D és una constant dependent del medi que s'anomena *coeficient de difusió*. Einstein descobrí també la senzilla relació entre el coeficient de difusió i el de fricció β :

$$D = \frac{2kT}{\beta}$$

on k és la constant de Boltzmann i T la temperatura absoluta.

La solució fonamental de l'equació (29) és (en $\mathbf{R}^3$)

$$f(t, x) = \frac{1}{(2\pi Dt)^{3/2}} \exp \left\{ -\frac{\|x\|^2}{2Dt} \right\}.$$

Aquesta densitat essencialment caracteritza el procés de Wiener (el procés de Wiener en $\mathbf{R}^3$ té per coordenades processos de Wiener reals independents). Més precisament, és la densitat de $X_t = \sqrt{D}W_t$, on W és un Procés de Wiener estàndard (tridimensional). Això justifica el nom de coeficient de difusió per a la constant σ^2 en el teorema 1.1.

Els articles d'Einstein sobre moviment brownià estan recollits en el llibre editat per Fürth [12]. Els articles clàssics d' Uhlenbeck i Ornstein, Chandrasekar, Wang i Uhlenbeck, i Doob, esmentats a l'apartat 1, es troben també en forma de llibre en Wax [29]. $\square$

Tornem ara per última vegada a l'equació de Langevin, l'equació que regeix la velocitat de la partícula browniana de l'exemple 1. Hem començat formulant unes certes hipòtesis físiques de tipus microscòpic (de comportament mecànic molecular) amb la idea que a partir d'elles es puguin predir de forma teòrica els comportaments macroscòpics observables. Aquest és el principi habitual d'elaboració d'una teoria física. Això ens ha portat a l'equació "física" (3). Posteriorment, la introducció a l'apartat 1 de les eines matemàtiques convenients ens ha donat la possibilitat de formular un model amb ple sentit matemàtic, (12), que podem expressar també amb el formalisme de les equacions diferencials estocàstiques (18) com

$$dv_t = -\beta v_t + \sigma dW_t.$$

L'únic element d'aquesta equació que les hipòtesis inicials no ens ha permès determinar és el coeficient σ . Deduirem què val σ en termes de constants físiques fent ús d'un important resultat de mecànica estadística.

De la fórmula per $\text{Cov}[v_t, v_s]$ de l'apartat 2, deduïm que $\text{Var}[v_t] = \frac{\sigma^2}{2\beta}(1 - e^{-2\beta t})$, i per tant, fent $t \rightarrow \infty$, obtenim una variància límit de $\frac{\sigma^2}{2\beta}$, sigui quina sigui la velocitat inicial v_0 . Anàlogament, s'observa que $\lim_{t \rightarrow \infty} E[v_t] = 0$. La llei d'equipartició de l'energia en mecànica estadística afirma que l'energia cinètica mitjana d'un sistema de partícules (molècules d'un fluid) en estat d'equilibri termodinàmic (vegeu Jou [15]) ha de valer $\frac{1}{2}kT$ per grau de llibertat del sistema. La partícula immersa en el fluid ha de compartir aquesta mateixa energia mitjana. L'esperança de l'energia cinètica es pot calcular, doncs, com $E[\frac{1}{2}v^2] = \frac{1}{2}\frac{\sigma^2}{2\beta}$ (on v representa la velocitat límit i continuem fent el conveni que la massa de la partícula val 1), i obtenim per tant la igualtat

$$\frac{1}{2}\frac{\sigma^2}{2\beta} = \frac{1}{2}kT.$$

Deduïm que, en el cas de dimensió 1, σ^2 és, en termes dels coeficients de fricció i de difusió,

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

L'equació de Langevin es pot expressar doncs

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

cosa que completa la seva formulació. La nova hipòtesi que hem introduït per poder determinar σ és que el sistema de molècules verifica l'equació tèrmica dels gasos ideals

$$pV = nkT$$

que és d'on es dedueix la llei d'equipartició. p és la pressió, V el volum i n el nombre total de molècules. $\square$

L'argument d'impossibilitat de la construcció d'una integral estocàstica útil com a extensió trivial de la integral ordinària està essencialment pres de Protter [25].

