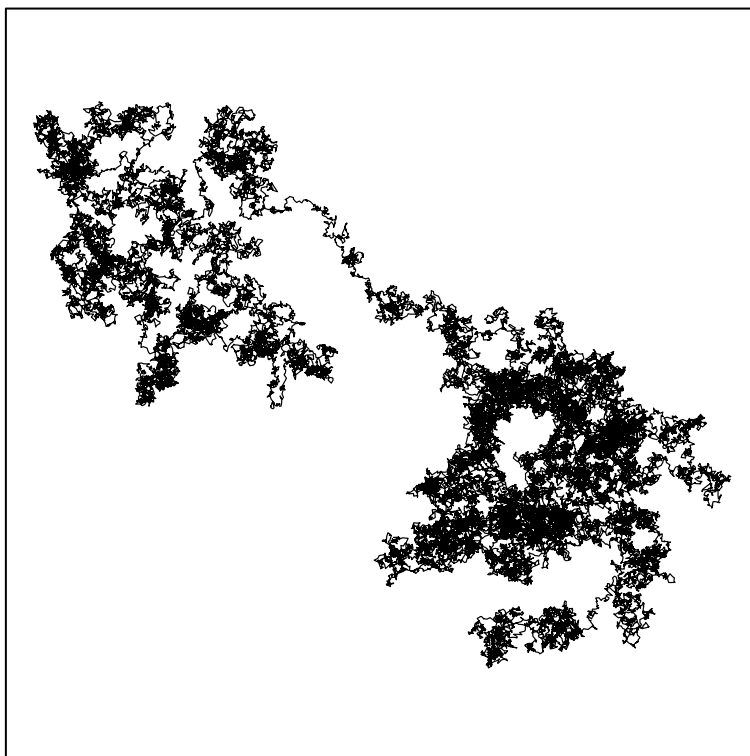


Moto browniano e analisi stocastica

FRANCESCO CARAVENNA
francesco.caravenna@unimib.it

<http://www.matapp.unimib.it/~fcaraven>

DIPARTIMENTO DI MATEMATICA E APPLICAZIONI
UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA
VIA COZZI 53, 20125 MILANO, ITALY



Versione 3.3

Ultima modifica: 9 giugno 2011.

SOMMARIO. Queste note nascono come dispense dei corsi che ho tenuto per la laurea magistrale in matematica presso le università di Padova (corso di “Analisi stocastica”, anni accademici 2008/09, 2009/10, 2010/11) e di Milano-Bicocca (corso di “Processi stocastici”, anno accademico 2010/11). Le parti in corpo minore indicano tipicamente argomenti non svolti a lezione, o soltanto accennati.

Le principali fonti di ispirazione sono stati i testi di Comets e Meyre [2006], Baldi [2000], Karatzas e Shreve [1998], Mörters e Peres [2009] e Le Gall [2008].

Segnalazioni di errori, osservazioni, suggerimenti e critiche sono molto graditi.

In copertina è raffigurato un segmento di traiettoria del moto browniano nel piano, ottenuto come interpolazione lineare di una passeggiata aleatoria con incrementi normali (50 000 passi).

INDICE

| | |
|---|-----------|
| Preludio | 1 |
| Capitolo 1. Richiami di probabilità | 5 |
| 1.1 Spazi misurabili | 5 |
| 1.2 Probabilità e variabili aleatorie | 7 |
| 1.3 Legge di una variabile aleatoria | 12 |
| 1.4 Indipendenza e misura prodotto | 13 |
| 1.5 Nozioni di convergenza | 16 |
| 1.6 Funzioni caratteristiche | 19 |
| 1.7 Leggi normali | 19 |
| 1.8 Questioni tecniche | 23 |
| Capitolo 2. Moto browniano | 25 |
| 2.1 Processi stocastici | 25 |
| 2.2 Il moto browniano | 27 |
| 2.3 Esistenza del moto browniano | 34 |
| 2.4 (Ir)regolarità delle traiettorie | 41 |
| 2.5 Processi e σ -algebre | 46 |
| 2.6 Moto browniano multidimensionale | 49 |
| 2.7 La misura di Wiener | 50 |
| Capitolo 3. Processi di Lévy e proprietà di Markov | 55 |
| 3.1 Processi stocastici e filtrazioni | 56 |
| 3.2 Processi di Lévy rispetto a una filtrazione | 61 |
| 3.3 Tempi d'arresto | 64 |
| 3.4 La proprietà di Markov forte | 67 |
| 3.5 Il principio di riflessione | 69 |
| Capitolo 4. Martingale | 71 |
| 4.1 Speranza condizionale | 71 |
| 4.2 Martingale a tempo discreto e continuo | 73 |
| 4.3 Variazione quadratica di una martingala continua | 79 |
| Capitolo 5. Integrale stocastico | 83 |
| 5.1 Prolungamento di isometrie | 83 |
| 5.2 L'integrale stocastico in $M^2[a, b]$ | 86 |
| 5.3 L'integrale stocastico come processo | 97 |
| 5.4 L'integrale stocastico in $M^2_{\text{loc}}[a, b]$ | 102 |

| | | |
|--|--|------------|
| 5.5 | Gli spazi \mathcal{M}^2 e $\mathcal{M}_{\text{loc}}^2$ | 108 |
| Capitolo 6. Calcolo stocastico e applicazioni | | 111 |
| 6.1 | Formula di Itô per il moto browniano | 111 |
| 6.2 | Processi di Itô e formula di Itô generale | 115 |
| 6.3 | Qualche esempio | 118 |
| 6.4 | Il caso multidimensionale | 120 |
| 6.5 | Moto browniano e laplaciano | 123 |
| 6.6 | Il teorema di Girsanov | 128 |
| Capitolo 7. Equazioni differenziali stocastiche | | 135 |
| 7.1 | Definizioni | 135 |
| 7.2 | Esistenza e unicità di soluzioni forti | 137 |
| 7.3 | La formula di Feynman-Kac | 142 |
| Capitolo 8. Rimorsi | | 147 |
| 8.1 | Le diffusioni come processi di Markov | 147 |
| 8.2 | L'integrale di Stratonovich | 149 |
| 8.3 | Teoremi di rappresentazione per martingale | 149 |

PRELUDIO

Questo corso è costituito da due parti. Nella prima parte definiamo il *moto browniano*, uno degli oggetti più importanti in teoria della probabilità, e ne studiamo le principali proprietà. Nella seconda parte mostriamo come sia possibile definire una teoria dell'integrazione rispetto al moto browniano, detta *integrazione stocastica*, che riveste un'importanza fondamentale da un punto di vista sia teorico sia applicativo. Prima di cominciare, cerchiamo di dare un'idea euristica del contenuto del corso.

MOTO BROWNIANO. Citando liberamente Mörters e Peres [2009], una parte importante della teoria della probabilità consiste nella descrizione del *quadro macroscopico* che emerge in sistemi caratterizzati da *effetti microscopici aleatori*. Da questo punto di vista, il moto browniano può essere visto come il processo macroscopico che descrive il moto di una particella sottoposta a urti microscopici casuali frequenti ma non troppo intensi.

Più precisamente, consideriamo il seguente modello microscopico per una particella vincolata a muoversi su una retta e sottoposta a urti frequenti: indicando con $\varepsilon > 0$ la distanza temporale (che supponiamo per semplicità costante) tra una coppia di urti successivi e con $z_i \in \mathbb{R}$ lo spostamento della particella provocato dall' i -esimo urto, la posizione $s(t) \in \mathbb{R}$ della particella all'istante $t > 0$ è data da

$$s(t) = s(0) + \sum_{i=0}^{\lfloor t/\varepsilon \rfloor} z_i,$$

dove $s(0) \in \mathbb{R}$ indica la posizione della particella all'istante iniziale. Supponiamo ora che gli spostamenti $(z_i)_{i \in \mathbb{N}}$ siano variabili aleatorie indipendenti e con la stessa distribuzione, con media nulla e con *varianza finita*. Sotto queste sole ipotesi, indipendentemente dai dettagli più fini della distribuzione delle variabili z_i , quando $\varepsilon > 0$ è piccolo il processo riscalato $\{\sqrt{\varepsilon} s(t)\}_{t \in [0, \infty)}$ è descritto da un *unico processo macroscopico* $\{B(t)\}_{t \in [0, \infty)}$, il moto browniano appunto. Il fatto che molteplici modelli microscopici (corrispondenti per esempio alle possibili scelte della distribuzione degli spostamenti z_i) diano origine allo stesso unico modello macroscopico è una proprietà molto interessante, che viene spesso indicata col nome di *universalità*.

Il moto browniano $\{B(t)\}_{t \in [0, \infty)}$ è un oggetto matematico estremamente ricco, di cui analizzeremo un certo numero di proprietà. Mostreremo ad esempio che le sue traiettorie $t \mapsto B(t)$ sono continue ma molto irregolari: non sono differenziabili in nessun punto (!) e hanno *variazione infinita*, nel senso che per ogni $T > 0$ si ha (con probabilità 1)

$$\sup_{N \in \mathbb{N}} \sum_{i=0}^{\lfloor NT \rfloor - 1} \left| B\left(\frac{i+1}{N}\right) - B\left(\frac{i}{N}\right) \right| = +\infty.$$

Questa irregolarità rende problematica la definizione di una teoria dell'integrazione rispetto alle traiettorie del moto browniano. Prima di discutere questo punto, forniamo qualche motivazione per la necessità di una tale teoria.

INTEGRALE STOCASTICO. Consideriamo una particella vincolata a muoversi lungo una retta su cui agisce un campo di forze. La posizione della particella all'istante $t \in [0, \infty)$ sarà indicata con $x(t) \in \mathbb{R}$, mentre $F(t, x)$ indicherà l'intensità della forza all'istante $t \geq 0$ nel punto $x \in \mathbb{R}$. Se la funzione $x(\cdot)$ è derivabile, il lavoro compiuto dalla forza sulla particella nell'intervallo di tempo $[0, T]$ è dato dall'integrale

$$\int_0^T F(t, x(t)) dx(t) := \int_0^T F(t, x(t)) x'(t) dt, \quad (0.1)$$

dove $x'(t) := \frac{dx(t)}{dt}$ indica la derivata prima di x . Lo stesso integrale ammette altre possibili interpretazioni. Supponiamo ad esempio che $x(t)$ indichi il valore di un certo titolo azionario all'istante t e che $F(t, x)$ rappresenti la mia strategia di investimento, ossia la quantità di azioni che possiedo all'istante t se il titolo vale x . L'integrale in (0.1) rappresenta allora il guadagno (con segno!) che ottengo delle azioni in mio possesso tra gli istanti 0 e T .

Questi argomenti spiegano l'interesse per lo studio e la generalizzazione di integrali del tipo (0.1). Consideriamo per esempio il caso in cui $x(\cdot)$ non sia derivabile: è possibile dare un senso all'integrale $\int_0^T F(t, x(t)) dx(t)$ anche in questa circostanza? Supponiamo per semplicità che la funzione $F(t, x) = F(t)$ non dipenda esplicitamente da x e che sia *continua*. Quando $x(\cdot)$ è derivabile con continuità, l'integrale $\int F(t) dx(t)$ definito in (0.1) può essere espresso come limite di somme di Riemann, nel modo seguente:

$$\int_0^T F(t) dx(t) = \lim_{N \rightarrow \infty} \sum_{i=0}^{\lfloor NT \rfloor - 1} F\left(\frac{i}{N}\right) \left(x\left(\frac{i+1}{N}\right) - x\left(\frac{i}{N}\right)\right). \quad (0.2)$$

Dato che in questa espressione non compare la derivata prima di $x(\cdot)$, si potrebbe sperare di usarla come *definizione* dell'integrale $\int_0^T F(t) dx(t)$ quando la funzione $x(\cdot)$ non è derivabile. Ebbene, è possibile mostrare (Teorema 56 del capitolo I in [Protter, 2004]) che data una funzione $x(\cdot)$ continua, o anche solo continua a destra, il limite in (0.2) esiste *per ogni* $F(\cdot)$ continua se e soltanto se $x(\cdot)$ è a variazione finita, cioè

$$\sup_{N \in \mathbb{N}} \sum_{i=0}^{\lfloor NT \rfloor - 1} \left| x\left(\frac{i+1}{N}\right) - x\left(\frac{i}{N}\right) \right| < \infty.$$

Il problema è che si vorrebbe estendere la definizione dell'integrale $\int_0^T F(t) dx(t)$ al caso in cui la funzione (aleatoria) $x(t) = B(t)$ è il moto browniano, che come abbiamo accennato più sopra ha variazione *infinita*. Un risultato fondamentale di Itô [1944] mostra che una tale estensione è effettivamente possibile, a patto di adottare un approccio probabilistico: l'*integrale stocastico* non è infatti definito puntualmente, per ogni traiettoria fissata

$t \mapsto B(t)$, ma solo come limite in probabilità; inoltre gli integrandi $F(t)$ ammessi, che possono essere essi stessi aleatori, sono soggetti a importanti restrizioni (devono essere *adattati* al moto browniano, in un senso che preciseremo). *Se le traiettorie $t \mapsto F(t)$ sono continue*, l'integrale stocastico può essere definito come limite *in probabilità* di somme di Riemann, analogamente a (0.2):

$$\int_0^T F(t) dB(t) := \lim_{N \rightarrow \infty} \sum_{i=0}^{\lfloor NT \rfloor - 1} F\left(\frac{i}{N}\right) \left(B\left(\frac{i+1}{N}\right) - B\left(\frac{i}{N}\right)\right). \quad (0.3)$$

Le proprietà di questo integrale sono peculiari. Per esempio, se nella somma di Riemann in (0.3) si sostituisce $F(\frac{i}{N})$ con $F(\frac{i+1}{N})$ — un cambio apparentemente innocuo, data la continuità di $t \mapsto F(t)$ — il valore dell'integrale risulta in generale diverso, a differenza di quanto accade per l'integrale ordinario. È importante dunque precisare quale definizione si usa. Noi ci concentreremo sull'*integrale di Itô*, definito proprio come in (0.3), che è quello più naturale per le applicazioni finanziarie e ha inoltre fondamentali legami con la teoria delle equazioni differenziali alle derivate parziali del secondo ordine. Altre definizioni risultano più convenienti in contesti diversi: accenneremo brevemente all'*integrale di Stratonovich*, usato nelle applicazioni fisiche e nella geometria stocastica, in cui il termine $F(\frac{i}{N})$ è sostituito da $\frac{1}{2}(F(\frac{i}{N}) + F(\frac{i+1}{N}))$.

Se $t \mapsto x(t)$ è una funzione derivabile, segue dal teorema fondamentale del calcolo che per ogni funzione $G : \mathbb{R} \rightarrow \mathbb{R}$ derivabile si ha

$$G(x(T)) - G(x(0)) = \int_0^T \frac{d}{dt} G(x(t)) dt =: \int_0^T G'(x(t)) dx(t).$$

Per l'integrale di Itô questa relazione non è più valida. Più precisamente, se $x(t) = B(t)$ è il moto browniano e se $G : \mathbb{R} \rightarrow \mathbb{R}$ è una funzione derivabile *due volte* con continuità, mostreremo che

$$G(B(T)) - G(B(0)) = \int_0^T G'(B(t)) dB(t) + \frac{1}{2} \int_0^T G''(B(t)) dt,$$

che può essere riscritta in forma differenziale come

$$dG(B(t)) = G'(B(t)) dB(t) + \frac{1}{2} G''(B(t)) dt.$$

Questa è la celebre *formula di Itô*, che costituisce il cuore del *calcolo stocastico*. Come vedremo, la presenza del termine aggiuntivo contenente G'' è dovuta precisamente al fatto che le traiettorie del moto browniano sono a variazione infinita.

NOTAZIONI. Indicheremo con $\mathbb{N} := \{1, 2, 3, \dots\}$ e con $\mathbb{N}_0 := \mathbb{N} \cup \{0\} = \{0, 1, 2, \dots\}$. Useremo gli aggettivi “crescente”, “decrescente”, “positivo” e “negativo” in senso debole: una funzione $f : \mathbb{R} \rightarrow \mathbb{R}$ è crescente se per ogni $x, y \in \mathbb{R}$ con $x < y$ si ha $f(y) \geq f(x)$, mentre è strettamente crescente se si ha $f(y) > f(x)$; analogamente, f è positiva se per ogni $x \in \mathbb{R}$ si ha $f(x) \geq 0$, mentre è strettamente positiva se si ha $f(x) > 0$. Dati $a, b \in \mathbb{R}$, useremo spesso le notazioni $a \wedge b := \min\{a, b\}$ e $a \vee b := \max\{a, b\}$.

1. RICHIAMI DI PROBABILITÀ

Forniamo un compendio delle nozioni basilari di probabilità che ci saranno utili. Per maggiori dettagli, si possono consultare i testi [Billingsley, 1995], [Williams, 1991].

1.1. SPAZI MISURABILI

1.1.1. σ -ALGEBRE. Una σ -algebra \mathcal{E} su un insieme non vuoto E è una famiglia di parti (sottoinsiemi) di E che contiene E e che sia chiusa per unioni numerabili e passaggio al complementare. La coppia (E, \mathcal{E}) è detta *spazio misurabile*. I due casi estremi di σ -algebre sono quella banale $\mathcal{E} = \{\emptyset, E\}$ e quella discreta $\mathcal{E} = \mathcal{P}(E)$.

Data una famiglia non vuota $I \subseteq \mathcal{P}(E)$ di parti di E , si indica con $\sigma(I)$ la più piccola σ -algebra che contenga I , cioè l'intersezione di tutte le σ -algebre che contengono I .[†] La famiglia I è detta un *generatore* di $\sigma(I)$. Se inoltre I è chiusa per intersezioni finite, I è detta una *base* di $\sigma(I)$. Date due famiglie $I, I' \subseteq \mathcal{P}(E)$, la più piccola σ -algebra che contiene sia I sia I' è $\sigma(I \cup I')$, che indicheremo con semplicità con $\sigma(I, I')$; analogamente, nel caso si abbiano più famiglie scriviamo $\sigma(I_j, j \in J)$ o $\sigma(\{I_j\}_{j \in J})$ invece di $\sigma(\bigcup_{j \in J} I_j)$.

ESEMPIO 1.1. Per $A \subseteq E$ si ha $\sigma(\{A\}) = \{\emptyset, A, A^c, E\}$. Più in generale, se $\{A_i\}_{i \in I}$ è una *partizione finita o numerabile* di E (cioè $\bigcup_{i \in I} A_i = E$ e $A_i \cap A_j = \emptyset$ per $i \neq j$, dove I è un insieme finito o numerabile), allora $\sigma(\{A_i\}_{i \in I}) = \{A = \bigcup_{j \in J} A_j, J \subseteq I\}$. \square

Se (E, τ) è uno spazio topologico, si dice *boreliana* la σ -algebra su E generata dagli insiemi aperti (equivalentemente, dagli insiemi chiusi), che indicheremo con $\mathcal{B}(E)$. L'esempio più importante è dato da $E = \mathbb{R}^d$, la cui σ -algebra boreliana $\mathcal{B}(\mathbb{R}^d)$ è generata dai *rettangoli* aperti $(a_1, b_1) \times \cdots \times (a_d, b_d)$, che ne costituiscono una base. Un altro esempio importante è la retta reale estesa $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$, in cui $\mathcal{B}(\overline{\mathbb{R}}) = \mathcal{B}(\mathbb{R}) \cup \{+\infty\} \cup \{-\infty\}$. Intenderemo sempre \mathbb{R}^d e $\overline{\mathbb{R}}$ come spazi misurabili, muniti della σ -algebra boreliana.

A differenza di quanto accade per la topologia, in generale non è possibile descrivere *esplicitamente* gli elementi della σ -algebra $\sigma(I)$ generata da una famiglia I . Per questa ragione, è necessario lavorare con generatori e basi (si veda il paragrafo 1.8.1).

1.1.2. APPLICAZIONI MISURABILI. Sia $X : E \rightarrow F$ una applicazione tra due spazi misurabili (E, \mathcal{E}) e (F, \mathcal{F}) (scriveremo sinteticamente $X : (E, \mathcal{E}) \rightarrow (F, \mathcal{F})$). L'applicazione X è detta *misurabile* se $X^{-1}(B) \in \mathcal{E}$ per ogni $B \in \mathcal{F}$. Se J è un generatore di \mathcal{F} , cioè se $\mathcal{F} = \sigma(J)$, è sufficiente richiedere che $X^{-1}(B) \in \mathcal{E}$ per ogni $B \in J$.

Data una qualunque applicazione $X : E \rightarrow F$ e una qualunque σ -algebra \mathcal{F} su F , è automaticamente definita la σ -algebra $\sigma(X) := \{X^{-1}(B) : B \in \mathcal{F}\}$ su E , detta σ -algebra *generata da X* : si tratta della più piccola σ -algebra su E che renda X misurabile.

[†]Si noti che l'intersezione di una famiglia *arbitraria* di σ -algebre è una σ -algebra, mentre in generale l'unione (anche finita) di σ -algebre non lo è.

Date due applicazioni X, Y definite entrambe su E , a valori in spazi misurabili (anche diversi), indicheremo con $\sigma(X, Y) := \sigma(\sigma(X), \sigma(Y))$ la σ -algebra da esse generata: si tratta della più piccola σ -algebra su E che rende misurabili sia X sia Y . Per una famiglia $\{X_j\}_{j \in J}$ di applicazioni, scriveremo analogamente $\sigma(\{X_j\}_{j \in J})$ invece di $\sigma(\bigcup_{j \in J} \sigma(X_j))$.

Si verifica facilmente che $\sigma(X)$ rende misurabile qualunque funzione di X : per ogni $g : (F, \mathcal{F}) \rightarrow (G, \mathcal{G})$ misurabile, la composizione $g(X) = g \circ X$ è $\sigma(X)$ -misurabile, cioè è misurabile come applicazione da $(E, \sigma(X))$ in (G, \mathcal{G}) . È interessante notare che vale un parziale viceversa, noto come lemma di misurabilità di Doob: se $X : E \rightarrow (F, \mathcal{F})$ è una applicazione generica e $Y : E \rightarrow \mathbb{R}^d$ è $\sigma(X)$ -misurabile, allora esiste $g : (F, \mathcal{F}) \rightarrow \mathbb{R}^d$ misurabile tale che $Y = g(X)$ (si veda il Lemma 1.18 alla fine del capitolo).

Se $X : (E, \mathcal{E}) \rightarrow (F, \mathcal{F})$ e $Y : (F, \mathcal{F}) \rightarrow (G, \mathcal{G})$ sono applicazioni misurabili, lo è la loro composizione $Y \circ X : (E, \mathcal{E}) \rightarrow (G, \mathcal{G})$. Se E, F sono spazi topologici, ogni applicazione $X : E \rightarrow F$ continua è misurabile rispetto alle σ -algebre boreliane $\mathcal{B}(E), \mathcal{B}(F)$. Segue che se $X, Y : (E, \mathcal{B}(E)) \rightarrow \mathbb{R}$ sono applicazioni misurabili, lo sono anche $X + Y, X \cdot Y, |X|, X^+ := \max(X, 0)$, ecc. Inoltre sono misurabili le applicazioni (a valori in $\overline{\mathbb{R}}$)

$$\sup_{n \in \mathbb{N}} X_n, \quad \inf_{n \in \mathbb{N}} X_n, \quad \limsup_{n \in \mathbb{N}} X_n, \quad \liminf_{n \in \mathbb{N}} X_n, \quad \sum_{n \in \mathbb{N}} |X_n|,$$

purché $X_n : (E, \mathcal{B}(E)) \rightarrow \mathbb{R}$ sia misurabile per ogni $n \in \mathbb{N}$. Sottolineiamo che per questi risultati è fondamentale che la famiglia $\{X_n\}_n$ sia (al più) numerabile.

La funzione indicatrice 1_B di un insieme B è definita da $1_B(x) = 1$ se $x \in B$ mentre $1_B(x) = 0$ se $x \notin B$. Chiaramente $1_B : (E, \mathcal{E}) \rightarrow \mathbb{R}$ è misurabile se e solo se $B \in \mathcal{E}$.

Un'applicazione reale $X : (E, \mathcal{E}) \rightarrow \mathbb{R}$ si dice *semplice* se si può scrivere $X = \sum_{i=1}^n c_i 1_{B_i}$ con $n \in \mathbb{N}$, $c_i \in \mathbb{R}$ e $B_i \in \mathcal{E}$ per $i = 1, \dots, n$. Ogni applicazione misurabile positiva $X : (E, \mathcal{E}) \rightarrow \mathbb{R}^+$ si può scrivere come limite puntuale crescente di funzioni semplici: $X(x) = \lim_{n \rightarrow \infty} X_n(x)$ per ogni $x \in E$, dove $X_n(x) := \max\{2^{-n} \lfloor 2^n X(x) \rfloor, 0\}$. Più esplicitamente:

$$X_n(x) := n \mathbf{1}_{\{x \in E: X(x) \geq n\}} + \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \mathbf{1}_{\{x \in E: \frac{k-1}{2^n} \leq X(x) < \frac{k}{2^n}\}}(x),$$

da cui è chiaro che, per ogni $x \in E$, si ha $X_n(x) \leq X_{n+1}(x)$ per ogni $n \in \mathbb{N}$ e $X_n(x) \rightarrow X(x)$.

1.1.3. SPAZI PRODOTTO. Dati due spazi misurabili $(F, \mathcal{F}), (G, \mathcal{G})$, sul loro prodotto cartesiano $F \times G$ si definisce la σ -algebra prodotto $\mathcal{F} \otimes \mathcal{G} := \sigma(\mathcal{F} \times \mathcal{G})$, cioè la σ -algebra generata dalla famiglia $\mathcal{F} \times \mathcal{G} := \{A \times B : A \in \mathcal{F}, B \in \mathcal{G}\}$ (si noti che in generale $\mathcal{F} \times \mathcal{G}$ non è una σ -algebra). Un'applicazione $X : (E, \mathcal{E}) \rightarrow (F \times G, \mathcal{F} \otimes \mathcal{G})$ si può sempre scrivere come $X = (X_1, X_2)$, con X_1, X_2 a valori in F, G rispettivamente, ed è *misurabile se e solo se lo sono le sue componenti* X_1 e X_2 , perché gli insiemi della forma $A \times G$ e $F \times B$ al variare di $A \in \mathcal{F}$ e $B \in \mathcal{G}$ sono un generatore di $\mathcal{F} \otimes \mathcal{G}$. Proprietà del tutto analoghe valgono per il prodotto di un numero finito di spazi misurabili. Il caso di una famiglia infinita di spazi misurabili sarà considerato nel paragrafo 2.1.1.

Sia $f : (F \times G, \mathcal{F} \otimes \mathcal{G}) \rightarrow \mathbb{R}$ una funzione reale misurabile; allora, per ogni $x \in F$ fissato, la funzione $y \mapsto f(x, y)$ è misurabile da (G, \mathcal{G}) in \mathbb{R} ; analogamente, per ogni $y \in G$ fissato, la funzione $x \mapsto f(x, y)$ è misurabile da (F, \mathcal{F}) in \mathbb{R} . Sottolineiamo che non vale il viceversa: per la misurabilità dell'applicazione f non è sufficiente che siano misurabili le applicazioni $x \mapsto f(x, y)$ e $y \mapsto f(x, y)$.

1.2. PROBABILITÀ E VARIABILI ALEATORIE

1.2.1. MISURE E PROBABILITÀ. Dato uno spazio misurabile (E, \mathcal{E}) , una misura μ è una funzione $\mu : \mathcal{E} \rightarrow [0, +\infty]$ tale che $\mu(\emptyset) = 0$ e con la proprietà di σ -additività, cioè $\mu(\bigcup_{n \in \mathbb{N}} A_n) = \sum_{n \in \mathbb{N}} \mu(A_n)$ per ogni successione $\{A_n\}_{n \in \mathbb{N}}$ di elementi di \mathcal{E} a due a due disgiunti ($A_n \cap A_m = \emptyset$ per $m \neq n$). La terna (E, \mathcal{E}, μ) è detta *spazio di misura* o *spazio misurato*. La misura μ è detta finita se $\mu(E) < \infty$ e σ -finita se si può scrivere $E = \bigcup_{n \in \mathbb{N}} A_n$ con $A_n \in \mathcal{E}$ e $\mu(A_n) < \infty$ per ogni $n \in \mathbb{N}$.

Un esempio classico di misura finita è dato dalla *delta di Dirac* δ_{x_0} , dove $x_0 \in E$ è un qualunque punto fissato, definita da $\delta_{x_0}(A) = 1$ se $x_0 \in A$ e $\delta_{x_0}(A) = 0$ altrimenti. L'esempio più importante di misura σ -finita è dato dalla misura di Lebesgue su $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$.

Una misura P su uno spazio misurabile (Ω, \mathcal{F}) tale che $P(\Omega) = 1$ è detta *probabilità* (o misura di probabilità o anche *legge*). La terna (Ω, \mathcal{F}, P) è detta *spazio di probabilità* e gli elementi di \mathcal{F} sono detti *eventi*. Un evento $A \in \mathcal{F}$ si dice *quasi certo* se $P(A) = 1$. Ricordiamo alcune importanti proprietà che ci saranno utili.

- Per ogni evento A si ha $P(A^c) = 1 - P(A)$.
- Per ogni coppia di eventi $A \subseteq B$ si ha $P(A) \leq P(B)$ [*monotonia*].
- Per ogni successione di eventi $\{A_n\}_{n \in \mathbb{N}}$ vale la relazione $P(\bigcup_{n \in \mathbb{N}} A_n) \leq \sum_{n \in \mathbb{N}} P(A_n)$ [*subadditività*].
- Per ogni successione di eventi $\{A_n\}_{n \in \mathbb{N}}$ crescente (risp. decrescente), cioè tale che $A_n \subseteq A_{n+1}$ (risp. $A_n \supseteq A_{n+1}$) per ogni $n \in \mathbb{N}$, indicando l'evento limite con $A = \bigcup_{n \in \mathbb{N}} A_n$ (risp. $A = \bigcap_{n \in \mathbb{N}} A_n$), si ha che $P(A_n) \uparrow P(A)$ (risp. $P(A_n) \downarrow P(A)$) [*continuità dal basso e dall'alto*].
- Se $\{A_n\}_{n \in \mathbb{N}}$ è una famiglia di eventi quasi certi, cioè $P(A_n) = 1$ per ogni $n \in \mathbb{N}$, anche $\bigcap_{n \in \mathbb{N}} A_n$ è un evento quasi certo.
- Se $\{B_n\}_{n \in \mathbb{N}}$ è q.c. una partizione dello spazio di probabilità, cioè se $P(\bigcup_{n \in \mathbb{N}} B_n) = 1$ e $P(B_n \cap B_m) = 0$ per $m \neq n$, vale la relazione $P(A) = \sum_{n \in \mathbb{N}} P(A \cap B_n)$, per ogni evento A [*formula di disintegrazione*].

Un importante risultato di unicità è il seguente: se I è una base di \mathcal{F} (cioè $\mathcal{F} = \sigma(I)$ e I è chiusa per intersezioni finite), due probabilità P, P' su (Ω, \mathcal{F}) che coincidono su I sono necessariamente uguali, cioè $P(A) = P'(A)$ per ogni $A \in \mathcal{F}$ (questo segue dal Lemma di Dynkin, cf. il paragrafo 1.8).

Un risultato analogo vale per misure σ -finite: più precisamente, se μ, μ' sono misure sullo spazio misurabile (E, \mathcal{E}) che coincidono su una base I di \mathcal{E} , e se si può scrivere $E = \bigcup_{n \in \mathbb{N}} A_n$ con $A_n \in I$ e $\mu(A_n) = \mu'(A_n) < \infty$ per ogni $n \in \mathbb{N}$, allora le misure μ, μ' coincidono su tutta la σ -algebra \mathcal{E} .

1.2.2. VARIABILI ALEATORIE. Consideriamo una applicazione $X : \Omega \rightarrow E$, dove (Ω, \mathcal{F}, P) è uno spazio di probabilità e (E, \mathcal{E}) è uno spazio misurabile (scriveremo sinteticamente $X : (\Omega, \mathcal{F}, P) \rightarrow (E, \mathcal{E})$). Se l'applicazione X è misurabile, essa è detta *variabile aleatoria*. Nel caso in cui lo spazio di arrivo E coincida con \mathbb{R} o con \mathbb{R}^d , si parla rispettivamente di variabile aleatoria reale (detta anche scalare) o di vettore aleatorio.

Data una variabile aleatoria $X : (\Omega, \mathcal{F}, P) \rightarrow (E, \mathcal{E})$ e un insieme $A \in \mathcal{E}$, è consuetudine indicare con $\{X \in A\}$ l'evento “ X assume valori in A ”, cioè

$$\{X \in A\} := X^{-1}(A) = \{\omega \in \Omega : X(\omega) \in A\}.$$

Analogamente, per una variabile aleatoria reale X si pone $\{X \geq a\} := \{X \in [a, \infty)\} = X^{-1}([a, \infty))$, ecc. *Useremo queste notazioni costantemente.* Osserviamo che $1_{\{X \in A\}} = 1_A \circ X$ per $A \in \mathcal{E}$ (si noti che $1_{\{X \in A\}}$ è definita su Ω mentre 1_A è definita su E).

Per definizione, la σ -algebra $\sigma(X)$ generata da X consiste esattamente degli eventi della forma $\{X \in A\}$ al variare di $A \in \mathcal{E}$. Intuitivamente, $\sigma(X)$ codifica l'*informazione associata alla variabile aleatoria X* : questa σ -algebra consta infatti degli *eventi che si possono esprimere tramite X* , ossia gli eventi per cui si può stabilire se si siano verificati oppure no conoscendo solo il valore assunto dalla variabile aleatoria X .

1.2.3. SPAZI DI PROBABILITÀ COMPLETI. Sebbene il contenuto di questo paragrafo si possa applicare a spazi di misura generali, considereremo per semplicità solo il caso degli spazi di probabilità. Uno spazio di probabilità (Ω, \mathcal{F}, P) si dice *completo* se, per ogni evento $C \in \mathcal{F}$ tale che $P(C) = 0$, si ha che ogni sottoinsieme $N \subseteq C$ è misurabile, cioè $N \in \mathcal{F}$ (in particolare segue che $P(N) = 0$). Equivalentemente, (Ω, \mathcal{F}, P) è *completo* se, per ogni evento $A \in \mathcal{F}$ tale che $P(A) = 1$, si ha che ogni insieme $B \supseteq A$ è misurabile, cioè $B \in \mathcal{F}$ (in particolare segue che $P(B) = 1$).

Se (Ω, \mathcal{F}, P) non è completo, è sempre possibile completarlo: più precisamente, si può costruire uno spazio di probabilità completo $(\Omega, \overline{\mathcal{F}}, \overline{P})$ tale che $\overline{\mathcal{F}} \supseteq \mathcal{F}$ e \overline{P} coincida con P su \mathcal{F} . Si definisce innanzitutto la famiglia \mathcal{N} degli insiemi *trascurabili* (o *P-trascurabili*, se si vuole enfatizzare la probabilità) ponendo

$$\mathcal{N} := \{N \subseteq \Omega : \exists C \in \mathcal{F} \text{ tale che } N \subseteq C \text{ e } P(C) = 0\}$$

e si estende la σ -algebra \mathcal{F} ponendo $\overline{\mathcal{F}} := \sigma(\mathcal{F}, \mathcal{N})$. Si può mostrare (esercizio) che $A \in \overline{\mathcal{F}}$ se e soltanto se esiste $A' \in \mathcal{F}$ tale che $A \triangle A' \in \mathcal{N}$, dove $A \triangle A' := (A \setminus A') \cup (A' \setminus A)$ indica la differenza simmetrica, e in questo caso si definisce $\overline{P}(A) := P(A')$. Si verifica che tale definizione è ben posta, cioè non dipende dalla scelta di A' , che \overline{P} definisce una probabilità su $\overline{\mathcal{F}}$ e che $(\Omega, \overline{\mathcal{F}}, \overline{P})$ è uno spazio di probabilità completo (esercizio).

Gli insiemi di $\overline{\mathcal{F}}$ di probabilità \overline{P} nulla sono esattamente gli elementi di \mathcal{N} . Una applicazione $X : (\Omega, \overline{\mathcal{F}}) \rightarrow (E, \mathcal{E})$ è misurabile se e solo se è \overline{P} -q.c. uguale a una applicazione misurabile $X' : (\Omega, \mathcal{F}) \rightarrow (E, \mathcal{E})$, cioè se e solo se vale che $\{\omega \in \Omega : X'(\omega) \neq X(\omega)\} \in \mathcal{N}$.

Sottolineiamo che il completamento $\overline{\mathcal{F}}$ della σ -algebra \mathcal{F} dipende fortemente dalla probabilità P . Per esempio, se sullo spazio misurabile $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ si sceglie la probabilità di Dirac δ_0 , la σ -algebra completata è l'insieme delle parti $\mathcal{P}(\mathbb{R})$, come si verifica facilmente.

OSSERVAZIONE 1.2 (IMPORTANTE). Dato uno spazio di probabilità (Ω, \mathcal{F}, P) , indichiamo genericamente con “[...]” una proprietà dipendente da $\omega \in \Omega$ (ad esempio $Y(\omega) \geq 0$, oppure $\limsup_{t \rightarrow \infty} X_t(\omega) = +\infty$, dove $Y, \{X_t\}_{t \in \mathbb{R}}$ sono funzioni reali definite su Ω). Introduciamo la notazione, di uso molto frequente,

“quasi certamente [...]” (abbreviato “q.c. [...]”)

intendendo con ciò

“esiste $A \in \mathcal{F}$ con $P(A) = 1$ tale che per ogni $\omega \in A$ [...]”.

Si potrebbe pensare che ciò sia equivalente a richiedere che $P(\{\omega \in \Omega : [\dots]\}) = 1$, ma questo non è corretto: infatti in generale non è detto che l'insieme $\{\omega \in \Omega : [\dots]\}$ sia un evento, cioè appartenga a \mathcal{F} . Scrivendo “q.c. [...]” si afferma soltanto che $\{\omega \in \Omega : [\dots]\}$ *contiene* un evento di probabilità 1. Si tratta tutto sommato di una sottigliezza, che si risolve immediatamente se lo spazio (Ω, \mathcal{F}, P) è completo: infatti in questo caso gli insiemi che contengono eventi di probabilità 1 sono automaticamente misurabili. Questa è una delle ragioni per cui risulta spesso conveniente lavorare con spazi completi.[†] \square

1.2.4. INTEGRALE E VALORE ATTESO. Dato uno spazio di misura (E, \mathcal{E}, μ) e una funzione misurabile *positiva* $g : E \rightarrow \mathbb{R}^+$, è sempre ben definito l'integrale $\int g d\mu = \int g(x) \mu(dx) \in [0, +\infty]$.[‡] Ricordiamo una proprietà importante:

$$\text{se } g \geq 0, \quad \int g d\mu = 0 \quad \text{se e solo se} \quad \mu(g > 0) = 0, \text{ cioè } g = 0 \text{ } \mu\text{-q.c.}$$

Un'arbitraria funzione misurabile reale g è detta *integrabile* se $\int |g| d\mu < \infty$ e in questo caso si definisce $\int g d\mu := \int g^+ d\mu - \int g^- d\mu$, dove si è posto $g^\pm(x) := \max\{\pm g(x), 0\}$, da cui $g = g^+ - g^-$. Si ha che $|\int g d\mu| \leq \int |g| d\mu$.

Specializzando queste definizioni al caso di uno spazio di probabilità (Ω, \mathcal{F}, P) , per una variabile aleatoria reale *positiva* $Y : (\Omega, \mathcal{F}, P) \rightarrow \mathbb{R}^+$ è sempre ben definito il valore atteso $E(Y) := \int Y(\omega) P(d\omega) \in [0, +\infty]$ e si ha che

$$E(Y) = 0 \quad \text{se e solo se} \quad P(Y = 0) = 1.$$

Una variabile aleatoria reale X è detta *integrabile* se $E(|X|) < \infty$, nel qual caso si definisce $E(X) := E(X^+) - E(X^-)$ e vale la relazione $|E(X)| \leq E(|X|)$. Più in generale, si pone $E(X) := E(X^+) - E(X^-) \in \mathbb{R}$ anche quando uno solo tra $E(X^+)$ e $E(X^-)$ è finito.

1.2.5. SPAZI L^p . Per ogni variabile aleatoria reale X definita su (Ω, \mathcal{F}, P) si definisce $\|X\|_p := (E(|X|^p))^{1/p} \in [0, +\infty]$, per $p \in [1, \infty)$, e si indica con $L^p = L^p(\Omega) = L^p(\Omega, \mathcal{F}, P)$ l'insieme delle variabili aleatorie reali X tali che $\|X\|_p < \infty$. Vale la disuguaglianza triangolare $\|X + Y\|_p \leq \|X\|_p + \|Y\|_p$ (*disuguaglianza di Minkowski*), da cui segue che L^p è uno spazio vettoriale su \mathbb{R} e $\|\cdot\|_p$ è una *seminorma* su L^p . Infatti $\|X\|_p = 0$ non implica che $X = 0$ ma soltanto che $P(X = 0) = 1$. Introducendo su L^p la relazione di equivalenza $X \sim Y$ se $P(X = Y) = 1$, $\|\cdot\|_p$ diventa una vera norma sull'insieme

[†]Una ragione analoga è la seguente: se $X : (\Omega, \mathcal{F}, P) \rightarrow \mathbb{R}$ è una variabile aleatoria e $X' : \Omega \rightarrow \mathbb{R}$ è una *funzione* tale che $X = X'$ q.c. (cioè esiste $A \in \mathcal{F}$ con $P(A) = 1$ tale che $X(\omega) = X'(\omega)$ per ogni $\omega \in A$), in generale non è detto che X' sia una funzione *misurabile*, dunque una variabile aleatoria; ciò è vero se lo spazio (Ω, \mathcal{F}, P) è completo, per quanto visto.

[‡]Si definisce innanzitutto $\int 1_A d\mu := \mu(A)$ per $A \in \mathcal{E}$ e si estende la definizione per linearità alle funzioni semplici $\sum_{i=1}^n c_i 1_{A_i}$, per $A_i \in \mathcal{F}$ e $c_i \in \mathbb{R}^+$. Per una funzione misurabile non negativa arbitraria, si definisce il valore atteso come il limite dei valori attesi di una qualunque successione crescente di funzioni semplici che converge alla variabile aleatoria (una tale successione esiste sempre, e il limite dei valori attesi non dipende dalla successione scelta).

delle classi di equivalenza L^p/\sim , che con tradizionale abuso di notazione viene indicato sempre con L^p (quando sarà importante distinguere tra funzioni e classi di equivalenza, lo sottolineeremo). In questo modo $(L^p, \|\cdot\|_p)$ diventa uno *spazio di Banach*, cioè uno spazio normato completo: tutte le successioni di Cauchy hanno limite.

Dalla disuguaglianza di Jensen, richiamata più in basso, segue che per ogni $p \geq q$ e per ogni variabile aleatoria X si ha $\|X\|_q \leq \|X\|_p$: di conseguenza $L^p \subseteq L^q$ (questa proprietà non vale se P è una misura non finita) e la convergenza in L^p implica quella in L^q .

Lo spazio più importante è L^2 , che è in effetti uno *spazio di Hilbert*, poiché la norma $\|\cdot\|_2$ è indotta dal prodotto scalare $\langle X, Y \rangle := E(XY)$. Per $X \in L^2$, la quantità $\text{Var}(X) := E[(X - E(X))^2] = E(X^2) - E(X)^2 \in [0, \infty)$ è detta *varianza* di X . Ricordiamo che $\text{Var}(X) = 0$ se e soltanto se esiste $c \in \mathbb{R}$ tale che $P(X = c) = 1$.

Per $X, Y \in L^2$, l'operatore bilineare $\text{Cov}(X, Y) := E[(X - E(X))(Y - E(Y))]$ = $E(XY) - E(X)E(Y) \in \mathbb{R}$ è detto *covarianza*. Si verifica facilmente che

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2 \text{Cov}(X, Y)$$

e inoltre $\text{Var}(X) = \text{Cov}(X, X)$.

Un vettore aleatorio $X = (X_1, \dots, X_d)$ a valori in \mathbb{R}^d è per definizione in L^p se e solo se lo sono tutte le sue componenti. Ponendo $\|X\|_p = (E(|X|^p))^{1/p}$, dove $|\cdot|$ indica la *norma euclidea* su \mathbb{R}^d , si ha che $X \in L^p$ se e solo se $\|X\|_p < \infty$. Inoltre $\|X\|_q \leq \|X\|_p$ se $p \geq q$, quindi $L^p \subseteq L^q$ anche nel caso vettoriale.

Dato un vettore aleatorio $X = (X_1, \dots, X_d)$ a valori in \mathbb{R}^d tale che $X \in L^1$, si definisce il vettore media ponendo $E(X) := (E(X_1), \dots, E(X_d)) \in \mathbb{R}^d$. Se inoltre $X \in L^2$, si definisce la *matrice delle covarianze* $K = K(X)$ data da $K_{ij} := \text{Cov}(X_i, X_j)$, per $1 \leq i, j \leq d$. La matrice K è simmetrica e *semi-definita positiva*: per ogni $u \in \mathbb{R}^d$ si ha $\langle u, Ku \rangle \geq 0$, dove indichiamo con Ku l'ordinario prodotto matrice-vettore, cioè $(Ku)_i = \sum_{j=1}^d K_{ij}u_j$, e dove $\langle x, y \rangle := \sum_{i=1}^d x_i y_i$ denota il prodotto scalare standard per $x, y \in \mathbb{R}^d$. Si ha infatti

$$\begin{aligned} \langle u, Ku \rangle &= \sum_{i=1}^d u_i (Ku)_i = \sum_{i,j=1}^d K_{ij} u_i u_j = \sum_{i,j=1}^d \text{Cov}(X_i, X_j) u_i u_j \\ &= \sum_{i,j=1}^d \text{Cov}(u_i X_i, u_j X_j) = \text{Cov}(\langle u, X \rangle, \langle u, X \rangle) = \text{Var}(\langle u, X \rangle) \geq 0, \end{aligned}$$

avendo usato la bilinearità dell'operatore $\text{Cov}(\cdot, \cdot)$. Vedremo nel paragrafo 1.7 che ogni matrice reale K simmetrica e semi-definita positiva è la matrice delle covarianze di un opportuno vettore aleatorio.

Ricordiamo infine che vettore media e matrice delle covarianze hanno un comportamento semplice sotto trasformazioni affini. Sia $X \in L^2$ un vettore aleatorio a valori in \mathbb{R}^d e $Y := AX + b$ è un vettore aleatorio in \mathbb{R}^m , con A matrice reale $m \times d$ e $b \in \mathbb{R}^m$, cioè $Y_i = \sum_{j=1}^d A_{ij}X_j + b_i$ per ogni $i = 1, \dots, m$. Allora anche $Y \in L^2$ e si verifica facilmente che $E(Y) = AE(X) + b$, $K(Y) = AK(X)A^*$, dove A^* indica la trasposta della matrice A .

1.2.6. TEOREMI DI CONVERGENZA. Richiamiamo di seguito i principali risultati di convergenza (validi in realtà anche per una misura P non di probabilità).

TEOREMA 1.3 (CONVERGENZA MONOTONA). Sia $\{X_n\}_{n \in \mathbb{N}}$ una successione q.c. *crescente* di variabili aleatorie reali *positive*, definite sullo stesso spazio di probabilità (Ω, \mathcal{F}, P) , che converge q.c. verso la variabile aleatoria X ; supponiamo cioè che $0 \leq X_n(\omega) \uparrow X(\omega)$ per q.o. $\omega \in \Omega$. Allora $E(X_n) \uparrow E(X)$.