La integral estocàstica fou definida rigorosament per primer cop per K. Itô (1944), essencialment de la mateixa manera que ha estat aquí exposada. Aquesta integració estocàstica té un precedent en el treball de N. Wiener, que definí la integració de funcions deterministes respecte al procés que porta el seu nom. La integral de Wiener és simplement la restricció natural de la d'Itô, obtenint-se un funcional

$$I: L^2(T) \rightarrow L^2(\Omega)$$

La integral estocàstica d'Itô admet diferents extensions. En primer lloc, el lector haurà observat que, contràriament al que s'afirma al principi de l'apartat 3, no tots els processos continus poden ser integrats amb la teoria vista, perquè a part de la restricció de mesurabilitat de l'integrand hem imposat també una altra petita condició, encara que no sobre les trajectòries: hem demanat que l'integrand fos de quadrat integrable en (t, ω) , cosa que implica que les variables aleatòries X_t que el constitueixen siguin de $L^2(\Omega)$, quasi per tot t respecte a la mesura de Lebesgue.

No obstant això, l'apartat 3 és reproduïble quasi exactament, canviant l'espai $L_a^2(T \times \Omega)$ per l'espai dels processos adaptats $\{X_t, t \in T\}$ que compleixen

$$\int_T |X_t(\omega)|^2 dt < \infty, \quad \forall \omega \in \Omega \quad (\text{o només q. s.}).$$

L'única diferència és que el nou espai de processos no és un espai normat. Hom pot definir la convergència de $X_t^{(n)}(\omega)$ a $X_t(\omega)$ en aquest

espai com la convergència a zero en mesura de les variables aleatòries $\int_T |X_t^{(n)}(\omega) - X_t(\omega)|^2 dt$ i la topologia que resulta és metrizable. La integral es defineix igual sobre el subconjunt dens de processos $\mathcal{S}$ i l'extensió al cas general es fa usant el teorema de prolongació de funcions contínues entre espais mètrics. No hem adoptat aquesta generalitat per fer més còmoda l'exposició.

Una generalització més interessant és la que es refereix als processos que es poden usar com a integradors. El procés de Wiener és un cas particular d'un tipus de procés estocàstic anomenat *semimartingala*. La teoria de la integració respecte a semimartingales és certament molt més complicada que l'exposada aquí i requereix aprofundir molt més en la teoria de processos estocàstics. Vegeu en Protter [25] una exposició que sembla especialment pedagògica. La teoria esdevé més senzilla si es consideren només integradors continus (com, per exemple, en Karatzas-Shreve [16]). Per a una exposició més completa de la integració només respecte al procés de Wiener, recomanem el llibre de Gard [13], on es discuteixen mètodes efectius de resolució analítica d'equacions, i també mètodes numèrics, amb un capítol dedicat a aplicacions a la biologia. Arnold [1] és una altra introducció elemental, on a més es dóna sentit, des del punt de vista distribucional, a la derivada del procés de Wiener $\frac{dW_t}{dt}$. El procés estocàstic derivada rep el nom de *soroll blanc*, i ha estat usat a bastament, almenys de manera formal, en enginyeria de la comunicació. $\square$

Segons les hem presentades, les equacions diferencials estocàstiques sorgeixen de manera natural en tenir en compte pertorbacions en sistemes que, en altre cas, serien modelables mitjançant equacions diferencials deterministes. Hem justificat l'adequació del model al final de l'apartat 3 usant com a base la pertorbació d'un problema de Cauchy per a equacions diferencials ordinàries. És natural demanar-se per l'aplicació de la teoria a equacions amb condicions a la frontera o a equacions amb derivades parcials.

Considereu per exemple

$$dx_t = f(t, x_t) dt, \quad t \in [0, 1], \quad h(x_0, x_1) = 0.$$

Si sumem a la banda dreta de l'equació el terme $g(t, x_t) dW_t$ obtenim formalment una equació diferencial estocàstica amb la solució sotmesa a una condició de contorn. És evident que les variables aleatòries x_0 i x_1 del procés solució no són independents, ja que la funció h lliga els seus valors. D'altra banda, el valor de x_1 també és clar que dependrà de l'evolució de

W en $[0, 1]$. En conclusió, la variable x_0 , i per tant totes les variables de la solució, no són independents dels increments futurs de W i l'integrand $g(t, x_t)$ no serà, en general, adaptat a cap filtració no-anticipativa. En aquest cas, la nostra teoria és inaplicable. Per aconseguir tractar aquest cas (o el del problema de Cauchy amb condició inicial no adaptada) amb el mateix formalisme (19), s'ha començat a desenvolupar molt recentment una integració i un càlcul estocàstic per a integrands no necessàriament adaptats. *Grosso modo*, això s'aconsegueix configurant Ω amb una estructura vectorial i topològica adequada i demanant a l'integrand una certa regularitat, però no respecte del paràmetre temporal t , sinó respecte del paràmetre aleatori $\omega \in \Omega$.