TEOREMA 1.4 (LEMMA DI FATOU). Sia $\{X_n\}_{n \in \mathbb{N}}$ una successione di variabili aleatorie reali, definite sullo stesso spazio di probabilità (Ω, \mathcal{F}, P) , tali che $X_n(\omega) \geq 0$ per q.o. $\omega \in \Omega$ (o più in generale $X_n(\omega) \geq Y(\omega)$, con Y variabile aleatoria reale integrabile). Allora $E(\liminf_n X_n) \leq \liminf_n E(X_n)$.

TEOREMA 1.5 (CONVERGENZA DOMINATA). Siano $\{X_n\}_{n \in \mathbb{N}}$, X variabili aleatorie reali, definite sullo stesso spazio di probabilità (Ω, \mathcal{F}, P) , tali che $X_n(\omega) \rightarrow X(\omega)$ per q.o. $\omega \in \Omega$. Supponiamo che esista una variabile aleatoria positiva Y *integrabile*, tale cioè che $E(Y) < \infty$, che soddisfi $|X_n(\omega)| \leq Y(\omega)$ per ogni $n \in \mathbb{N}$ e per q.o. $\omega \in \Omega$. Allora $E(|X_n - X|) \rightarrow 0$, cioè $X_n \rightarrow X$ in L^1 ; in particolare $E(X_n) \rightarrow E(X)$.

1.2.7. DISUGUAGLIANZE. Ricordiamo la *disuguaglianza di Markov*: per ogni variabile aleatoria *positiva* X e per ogni $\delta > 0$ si ha che

$$P(X \geq \delta) \leq \frac{E(X)}{\delta}, \quad (1.1)$$

da cui segue la *disuguaglianza di Chebychev*: per ogni variabile aleatoria $X \in L^2$ e per ogni $\delta > 0$ si ha che

$$P(|X - E(X)| > \delta) \leq \frac{\text{Var}(X)}{\delta^2}. \quad (1.2)$$

La *disuguaglianza di Jensen* afferma che per ogni variabile aleatoria reale $X \in L^1$ e per ogni funzione convessa $\varphi : \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ tale che $\varphi(X) \in L^1$ si ha

$$\varphi(E(X)) \leq E(\varphi(X)). \quad (1.3)$$

Date due variabili aleatorie $X \in L^p$ e $Y \in L^q$, con $\frac{1}{p} + \frac{1}{q} = 1$, la *disuguaglianza di Hölder* afferma che $XY \in L^1$ e si ha $\|XY\|_1 \leq \|X\|_p \|Y\|_q$, o più esplicitamente

$$E(|XY|) \leq E(|X|^p)^{1/p} E(|Y|^q)^{1/q}. \quad (1.4)$$

Nel caso speciale $p = q = \frac{1}{2}$ si ha la *disuguaglianza di Cauchy-Schwarz*.

1.3. LEGGE DI UNA VARIABILE ALEATORIA

1.3.1. DEFINIZIONE DI LEGGE. Una variabile aleatoria $X : (\Omega, \mathcal{F}, P) \rightarrow (E, \mathcal{E})$ induce su (E, \mathcal{E}) una probabilità μ_X , detta *legge* (o *misura immagine*) di X , definita da

$$\mu_X(A) := P(X^{-1}(A)) = P(X \in A).$$

La legge μ_X descrive la probabilità con cui i possibili valori di X vengono assunti ed è talvolta indicata con $P \circ X^{-1}$.

Si noti che, data una probabilità μ sullo spazio misurabile (E, \mathcal{E}) , la funzione identità $id : (E, \mathcal{E}, \mu) \rightarrow (E, \mathcal{E})$ è una variabile aleatoria di legge μ . Questo è il procedimento canonico per costruire una variabile aleatoria a valori in (E, \mathcal{E}) di legge μ assegnata. In particolare, data una variabile aleatoria $X : (\Omega, \mathcal{F}, P) \rightarrow (E, \mathcal{E})$, la funzione identità $id : (E, \mathcal{E}, \mu_X) \rightarrow (E, \mathcal{E})$ è una variabile aleatoria che ha la stessa legge di X .

Se le variabili aleatorie $X, Y : (\Omega, \mathcal{F}, P) \rightarrow (E, \mathcal{E})$ sono q.c. uguali, cioè $P(X = Y) = 1$, allora hanno la stessa legge: infatti da $X = Y$ q.c. segue che $1_{\{X \in A\}} = 1_{\{Y \in A\}}$ q.c. e dunque $\mu_X(A) = E(1_{\{X \in A\}}) = E(1_{\{Y \in A\}}) = \mu_Y(A)$, per ogni $A \in \mathcal{E}$.

Fondamentale è il seguente teorema, noto come formula del *cambio di variabili* (o *passaggio alla misura immagine*).

TEOREMA 1.6 (CAMBIO DI VARIABILI). Sia $X : (\Omega, \mathcal{F}, P) \rightarrow (E, \mathcal{E})$ una variabile aleatoria e $f : (E, \mathcal{E}) \rightarrow \mathbb{R}$ un'applicazione misurabile. La variabile aleatoria $f(X) : (\Omega, \mathcal{F}, P) \rightarrow \mathbb{R}$ è integrabile se e solo se lo è la variabile aleatoria $f : (E, \mathcal{E}, \mu_X) \rightarrow \mathbb{R}$, nel qual caso si ha

$$E(f(X)) = \int_{\Omega} f(X(\omega)) P(d\omega) = \int_E f(x) \mu_X(dx).$$

Questa formula è sempre valida se $f \geq 0$.

1.3.2. ASSOLUTA CONTINUITÀ. Date due misure μ, ν sullo stesso spazio misurabile (E, \mathcal{E}) , si dice che μ è assolutamente continua rispetto a ν se esiste una funzione misurabile positiva $f : (E, \mathcal{E}) \rightarrow [0, \infty)$, detta *densità* (o *derivata di Radon-Nikodym*), tale che $\mu(A) = \int_E 1_A(x) f(x) \nu(dx)$ per ogni $A \in \mathcal{E}$, o equivalentemente tale che $\int_E g(x) \mu(dx) = \int_E g(x) f(x) \nu(dx)$ per ogni $g : (E, \mathcal{E}) \rightarrow \mathbb{R}^+$ misurabile positiva. L'assoluta continuità di μ rispetto a ν si indica con la notazione $\mu \ll \nu$; se si vuole indicare esplicitamente la densità, si scrive $\mu = f \cdot \nu$, $d\mu = f \cdot d\nu$ o anche $\frac{d\mu}{d\nu} = f$.

Chiaramente se $\mu \ll \nu$ allora per ogni A tale che $\nu(A) = 0$ si ha $\mu(A) = 0$. Il celebre teorema di Radon-Nikodym afferma che anche il viceversa è vero.

1.3.3. LEGGI SU \mathbb{R}^d . Il caso che a noi interessa di più è dato dalle leggi su \mathbb{R}^d che sono assolutamente continue rispetto alla misura di Lebesgue, dette leggi *assolutamente continue* tout court. Per una tale legge μ esiste cioè una densità $f : \mathbb{R}^d \rightarrow [0, \infty)$ tale che $\mu(A) = \int_{\mathbb{R}^d} 1_A(x) f(x) dx$ per ogni $A \in \mathcal{B}(\mathbb{R})$, dove dx indica la misura di Lebesgue su \mathbb{R}^d . Un vettore aleatorio X a valori in \mathbb{R}^d è detto assolutamente continuo se la sua legge μ_X lo è. Indicando con f_X la sua densità, dalla formula del cambio di variabili si ottiene $E(g(X)) = \int_{\mathbb{R}^d} g(x) f_X(x) dx$, per ogni funzione misurabile e limitata $g : \mathbb{R}^d \rightarrow \mathbb{R}$.

Una legge μ su \mathbb{R}^d (o, più in generale, su un insieme arbitrario) è detta *discreta* se è una sovrapposizione di misure di Dirac, cioè se esistono $\{x_i\}_{i \in \mathbb{N}}$, $\{p_i\}_{i \in \mathbb{N}}$, con $x_i \in \mathbb{R}^d$ e $p_i \in [0, 1]$, tali che $\mu = \sum_{i \in \mathbb{N}} p_i \delta_{x_i}$ (chiaramente $\sum_{i \in \mathbb{N}} p_i = 1$). Un vettore aleatorio X a valori in \mathbb{R}^d è detto discreto se lo è la sua legge μ_X . In questo caso la formula del cambio di variabili diventa $E(g(X)) = \int_{\mathbb{R}^d} g(x) \mu_X(dx) = \sum_{i \in \mathbb{N}} p_i g(x_i)$ per ogni $g : \mathbb{R}^d \rightarrow \mathbb{R}^+$. La funzione $p_X : \mathbb{R} \rightarrow [0, 1]$ definita da $p_X(x) := \sum_{i \in \mathbb{N}} p_i \mathbf{1}_{\{x_i\}}(x)$ (cioè $p_X(x) = p_i$ se $x = x_i$ per qualche $i \in \mathbb{N}$ e $p_X(x) = 0$ altrimenti) è detta *densità discreta* o funzione di massa.

Sia X un vettore aleatorio d -dimensionale assolutamente continuo, con densità f_X . Se A è una matrice $d \times d$ invertibile e $b \in \mathbb{R}^d$, il vettore aleatorio $Y := AX + b$ è ancora assolutamente continuo, con densità $f_Y(y) := |\det A|^{-1} f_X(A^{-1}(y - b))$. Questa formula si può generalizzare al caso in cui la trasformazione affine $x \mapsto Ax + b$ sia sostituita da un diffeomorfismo, ma non ne avremo bisogno.

Ricordiamo infine che le misure di probabilità μ su \mathbb{R} sono in corrispondenza biunivoca con le funzioni $F : \mathbb{R} \rightarrow [0, 1]$ crescenti, continue a destra e tali che $\lim_{x \rightarrow -\infty} F(x) = 0$ e $\lim_{x \rightarrow +\infty} F(x) = 1$: la corrispondenza è semplicemente quella che associa a una probabilità μ la sua *funzione di ripartizione* $F(x) := \mu((-\infty, x])$.

1.4. INDIPENDENZA E MISURA PRODOTTO

1.4.1. PROBABILITÀ CONDIZIONATA E INDIPENDENZA. Fissiamo uno spazio di probabilità (Ω, \mathcal{F}, P) . Dati due eventi $A, B \in \mathcal{F}$ con $P(B) > 0$, si definisce *probabilità condizionata di A dato B* la quantità $P(A|B) := P(A \cap B)/P(B)$.

Intuitivamente, se si ha $P(A|B) = P(A)$, l'evento A si dice indipendente dall'evento B . Si noti che la condizione $P(A|B) = P(A)$ si può riscrivere come $P(A \cap B) = P(A)P(B)$, che ha il vantaggio di essere simmetrica in A e B e ben posta anche se $P(B) = 0$. Si preferisce quindi dire che due eventi A e B sono indipendenti se e solo se $P(A \cap B) = P(A)P(B)$.

Normalmente il passo successivo è di estendere questa nozione a più di due eventi e successivamente di definire l'indipendenza di variabili aleatorie. Da un punto di vista concettuale risulta però più conveniente definire innanzitutto l'indipendenza di σ -algebre, perché l'indipendenza di eventi e di variabili aleatorie segue come caso particolare. Procediamo dunque in questo modo.

- Date le σ -algebre $\mathcal{F}_1, \dots, \mathcal{F}_n$ contenute in \mathcal{F} , esse si dicono indipendenti se per ogni scelta di eventi $A_1 \in \mathcal{F}_1, \dots, A_n \in \mathcal{F}_n$, si ha che

$$P(A_1 \cap \dots \cap A_n) = \prod_{j=1}^n P(A_j). \quad (1.5)$$

- Date le variabili aleatorie X_1, \dots, X_n , definite su Ω a valori rispettivamente negli spazi $(E_1, \mathcal{E}_1), \dots, (E_n, \mathcal{E}_n)$, esse si dicono indipendenti se lo sono le σ -algebre da loro generate $\sigma(X_1), \dots, \sigma(X_n)$. Ricordando che $\sigma(X) = \{\{X \in B\}, B \in \mathcal{E}\}$, ciò significa che per ogni scelta di eventi $B_1 \in \mathcal{E}_1, \dots, B_n \in \mathcal{E}_n$ si deve avere

$$P(X_1 \in B_1, \dots, X_n \in B_n) = \prod_{j=1}^n P(X_j \in B_j).$$

- Dati gli eventi A_1, \dots, A_n di \mathcal{F} , essi si dicono indipendenti se lo sono le σ -algebre $\sigma(\{A_1\}), \dots, \sigma(\{A_n\})$ da essi generate (o, equivalentemente, se lo sono le variabili aleatorie $1_{A_1}, \dots, 1_{A_n}$). Ricordando che $\sigma(\{A\}) = \{\emptyset, A, A^c, \Omega\}$, questo equivale a richiedere che per ogni scelta di $\sigma_1, \dots, \sigma_n \in \{1, c\}$

$$P(A_1^{\sigma_1} \cap \dots \cap A_n^{\sigma_n}) = \prod_{j=1}^n P(A_j^{\sigma_j}),$$

dove poniamo $A_i^1 := A_i$. Si può mostrare che ciò corrisponde alla definizione classica di indipendenza di eventi, cioè che per ogni sottoinsieme $J \subseteq \{1, \dots, n\}$ si abbia

$$P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j).$$

Sono anche possibili enunciati misti, sempre riducendosi alle σ -algebre: per esempio, una variabile aleatoria X e una σ -algebra \mathcal{G} si dicono indipendenti se sono indipendenti le σ -algebre $\{\sigma(X), \mathcal{G}\}$, ecc.

Un'osservazione molto utile in pratica è la seguente: date le σ -algebre $\mathcal{F}_1, \dots, \mathcal{F}_n$, e data, per ogni $1 \leq i \leq n$, una *base* J_i di \mathcal{F}_i , le σ -algebre $\mathcal{F}_1, \dots, \mathcal{F}_n$ sono indipendenti se e solo vale la relazione (1.5) per $A_1 \in J_1, \dots, A_n \in J_n$ (questo segue dal Lemma di Dynkin; si veda la sezione 1.8.1).

Abbiamo definito l'indipendenza per un numero finito di σ -algebre $\mathcal{F}_1, \dots, \mathcal{F}_n$. L'estensione a una famiglia infinita è immediata: le σ -algebre $\{\mathcal{F}_i\}_{i \in I}$, con I insieme arbitrario, si dicono indipendenti se lo è ogni sottofamiglia finita, cioè se le σ -algebre $\mathcal{F}_{i_1}, \dots, \mathcal{F}_{i_n}$ sono indipendenti per ogni $n \in \mathbb{N}$ e per ogni scelta di $i_1, \dots, i_n \in I$. In modo analogo si definisce l'indipendenza di famiglie arbitrarie di variabili aleatorie $\{X_i\}_{i \in I}$ ed eventi $\{A_i\}_{i \in I}$.

Se le σ -algebre $\{\mathcal{F}_i\}_{i \in I}$ sono indipendenti e se, per ogni $i \in I$, si considera una σ -algebra $\mathcal{G}_i \subseteq \mathcal{F}_i$, è chiaro che anche le σ -algebre $\{\mathcal{G}_i\}_{i \in I}$ sono indipendenti. Ricordando che una variabile aleatoria X_i è \mathcal{F}_i -misurabile se e solo se $\sigma(X_i) \subseteq \mathcal{F}_i$, si ha allora l'utile osservazione seguente: se le σ -algebre $\{\mathcal{F}_i\}_{i \in I}$ sono indipendenti e se X_i è una variabile aleatoria \mathcal{F}_i -misurabile, per ogni $i \in I$, le variabili aleatorie $\{X_i\}_{i \in I}$ sono indipendenti.

Concludiamo enunciando un risultato intuitivo (di frequente uso implicito): se le σ -algebre $\{\mathcal{F}_i\}_{i \in I}$ sono indipendenti e se $\{I_k\}_{k \in K}$ è una partizione dell'insieme di indici I (cioè $\bigcup_{k \in K} I_k = I$ e $I_k \cap I_{k'} = \emptyset$ per $k \neq k'$), allora le σ -algebre $\{\mathcal{H}_k := \sigma(\mathcal{F}_i, i \in I_k)\}_{k \in K}$ sono indipendenti.

1.4.2. INDIPENDENZA E SCORRELAZIONE. Se X e Y sono variabili aleatorie reali integrabili indipendenti, allora il prodotto XY è integrabile e si ha $E(XY) = E(X)E(Y)$; dunque $\text{Cov}(X, Y) = 0$, cioè le variabili sono *scorrelate*. In particolare, se $X, Y \in L^2$ sono indipendenti si ha che $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$. Il viceversa è falso in generale. Sottolineiamo che se X e Y sono variabili aleatorie reali integrabili non indipendenti, non è detto che XY sia integrabile.[†]

[†]È sufficiente considerare $Y = X$ con $X \in L^1 \setminus L^2$.

Notiamo anche che se X e Y sono variabili aleatorie indipendenti, lo sono anche $\varphi(X)$ e $\psi(Y)$, qualunque siano le applicazioni misurabili φ, ψ . In particolare, se φ, ψ sono funzioni reali misurabili e limitate, si ha $E(\varphi(X)\psi(Y)) = E(\varphi(X))E(\psi(Y))$.

1.4.3. MISURA PRODOTTO E TEOREMA DI FUBINI. Dati due spazi di misura $(E_1, \mathcal{E}_1, \mu_1)$, $(E_2, \mathcal{E}_2, \mu_2)$ finiti o σ -finiti, esiste un'unica misura $\mu = \mu_1 \otimes \mu_2$ sullo spazio prodotto $(E_1 \times E_2, \mathcal{E}_1 \otimes \mathcal{E}_2)$, detta *misura prodotto*, con la proprietà $\mu(A \times B) = \mu_1(A)\mu_2(B)$ per ogni $A \in \mathcal{E}_1$ e $B \in \mathcal{E}_2$.[†] Se $E_1 = E_2 = \mathbb{R}$ e le misure μ_1, μ_2 sono assolutamente continue, con densità rispettive f_1, f_2 , la misura prodotto è anch'essa assolutamente continua, con densità $f(x_1, x_2) := f_1(x_1) \cdot f_2(x_2)$. Questi risultati si estendono al prodotto di un numero finito di spazi.

Richiamiamo ora il Teorema di Fubini. Siano μ_1 e μ_2 misure σ -finite su (E_1, \mathcal{E}_1) e (E_2, \mathcal{E}_2) rispettivamente e sia $\mu := \mu_1 \otimes \mu_2$ la misura prodotto. Sia $f : (E_1 \times E_2, \mathcal{E}_1 \otimes \mathcal{E}_2) \rightarrow \mathbb{R}$ una funzione reale misurabile. Se f è μ -integrabile (cioè $\int_{E_1 \times E_2} |f| d\mu < \infty$), oppure se $f \geq 0$, vale che

$$\begin{aligned} \int_{E_1 \times E_2} f d\mu &= \int_{E_1} \left(\int_{E_2} f(x_1, x_2) \mu_2(dx_2) \right) \mu_1(dx_1) \\ &= \int_{E_2} \left(\int_{E_1} f(x_1, x_2) \mu_1(dx_1) \right) \mu_2(dx_2). \end{aligned} \quad (1.6)$$

ESERCIZIO 1.7. Si dimostri la formula $E(Y) = \int_0^\infty P(Y > t) dt$, valida per ogni variabile aleatoria reale *positiva* Y . [Sugg.: Si noti che $Y(\omega) = \int_0^\infty 1_{\{t < Y(\omega)\}} dt$.]

Date due variabili aleatorie X_1, X_2 definite su (Ω, \mathcal{F}, P) a valori rispettivamente in (E_1, \mathcal{E}_1) , (E_2, \mathcal{E}_2) , indichiamo le loro leggi con μ_{X_1}, μ_{X_2} . La coppia $X = (X_1, X_2)$ è una variabile aleatoria a valori in $(E_1 \times E_2, \mathcal{E}_1 \otimes \mathcal{E}_2)$, la cui legge indichiamo con μ_X . È facile vedere che X_1 e X_2 sono indipendenti se e soltanto se $\mu_X = \mu_{X_1} \otimes \mu_{X_2}$. Lo stesso vale per un numero finito di variabili aleatorie X_1, \dots, X_n a valori negli spazi (E_i, \mathcal{E}_i) : le variabili sono indipendenti se e soltanto se la loro legge congiunta su $(\times_{i=1}^n E_i, \otimes_{i=1}^n \mathcal{E}_i)$ è data dal prodotto delle leggi marginali.

1.4.4. SUCCESSIONI INDIPENDENTI. È noto che, assegnata un'arbitraria successione di probabilità $\{\mu_n\}_{n \in \mathbb{N}}$ su \mathbb{R} , esiste un opportuno spazio di probabilità (Ω, \mathcal{F}, P) su cui è definita una successione $\{X_n\}_{n \in \mathbb{N}}$ di variabili aleatorie reali *indipendenti* tali che la legge di X_n sia μ_n . Una costruzione tipica è richiamata nella sezione 1.8.2.

1.4.5. LEMMA DI BOREL-CANTELLI. Data una successione di eventi $\{A_n\}_{n \in \mathbb{N}}$ di uno spazio di probabilità (Ω, \mathcal{F}, P) , si definisce l'evento

$$\limsup_{n \rightarrow \infty} A_n := \bigcap_{k \in \mathbb{N}} \bigcup_{n \geq k} A_n = \{\omega \in \Omega : \omega \in A_n \text{ per infiniti } n\} = \left\{ \sum_{n \in \mathbb{N}} 1_{A_n} = \infty \right\}.$$

Si ha allora l'utilissimo

[†]L'unicità segue dal fatto che gli insiemi della forma $A \times B$, al variare di $A \in \mathcal{E}_1$ e $B \in \mathcal{E}_2$, costituiscono una base di $\mathcal{E}_1 \otimes \mathcal{E}_2$.

LEMMA 1.8 (BOREL-CANTELLI). Sia $\{A_n\}_{n \in \mathbb{N}}$ una successione di eventi di uno spazio di probabilità (Ω, \mathcal{F}, P) .

- Se $\sum_{n \in \mathbb{N}} P(A_n) < \infty$, allora $P(\limsup_{n \rightarrow \infty} A_n) = 0$.
- Se $\sum_{n \in \mathbb{N}} P(A_n) = \infty$ e inoltre se A_i e A_j sono indipendenti per ogni $i \neq j$, allora $P(\limsup_{n \rightarrow \infty} A_n) = 1$.

ESERCIZIO 1.9. Se $\{X_n\}_{n \in \mathbb{N}}$ sono variabili aleatorie i.i.d. con $X_n \sim \text{Exp}(\lambda)$, allora q.c. si ha $\limsup_{n \rightarrow \infty} X_n / \log n = \lambda^{-1}$.

1.4.6. CONVOLUZIONE. Date due probabilità μ, ν su \mathbb{R}^d e due variabili aleatorie X e Y indipendenti, le cui leggi siano rispettivamente μ e ν , la *convoluzione* di μ e ν , indicata con $\mu * \nu$, è per definizione la legge della variabile aleatoria $X + Y$. Per ogni insieme A boreliano di \mathbb{R}^d si ha $\mu * \nu(A) = \int_{\mathbb{R}^d} \mu(A - y) \nu(dy) = \int_{\mathbb{R}^d} \nu(A - y) \mu(dy)$, che mostra tra l'altro come $\mu * \nu$ dipenda solo da μ e ν e non dalle variabili X e Y .

Il caso più importante è quello in cui le leggi μ e ν siano assolutamente continue, con densità rispettivamente f e g . In questo caso la legge di $\mu * \nu$ è anch'essa assolutamente continua, con densità $h(x) = \int_{\mathbb{R}^d} f(x - y)g(y)dy = \int_{\mathbb{R}^d} g(x - y)f(y)dy$, detta convoluzione di f e g e indicata con $h = f * g$.

1.5. NOZIONI DI CONVERGENZA

1.5.1. CONVERGENZA DI MISURE. Sia $(E, \mathcal{B}(E))$ uno spazio metrico, con distanza $d(\cdot, \cdot)$, munito della σ -algebra boreliana. Il caso tipico è dato da \mathbb{R}^d , con la distanza indotta dalla norma euclidea: $d(x, y) = |x - y| = \sqrt{\sum_{i=1}^d (x_i - y_i)^2}$. Data una successione di probabilità $\{\mu_n\}_{n \in \mathbb{N}}$ su E , si dice che essa converge *debolmente* verso la probabilità μ su E se per ogni funzione $f : E \rightarrow \mathbb{R}$ continua e limitata si ha che $\int f d\mu_n \rightarrow \int f d\mu$. Sebbene esistano altre nozioni di convergenza per successioni di *measure*, questa è la più importante e sarà l'unica che considereremo.

1.5.2. CONVERGENZA DI VARIABILI ALEATORIE. Consideriamo una famiglia di variabili aleatorie $X_n : (\Omega_n, \mathcal{F}_n, P_n) \rightarrow (E, \mathcal{B}(E))$, per $n \in \mathbb{N}$, e $X : (\Omega, \mathcal{F}, P) \rightarrow (E, \mathcal{B}(E))$, definite non necessariamente sullo stesso spazio di probabilità, ma tutte a valori nello stesso spazio metrico E .

- Diremo che la successione $\{X_n\}_{n \in \mathbb{N}}$ converge *in legge* (o in distribuzione) verso X se la successione delle leggi μ_{X_n} di X_n converge debolmente verso la legge μ_X di X . Usando la formula del cambio di variabili (Teorema 1.6), ciò è equivalente a richiedere che $E_n(f(X_n)) \rightarrow E(f(X))$ per ogni funzione $f : E \rightarrow \mathbb{R}$ continua e limitata.

Supponiamo ora che le variabili aleatorie $\{X_n\}_{n \in \mathbb{N}}$, X siano tutte definite sullo stesso spazio di probabilità (Ω, \mathcal{F}, P) e assumano valori nello spazio metrico $(E, \mathcal{B}(E))$.

- Diremo che la successione $\{X_n\}_{n \in \mathbb{N}}$ converge *in probabilità* verso X se per ogni $\varepsilon > 0$ si ha che $P(d(X_n, X) > \varepsilon) \rightarrow 0$.

- Diremo che la successione $\{X_n\}_{n \in \mathbb{N}}$ converge *quasi certamente* (*q.c.*) verso X se esiste $A \in \mathcal{F}$ con $P(A) = 1$ tale che per ogni $\omega \in A$ si ha $X_n(\omega) \rightarrow X(\omega)$, cioè $d(X_n(\omega), X(\omega)) \rightarrow 0$.

Consideriamo infine il caso in cui le variabili aleatorie $\{X_n\}_{n \in \mathbb{N}}$, X siano definite sullo stesso spazio (Ω, \mathcal{F}, P) e assumano valori in \mathbb{R}^d .

- Diremo che la successione $\{X_n\}_{n \in \mathbb{N}}$ converge verso X in L^p se $\|X_n - X\|_p \rightarrow 0$, cioè se $E(|X_n - X|^p) \rightarrow 0$, dove $|\cdot|$ indica la norma euclidea su \mathbb{R}^d .

Si noti che, essendo $\|X_n - X\|_q \leq \|X_n - X\|_p$ se $p \geq q$ (Jensen), la convergenza di X_n verso X in L^p implica quella in L^q . Dalla disuguaglianza triangolare si ha inoltre che $|\|X_n\|_p - \|X\|_p| \leq \|X_n - X\|_p$, da cui si ricava che la convergenza in L^p implica quella del momento p -esimo. In definitiva,

$$X_n \rightarrow X \text{ in } L^p \quad \implies \quad E(|X_n|^q) \rightarrow E(|X|^q), \quad \text{per ogni } 1 \leq q \leq p. \quad (1.7)$$

PROPOSIZIONE 1.10. Date le variabili aleatorie X_n, X a valori in uno spazio metrico E , valgono le seguenti relazioni:

- se $X_n \rightarrow X$ q.c., allora $X_n \rightarrow X$ in probabilità;
- se $X_n \rightarrow X$ in L^p , allora $X_n \rightarrow X$ in probabilità;
- se $X_n \rightarrow X$ in probabilità, allora esiste una sottosuccessione $\{n_k\}_{k \in \mathbb{N}}$ tale che $X_{n_k} \rightarrow X$ q.c.;
- se $X_n \rightarrow X$ in probabilità, allora $X_n \rightarrow X$ in legge.

DIMOSTRAZIONE. Se $X_n \rightarrow X$ q.c., si ha $d(X_n, X) \rightarrow 0$ q.c. e dunque $1_{\{d(X_n, X) > \varepsilon\}} \rightarrow 0$ q.c., per ogni $\varepsilon > 0$. Per convergenza dominata si ottiene dunque $P(d(X_n, X) > \varepsilon) = E(1_{\{d(X_n, X) > \varepsilon\}}) \rightarrow 0$, poiché $|1_{\{d(X_n, X) > \varepsilon\}}| \leq 1$. Di conseguenza $X_n \rightarrow X$ in probabilità.

Supponiamo ora che $X_n \rightarrow X$ in L^p . In questo caso $E = \mathbb{R}^d$ e $d(x, y) = |x - y|$. Applicando la disuguaglianza di Markov, si ha $P(d(X_n, X) > \varepsilon) = P(|X_n - X| > \varepsilon) \leq \varepsilon^{-p} E(|X_n - X|^p) \rightarrow 0$ per ogni $\varepsilon > 0$, dunque $X_n \rightarrow X$ in probabilità.

Facciamo ora l'ipotesi che $X_n \rightarrow X$ in probabilità. Fissiamo arbitrariamente una successione $\{\varepsilon_k\}_{k \in \mathbb{N}}$ positiva e infinitesima, per esempio $\varepsilon_k := \frac{1}{k}$. Per ogni k fissato si ha $P(d(X_n, X) > \varepsilon_k) \rightarrow 0$ per $n \rightarrow \infty$, quindi possiamo definire n_k come il più piccolo valore di $n \in \mathbb{N}$ per cui $P(d(X_n, X) > \varepsilon_k) \leq \frac{1}{2^k}$. Per costruzione $\sum_{k \in \mathbb{N}} P(d(X_{n_k}, X) > \varepsilon_k) \leq \sum_{k \in \mathbb{N}} \frac{1}{2^k} < \infty$, quindi per il Lemma di Borel-Cantelli si ha che q.c. $d(X_{n_k}, X) \leq \varepsilon_k = \frac{1}{k}$ per k grande, da cui segue che $d(X_{n_k}, X) \rightarrow 0$ q.c. per $k \rightarrow \infty$. Abbiamo dunque determinato una sottosuccessione $(n_k)_{k \in \mathbb{N}}$ per cui $X_{n_k} \rightarrow X$ q.c..

Supponiamo infine che $X_n \rightarrow X$ in probabilità e sia $f : E \rightarrow \mathbb{R}$ una qualunque funzione continua e limitata. Vogliamo mostrare che $E(f(X_n)) \rightarrow E(f(X))$, da cui segue che $X_n \rightarrow X$ in legge. Per un argomento classico (vedi Lemma 1.11 più giù), è sufficiente mostrare che per ogni sottosuccessione $\{n_k\}_{k \in \mathbb{N}}$ esiste una sotto-sottosuccessione $\{n'_k\}_{k \in \mathbb{N}}$ tale che $E(f(X_{n'_k})) \rightarrow E(f(X))$. Visto che per ipotesi $X_n \rightarrow X$ in probabilità, anche $X_{n_k} \rightarrow X$ in probabilità. Per quanto visto sopra, possiamo dunque estrarre una

sottosuccessione $\{n'_k\}_{k \in \mathbb{N}}$ di $\{n_k\}_{k \in \mathbb{N}}$ tale che $X_{n'_k} \rightarrow X$ q.c.. Di conseguenza anche $f(X_{n'_k}) \rightarrow f(X)$ q.c., perché f è continua, e la convergenza $E(f(X_{n'_k})) \rightarrow E(f(X))$ segue dal teorema di convergenza dominata, poiché f è limitata. \square

LEMMA 1.11. Sia $\{x_n\}_{n \in \mathbb{N}}$ una successione in uno spazio topologico E . Supponiamo esista $\bar{x} \in E$ con la seguente proprietà: per ogni sottosuccessione $\{x_{n_k}\}_{k \in \mathbb{N}}$ esiste una sotto-sottosuccessione $\{x_{n'_k}\}_{k \in \mathbb{N}}$ di $\{x_{n_k}\}_{k \in \mathbb{N}}$ che converge verso \bar{x} . Allora la successione completa $\{x_n\}_{n \in \mathbb{N}}$ converge verso \bar{x} .

DIMOSTRAZIONE. La convergenza di $\{x_n\}_{n \in \mathbb{N}}$ verso \bar{x} significa per definizione che per ogni aperto $A \ni \bar{x}$ esiste $n_0 < \infty$ tale che $x_n \in A$ per ogni $n \geq n_0$. Da ciò segue che, se $\{x_n\}_{n \in \mathbb{N}}$ non convergesse verso \bar{x} , esisterebbe un aperto $A \ni \bar{x}$ tale che $x_{n_k} \notin A$ per un insieme infinito di indici $\{n_k\}_{k \in \mathbb{N}}$, che possiamo supporre crescente; ma allora dalla sottosuccessione $\{x_{n_k}\}_{k \in \mathbb{N}}$ non si potrebbe estrarre nessuna sotto-sottosuccessione che converga a \bar{x} , contro l'ipotesi. \square

COROLLARIO 1.12. Siano $X, \{X_n\}_{n \in \mathbb{N}}$ variabili aleatorie reali tali che, per ogni sottosuccessione di $\{X_n\}_{n \in \mathbb{N}}$, è possibile estrarre una sotto-sottosuccessione che converge a X in L^p (risp. in probabilità). Allora $\{X_n\}_{n \in \mathbb{N}}$ converge a X in L^p (risp. in probabilità).

DIMOSTRAZIONE. È sufficiente applicare il Lemma 1.11 alla successione $\{X_n\}_{n \in \mathbb{N}} \in L^p$, oppure alla successione reale $\|X_n - X\|_p$ (risp. alla successione reale $P(d(X_n, X) > \varepsilon)$, per $\varepsilon > 0$ fissato). \square

OSSERVAZIONE 1.13. Consideriamo uno spazio di probabilità (Ω, \mathcal{F}, P) per cui le nozioni di convergenza in probabilità e convergenza q.c. siano distinte, su cui si possa cioè definire una successione di variabili aleatorie $\{X_n\}_{n \in \mathbb{N}}$ che converge in probabilità ma non converge q.c. (è il caso tipico di uno spazio di probabilità senza atomi).[†] La Proposizione 1.10 e il Lemma 1.11 mostrano che in questo caso *non esiste nessuna topologia sullo spazio delle variabili aleatorie definite su (Ω, \mathcal{F}, P) che induca la nozione di convergenza quasi certa*. Infatti, grazie alla Proposizione 1.10, sappiamo che da ogni sottosuccessione di $\{X_n\}_{n \in \mathbb{N}}$ si può estrarre una sotto-sottosuccessione che converge q.c.; se la convergenza q.c. fosse indotta da una topologia, per il Lemma 1.11 si dovrebbe avere che l'intera successione $\{X_n\}_{n \in \mathbb{N}}$ converge q.c., cosa che abbiamo escluso per ipotesi.

La convergenza in probabilità è invece indotta da una topologia, anzi da una pseudometrica:[‡] introducendo la pseudodistanza $\delta(X, Y) := E(|X - Y|/(1 + |X - Y|))$ tra variabili aleatorie, non è difficile vedere che $X_n \rightarrow X$ in probabilità se e solo se $\delta(X_n, X) \rightarrow 0$. \square

1.5.3. ULTERIORI OSSERVAZIONI. Se $X_n \rightarrow X$ in legge e lo spazio d'arrivo è *polacco* (cioè metrico completo e separabile), è possibile definire su un opportuno spazio di probabilità (Ω, \mathcal{F}, P) variabili aleatorie $\{\tilde{X}_n\}_{n \in \mathbb{N}}$ e \tilde{X} , con la stessa legge rispettivamente di X_n e X , tali che $\tilde{X}_n \rightarrow \tilde{X}$ q.c. (teorema di Skorokod).

Date leggi μ_n, μ su \mathbb{R} con funzioni di ripartizione rispettivamente $F_n(\cdot), F(\cdot)$, la convergenza debole di μ_n verso μ è equivalente alla convergenza di $F_n(x)$ verso $F(x)$ per ogni $x \in \mathbb{R}$ in cui $F(\cdot)$ è continua.

Ricordiamo infine l'enunciato del Teorema Limite Centrale: se $\{X_n\}_{n \in \mathbb{N}}$ è una successione i.i.d. di variabili aleatorie reali con $E(X_n) = 0$, $E(X_n^2) = 1$, allora $P(X_1 + \dots + X_n \leq x\sqrt{n}) \rightarrow \Phi(x)$ per ogni

[†]Se (Ω, \mathcal{F}, P) è uno spazio di probabilità in cui Ω è un insieme numerabile, è facile vedere che ogni successione convergente in probabilità converge anche q.c..

[‡]Per la definizione di spazio pseudometrico, si veda il paragrafo 5.1 del capitolo 5.

$x \in \mathbb{R}$, dove $\Phi(\cdot)$ indica la funzione di ripartizione della legge normale standard (si noti che $\Phi(\cdot)$ è continua in ogni $x \in \mathbb{R}$). Possiamo dunque riformulare il Teorema Limite Centrale nel modo seguente: la legge della variabile aleatoria $(X_1 + \dots + X_n)/\sqrt{n}$ converge debolmente verso la legge normale standard.

1.6. FUNZIONI CARATTERISTICHE

Data una vettore aleatorio X in \mathbb{R}^d e detta μ la sua legge, la funzione caratteristica (o trasformata di Fourier) di μ (o, per estensione, di X) è la funzione $\hat{\mu} : \mathbb{R}^d \rightarrow \mathbb{C}$ definita da

$$\hat{\mu}(\vartheta) := E(e^{i\langle \vartheta, X \rangle}) = \int_{\mathbb{R}^d} e^{i\langle \vartheta, x \rangle} \mu(dx),$$

dove ricordiamo che $\langle a, b \rangle := \sum_{i=1}^d a_i b_i$ indica il prodotto scalare standard su \mathbb{R}^d . È facile verificare che $\hat{\mu}(\cdot)$ è una funzione uniformemente continua su \mathbb{R}^d e che $|\hat{\mu}(\cdot)| \leq 1$.

Le proprietà fondamentali delle funzioni caratteristiche sono le seguenti:

- La funzione caratteristica identifica la legge, cioè se due leggi μ, ν su \mathbb{R}^d sono tali che $\hat{\mu}(\vartheta) = \hat{\nu}(\vartheta)$ per ogni $\vartheta \in \mathbb{R}^d$, allora $\mu = \nu$.
- Siano X_1, \dots, X_d variabili casuali reali, con legge rispettivamente μ_1, \dots, μ_d ; indichiamo con μ la legge del vettore aleatorio (X_1, \dots, X_d) su \mathbb{R}^d . Allora le variabili X_1, \dots, X_d sono indipendenti se e solo se $\hat{\mu}(\vartheta) = \hat{\mu}_1(\vartheta_1) \cdots \hat{\mu}_d(\vartheta_d)$ per ogni $\vartheta = (\vartheta_1, \dots, \vartheta_d) \in \mathbb{R}^d$.[†]
- Se una successione $\{\mu_n\}_{n \in \mathbb{N}}$ di leggi su \mathbb{R}^d converge debolmente verso la legge μ , si ha naturalmente $\hat{\mu}_n(\vartheta) \rightarrow \hat{\mu}(\vartheta)$ per ogni $\vartheta \in \mathbb{R}^d$.

Viceversa, se $\hat{\mu}_n(\vartheta) \rightarrow \psi(\vartheta)$ per ogni $\vartheta \in \mathbb{R}^d$ e se $\psi(\cdot)$ è continua in zero, allora $\psi(\cdot)$ è la funzione caratteristica di una probabilità μ su \mathbb{R}^d e $\mu_n \rightarrow \mu$ debolmente (teorema di convergenza di Lévy).

1.7. LEGGI NORMALI

1.7.1. LEGGI NORMALI UNIVARIATE. Dati $\mu \in \mathbb{R}$ e $\sigma \in (0, \infty)$, la *legge normale* (o gaussiana) di media μ e varianza σ^2 , indicata con $\mathcal{N}(\mu, \sigma^2)$, è la probabilità su \mathbb{R} assolutamente continua con densità

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Si verifica che effettivamente la media e la varianza di questa legge valgono rispettivamente μ e σ^2 , mentre la funzione caratteristica vale

$$\int_{\mathbb{R}} e^{i\vartheta x} f(x) dx = e^{i\vartheta\mu - \frac{1}{2}\sigma^2\vartheta^2}.$$

[†]Se X_1, \dots, X_d sono indipendenti, è immediato verificare che $\hat{\mu}(\vartheta) = \hat{\mu}_1(\vartheta_1) \cdots \hat{\mu}_d(\vartheta_d)$ per ogni $\vartheta \in \mathbb{R}^d$: questo mostra che $\hat{\mu}_1(\vartheta_1) \cdots \hat{\mu}_d(\vartheta_d)$ è la funzione caratteristica della legge prodotto $\mu_1 \otimes \cdots \otimes \mu_d$. Viceversa, se supponiamo che $\hat{\mu}(\vartheta) = \hat{\mu}_1(\vartheta_1) \cdots \hat{\mu}_d(\vartheta_d)$ per ogni $\vartheta \in \mathbb{R}^d$, dal fatto che la funzione caratteristica identifica la legge segue che $\mu = \mu_1 \otimes \cdots \otimes \mu_d$, dunque X_1, \dots, X_d sono indipendenti.

Una variabile aleatoria reale X è detta normale di media $\mu \in \mathbb{R}$ e varianza $\sigma^2 \geq 0$, e scriveremo $X \sim \mathcal{N}(\mu, \sigma^2)$, se lo è la sua legge, cioè se

$$\mathbb{E}(e^{i\vartheta X}) = e^{i\vartheta\mu - \frac{1}{2}\sigma^2\vartheta^2}. \quad (1.8)$$

Per estensione, quando $\sigma^2 = 0$ definiremo la legge $\mathcal{N}(\mu, 0)$ come la misura di Dirac concentrata nel punto μ . Analogamente, per una variabile aleatoria X scriviamo $X \sim \mathcal{N}(\mu, 0)$ se $\mathbb{P}(X = \mu) = 1$. Si noti che media, varianza e funzione caratteristica sono consistenti con la notazione. Quando $\mu = 0$ e $\sigma^2 = 1$ parleremo di legge *normale standard*.

Se $X \sim \mathcal{N}(\mu, \sigma^2)$, allora $aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$, come si verifica facilmente.

Se $X \sim \mathcal{N}(\mu_x, \sigma_x^2)$ e $Y \sim \mathcal{N}(\mu_y, \sigma_y^2)$ sono variabili aleatorie indipendenti, per $u, v \in \mathbb{R}$ si calcola facilmente usando (1.8)

$$\mathbb{E}(e^{i\vartheta(uX+vY)}) = \mathbb{E}(e^{i\vartheta uX}) \mathbb{E}(e^{i\vartheta vY}) = e^{i\vartheta u\mu_x - \frac{1}{2}\vartheta^2 u^2 \sigma_x^2} e^{i\vartheta v\mu_y - \frac{1}{2}\vartheta^2 v^2 \sigma_y^2},$$

da cui segue che $uX + vY \sim \mathcal{N}(u\mu_x + v\mu_y, u^2\sigma_x^2 + v^2\sigma_y^2)$. Analogamente, se X_1, \dots, X_n sono variabili aleatorie indipendenti con $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, per ogni $u \in \mathbb{R}^n$ si ha che $\sum_{i=1}^n u_i X_i \sim \mathcal{N}(\sum_{i=1}^n u_i \mu_i, \sum_{i=1}^n u_i^2 \sigma_i^2)$. Questo mostra in particolare che ogni combinazione lineare di variabili normali *indipendenti* è normale.

1.7.2. LEGGI NORMALI MULTIVARIATE. Un vettore aleatorio $X = (X_1, \dots, X_d)$ a valori in \mathbb{R}^d è detto normale (o gaussiano) se ogni combinazione lineare $\langle u, X \rangle := \sum_{i=1}^d u_i X_i$ delle sue componenti, dove $u \in \mathbb{R}^d$, è una variabile aleatoria reale normale. Una probabilità su \mathbb{R}^d è detta normale se è la legge di un vettore aleatorio normale.

Un esempio importante: se X_1, \dots, X_d sono variabili aleatorie reali normali *indipendenti*, allora $X = (X_1, \dots, X_d)$ è un vettore aleatorio normale: infatti, come abbiamo visto, ogni combinazione lineare delle variabili X_1, \dots, X_d è normale.

In generale, se $X = (X_1, \dots, X_d)$ è un vettore aleatorio normale, segue dalla definizione che ciascuna componente X_i è una variabile aleatoria reale normale.[†] In particolare, $X_i \in L^2$ e sono dunque ben definiti il vettore media $\mu = \mathbb{E}(X) = (\mathbb{E}(X_1), \dots, \mathbb{E}(X_d))$ e la matrice delle covarianze $K_{ij} = K(X)_{ij} := \text{Cov}(X_i, X_j)$ di X . Diremo allora che il vettore aleatorio X su \mathbb{R}^d (e, per estensione, la sua legge) è *normale di media μ e matrice delle covarianze K* e scriveremo $X \sim \mathcal{N}(\mu, K)$. La funzione caratteristica di X si calcola facilmente: per definizione $\langle \vartheta, X \rangle$ è una variabile aleatoria reale normale, per ogni $\vartheta \in \mathbb{R}^d$, per cui applicando (1.8) si ottiene

$$\mathbb{E}(e^{i\langle \vartheta, X \rangle}) = e^{i\mathbb{E}(\langle \vartheta, X \rangle) - \frac{1}{2}\text{Var}(\langle \vartheta, X \rangle)} = e^{i\langle \vartheta, \mu \rangle - \frac{1}{2}\langle \vartheta, K\vartheta \rangle}, \quad (1.9)$$

poiché $\mathbb{E}(\langle \vartheta, X \rangle) = \langle \vartheta, \mu \rangle$ e $\text{Var}(\langle \vartheta, X \rangle) = \langle \vartheta, K\vartheta \rangle$. Questo mostra che, se un vettore aleatorio X è normale, la sua funzione caratteristica (e dunque la sua legge) è determinata dal vettore media μ e dalla matrice delle covarianze K .

[†]Il viceversa non è vero. Ad esempio, se X e σ sono variabili aleatorie reali indipendenti con $X \sim \mathcal{N}(0, 1)$ e $\mathbb{P}(\sigma = +1) = \mathbb{P}(\sigma = -1) = \frac{1}{2}$, allora $Y := \sigma X \sim \mathcal{N}(0, 1)$ ma (X, Y) non è un vettore normale, perché $X + Y = (1 + \sigma)X$ non è normale (infatti $\mathbb{P}(X + Y = 0) = \mathbb{P}(\sigma = -1) = \frac{1}{2}$). Di conseguenza, X e Y non sono indipendenti, ma $\text{Cov}(X, Y) = 0$ (cf. il Lemma 1.14).