El cas de les equacions amb derivades parcials requereix l'ús de la teoria dels processos estocàstics amb paràmetre multidimensional ($t \in T \subset \mathbf{R}^n$), on una primera dificultat ve donada per les diferents possibles nocions de "passat" i "futur" i, per tant, de filtració no-anticipativa.

És possible definir altres tipus d'integral estocàstica diferent de la d'Itô i les seves extensions. Particularment, Stratonovich (1966) donà una definició de la integral tal que el càlcul estocàstic resultant obeeix les regles

clàssiques del càlcul newtonià. Per exemple, $\int_0^t W_s dW_s = \frac{1}{2}W_t^2$. L'in-

convenient de la integral de Stratonovich és que els processos integrals no tenen tan bones propietats des del punt de vista probabilístic com els que resulten de la integral d'Itô. Conseqüentment, la facilitat en el càlcul es paga al preu d'una més gran dificultat a l'hora d'analitzar o utilitzar els resultats. En tot cas, la possibilitat de diferents definicions del símbol

$\int_T g(t, X_t) dW_t$ ens condueix a la qüestió de l'ambigüitat en la modelació.

Una mateixa equació té solucions diferents segons interpretem la integral estocàstica en el sentit d'Itô o de Stratonovich. La interpretació correcta ha de venir d'alguna manera donada per l'heurística emprada per arribar a l'equació estocàstica a partir del problema real o d'una equació prèvia no pertorbada. Una extensa discussió sobre els dos tipus d'integral i els problemes de modelació pot trobar-se en el llibre de Øksendal [24], amb exemples apropiats. Sense abandonar un nivell elemental, el llibre tracta també del problema del filtratge (exemple 3 de l'apartat 0) i de les aplicacions del càlcul estocàstic a les equacions amb derivades parcials (deterministes). També hi ha força comentaris sobre el tema de la modelació, amb moltes referències a altres treballs, en Gard [13]. Aquesta qüestió és precisament la motivació inicial en McShane [19] i [18], que intenta una teoria unificadora de les integrals ordinàries i estocàstiques.

Referències

- [1] L. Arnold: *Stochastic Differential Equations: Theory and Applications*, John Wiley & Sons, 1974.
- [2] R.B. Ash: *Real Analysis and Probability*, Academic Press, 1972.
- [3] S.N. Bernštein: *Principis de la teoria de les equacions diferencials estocàstiques*, Institut Steklov, **5** (1934) [en rus].
- [4] E. Bonet: *Espais de Probabilitat Finita*, segona edició, Teide, 1975.
- [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).
- [8] J.L. Doob: *Stochastic Processes*, John Wiley & Sons, 1953.
- [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).
- [11] J. Favard: *Cours d'Analyse de l'École Polytechnique*, Gauthier-Villars, 1968.
- [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.

- [15] D. Jou: *Mecànica Estadística i Biologia Molecular*, Publicacions de la Universitat Autònoma de Barcelona, 1989.
- [16] I. Karatzas, S.E. Shreve: *Brownian Motion and Stochastic Calculus*, Springer-Verlag, 1988.
- [17] P. Langevin: *Comptes Rendus*, **146** (1908).
- [18] E.J. McShane: *A Riemann-type integral that includes Lebesgue-Stieltjes, Bochner and Stochastic Integrals*, *Memoirs of the A.M.S.*, **68** (1969).
- [19] E.J. McShane: *Stochastic Calculus and Stochastic Models*, Academic Press, 1974.
- [20] P.A. Meyer: *Geometrie Différentielle Stochastique*, Colloque en l'Honneur de Laurent Schwartz, *Asterisque* **131**, 1985.
- [21] R.E. von Mises: *Probability, Statistics and Truth*, Dover.
- [22] E. Nelson: *Dynamical Theories of Brownian Motion*, Princeton University Press, 1967.
- [23] D. Nualart, M. Sanz: *Curs de Probabilitats*, Promociones i Publicaciones Universitarias, 1990.
- [24] B. Øksendal: *Stochastic Differential Equations*, second edition, Springer-Verlag, 1989.
- [25] P. Protter: *Stochastic Integration and Differential Equations*, Springer-Verlag, 1990.
- [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.