Mostriamo ora che, per ogni $\mu \in \mathbb{R}^d$ e per ogni matrice K $d \times d$ simmetrica e semi-definita positiva, è effettivamente possibile costruire un vettore aleatorio $X \sim \mathcal{N}(\mu, K)$. Consideriamo innanzitutto d variabili aleatorie reali Z_1, \dots, Z_d indipendenti, ciascuna normale standard. Introducendo il vettore $Z = (Z_1, \dots, Z_d)$, per $\vartheta \in \mathbb{R}^d$ si ha

$$\mathbb{E}(e^{i\langle \vartheta, Z \rangle}) = \prod_{i=1}^d \mathbb{E}(e^{i\vartheta_i Z_i}) = \prod_{i=1}^d e^{-\frac{1}{2}\vartheta_i^2} = e^{-\frac{1}{2}\sum_{i=1}^d \vartheta_i^2} = e^{-\frac{1}{2}\langle \vartheta, \vartheta \rangle}. \quad (1.10)$$

Questo significa, in base a (1.9), che $Z \sim \mathcal{N}(0, I_d)$, dove I_d indica la matrice identità: $(I_d)_{ij} := \delta_{ij}$ per $1 \leq i, j \leq d$. Si dice che Z è un vettore aleatorio normale standard.

Data una matrice K simmetrica $d \times d$ e semi-definita positiva, essa ammette, per il teorema spettrale, una base ortonormale $\{v_1, \dots, v_d\}$ di autovettori: $\langle v_i, v_j \rangle = \delta_{ij}$ e $Kv_i = \lambda_i v_i$, dove $\lambda_i \geq 0$ perché K è semi-definita positiva. Ricordiamo che A^* indica la trasposta della matrice A . Definendo l'operatore lineare $K^{1/2}$ tramite $K^{1/2}v_i := \sqrt{\lambda_i}v_i$, si verifica facilmente che $K^{1/2} = (K^{1/2})^*$ e $K^{1/2}(K^{1/2})^* = (K^{1/2})^2 = K$. Se ora poniamo $X := K^{1/2}Z + \mu$, cioè $X_i := \sum_{j=1}^d (K^{1/2})_{ij}Z_j + \mu_i$, grazie a (1.10) si ha

$$\mathbb{E}(e^{i\langle \vartheta, X \rangle}) = e^{i\langle \vartheta, \mu \rangle} \mathbb{E}(e^{i\langle (K^{1/2})^* \vartheta, Z \rangle}) = e^{i\langle \vartheta, \mu \rangle - \frac{1}{2}\langle (K^{1/2})^* \vartheta, (K^{1/2})^* \vartheta \rangle} = e^{i\langle \vartheta, \mu \rangle - \frac{1}{2}\langle \vartheta, K \vartheta \rangle}.$$

Grazie a (1.9), si ha dunque che $X \sim \mathcal{N}(\mu, K)$.

1.7.3. PROPRIETÀ DELLE LEGGI NORMALI. Segue immediatamente dalla definizione che una trasformazione affine di un vettore normale è ancora un vettore normale (ogni combinazione lineare delle componenti di Y è normale perché si può scrivere come una costante più una combinazione lineare delle componenti di X , che è normale per ipotesi). Più precisamente, sia $X \sim \mathcal{N}(\mu, K)$ un vettore aleatorio normale in \mathbb{R}^d e sia $Y := AX + b$, con A matrice reale $m \times d$ e $b \in \mathbb{R}^m$; dalla legge di trasformazione di media e matrice delle covarianze segue allora che $Y \sim \mathcal{N}(\mathbb{E}(Y), K(Y)) = \mathcal{N}(A\mu + b, AK A^*)$.[†]

Esponiamo ora una proprietà fondamentale delle variabili normali. Supponiamo che X, Y siano due variabili aleatorie reali normali tali che il vettore aleatorio bidimensionale (X, Y) sia normale. Diremo in questo caso che X e Y sono *congiuntamente normali*. Allora X e Y sono indipendenti se (e solo se) sono scorrelate, cioè se (e solo se) $\text{Cov}(X, Y) = 0$. Infatti in questo caso la matrice di covarianza K del vettore (X, Y) è diagonale. Ricordando la relazione (1.9), possiamo allora scrivere

$$\mathbb{E}(e^{i(\vartheta_1 X + \vartheta_2 Y)}) = e^{i(\vartheta_1 \mu_X + \vartheta_2 \mu_Y)} e^{-\frac{1}{2}(\text{Var}(X)\vartheta_1^2 + \text{Var}(Y)\vartheta_2^2)} = \mathbb{E}(e^{i\vartheta_1 X}) \mathbb{E}(e^{i\vartheta_2 Y}),$$

e l'indipendenza segue dalle proprietà delle funzioni caratteristiche. L'estensione al caso in cui le variabili siano più di due è immediata.

[†]In alternativa, si può mostrare che $Y \sim \mathcal{N}(A\mu + b, AK A^*)$ direttamente dall'equazione (1.9): per ogni $\vartheta \in \mathbb{R}^m$ si ha infatti $\langle \vartheta, Y \rangle = \langle \vartheta, b \rangle + \langle A^* \vartheta, X \rangle$ e possiamo scrivere

$$\mathbb{E}(e^{i\langle \vartheta, Y \rangle}) = e^{i\langle \vartheta, b \rangle} \mathbb{E}(e^{i\langle A^* \vartheta, X \rangle}) = e^{i\langle \vartheta, b \rangle} e^{i\langle A^* \vartheta, \mu \rangle - \frac{1}{2}\langle A^* \vartheta, K A^* \vartheta \rangle} = e^{i\langle \vartheta, A\mu + b \rangle - \frac{1}{2}\langle \vartheta, AK A^* \vartheta \rangle}.$$

LEMMA 1.14. Date X_1, \dots, X_n variabili aleatorie reali *congiuntamente normali*, tali cioè che il vettore aleatorio $X = (X_1, \dots, X_n)$ a valori in \mathbb{R}^n sia normale, esse sono indipendenti se e solo se sono scorrelate, cioè $\text{Cov}(X_i, X_j) = 0$ per ogni $i \neq j$.

Sottolineiamo che questa proprietà in generale non è valida se le variabili normali X_1, \dots, X_n non sono *congiuntamente* normali.

Il Lemma 1.14 può essere riformulato nel modo seguente: dato un vettore aleatorio *normale* $X = (X_1, \dots, X_n)$ a valori in \mathbb{R}^n , le sue componenti X_1, \dots, X_n sono indipendenti se e solo se la matrice di covarianza di X è diagonale, cioè $K_{ij} = 0$ per ogni $i \neq j$.

OSSERVAZIONE 1.15. Si può estendere il Lemma 1.14 al caso in cui X_1, \dots, X_n siano *vettori* aleatori congiuntamente normali. Più precisamente, supponiamo che $\{X_i = (X_{i,1}, \dots, X_{i,d_i})\}_{i=1, \dots, n}$ siano vettori normali in \mathbb{R}^{d_i} , con $d_i \in \mathbb{N}$, tali che $X = \{X_{i,j}\}_{i=1, \dots, n, j=1, \dots, d_i}$ sia un vettore normale in $\mathbb{R}^{d_1 + \dots + d_n}$. Allora X_1, \dots, X_n sono indipendenti se e solo se le *rispettive* componenti sono scorrelate, cioè $\text{Cov}(X_{i,a}, X_{j,b}) = 0$ per ogni $i \neq j$ e per ogni $1 \leq a \leq d_i$ e $1 \leq b \leq d_j$. Sottolineiamo che *non* si richiede che $\text{Cov}(X_{i,a}, X_{i,b}) = 0$. \square

Se $X \sim \mathcal{N}(\mu, K)$ è un vettore aleatorio normale in \mathbb{R}^d , nel caso in cui $\det(K) \neq 0$ il vettore X è assolutamente continuo con densità

$$f_X(x) = \frac{1}{(2\pi)^{d/2} \sqrt{|\det(K)|}} e^{-\frac{1}{2} \langle x - \mu, K^{-1}(x - \mu) \rangle}, \quad (1.11)$$

come segue dalla già menzionata rappresentazione $X = K^{1/2}Z + \mu$ con $Z \sim \mathcal{N}(0, I_d)$, applicando la legge di trasformazione $f_X(x) = |\det(K)|^{-1/2} f_Z(K^{-1/2}(x - \mu))$ e osservando che $f_Z(z) = (2\pi)^{-d/2} e^{-\frac{1}{2} \langle z, z \rangle}$.[†]

Enunciamo infine un utile risultato di convergenza.

PROPOSIZIONE 1.16. Sia $\{X_n\}_{n \in \mathbb{N}}$ una successione di vettori aleatori che converge in legge verso un vettore aleatorio X . Se X_n è normale per ogni $n \in \mathbb{N}$, $X_n \sim \mathcal{N}(\mu_n, K_n)$, allora anche X è normale, $X \sim \mathcal{N}(\mu, K)$, e si ha $\mu = \lim_{n \rightarrow \infty} \mu_n$ e $K = \lim_{n \rightarrow \infty} K_n$.

DIMOSTRAZIONE. Se lo spazio di arrivo di X_n e X è \mathbb{R}^d , fissiamo arbitrariamente $u \in \mathbb{R}^d$ e poniamo $Y_n := \langle u, X_n \rangle$ e $Y := \langle u, X \rangle$. Segue dalle ipotesi che $Y_n \sim \mathcal{N}(\nu_n, \sigma_n^2)$, con $\nu_n = \langle u, \mu \rangle$ e $\sigma_n^2 = \langle u, K_n u \rangle$. Inoltre $Y_n \rightarrow Y$ in legge, come si verifica immediatamente. Indicando con $\varphi(\vartheta) := E(e^{i\vartheta Y})$ la funzione caratteristica di Y , dalla relazione (1.8) si ha

$$\lim_{n \rightarrow \infty} e^{i\vartheta \nu_n - \frac{1}{2} \vartheta^2 \sigma_n^2} = \varphi(\vartheta), \quad \text{per ogni } \vartheta \in \mathbb{R}. \quad (1.12)$$

Prendendo i moduli, si ottiene che $e^{-\vartheta^2 \sigma_n^2/2} \rightarrow |\varphi(\vartheta)|$ per ogni $\vartheta \in \mathbb{R}$. Se la successione σ_n^2 non fosse limitata, esisterebbe una sottosuccessione che diverge e si avrebbe $|\varphi(\vartheta)| =$

[†]Se X è un generico vettore aleatorio (non necessariamente normale) con matrice delle covarianze singolare, ossia $\det(K(X)) = 0$, allora X *non* è assolutamente continuo. Basta infatti considerare $u \neq 0$ tale che $Ku = 0$ e si ha $E(\langle u, (X - E(X)) \rangle^2) = \langle u, Ku \rangle = 0$, da cui segue che q.c. $X - E(X)$ appartiene al sottospazio ortogonale a u .

$1_{\{0\}}(\vartheta)$, che non è possibile perché $\varphi(\cdot)$ è una funzione continua. Questo mostra che la successione σ_n^2 è limitata. Dato che $e^{-\sigma_n^2/2} \rightarrow \varphi(1) > 0$, passando ai logaritmi si ottiene che $\sigma_n^2 \rightarrow \sigma^2 := -2 \log \varphi(1) \in [0, \infty)$, dunque σ_n^2 ha limite finito.

Mostriamo ora che anche la successione ν_n è limitata. In caso contrario, dovrebbe esistere una sottosuccessione $\nu_{n_k} \rightarrow +\infty$ (il caso $\nu_{n_k} \rightarrow -\infty$ è analogo). Indichiamo con $f_a(\cdot)$ una funzione continua e limitata tale che $1_{[a+1, \infty)} \leq f_a \leq 1_{[a, \infty)}$; per esempio, possiamo scegliere $f_a(x) := (x - a)1_{[a, a+1)}(x) + 1_{[a+1, \infty)}(x)$. Dalla convergenza debole di Y_n verso Y segue che $E(f_a(Y_n)) \rightarrow E(f_a(Y))$ per ogni $a \in \mathbb{R}$. Ma se $\nu_{n_k} > a + 1$ si ha $E(f_a(Y_{n_k})) \geq E(1_{[a+1, \infty)}(Y_{n_k})) = P(Y_{n_k} \geq a + 1) \geq P(Y_{n_k} \geq \nu_{n_k}) \geq \frac{1}{2}$ (si ricordi che $Y_{n_k} \sim \mathcal{N}(\nu_{n_k}, \sigma_{n_k}^2)$). D'altro canto $E(f_a(Y)) \leq E(1_{[a, \infty)}(Y)) = P(Y \geq a)$, da cui si ottiene $P(Y \geq a) \geq \frac{1}{2}$ per ogni $a \in \mathbb{R}$ e dunque $P(Y = +\infty) = \lim_{m \rightarrow \infty} P(Y \geq m) \geq \frac{1}{2}$, che è impossibile.

La successione ν_n è dunque limitata. Se consideriamo due arbitrarie sottosuccessioni convergenti $\nu_{n'} \rightarrow \nu'$ e $\nu_{n''} \rightarrow \nu''$, dalla relazione (1.12) e dal fatto che $\sigma_n^2 \rightarrow \sigma^2 \in [0, \infty)$ segue che $e^{i\vartheta\nu'} = e^{i\vartheta\nu''}$ per ogni $\vartheta \in \mathbb{R}$; derivando rispetto a ϑ e ponendo $\vartheta = 0$ si ottiene allora $\nu' = \nu''$. Questo mostra che ogni sottosuccessione convergente di ν_n ha lo stesso limite, quindi l'intera successione ν_n converge verso un limite finito, che chiamiamo ν .

Infine, ora che sappiamo che $\nu_n \rightarrow \nu \in (-\infty, +\infty)$ e $\sigma_n^2 \rightarrow \sigma^2 \in [0, \infty)$, segue dalla relazione (1.12) che $E(e^{i\vartheta Y}) =: \varphi(\vartheta) = e^{i\vartheta\nu - \vartheta^2\sigma^2/2}$, ossia $Y \sim \mathcal{N}(\nu, \sigma^2)$. Ricordando che $Y := \langle u, X \rangle$ e $u \in \mathbb{R}^d$ è arbitrario, segue che X è un vettore normale.

Abbiamo mostrato che $E(\langle u, X \rangle) = \lim_{n \rightarrow \infty} E(\langle u, X_n \rangle)$ per ogni $u \in \mathbb{R}^d$, da cui segue in particolare $E(X^{(i)}) = \lim_{n \rightarrow \infty} E(X_n^{(i)})$ per ogni $i = 1, \dots, d$. Analogamente, sappiamo che $\text{Var}(\langle u, X \rangle) = \lim_{n \rightarrow \infty} \text{Var}(\langle u, X_n \rangle)$ per ogni $u \in \mathbb{R}^d$; dato che $\text{Cov}(Z, W) = \frac{1}{4}(\text{Var}(Z + W) - \text{Var}(Z - W))$ per ogni coppia di variabili aleatorie $Z, W \in L^2$, segue che $\text{Cov}(X^{(i)}, X^{(j)}) = \lim_{n \rightarrow \infty} \text{Cov}(X_n^{(i)}, X_n^{(j)})$ per ogni $i, j = 1, \dots, d$. \square

1.8. QUESTIONI TECNICHE

1.8.1. CLASSI DI DYNKIN. A differenza di quanto accade con la topologia, la σ -algebra $\mathcal{E} = \sigma(I)$ generata da una famiglia I non ammette una descrizione esplicita. Si potrebbe pensare di considerare la famiglia $I^{(1)}$ contenente gli elementi di I , i loro complementari e le loro unioni numerabili, ma non è detto che $I^{(1)}$ sia una σ -algebra. Aggiungendo agli elementi di $I^{(1)}$ i loro complementari e le loro unioni numerabili, si ottiene una famiglia più ampia $I^{(2)}$, e iterando la procedura si definisce $I^{(n)}$ per $n \in \mathbb{N}$. Chiaramente $I^{(1)} \subseteq I^{(2)} \subseteq \dots$ e si potrebbe sperare che $\sigma(I) = I^{(n)}$ per qualche $n \in \mathbb{N}$, o per lo meno che $\sigma(I) = \bigcup_{n \in \mathbb{N}} I^{(n)}$. Purtroppo questo è falso in generale: per esempio, quando $E = \mathbb{R}$ e I è la famiglia degli intervalli aperti, $\sigma(I)$ è strettamente più grande di $\bigcup_{n \in \mathbb{N}} I^{(n)}$ (cf. la fine della sezione 2 nel capitolo 1 in [Billingsley, 1995], pagg. 30 e seguenti).

Non essendo disponibile una descrizione esplicita degli elementi di $\mathcal{E} = \sigma(I)$, si rendono necessarie tecniche per dimostrare che una certa proprietà è verificata per ogni elemento di \mathcal{E} . Dato un insieme E , una famiglia $\mathcal{D} \subseteq \mathcal{P}(E)$ di parti di E è detta una *classe di Dynkin* (o anche *classe monotona*) se contiene E e se è chiusa per unioni numerabili disgiunte e per passaggio al complementare:

$$A \in \mathcal{D} \implies A^c \in \mathcal{D}; \quad A_n \in \mathcal{D} \forall n \in \mathbb{N}, \quad A_i \cap A_j = \emptyset \text{ se } i \neq j \implies \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{D}.$$

Una definizione alternativa equivalente è che, oltre a contenere E , la classe \mathcal{D} sia chiusa per differenze proprie e per unioni numerabili crescenti, cioè

$$A, B \in \mathcal{D}, A \subseteq B \implies B \setminus A \in \mathcal{D}; \quad A_n \in \mathcal{D} \forall n \in \mathbb{N}, A_n \subseteq A_{n+1} \implies \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{D}.$$

Si ha allora il seguente risultato (cf. §A1.3 in [Williams, 1991]).

LEMMA 1.17 (DYNKIN). Sia \mathcal{D} una classe di Dynkin su un insieme E e sia $I \subseteq \mathcal{P}(E)$ una famiglia chiusa per intersezioni finite. Se \mathcal{D} contiene I , allora \mathcal{D} contiene $\sigma(I)$.

Di conseguenza, per dimostrare che una certa proprietà è soddisfatta da tutti gli elementi di una σ -algebra \mathcal{E} , si dimostra innanzitutto che gli insiemi che hanno questa proprietà formano una classe di Dynkin, quindi si mostra che questa proprietà è soddisfatta dagli elementi di una base I di \mathcal{E} .

Ricordiamo infine il *lemma di misurabilità di Doob* (cf. Problem 13.3 in [Billingsley, 1995]).

LEMMA 1.18. Siano $X : (E, \mathcal{E}) \rightarrow (F, \mathcal{F})$ e $Y : (E, \mathcal{E}) \rightarrow \mathbb{R}^d$ applicazioni misurabili e sia $\sigma(X)$ la σ -algebra su E generata da X . L'applicazione Y è misurabile rispetto a $\sigma(X)$ se e soltanto se esiste un'applicazione misurabile $g : (F, \mathcal{F}) \rightarrow \mathbb{R}^d$ tale che $Y = g(X)$.

1.8.2. COSTRUZIONE DI SUCCESSIONI INDIPENDENTI. Mostriamo che è sempre possibile costruire una successione di variabili aleatorie reali indipendenti con leggi assegnate $\{\mu_k\}_{k \in \mathbb{N}}$. Utilizzeremo come spazio di probabilità $([0, 1], \mathcal{B}[0, 1], dx)$, dove dx indica la misura di Lebesgue. Per $\omega \in [0, 1]$, indichiamo con $Y_n(\omega)$ l' n -esima cifra nello sviluppo in base due, cioè $\omega = 0.Y_1(\omega)Y_2(\omega)Y_3(\omega)\dots$ con $Y_n(\omega) \in \{0, 1\}$; nei casi ambigui, per es. $0.0\bar{1} = 0.1$, scegliamo lo sviluppo finito. Più precisamente, poniamo per $\omega \in [0, 1]$

$$Y_n(\omega) := \sum_{k=1}^{2^{n-1}} 1_{\left[\frac{2k-1}{2^n}, \frac{2k}{2^n}\right)}(\omega).$$

È facile verificare che le variabili aleatorie $\{Y_n\}_{n \in \mathbb{N}}$ sono indipendenti e hanno la stessa legge $P(Y_n = 0) = P(Y_n = 1) = \frac{1}{2}$. Essendo Y_n l' n -esima cifra nello sviluppo in base due, si ha che per ogni $\omega \in [0, 1]$

$$\sum_{n=1}^{\infty} \frac{Y_n(\omega)}{2^n} = \omega, \quad \text{cioè} \quad \sum_{n=1}^{\infty} \frac{Y_n}{2^n} = id,$$

dove id indica l'identità su $[0, 1]$. In particolare, la variabile aleatoria $Z := \sum_{n \in \mathbb{N}} \frac{Y_n}{2^n}$ è uniformemente distribuita su $[0, 1]$, cioè ha come legge la misura di Lebesgue su $[0, 1]$.

Indicando con $\{p_i\}_{i \in \mathbb{N}}$ la successione dei numeri primi, poniamo per $k \in \mathbb{N}$

$$X_k := \sum_{n=1}^{\infty} \frac{Y_{(p_k)^n}}{2^n}.$$

Per k fissato, le variabili $\{Y_{(p_k)^n}\}_{n \in \mathbb{N}}$ sono i.i.d. con legge marginale $P(Y_{(p_k)^n} = 0) = P(Y_{(p_k)^n} = 1) = \frac{1}{2}$, esattamente come la successione originale $\{Y_n\}_{n \in \mathbb{N}}$. Di conseguenza, la variabile X_k ha la stessa legge della variabile Z , cioè è uniformemente distribuita sull'intervallo $[0, 1]$. (Si noti che non è vero che $\sum_{n=1}^{\infty} \frac{Y_{(p_k)^n}}{2^n} = id$.) Dato che, per p, p' primi distinti, le successioni $\{p^n\}_n$ e $\{(p')^n\}_n$ sono disgiunte, segue che per k_1, \dots, k_n distinti le variabili $\{X_{k_1}, \dots, X_{k_n}\}$ sono indipendenti (intuitivamente, sono costruite a partire da famiglie disgiunte di variabili Y_i). Di conseguenza abbiamo costruito una successione $\{X_k\}_{k \in \mathbb{N}}$ di variabili aleatorie reali indipendenti, ciascuna con legge uniforme sull'intervallo $[0, 1]$.

Infine, basta osservare che una variabile aleatoria reale con legge assegnata μ si può sempre ottenere come immagine di una variabile aleatoria uniforme su $[0, 1]$. Più precisamente, se Z è una variabile aleatoria uniforme su $[0, 1]$ e se $F(x) = \mu((-\infty, x])$ è la funzione di ripartizione di μ , indicando con $G(y) := \inf\{x : F(x) \geq y\}$ la pseudo-inversa di $F(\cdot)$, definita per $y \in [0, 1]$, si verifica facilmente che $G(Z)$ è una variabile aleatoria con funzione di ripartizione $F(\cdot)$, cioè con legge μ . Se indichiamo con $F_k(\cdot)$ la funzione di ripartizione di μ_k e con $G_k(\cdot)$ la corrispondente pseudo-inversa, abbiamo che $\{W_k := G_k(Y_k)\}_{k \in \mathbb{N}}$ è una successione di variabili aleatorie reali indipendenti con leggi marginali $\{\mu_k\}_{k \in \mathbb{N}}$.

2. MOTO BROWNIANO

In questo capitolo sviluppiamo la trattazione matematica del *moto browniano*. Questo processo prende il nome dal botanico scozzese Robert Brown, che nel 1827 descrisse il movimento frenetico dei granelli di polline in sospensione nell'acqua. La teoria fisica del moto browniano fu sviluppata all'inizio del ventesimo secolo indipendentemente da Albert Einstein e Marian Smoluchowski, mentre i pionieri della trattazione matematica sono stati Louis Bachelier, Norbert Wiener e Paul Lévy.

2.1. PROCESSI STOCASTICI

2.1.1. SPAZIO DELLE TRAIETTORIE. Abbiamo visto nel paragrafo 1.1.3 il prodotto di un numero finito di spazi misurabili. Più in generale, è possibile considerare il prodotto di una famiglia di spazi misurabili $\{(E_i, \mathcal{E}_i)\}_{i \in I}$ indicizzata da un insieme di indici I arbitrario, anche infinito. Ci limitiamo per semplicità al caso fondamentale in cui tutti gli spazi sono uguali a uno spazio fissato, ossia $(E_i, \mathcal{E}_i) = (E, \mathcal{E})$ per ogni $i \in I$.

Dato uno spazio misurabile (E, \mathcal{E}) e un insieme I , il *prodotto cartesiano* $E^I = \times_{i \in I} E$ è per definizione lo spazio delle funzioni definite su I a valori in E . Gli elementi di E^I saranno indicati con $x = \{x_i\}_{i \in I}$, dove $x_i \in E$ per ogni $i \in I$. Si dicono *cilindrici* i sottoinsiemi di E^I definiti specificando un numero finito di componenti, ossia quelli della forma $\{x \in E^I : x_{i_1} \in A_{i_1}, \dots, x_{i_k} \in A_{i_k}\}$, al variare di $k \in \mathbb{N}$, $i_1, \dots, i_k \in I$ e $A_{i_1}, \dots, A_{i_k} \in \mathcal{E}$. Indicata con \mathcal{C} la famiglia dei sottoinsiemi cilindrici di E^I , si definisce *σ -algebra prodotto* $\mathcal{E}^I = \otimes_{i \in I} \mathcal{E}$ la più piccola σ -algebra su E^I che contiene i sottoinsiemi cilindrici, ossia $\mathcal{E}^I := \sigma(\mathcal{C})$. Un'applicazione $X = \{X_i\}_{i \in I} : (\Omega, \mathcal{F}) \rightarrow (E^I, \mathcal{E}^I)$ è misurabile se e solo se lo sono le singole componenti $X_i : (\Omega, \mathcal{F}) \rightarrow (E, \mathcal{E})$ per ogni $i \in I$, come si verifica facilmente usando gli insiemi cilindrici.

2.1.2. PROCESSI STOCASTICI. Indichiamo con I un arbitrario insieme di indici (sarà tipicamente un sottoinsieme di \mathbb{R}).

DEFINIZIONE 2.1. Una famiglia di variabili aleatorie $\{X_t\}_{t \in I}$, definite sullo stesso spazio di probabilità (Ω, \mathcal{F}, P) a valori nello stesso spazio misurabile (E, \mathcal{E}) , è detta *processo stocastico* (o semplicemente *processo*). Le leggi dei vettori $(X_{t_1}, \dots, X_{t_k})$ su E^k , al variare di $k \in \mathbb{N}$ e $t_1, \dots, t_k \in I$, sono dette *leggi finito-dimensionali* del processo. Nel caso in cui $E = \mathbb{R}$ (risp. $E = \mathbb{R}^n$), il processo stocastico è detto reale (risp. vettoriale).

Abbiamo visto sopra che un'applicazione a valori nello spazio prodotto (E^I, \mathcal{E}^I) è misurabile se e solo se lo è ciascuna componente. Di conseguenza, oltre che come famiglia di variabili aleatorie, un processo stocastico $X = \{X_t\}_{t \in I}$ a valori in (E, \mathcal{E}) può essere visto come *un'unica variabile aleatoria* $X : (\Omega, \mathcal{F}, P) \rightarrow (E^I, \mathcal{E}^I)$ a valori nello spazio

prodotto (E^I, \mathcal{E}^I) , che è detto talvolta *spazio delle traiettorie* del processo X . Come ogni variabile aleatoria, X induce sullo spazio d'arrivo (E^I, \mathcal{E}^I) la sua legge μ_X : questa probabilità è detta *legge del processo*.

Se $C = \{x \in E^I : x_{t_1} \in A_1, \dots, x_{t_k} \in A_k\}$ è un sottoinsieme cilindrico di E^I , si ha $\mu_X(C) = P(X \in C) = P((X_{t_1}, \dots, X_{t_k}) \in A_1 \times \dots \times A_k)$, dunque la probabilità $\mu_X(C)$ può essere calcolata conoscendo le leggi finito-dimensionali di X . Ricordando che i sottoinsiemi cilindrici sono una base della σ -algebra \mathcal{E}^I , segue che la legge μ_X del processo X sullo spazio delle traiettorie (E^I, \mathcal{E}^I) è determinata dalle leggi finito-dimensionali di X . (Per questa ragione, con il termine legge del processo X si indica talvolta la famiglia delle leggi finito dimensionali.) In particolare, due processi $X = \{X_t\}_{t \in I}$, $X' = \{X'_t\}_{t \in I}$ con lo stesso insieme degli indici I e a valori nello stesso spazio (E, \mathcal{E}) hanno la stessa legge se e solo se hanno le stesse leggi finito-dimensionali.

2.1.3. PROCESSI GAUSSIANI. Un processo vettoriale $X = \{X_t\}_{t \in I}$ a valori in \mathbb{R}^d , con $X_t = (X_t^{(1)}, \dots, X_t^{(d)})$, può essere sempre visto come un processo stocastico reale a patto di ampliare l'insieme degli indici: infatti basta scrivere $X = \{X_t^{(i)}\}_{(i,t) \in \{1, \dots, d\} \times I}$. Per questa ragione, quando risulta conveniente, è possibile limitare la trattazione ai processi reali, senza perdita di generalità. Questo è quello che faremo sempre nel caso dei *processi gaussiani*, che ora definiamo.

DEFINIZIONE 2.2. Un processo stocastico reale $X = \{X_t\}_{t \in I}$ è detto *gaussiano* se, per ogni scelta di $t_1, \dots, t_n \in I$, il vettore aleatorio $(X_{t_1}, \dots, X_{t_n})$ è normale, cioè se qualunque combinazione lineare finita delle X_t è una variabile aleatoria normale.

I processi gaussiani costituiscono una generalizzazione dei vettori aleatori normali. Si noti infatti che, quando $I = \{t_1, \dots, t_k\}$ è un insieme finito, un processo gaussiano $X = \{X_t\}_{t \in I} = (X_{t_1}, \dots, X_{t_k})$ non è altro che un vettore aleatorio normale a valori in \mathbb{R}^k .

Come per i vettori normali, dato un processo gaussiano $X = \{X_t\}_{t \in I}$ introduciamo le funzioni *media* $\mu(t) := E(X_t)$ e *covarianza* $K(s, t) := \text{Cov}(X_s, X_t)$, ben definite in quanto $X_t \in L^2$ per ogni $t \in I$ (perché?). Si noti che la funzione $K(\cdot, \cdot)$ è simmetrica e *semi-definita positiva*, nel senso seguente: per ogni scelta di $n \in \mathbb{N}$, $t_1, \dots, t_n \in I$ e di $u \in \mathbb{R}^n$ si ha $\sum_{i,j=1}^n K(t_i, t_j) u_i u_j \geq 0$; infatti $\{K_{ij} := K(t_i, t_j)\}_{1 \leq i,j \leq n}$ è la matrice di covarianza del vettore $(X_{t_1}, \dots, X_{t_n})$. Si può mostrare (non lo faremo) che, assegnate arbitrariamente due funzioni $\mu : I \rightarrow \mathbb{R}$ e $K : I \times I \rightarrow \mathbb{R}$, con K simmetrica e semi-definita positiva, esiste un processo gaussiano $\{X_t\}_{t \in I}$ che ha μ e K come funzioni media e covarianza.

Una proprietà fondamentale è che le leggi finito-dimensionali di un processo gaussiano sono univocamente determinate dalle sue funzioni media $\mu(\cdot)$ e covarianza $K(\cdot, \cdot)$. Questo segue immediatamente dal fatto che ogni vettore della forma $(X_{t_1}, \dots, X_{t_k})$ è per definizione normale a valori in \mathbb{R}^k e dunque la sua funzione caratteristica, espressa dalla formula (1.9), è una funzione del vettore $(\mu(t_1), \dots, \mu(t_k))$ e della matrice $\{K_{ij} := K(t_i, t_j)\}_{1 \leq i,j \leq k}$.

Anche la proprietà basilare per cui variabili *congiuntamente* normali sono indipendenti se e solo se sono scorrelate, cf. il Lemma 1.14, si estende ai processi gaussiani. Rimandiamo

la formalizzazione precisa di questo risultato alla Proposizione 2.25, dopo che avremo discusso la nozione di σ -algebra associata a un processo.

2.2. IL MOTO BROWNIANO

Ricordiamo l'Osservazione 1.2: fissato uno spazio di probabilità (Ω, \mathcal{F}, P) , scriveremo “q.c. [...]” come abbreviazione di “*esiste* $A \in \mathcal{F}$, con $P(A) = 1$, tale che per ogni $\omega \in A$ [...]”.

Definiamo ora il moto browniano, detto anche *processo di Wiener*, che costituisce l'oggetto centrale di questo corso. Si tratta dell'esempio più importante di processo stocastico a tempo continuo. Esso può essere visto come l'analogo a tempo continuo di una passeggiata aleatoria reale con incrementi gaussiani. In effetti, come discuteremo più avanti, il moto browniano può essere ottenuto come un opportuno limite di qualunque passeggiata aleatoria con incrementi di varianza finita (cf. il sottoparagrafo 2.7.1).

DEFINIZIONE 2.3 (MOTO BROWNIANO). Si dice *moto browniano* qualunque processo stocastico reale $B = \{B_t\}_{t \in [0, \infty)}$ che soddisfa le seguenti proprietà:

- (a) $B_0 = 0$ q.c.;
- (b) B ha *incrementi indipendenti*, cioè per ogni scelta di $k \geq 2$ e $0 \leq t_0 < t_1 < \dots < t_k < \infty$ le variabili aleatorie $\{B_{t_i} - B_{t_{i-1}}\}_{1 \leq i \leq k}$ sono indipendenti;
- (c) B ha *incrementi stazionari gaussiani centrati*: più precisamente, per ogni scelta di $t > s \geq 0$ si ha $(B_t - B_s) \sim \mathcal{N}(0, t - s)$;
- (d) q.c. B ha *traiettorie continue*, cioè q.c. la funzione $t \mapsto B_t$ è continua.

Nella definizione è sottinteso lo spazio di probabilità (Ω, \mathcal{F}, P) su cui è definito il processo B , per cui si ha $B_t = B_t(\omega)$ con $\omega \in \Omega$. La dipendenza da ω verrà quasi sempre omessa, ma è importante essere in grado di esplicitarla quando è necessario. Per esempio, la proprietà (d) si può riformulare nel modo seguente: esiste $A \in \mathcal{F}$ con $P(A) = 1$ tale che per ogni $\omega \in A$ la funzione $t \mapsto B_t(\omega)$ è continua. Oltre a essere una richiesta molto naturale dal punto di vista fisico, la continuità delle traiettorie è una proprietà di basilare importanza anche da un punto di vista matematico (si veda il sottoparagrafo § 2.2.2).

Talvolta parleremo di moto browniano con insieme dei tempi ristretto a un intervallo $\mathbb{T} = [0, t_0]$, dove $t_0 \in (0, \infty)$ è fissato, intendendo naturalmente con ciò un processo $\{B_t\}_{t \in \mathbb{T}}$ che soddisfa le condizioni della Definizione 2.3 per t ristretto a \mathbb{T} .

Nella Figura 2.1 sono mostrate tre traiettorie illustrative del moto browniano.

Veniamo ora al primo risultato fondamentale sul moto browniano, dimostrato per la prima volta da Norbert Wiener nel 1923. A dispetto delle apparenze, si tratta di un risultato non banale.

TEOREMA 2.4 (WIENER). Il moto browniano esiste.

Sono possibili diverse dimostrazioni di questo teorema. Un metodo standard, basato su un teorema molto generale dovuto a Kolmogorov, consiste nel costruire sullo spazio

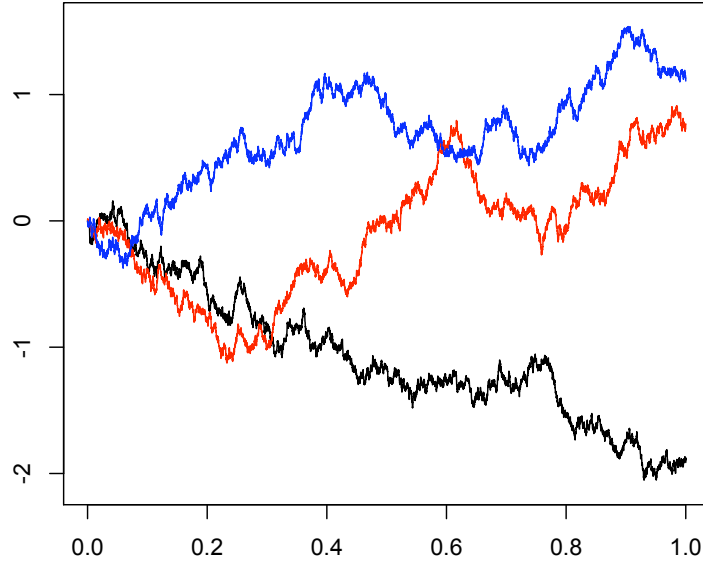


FIGURA 2.1. Tre traiettorie simulate del moto browniano, ottenute mediante interpolazione lineare e riscaldamento diffusivo di 10^4 passi di una passeggiata aleatoria con incrementi gaussiani (si veda il paragrafo 2.7.1). Le scale sui due assi sono diverse.

delle traiettorie $(\mathbb{R}^{[0,\infty)}, \mathcal{B}(\mathbb{R})^{[0,\infty)})$ la legge μ del moto browniano, che è identificata univocamente dalla Definizione 2.3 (cf. l'Osservazione 2.7 più sotto). Questa costruzione fornisce un processo che soddisfa le proprietà (a), (b) e (c) della Definizione 2.3, e con un'opportuna *modificazione* (cf. la Definizione 3.1 nel capitolo 3) si ottiene che anche la proprietà (d) sia soddisfatta (cf. per esempio [Baldi, 2000]). La dimostrazione che presentiamo nel paragrafo 2.3, dovuta a Paul Lévy, è certamente meno generale, ma ha il pregio di essere molto esplicita e relativamente elementare.

OSSERVAZIONE 2.5. Supponiamo di rimpiazzare la condizione (c) nella Definizione 2.3 con la richiesta più debole che gli incrementi siano soltanto *stazionari*, cioè che le variabili $B_{t+h} - B_{s+h}$ e $B_t - B_s$ abbiano la stessa legge per ogni $s, t, h \geq 0$. Si può allora mostrare che il processo risultante è necessariamente della forma $\{a\beta_t + bt\}_{t \in [0,\infty)}$, dove $\{\beta_t\}_{t \in [0,\infty)}$ è un moto browniano secondo la Definizione 2.3 e $a, b \in \mathbb{R}$. In altri termini, a meno di fattori di scala e addizione di una funzione lineare, *il moto browniano è l'unico processo stocastico nullo al tempo zero, con incrementi indipendenti e stazionari e traiettorie q.c. continue*. Si noti che la legge normale non è neppure menzionata in questa caratterizzazione! Per una dimostrazione di questo risultato (non banale), si veda ad esempio il Lemma 1.8 nel capitolo IX in [Asmussen, 2003]. \square

2.2.1. PRIME PROPRIETÀ DEL MOTO BROWNIANO. Per cominciare, forniamo un'espressione esplicita delle leggi finito-dimensionali del moto browniano.

PROPOSIZIONE 2.6. Dato un moto browniano $B = \{B_t\}_{t \in [0, \infty)}$, il vettore aleatorio $(B_{t_1}, \dots, B_{t_k})$ a valori in \mathbb{R}^k è normale, per ogni scelta di $k \in \mathbb{N}$ e $0 \leq t_1 < \dots < t_k < \infty$. Tale vettore è assolutamente continuo se e solo se $t_1 > 0$, nel qual caso la sua densità nel punto $x = (x_1, \dots, x_k) \in \mathbb{R}^k$ è data da

$$f_{t_1, \dots, t_k}^{(k)}(x_1, \dots, x_k) := \frac{\exp \left\{ -\frac{1}{2} \sum_{i=1}^k \left(\frac{x_i^2}{t_i} + \frac{(x_2 - x_1)^2}{t_2 - t_1} + \dots + \frac{(x_k - x_{k-1})^2}{t_k - t_{k-1}} \right) \right\}}{(2\pi)^{k/2} \sqrt{t_1(t_2 - t_1) \cdots (t_k - t_{k-1})}}. \quad (2.1)$$

DIMOSTRAZIONE. Fissiamo $0 \leq t_1 < \dots < t_k < \infty$ e poniamo $Y_i := B_{t_i} - B_{t_{i-1}}$ per $i = 1, \dots, k$ (con $t_0 := 0$). Per le proprietà (b) e (c) della Definizione 2.3, il vettore $Y := (Y_1, \dots, Y_k)$ ha componenti indipendenti e normali (più precisamente $Y_i \sim \mathcal{N}(0, t_i - t_{i-1})$), quindi Y è un vettore aleatorio normale. Dato che $B_{t_i} = Y_1 + \dots + Y_i$, segue che anche il vettore aleatorio $(B_{t_1}, \dots, B_{t_k})$, trasformazione lineare di Y , è normale.

Se $t_1 = 0$ allora $B_{t_1} = 0$ q.c. per la proprietà (a) della Definizione 2.3. Dato che se un vettore aleatorio è assolutamente continuo lo è anche ogni sua componente, grazie al Teorema di Fubini, segue che $(B_{t_1}, \dots, B_{t_k})$ non è assolutamente continuo.

Se $t_1 > 0$, le variabili normali $Y_i \sim \mathcal{N}(0, t_i - t_{i-1})$ hanno varianze strettamente positive e sono quindi assolutamente continue: $f_{Y_i}(y_i) = (2\pi(t_i - t_{i-1}))^{-1/2} \exp(-\frac{1}{2}y_i^2/(t_i - t_{i-1}))$. Essendo indipendenti, segue che il vettore $Y = (Y_1, \dots, Y_k)$ è assolutamente continuo con densità $f_Y(y_1, \dots, y_k) = \prod_{i=1}^k f_{Y_i}(y_i)$, ossia

$$f_Y(y_1, \dots, y_k) = \frac{1}{(2\pi)^{k/2} \sqrt{t_1(t_2 - t_1) \cdots (t_k - t_{k-1})}} \exp \left(-\frac{1}{2} \sum_{i=1}^k \frac{y_i^2}{t_i - t_{i-1}} \right). \quad (2.2)$$

Ricordando che $(B_{t_1}, \dots, B_{t_k}) = F(Y)$ con F lineare, la densità del vettore $(B_{t_1}, \dots, B_{t_k})$ si ottiene applicando la formula del cambiamento di variabili: dato che $Y_i = B_{t_i} - B_{t_{i-1}}$, si ha $\det(F^{-1}) = \det(F) = 1$ e ponendo $y_i = x_i - x_{i-1}$ in (2.2) si ottiene (2.1). \square

OSSERVAZIONE 2.7. La Proposizione 2.6 mostra che la Definizione 2.3 determina le *leggi finito dimensionali* del moto browniano (bastano anzi le proprietà (a), (b) e (c)). Infatti, dalla relazione (2.1) e dal fatto che $B_0 = 0$ q.c. segue che, per ogni scelta di $k \in \mathbb{N}$ e $0 \leq t_1 < \dots < t_k < \infty$, la legge $\mu_{t_1, \dots, t_k}^{(k)}$ del vettore aleatorio $(B_{t_1}, \dots, B_{t_k})$ è data da

$$\mu_{t_1, \dots, t_k}^{(k)}(dx_1, \dots, dx_k) = \begin{cases} f_{t_1, \dots, t_k}^{(k)}(x_1, \dots, x_k) dx_1 \cdots dx_k & \text{se } t_1 > 0 \\ \delta_0(dx_1) \cdot f_{t_2, \dots, t_k}^{(k-1)}(x_2, \dots, x_k) dx_2 \cdots dx_k & \text{se } t_1 = 0 \end{cases}. \quad (2.3)$$

Dato che le leggi finito-dimensionali determinano la legge di un processo, la *Definizione 2.3* identifica univocamente la legge del moto browniano sullo spazio delle traiettorie $\mathbb{R}^{[0, \infty)}$. \square

OSSERVAZIONE 2.8. Le proprietà (a), (b) e (c) della Definizione 2.3 sono proprietà delle leggi finito dimensionali. Di conseguenza, affinché un processo $X = \{X_t\}_{t \geq 0}$ sia un moto browniano è sufficiente (oltre che necessario) che le sue leggi finito-dimensionali siano date da (2.3), (2.1) e che q.c. abbia traiettorie continue. \square

Diamo ora una caratterizzazione alternativa del moto browniano di cruciale importanza.

TEOREMA 2.9. Un processo stocastico reale $B = \{B_t\}_{t \in [0, \infty)}$ è un moto browniano se e soltanto se è un processo gaussiano di media nulla e di covarianza $\text{Cov}(B_s, B_t) = \min\{s, t\}$, con traiettorie q.c. continue.

DIMOSTRAZIONE. Come nella dimostrazione della Proposizione 2.6, per ogni scelta di $0 \leq t_1 < \dots < t_k < \infty$ poniamo $Y_1 := B_{t_1}$ e $Y_i := B_{t_i} - B_{t_{i-1}}$ per $i = 2, \dots, k$. Il vettore $Y := (Y_1, \dots, Y_k)$ ha componenti indipendenti e normali, per le proprietà (b) e (c) della Definizione 2.3, quindi è un vettore aleatorio normale; di conseguenza, anche il vettore aleatorio $(B_{t_1}, \dots, B_{t_k})$, ottenuto da Y mediante una trasformazione lineare, è normale. Questo mostra che B è un processo gaussiano. Dalla proprietà (c) della Definizione 2.3 segue che $B_t \sim \mathcal{N}(0, t)$ e dunque $E(B_t) = 0$ per ogni $t \geq 0$. Per quanto riguarda la covarianza delle variabili B_s e B_t , assumendo senza perdita di generalità che $s < t$ si ha

$$\text{Cov}(B_s, B_t) = \text{Cov}(B_s, (B_t - B_s) + B_s) = \text{Cov}(B_s, B_t - B_s) + \text{Cov}(B_s, B_s) = s,$$

dove si è usato che le variabili B_s e $(B_t - B_s)$ sono indipendenti e che $B_s \sim \mathcal{N}(0, s)$, per le proprietà (b) e (c) della Definizione 2.3, da cui segue che $\text{Cov}(B_s, B_t - B_s) = 0$ e $\text{Cov}(B_s, B_s) = \text{Var}(B_s) = s$.

Viceversa, assumiamo che valgano le ipotesi di questa Proposizione e mostriamo che valgono le proprietà della Definizione 2.3. La proprietà (a) è immediata: B_0 è una variabile normale in quanto componente di un processo gaussiano; dato che $E(B_0) = 0$ e $\text{Var}(B_0) = \text{Cov}(B_0, B_0) = \min\{0, 0\} = 0$, segue che $B_0 \sim \mathcal{N}(0, 0)$ e dunque $B_0 = 0$ q.c.. Anche la proprietà (c) è semplice: sempre dal fatto che B è un processo gaussiano segue che $B_t - B_s$ è una variabile normale, per ogni $s < t$, con media $E(B_t - B_s) = E(B_t) - E(B_s) = 0$ e varianza data da

$$\begin{aligned} \text{Var}(B_t - B_s) &= \text{Cov}(B_t - B_s, B_t - B_s) = \text{Cov}(B_t, B_t) - 2\text{Cov}(B_s, B_t) + \text{Cov}(B_s, B_s) \\ &= t - 2s + s = t - s. \end{aligned}$$

Infine, dati $0 < t_1 < \dots < t_k < \infty$, dall'ipotesi che B è un processo gaussiano segue che il vettore degli incrementi $(B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_k} - B_{t_{k-1}})$ è normale (perché?). Per mostrare che le sue componenti sono indipendenti, basta dunque mostrare che sono a due a due scorrelate. Per $1 \leq i < j \leq k$ si ha

$$\begin{aligned} \text{Cov}(B_{t_j} - B_{t_{j-1}}, B_{t_i} - B_{t_{i-1}}) &= \text{Cov}(B_{t_j}, B_{t_i}) + \text{Cov}(B_{t_{j-1}}, B_{t_{i-1}}) \\ &\quad - \text{Cov}(B_{t_j}, B_{t_{i-1}}) - \text{Cov}(B_{t_{j-1}}, B_{t_i}) = t_i + t_{i-1} - t_{i-1} - t_i = 0, \end{aligned}$$

poiché $t_{i-1} < t_i \leq t_{j-1} < t_j$. La dimostrazione è conclusa. \square

Mostriamo che il moto browniano ha diverse interessanti proprietà di invarianza.

PROPOSIZIONE 2.10. Se $B = \{B_t\}_{t \in [0, \infty)}$ è un moto browniano, anche i seguenti processi $X = \{X_t\}_{t \in \mathbb{T}}$ lo sono:

- (a) $X_t := -B_t$, $\mathbb{T} = [0, \infty)$ (*riflessione spaziale*);
- (b) $X_t := B_{t_0+t} - B_{t_0}$, $\mathbb{T} = [0, \infty)$, per ogni $t_0 \geq 0$ fissato (*traslazione temporale*);
- (c) $X_t := B_{t_0-t} - B_{t_0}$, $\mathbb{T} = [0, t_0]$, per ogni $t_0 \geq 0$ fissato (*riflessione temporale*);
- (d) $X_t := \frac{1}{\sqrt{c}} B_{ct}$, $\mathbb{T} = [0, \infty)$, per ogni $c > 0$ fissato (*riscaldamento diffusivo*);
- (e) $X_t := tB_{1/t}$ per $t > 0$ e $X_0 := 0$, $\mathbb{T} = [0, \infty)$ (*inversione temporale*).

DIMOSTRAZIONE. Conviene utilizzare la caratterizzazione data nel Teorema 2.9. Infatti in tutti i casi è immediato verificare che $\{X_t\}_{t \in [0, \infty)}$ è un processo gaussiano, in quanto le sue componenti sono funzioni lineari delle componenti del processo gaussiano $\{B_t\}_{t \in [0, \infty)}$. Anche le relazioni $E(X_t) = 0$ e $\text{Cov}(X_s, X_t) = \min\{s, t\}$ si verificano facilmente (esercizio). Per esempio, nel caso (e) per $s < t$ si ha

$$\text{Cov}(sB_{1/s}, tB_{1/t}) = st \text{Cov}(B_{1/s}, B_{1/t}) = st \min\left\{\frac{1}{s}, \frac{1}{t}\right\} = st \frac{1}{t} = s.$$

Consideriamo ora la continuità delle traiettorie. Dato che q.c. le traiettorie di $\{B_t\}_{t \in [0, \infty)}$ sono continue, nei casi (a), (b), (c), (d) lo stesso vale per il processo $\{X_t\}_{t \in [0, \infty)}$, le cui traiettorie sono ottenute componendo le traiettorie di $\{B_t\}_{t \in [0, \infty)}$ con funzioni continue.

Resta da verificare la continuità delle traiettorie nel caso (e), per il quale solo la continuità in $t = 0$ non è evidente. In effetti, poiché la funzione $t \mapsto 1/t$ è continua per $t > 0$, q.c. le traiettorie di $\{X_t\}_{t \in [0, \infty)}$ sono continue in $(0, \infty)$; esiste cioè $A \in \mathcal{F}$, con $P(A) = 1$, tale che per ogni $\omega \in A$ la funzione $t \mapsto X_t(\omega)$ è continua in ogni punto $t > 0$. Introduciamo ora l'evento

$$D := \left\{ \omega \in \Omega : \lim_{t \downarrow 0, t \in \mathbb{Q}} X_t(\omega) = 0 \right\} \quad (2.4)$$

e assumiamo per il momento che $P(D) = 1$; di conseguenza anche $P(A \cap D) = 1$. Per costruzione, per ogni $\omega \in A \cap D$ la funzione $f(t) := X_t(\omega)$ è continua in ogni $t \in (0, \infty)$ e inoltre $\lim_{t \downarrow 0, t \in \mathbb{Q}} f(t) = f(0)$. Ma è immediato verificare che ogni funzione $f(t)$ con tali proprietà è necessariamente continua anche in $t = 0$. Infatti, per ogni $\varepsilon > 0$ sia $\delta > 0$ tale che $|f(t) - f(0)| \leq \varepsilon$ per ogni $t \in (0, \delta] \cap \mathbb{Q}$. Preso $t \in (0, \delta] \setminus \mathbb{Q}$, se $\{t_n\}_{n \in \mathbb{N}}$ è una successione in $(0, \delta] \cap \mathbb{Q}$ tale che $t_n \rightarrow t$, si ha $|f(t) - f(0)| = \lim_{n \rightarrow \infty} |f(t_n) - f(0)|$ (perché f è continua in t); dato che $|f(t_n) - f(0)| \leq \varepsilon$ per ogni $n \in \mathbb{N}$ (perché $t_n \in (0, \delta] \cap \mathbb{Q}$) segue che $|f(t) - f(0)| \leq \varepsilon$. In definitiva, si ha $|f(t) - f(0)| \leq \varepsilon$ per ogni $t \in (0, \delta]$, cioè $f(\cdot)$ è continua (anche) in zero. Questo mostra che q.c. $\lim_{t \downarrow 0} X_t = 0$.

Resta solo da mostrare che $P(D) = 1$. Possiamo scrivere $D = \{\omega \in \Omega : X(\omega) \in G\}$, dove $X = X(\omega)$ indica l'intero processo $\{X_t\}_{t \in [0, \infty)}$, pensato come variabile aleatoria a valori nello spazio delle traiettorie $\mathbb{R}^{[0, \infty)}$, mentre $G := \{x \in \mathbb{R}^{[0, \infty)} : \lim_{t \downarrow 0, t \in \mathbb{Q}} x_t = 0\}$. Dato che l'insieme G è definito in termini di una famiglia numerabile di componenti, esso appartiene alla σ -algebra prodotto $\mathcal{B}(\mathbb{R})^{[0, \infty)}$ e dunque $P(D) = P(X \in G) = \mu_X(G)$, dove μ_X indica la legge del processo X su $(\mathbb{R}^{[0, \infty)}, \mathcal{B}(\mathbb{R})^{[0, \infty)})$. Ricordando che le leggi finite

dimensionali di un processo ne caratterizzano la legge, segue in particolare che *il valore della probabilità $P(D)$ è determinato dalle leggi finito-dimensionali del processo X .*

È istruttivo dare una dimostrazione più concreta di questo fatto. Notiamo che $\omega \in D$ se e solo se per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che $|X_{k/n}(\omega)| \leq \varepsilon$ per ogni $k, n \in \mathbb{N}$ con $0 \leq k/n \leq \delta$. Di conseguenza, ponendo $\varepsilon_l := \frac{1}{l}$ e $\delta_m := \frac{1}{m}$, possiamo scrivere

$$D = \bigcap_{l \in \mathbb{N}} \bigcup_{m \in \mathbb{N}} \bigcap_{(n,k) \in \mathbb{N} \times \mathbb{N}: 0 \leq \frac{k}{n} \leq \delta_m} \{|X_{k/n}| \leq \varepsilon_l\}. \quad (2.5)$$

L'intersezione in l è decrescente e l'unione in m crescente, per cui possiamo scrivere

$$\begin{aligned} P(D) &= \lim_{l \rightarrow \infty} \lim_{m \rightarrow \infty} P(|X_{k/n}| \leq \varepsilon_l, \forall n \in \mathbb{N}, \forall k \in \{1, \dots, n\delta_m\}) \\ &= \lim_{l \rightarrow \infty} \lim_{m \rightarrow \infty} \lim_{N \rightarrow \infty} P(|X_{k/n}| \leq \varepsilon_l, \forall n \leq N, \forall k \in \{1, \dots, n\delta_m\}), \end{aligned}$$

avendo usato la monotonia della probabilità. Chiaramente l'ultima probabilità si esprime mediante le leggi finito-dimensionali del processo X .

Ma le leggi finito-dimensionali del processo X coincidono con quelle moto browniano $\{B_t\}_{t \in [0, \infty)}$, perché entrambi i processi sono gaussiani e hanno le stesse media e covarianza. Di conseguenza, la probabilità dell'evento D non cambia se nella sua definizione (2.4) si sostituisce il processo $\{X_t\}_{t \in [0, \infty)}$ con il moto browniano $\{B_t\}_{t \in [0, \infty)}$. Dato che q.c. le traiettorie di $\{B_t\}_{t \in [0, \infty)}$ sono continue in zero per definizione, segue che $P(D) = 1$, come dovevasi dimostrare. \square

Come semplice corollario, otteniamo un risultato interessante.

TEOREMA 2.11 (LEGGE DEI GRANDI NUMERI PER IL MOTO BROWNIANO). Se $\{B_t\}_{t \in [0, \infty)}$ è un moto browniano, q.c. si ha $\lim_{t \rightarrow \infty} B_t/t = 0$.

DIMOSTRAZIONE. Definendo $X_s := sB_{1/s}$ per $s > 0$ e $X_0 := 0$, il processo $\{X_s\}_{s \in [0, \infty)}$ è un moto browniano per il punto (e) della Proposizione 2.10. Per definizione di moto browniano, q.c. si ha $\lim_{s \rightarrow 0} X_s = 0$ e ponendo $s = 1/t$ possiamo riscrivere questa relazione come q.c. $\lim_{t \rightarrow \infty} \frac{1}{t} B_t = 0$. \square

2.2.2. CONTINUITÀ DELLE TRAIETTORIE E COMPLETEZZA. Dato un processo reale $B = \{B_t\}_{t \geq 0}$, definito su uno spazio di probabilità (Ω, \mathcal{F}, P) , poniamo

$$C := \{\omega \in \Omega : \text{la funzione } t \mapsto B_t(\omega) \text{ è continua}\}. \quad (2.6)$$

Ricordando la Definizione 2.3 del moto browniano, si potrebbe essere tentati di riformulare la proprietà (d) come $P(C) = 1$. Questo tuttavia non è corretto: infatti l'insieme C è definito in termini di una famiglia più che numerabile di variabili aleatorie e di conseguenza *in generale* non è detto che $C \in \mathcal{F}$. Una riformulazione corretta della proprietà (d) consiste nel richiedere che C *contenga* un evento $A \in \mathcal{F}$ tale che $P(A) = 1$.

In un certo senso, questa è una sottigliezza. Basta infatti supporre che lo spazio di probabilità (Ω, \mathcal{F}, P) sia *completo* (eventualmente provvedendo a completarlo, come

descritto nel paragrafo 1.2.3 del capitolo 1) e dall'informazione che C contiene un evento quasi certo segue automaticamente che $C \in \mathcal{F}$, per cui è lecito scrivere $P(C) = 1$.

La continuità delle traiettorie e la completezza dello spazio sono collegate ad altre questioni interessanti legate alla misurabilità. Per fissare le idee, supponiamo che su uno spazio di probabilità (Ω, \mathcal{F}, P) sia definito un processo stocastico reale $B = \{B_t\}_{t \in [0, \infty)}$. È naturale interessarsi a funzioni quali

$$\sup_{0 \leq t \leq 1} |B_t(\omega)|, \quad \int_0^1 B_t(\omega) dt, \quad \inf\{t > 0 : B_t(\omega) = 0\}, \quad (2.7)$$

ma *in generale* non c'è alcuna ragione per cui queste espressioni, definite in funzione di una quantità più che numerabile di variabili B_t , siano funzioni misurabili da Ω in \mathbb{R} . L'integrale non è nemmeno ben definito, se non si hanno informazioni sulla misurabilità della funzione $t \mapsto B_t(\omega)$.

È a questo proposito che la continuità delle traiettorie di B assume grande importanza. Infatti, per ogni ω per cui $t \mapsto B_t(\omega)$ è continua, cioè per ogni $\omega \in C$, definito in (2.6), possiamo riscrivere le quantità in (2.7) rispettivamente come

$$\begin{aligned} \sup_{t \in [0, 1] \cap \mathbb{Q}} |B_t(\omega)|, \quad \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} B_{k/n}(\omega), \\ \limsup_{n \rightarrow \infty} \limsup_{k \rightarrow \infty} \inf \left\{ t \in \left[\frac{1}{n}, \infty \right) \cap \mathbb{Q} : |B_t(\omega)| \leq \frac{1}{k} \right\}, \end{aligned} \quad (2.8)$$

avendo usato per l'integrale l'approssimazione mediante somme di Riemann. Queste nuove espressioni sono ben definite per ogni $\omega \in \Omega$ e determinano funzioni *misurabili* da Ω in \mathbb{R} , dunque variabili aleatorie, perché coinvolgono una quantità numerabile di variabili B_t .

Di conseguenza, se il processo B ha traiettorie q.c. continue, le espressioni in (2.7) sono q.c. uguali alle variabili aleatorie in (2.8). Se supponiamo inoltre che lo spazio (Ω, \mathcal{F}, P) sia completo, le espressioni in (2.7) sono esse stesse variabili aleatorie (una volta definito l'integrale per ogni $\omega \in \Omega$, per esempio ponendolo uguale a zero per $\omega \notin C$).

Queste sono alcune delle ragioni per cui ci interesseremo sempre alla continuità delle traiettorie per i processi che incontreremo nel seguito e, quando risulta conveniente, assumeremo che lo spazio di probabilità su cui lavoriamo sia completo.

2.2.3. ANCORA SULLA CONTINUITÀ DELLE TRAIETTORIE. Ci si può chiedere se la proprietà (d) nella Definizione 2.3 di moto browniano non sia una conseguenza delle proprietà precedenti. In altre parole, se un processo $X = \{X_t\}_{t \in [0, \infty)}$ definito su (Ω, \mathcal{F}, P) soddisfa le proprietà (a), (b), (c), esiste necessariamente $A \in \mathcal{F}$, con $P(A) = 1$, tale che per ogni $\omega \in A$ la traiettoria $t \mapsto X_t(\omega)$ sia continua? *La risposta è negativa.* Infatti è possibile costruire sullo stesso spazio (Ω, \mathcal{F}, P) un processo $B' = \{B'_t\}_{t \in [0, \infty)}$ con traiettorie q.c. discontinue ma con le stesse leggi finito-dimensionali di B , cf. l'Esempio 3.2 e l'Osservazione 3.3 nel capitolo 3. Di conseguenza, B' soddisfa le proprietà (a), (b), (c) della Definizione 2.3 che, come abbiamo già ricordato, sono proprietà delle leggi finito-dimensionali. Questo mostra che la continuità delle traiettorie è una proprietà che va oltre le leggi finito-dimensionali. Ritorneremo su questo tipo di problemi nel prossimo capitolo.

2.3. ESISTENZA DEL MOTO BROWNIANO

«[The construction of Brownian motion], like the birth of a child, is messy and painful, but after a while we will be able to have fun with our new arrival.»

Richard Durrett

Dimostriamo ora il Teorema 2.4 per mezzo di una costruzione proposta da Paul Lévy (seguiamo parzialmente la trattazione in [Karatzas e Shreve, 1998, § 2.3]). Per semplicità, ci limitiamo a costruire un moto browniano $\{B_t\}_{t \in [0,1]}$ in cui l'insieme dei tempi è ristretto all'intervallo $[0, 1]$. Per ottenere un moto browniano con insieme dei tempi $[0, \infty)$, è sufficiente considerare una successione indipendente di moti browniani con insieme dei tempi $[0, 1]$ e “incollarli uno dopo l'altro” usando la proprietà (b) della Proposizione 2.10, come mostra l'esercizio seguente.

ESERCIZIO 2.12. Siano $\{B^{(m)} = \{B_t^{(m)}\}_{t \in [0,1]}\}_{m \in \mathbb{N}_0}$ moti browniani indipendenti con insieme dei tempi $[0, 1]$. Si mostri che definendo $B_t := B_{t - [t]}^{([t])} + \sum_{m=0}^{[t]-1} B_1^{(m)}$ per $t \in [0, \infty)$ (la somma è per definizione nulla se $[t] = 0$) si ottiene un moto browniano $B = \{B_t\}_{t \in [0, \infty)}$.

Sia $\{\xi_k^{(n)}\}_{n,k \in \mathbb{N}_0}$ una famiglia (numerabile) di variabili aleatorie reali indipendenti normali standard, definite su un opportuno spazio di probabilità (Ω, \mathcal{F}, P) .[†] Costruiremo il moto browniano su questo spazio di probabilità. Definiamo $\tau_n := \{\frac{k}{2^n} : 0 \leq k \leq 2^n\}$. Si noti che $\tau_n \supseteq \tau_{n-1}$ e $\tau_n \setminus \tau_{n-1} = \{\frac{k}{2^n} : k \in I(n)\}$, dove $I(n)$ indica, per $n \in \mathbb{N}_0$, l'insieme degli interi dispari compresi tra 0 e 2^n : $I(0) = I(1) = \{1\}$, $I(2) = \{1, 3\}$, ecc.. L'insieme $\tau := \bigcup_{n \geq 0} \tau_n$ è costituito dai *razionali diadici*, il cui denominatore (una volta ridotta la frazione ai minimi termini) è una potenza di 2. Si osservi che τ è denso in $[0, 1]$.

L'idea è di costruire una successione di processi $B^{(n)} = \{B_t^{(n)}\}_{t \in [0,1]}$ che converge per $n \rightarrow \infty$ verso un processo limite $\{B_t\}_{t \in [0,1]}$, che sarà il moto browniano cercato. Fissati $n \in \mathbb{N}_0$ e $\omega \in \Omega$, la traiettoria $\{B_t^{(n)}(\omega)\}_{t \in [0,1]}$ sarà lineare a tratti: più precisamente, sarà innanzitutto definita sui punti del reticolo τ_n e verrà poi prolungata su tutto l'intervallo $[0, 1]$ mediante interpolazione lineare. Inoltre, i valori di $B_t^{(n)}$ per $t \in \tau_n$ estenderanno i valori di $B_t^{(n-1)}$: porremo cioè $B_t^{(n)} := B_t^{(n-1)}$ per $t \in \tau_{n-1} \subseteq \tau_n$, per cui resterà solo da definire $B_t^{(n)}$ per $t \in \tau_n \setminus \tau_{n-1} = \{\frac{k}{2^n} : k \in I(n)\}$ (si veda la Figura 2.2).

In altri termini, ad ogni passo definiremo il valore $B_{k/2^n}^{(n)}$ nel punto medio $\frac{k}{2^n}$ di ogni intervallo $[\frac{k-1}{2^n}, \frac{k+1}{2^n}]$, con $k \in I(n)$, avendone già definito i valori $B_{(k-1)/2^n}^{(n)}$ e $B_{(k+1)/2^n}^{(n)}$ negli estremi dell'intervallo. Anche se non è strettamente necessario per la dimostrazione, per motivare la definizione che daremo è utile porsi la seguente domanda: dato un moto browniano B e fissati gli istanti $s < t$, conoscendo i valori B_s e B_t , qual è la distribuzione (condizionale) di $B_{(s+t)/2}$? La risposta è $\mathcal{N}(\mu, \frac{t-s}{4})$ dove $\mu = \frac{1}{2}(B_s + B_t)$, come mostra esplicitamente l'esercizio seguente.

[†]Per esempio, è possibile scegliere come spazio di probabilità l'intervallo $[0, 1]$ munito della misura di Lebesgue, come spiegato nel sottoparagrafo 1.8.2 del capitolo 1.

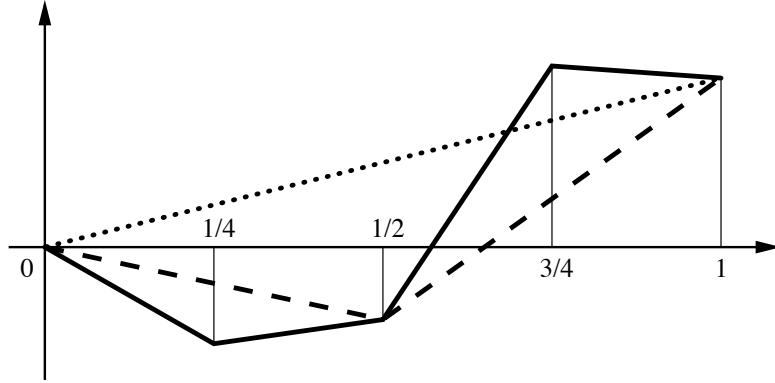


FIGURA 2.2. Un esempio di traiettoria di $B^{(0)}$ (linea puntata), $B^{(1)}$ (linea tratteggiata) e $B^{(2)}$ (linea piena).

ESERCIZIO 2.13. Sia $B = \{B_t\}_{t \in [0, \infty)}$ un moto browniano e siano $0 \leq s < t < \infty$ istanti fissati. Definendo $Z := \frac{2}{\sqrt{t-s}}(B_{(s+t)/2} - \frac{1}{2}(B_s + B_t))$, possiamo scrivere

$$B_{(s+t)/2} = \frac{B_s + B_t}{2} + \frac{\sqrt{t-s}}{2}Z. \quad (2.9)$$

Si mostri che Z ha legge $\mathcal{N}(0, 1)$ ed è scorrelata, quindi indipendente (perché?), dalle variabili aleatorie B_s, B_t .

Passiamo finalmente alla definizione dei processi $B^{(n)} = \{B_t^{(n)}\}_{t \in [0, 1]}$. Cominciamo a definire $B_t^{(0)}$ per $t \in \tau_0 = \{0, 1\}$, ponendo

$$B_0^{(0)} := 0, \quad B_1^{(0)} := \xi_1^{(0)} \quad (\text{cioè } B_1^{(0)}(\omega) := \xi_1^{(0)}(\omega)), \quad (2.10)$$

ed estendiamo poi $B_t^{(0)}$ per ogni $t \in [0, 1]$ mediante interpolazione lineare.

Per chiarezza, trattiamo esplicitamente anche il caso $n = 1$. Cominciamo definendo $B_t^{(1)}$ per $t \in \tau_1 = \{0, \frac{1}{2}, 1\}$. Non modifichiamo i valori già assegnati per $t \in \tau_0 = \{0, 1\}$, poniamo cioè $B_0^{(1)} := B_0^{(0)}$ e $B_1^{(1)} := B_1^{(0)}$. Resta solo da definire $B_{1/2}^{(1)}$: prendendo ispirazione dalla relazione (2.9), poniamo

$$B_{1/2}^{(1)} := \frac{B_0^{(1)} + B_1^{(1)}}{2} + \frac{1}{2}\xi_1^{(1)}. \quad (2.11)$$

Si noti che $\frac{1}{2}(B_0^{(1)} + B_1^{(1)}) = \frac{1}{2}(B_0^{(0)} + B_1^{(0)}) = B_{1/2}^{(0)}$, perché $B_t^{(0)}$ è per definizione lineare per $t \in [0, 1]$. Possiamo quindi riformulare (2.11) come

$$B_{1/2}^{(1)} := B_{1/2}^{(0)} + \frac{1}{2}\xi_1^{(1)}. \quad (2.12)$$

Estendiamo quindi la definizione di $B_t^{(1)}$ a ogni $t \in [0, 1]$ mediante interpolazione lineare.

Possiamo ora procedere in modo ricorsivo: una volta che $B^{(n-1)}$ è stato costruito, definiamo $B^{(n)}$ ponendo $B_t^{(n)} := B_t^{(n-1)}$ per $t \in \tau_{n-1}$, mentre per $t \in \tau_n \setminus \tau_{n-1} = \{\frac{k}{2^n} : k \in I(n)\}$ definiamo $B_t^{(n)}$ in accordo con (2.9): più precisamente, poniamo

$$B_{k/2^n}^{(n)} := \frac{B_{(k-1)/2^n}^{(n)} + B_{(k+1)/2^n}^{(n)}}{2} + \frac{1}{2^{(n+1)/2}} \xi_k^{(n)}, \quad \forall k \in I(n). \quad (2.13)$$

Per costruzione $B_{(k\pm 1)/2^n}^{(n)} = B_{(k\pm 1)/2^n}^{(n-1)}$, perché $\frac{k\pm 1}{2^n} \in \tau_{n-1}$ per $k \in I(n)$. Dato che $B_t^{(n-1)}$ è lineare nell'intervallo $[\frac{k-1}{2^n}, \frac{k+1}{2^n}]$, possiamo scrivere

$$B_{k/2^n}^{(n)} = B_{k/2^n}^{(n-1)} + \frac{1}{2^{(n+1)/2}} \xi_k^{(n)}, \quad \forall k \in I(n). \quad (2.14)$$

Estendiamo quindi $B_t^{(n)}$ per ogni $t \in [0, 1]$ mediante interpolazione lineare. Questo completa la definizione di $\{B_t^{(n)}\}_{t \in [0, 1]}$.

Prima di procedere, è conveniente dare una descrizione più esplicita del procedimento di interpolazione lineare usato nella definizione di $B_t^{(n)}$. Per costruzione, per ogni $n \in \mathbb{N}$, la traiettoria $t \mapsto \Delta_t^{(n)} := B_t^{(n)} - B_t^{(n-1)}$ è lineare tra ogni coppia di punti consecutivi del reticolo τ_n e si annulla per ogni punto $t \in \tau_{n-1} \subseteq \tau_n$. Ricordando che $\tau_n \setminus \tau_{n-1} = \{\frac{k}{2^n} : k \in I(n)\}$, ciò significa che, per ogni $k \in I(n)$, il grafico di $\Delta_t^{(n)}$ per $t \in [\frac{k-1}{2^n}, \frac{k+1}{2^n}]$ è un triangolo isoscele, nullo agli estremi $\frac{(k\pm 1)}{2^n}$ dell'intervallo e con vertice in corrispondenza del punto medio $\frac{k}{2^n}$, di altezza pari a $\frac{1}{2^{(n+1)/2}} \xi_k^{(n)}$, cf. (2.14).

Definiamo allora, per $n \in \mathbb{N}$ e $k \in I(n)$, la *funzione di Schauder* $S_k^{(n)}(\cdot) : [0, 1] \rightarrow \mathbb{R}$ il cui grafico è un triangolo isoscele di base $[\frac{k-1}{2^n}, \frac{k+1}{2^n}]$ e di altezza $\frac{1}{2^{(n+1)/2}}$ (cf. la Figura 2.3). Volendo essere formali:

$$S_k^{(n)}(t) := \begin{cases} 0 & \text{se } 0 \leq t \leq \frac{k-1}{2^n} \\ 2^{(n-1)/2}(t - \frac{k-1}{2^n}) & \text{se } \frac{k-1}{2^n} \leq t \leq \frac{k}{2^n} \\ 2^{(n-1)/2}(\frac{k+1}{2^n} - t) & \text{se } \frac{k}{2^n} \leq t \leq \frac{k+1}{2^n} \\ 0 & \text{se } \frac{k+1}{2^n} \leq t \leq 1 \end{cases}, \quad n \in \mathbb{N}, k \in I(n). \quad (2.15)$$

Per le considerazioni sopra esposte, possiamo scrivere $\Delta_t^{(n)} := B_t^{(n)} - B_t^{(n-1)} = S_k^{(n)}(t) \cdot \xi_k^{(n)}$ per ogni $k \in I(n)$ e per ogni $t \in [\frac{k-1}{2^n}, \frac{k+1}{2^n}]$. Dato che le funzioni $S_k^{(n)}(\cdot)$ e $S_{k'}^{(n)}(\cdot)$ hanno chiaramente supporti disgiunti per $k \neq k'$, otteniamo la seguente relazione:

$$B_t^{(n)} - B_t^{(n-1)} = \sum_{k \in I(n)} \xi_k^{(n)} S_k^{(n)}(t), \quad \forall t \in [0, 1], \forall n \in \mathbb{N}. \quad (2.16)$$

Infatti ogni $t \in [0, 1]$ cade al più in un solo intervallo della forma $(\frac{k-1}{2^n}, \frac{k+1}{2^n})$ con $k \in I(n)$, di conseguenza non più di un termine della somma è diverso da zero per ogni $t \in [0, 1]$.

Per completare la costruzione, introduciamo un'ultima funzione di Schauder

$$S_1^{(0)}(t) := t \mathbf{1}_{[0, 1]}(t), \quad (2.17)$$

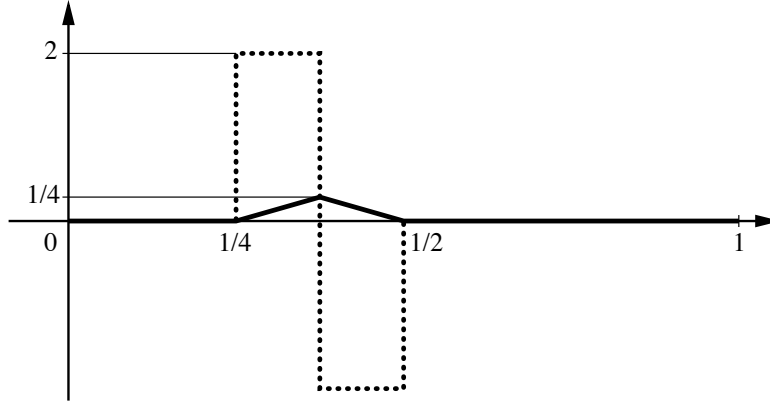


FIGURA 2.3. Grafico delle funzioni $S_3^{(3)}(t)$ (linea piena) e $H_3^{(3)}(t)$ (linea puntata), definite in (2.15) e (2.22). Le unità di misura sui due assi sono diverse.

così che per (2.10) possiamo scrivere $B_t^{(0)} = \xi_1^{(0)} S_1^{(0)}(t)$. Ricordando (2.16) ed esplicitando la dipendenza da $\omega \in \Omega$, otteniamo infine la seguente espressione:

$$B_t^{(n)}(\omega) = \sum_{m=0}^n \sum_{k \in I(m)} \xi_k^{(m)}(\omega) S_k^{(m)}(t), \quad \forall n \in \mathbb{N}_0, \quad t \in [0, 1], \quad \omega \in \Omega. \quad (2.18)$$

Questa è la formula chiave. In effetti, avremmo potuto *definire* $B_t^{(n)}(\omega)$ direttamente in questo modo, senza alcun riferimento alla costruzione grafica sopra descritta.

Siamo giunti al cuore della dimostrazione. Mostriamo ora che q.c. il processo $t \mapsto B_t^{(n)}$ converge per $n \rightarrow \infty$ verso un processo limite $t \mapsto B_t$, che sarà il moto browniano cercato.

LEMMA 2.14. Esiste un evento $A \in \mathcal{F}$ con $P(A) = 1$ tale che, per ogni $\omega \in A$, la successione di funzioni $\{t \mapsto B_t^{(n)}(\omega)\}_{n \in \mathbb{N}}$ converge per $n \rightarrow \infty$ verso una funzione continua, che indicheremo con $\{t \mapsto B_t(\omega)\}$.

Prima di passare alla dimostrazione, ricordiamo che lo spazio $C([0, 1], \mathbb{R})$ delle funzioni continue definite su $[0, 1]$ a valori in \mathbb{R} , munito della norma uniforme $\|f\|_\infty := \sup_{t \in [0, 1]} |f(t)|$, è uno *spazio di Banach*. Ciò significa che, data una successione $\{f_n\}_{n \in \mathbb{N}}$ in $C([0, 1], \mathbb{R})$ che sia *di Cauchy* (per ogni $\varepsilon > 0$ esiste $n_0 < \infty$ tale che $\|f_n - f_k\|_\infty < \varepsilon$ per ogni $n, k \geq n_0$), la successione converge, cioè esiste $f \in C([0, 1], \mathbb{R})$ tale che $\|f_n - f\|_\infty \rightarrow 0$ per $n \rightarrow \infty$. Sottolineiamo che la funzione limite f è *continua*.

Ci interesseremo in particolare a successioni $\{f_n\}_{n \in \mathbb{N}}$ della forma $f_n(t) = \sum_{m=0}^n g_m(t)$, dove $g_n \in C([0, 1], \mathbb{R})$ per ogni $n \in \mathbb{N}_0$. Usando la disuguaglianza triangolare, otteniamo la semplice stima $\|f_n - f_k\|_\infty \leq \sum_{m=k+1}^n \|g_m\|_\infty$, valida per ogni $n > k$. Da ciò segue che, se la serie delle norme $\sum_{m=0}^\infty \|g_m\|_\infty$ è convergente, la successione $\{f_n\}_{n \in \mathbb{N}}$ è di Cauchy in $C([0, 1], \mathbb{R})$ e dunque, per quanto detto sopra, ha limite in $C([0, 1], \mathbb{R})$.

DIMOSTRAZIONE DEL LEMMA 2.14. Se $Z \sim \mathcal{N}(0, 1)$, per $a \geq 1$ si ha la stima

$$P(|Z| > a) = 2 \int_a^\infty \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx \leq 2 \int_a^\infty \frac{x e^{-x^2/2}}{\sqrt{2\pi}} dx = \frac{2}{\sqrt{2\pi}} e^{-a^2/2} \leq e^{-a^2/2},$$

poiché $1 \leq a \leq x$ nel dominio di integrazione. Poniamo ora $\Xi_n(\omega) := \max_{k \in I(n)} |\xi_k^{(n)}(\omega)|$. Per $n \in \mathbb{N}$ possiamo scrivere

$$P(\Xi_n > n) = P\left(\bigcup_{k \in I(n)} \{|\xi_k^{(n)}| > n\}\right) \leq \sum_{k \in I(n)} P(|\xi_k^{(n)}| > n) \leq \frac{2^n}{2} e^{-n^2/2},$$

quindi $\sum_{n \geq 0} P(\Xi_n > n) < \infty$. Introducendo l'evento $A := (\limsup_n \{\Xi_n > n\})^c$, per il lemma di Borel-Cantelli si ha $P(A) = 1$; inoltre, per definizione di A , per ogni $\omega \in A$ si ha $\Xi_n(\omega) > n$ solo per un numero finito di $n \in \mathbb{N}$, cioè esiste $n_0(\omega) < \infty$ tale che $\Xi_n(\omega) \leq n$ per ogni $n \geq n_0(\omega)$.

D'ora in avanti fissiamo $\omega \in A$. La relazione (2.18) si può riscrivere come

$$B_t^{(n)}(\omega) = \sum_{m=0}^n g^{(m)}(\omega, t), \quad \text{dove} \quad g^{(m)}(\omega, t) := \sum_{k \in I(m)} \xi_k^{(m)}(\omega) S_k^{(m)}(t),$$

e osserviamo che $t \mapsto g^{(m)}(\omega, t)$ è una funzione continua, per ogni $m \in \mathbb{N}_0$. Mostriamo ora la convergenza della serie $\sum_{m=0}^\infty \|g^{(m)}(\omega, \cdot)\|_\infty$. Abbiamo già notato che, per ogni $m \in \mathbb{N}_0$ fissato, le funzioni $\{S_k^{(m)}(\cdot)\}_{k \in I(m)}$ hanno supporti disgiunti, cioè per ogni $t \in [0, 1]$ esiste al più un solo $k \in I(m)$ tale che $S_k^{(m)}(t) \neq 0$ (cf. (2.15)). Dato che per ogni $m \geq 1$ si ha $\|S_k^{(m)}\|_\infty = S_k^{(m)}(\frac{k}{2^m}) = 2^{-(m+1)/2}$, segue che $\|\sum_{k \in I(m)} S_k^{(m)}(\cdot)\|_\infty = 2^{-(m+1)/2}$. Ricordando che $\Xi_m(\omega) \leq m$ per $m \geq n_0(\omega)$, possiamo dunque stimare

$$\begin{aligned} \sum_{m=n_0(\omega)}^\infty \|g^{(m)}(\omega, \cdot)\|_\infty &= \sum_{m=n_0(\omega)}^\infty \left\| \sum_{k \in I(m)} \xi_k^{(m)}(\omega) S_k^{(m)}(t) \right\|_\infty \\ &\leq \sum_{m=n_0(\omega)}^\infty \Xi_m(\omega) \left\| \sum_{k \in I(m)} S_k^{(m)}(t) \right\|_\infty \leq \sum_{m=n_0(\omega)}^\infty m \cdot \frac{1}{2^{(m+1)/2}} < \infty, \end{aligned}$$

da cui discende che anche l'intera serie delle norme $\sum_{m=0}^\infty \|g^{(m)}(\omega, \cdot)\|_\infty$ converge (abbiamo tralasciato un numero finito $n_0(\omega)$ di termini). Per quanto detto sopra segue allora che, per ogni $\omega \in A$, la successione di funzioni continue $\{t \mapsto B_t^{(n)}(\omega)\}_{n \in \mathbb{N}}$ converge uniformemente per $n \rightarrow \infty$ verso una funzione limite continua, che indicheremo con $t \mapsto B_t(\omega)$, data da

$$B_t(\omega) = \lim_{n \rightarrow \infty} B_t^{(n)}(\omega) = \sum_{m=0}^\infty \sum_{k \in I(m)} \xi_k^{(m)}(\omega) S_k^{(m)}(t).$$

Definendo per completezza $B_t(\omega) := 0$ quando $\omega \notin A$, la dimostrazione è conclusa. \square

Resta infine da dimostrare che il processo ottenuto $\{B_t\}_{t \in [0,1]}$ è un moto browniano. Grazie al Teorema 2.9, basta mostrare che $\{B_t\}_{t \in [0,1]}$ è un processo gaussiano con $E(B_t) = 0$ e $\text{Cov}(B_s, B_t) = \min\{s, t\}$. Si noti che, per costruzione, già sappiamo che $\{B_t\}_{t \in [0,1]}$ ha traiettorie continue.

Per verificare che $B = \{B_t\}_{t \in [0,1]}$ è un processo gaussiano, basta mostrare che ogni combinazione lineare finita $Z := \vartheta_1 B_{t_1} + \dots + \vartheta_k B_{t_k}$ di componenti di B è normale. Sappiamo che $B_t^{(n)} \rightarrow B_t$ q.c. per $n \rightarrow \infty$, per ogni $t \in [0, 1]$, quindi $Z = \lim_{n \rightarrow \infty} Z^{(n)}$ q.c., dove $Z^{(n)} := \vartheta_1 B_{t_1}^{(n)} + \dots + \vartheta_k B_{t_k}^{(n)}$. Dall'equazione (2.18) è chiaro che $B_t^{(n)}$, e dunque anche $Z^{(n)}$, è una combinazione lineare finita delle variabili normali indipendenti $\{\xi_k^{(m)}\}_{0 \leq m \leq n, k \in I(m)}$, quindi $Z^{(n)}$ è normale. Segue allora dalla Proposizione 1.10 che anche Z è normale, come limite quasi certo di variabili normali.

Sempre dall'equazione (2.18) è chiaro che $E(B_t^{(n)}) = 0$ per ogni $t \in [0, 1]$ e $n \in \mathbb{N}$, poiché $E(\xi_k^{(m)}) = 0$. Dato che $B_t^{(n)}$ è normale e $B_t^{(n)} \rightarrow B_t$ q.c., deduciamo dalla Proposizione 1.10 che $E(B_t) = \lim_{n \rightarrow \infty} E(B_t^{(n)}) = 0$ per ogni $t \in [0, 1]$.

Mostriamo infine che $\text{Cov}(B_s, B_t) = \min\{s, t\}$. La dimostrazione che forniamo è esplicita e diretta, ma un po' tecnica. Una dimostrazione più veloce, che sfrutta l'*identità di Parseval*, è descritta nell'Osservazione 2.15 più sotto.

Supponiamo di aver dimostrato che $\text{Cov}(B_s, B_t) = \min\{s, t\}$ per ogni $s, t \in \tau$ (o, più in generale, in un generico sottoinsieme denso di $[0, 1]$). Dati $s, t \in [0, 1]$ arbitrari, siano $\{s_n\}_{n \in \mathbb{N}}$ e $\{t_n\}_{n \in \mathbb{N}}$ successioni in τ che convergono verso s e t rispettivamente. Per la continuità delle traiettorie del processo B , si ha la convergenza q.c. del vettore aleatorio normale (B_{s_n}, B_{t_n}) verso (B_s, B_t) . Grazie alla Proposizione 1.10 si ha dunque $\text{Cov}(B_s, B_t) = \lim_{n \rightarrow \infty} \text{Cov}(B_{s_n}, B_{t_n}) = \lim_{n \rightarrow \infty} \min\{s_n, t_n\} = \min\{s, t\}$.

Resta solo da mostrare che $\text{Cov}(B_s, B_t) = \min\{s, t\}$ per ogni $s, t \in \tau = \bigcup_{n \in \mathbb{N}} \tau_n$. Senza perdita di generalità, sia $n \in \mathbb{N}_0$ tale che $s, t \in \tau_n$ e assumiamo che $s \leq t$. Si ha allora $B_s = B_s^{(n)}$ e $B_t = B_t^{(n)}$, perché per costruzione, per ogni $u \in \tau_n$, il valore di $B_u^{(k)}(\omega)$ è costante per $k \geq n$. Introduciamo le variabili $Y_i^{(n)} := B_{i/2^n}^{(n)} - B_{(i-1)/2^n}^{(n)}$ per $1 \leq i \leq 2^n$. Mostriamo tra poco che $\text{Cov}(Y_i^{(n)}, Y_j^{(n)}) = 2^{-n} \delta_{ij}$. Scrivendo $s = k/2^n$ e $t = m/2^n$, con $k \leq m$, si ha allora $B_s^{(n)} = \sum_{i=1}^k Y_i^{(n)}$ e $B_t^{(n)} = \sum_{j=1}^m Y_j^{(n)}$, da cui

$$\text{Cov}(B_s, B_t) = \sum_{i=1}^k \sum_{j=1}^m \text{Cov}(Y_i^{(n)}, Y_j^{(n)}) = \frac{1}{2^n} \sum_{i=1}^k \sum_{j=1}^m \delta_{ij} = \frac{k}{2^n} = s = \min\{s, t\}.$$

Concludiamo la dimostrazione mostrando che effettivamente $\text{Cov}(Y_i^{(n)}, Y_j^{(n)}) = 2^{-n} \delta_{ij}$, per ogni $n \in \mathbb{N}_0$ e $1 \leq i, j \leq 2^n$. Segue dalla relazione (2.10) che l'affermazione è vera per $n = 0$, poiché $Y_1^{(0)} = \xi_1^{(0)} \sim \mathcal{N}(0, 1)$. Procediamo per induzione: sia $n \geq 1$ e notiamo che per $k \in I(n)$ — l'insieme dei numeri dispari tra 1 e 2^n — si ha $\frac{k \pm 1}{2^n} \in \tau_{n-1}$ e dunque $B_{(k \pm 1)/2^n}^{(n)} = B_{(k \pm 1)/2^n}^{(n-1)}$ (ricordiamo che per costruzione $B_t^{(n)} = B_t^{(n-1)}$ per $t \in \tau_{n-1}$). Dalla

relazione (2.13) si ha allora

$$\begin{aligned} Y_k^{(n)} &= \frac{1}{2} (B_{(k+1)/2^n}^{(n-1)} - B_{(k-1)/2^n}^{(n-1)}) + \frac{1}{2^{(n+1)/2}} \xi_k^{(n)} = \frac{1}{2} Y_{(k+1)/2}^{(n-1)} + \frac{1}{2^{(n+1)/2}} \xi_k^{(n)}, \\ Y_{k+1}^{(n)} &= \frac{1}{2} (B_{(k+1)/2^n}^{(n-1)} - B_{(k-1)/2^n}^{(n-1)}) - \frac{1}{2^{(n+1)/2}} \xi_k^{(n)} = \frac{1}{2} Y_{(k+1)/2}^{(n-1)} - \frac{1}{2^{(n+1)/2}} \xi_k^{(n)}. \end{aligned} \quad (2.19)$$

Se poniamo

$$X_\ell := \frac{1}{2} Y_\ell^{(n-1)}, \quad W_\ell := \frac{1}{2^{(n+1)/2}} \xi_{2\ell-1}^{(n)}, \quad \text{per } 1 \leq \ell \leq 2^{n-1}, \quad (2.20)$$

grazie alle relazioni in (2.19) possiamo allora scrivere

$$\{Y_i^{(n)}\}_{1 \leq i \leq 2^n} = \{Y_k^{(n)}, Y_{k+1}^{(n)}\}_{k \in I(n)} = \{X_\ell \pm W_\ell\}_{1 \leq \ell \leq 2^{n-1}}. \quad (2.21)$$

Ora affermiamo che per ogni $\ell, m \in \{1, \dots, 2^{n-1}\}$

$$\text{Cov}(X_\ell, X_m) = \frac{1}{2^{n+1}} \delta_{\ell m}, \quad \text{Cov}(W_\ell, W_m) = \frac{1}{2^{n+1}} \delta_{\ell m}, \quad \text{Cov}(X_\ell, W_m) = 0.$$

Ricordando le definizioni in (2.20), la prima relazione è vera per ipotesi induttiva, la seconda segue dal fatto che le variabili aleatorie $\{\xi_k^{(m)}\}_{k,m \in \mathbb{N}_0}$ sono i.i.d. $\mathcal{N}(0, 1)$, mentre la terza è valida perché $Y_\ell^{(n-1)} = B_{\ell/2^n}^{(n-1)} - B_{(\ell-1)/2^n}^{(n-1)}$ e il processo $B_t^{(n-1)}$ è costruito usando solo le variabili $\xi_k^{(m)}$ con $m \leq n-1$, come è chiaro dalla relazione (2.18). Di conseguenza

$$\text{Var}(X_\ell \pm W_\ell) = \text{Var}(X_\ell) + \text{Var}(W_\ell) = \frac{1}{2^{n+1}} + \frac{1}{2^{n+1}} = \frac{1}{2^n}, \quad \forall \ell \in \{1, \dots, 2^{n-1}\},$$

mentre per $\ell \neq m$ si ha $\text{Cov}(X_\ell \pm W_\ell, X_m \pm W_m) = \text{Cov}(X_\ell \pm W_\ell, X_m \mp W_m) = 0$. Ricordando (2.21), segue che $\text{Cov}(Y_i^{(n)}, Y_j^{(n)}) = 2^{-n} \delta_{ij}$, per ogni $1 \leq i, j \leq 2^n$. \square

OSSERVAZIONE 2.15. Introduciamo le *funzioni di Haar* $\{H_k^{(n)}(\cdot)\}_{n \in \mathbb{N}_0, k \in I(n)}$, definite da

$$H_k^{(n)}(t) := 2^{(n-1)/2} 1_{[\frac{k-1}{2^n}, \frac{k}{2^n})}(t) - 2^{(n-1)/2} 1_{[\frac{k}{2^n}, \frac{k+1}{2^n})}(t), \quad n \in \mathbb{N}, \quad k \in I(n), \quad (2.22)$$

mentre $H_1^{(0)}(t) := 1_{[0,1]}(t)$ (cf. Figura 2.3). È facile verificare che queste funzioni costituiscono un set ortonormale in $L^2([0, 1], dt)$, cioè $\int_0^1 H_k^{(n)}(t) H_{k'}^{(n')}(t) dt = \delta_{n,n'} \delta_{k,k'}$, e si può mostrare che sono anche un set *completo*. Si noti che le *funzioni di Schauder* $S_k^{(n)}(\cdot)$, definite in (2.15) e (2.17), non sono altro che le primitive delle funzioni di Haar:

$$S_k^{(n)}(t) = \int_0^t H_k^{(n)}(s) ds, \quad \forall n \in \mathbb{N}_0, \quad k \in I(n), \quad t \in [0, 1]. \quad (2.23)$$

Possiamo allora fornire una dimostrazione alternativa della relazione $\text{Cov}(B_s, B_t) = \min\{s, t\}$, sfruttando la completezza del set di funzioni $\{H_k^{(m)}\}_{m \in \mathbb{N}_0, k \in I(m)}$ in $L^2([0, 1])$. Notiamo infatti che, poiché $\text{Cov}(\xi_k^{(m)}, \xi_{k'}^{(m')}) = \delta_{k,k'} \delta_{m,m'}$, dalla relazione (2.18) si ha

$$\text{Cov}(B_s^{(n)}, B_t^{(n)}) = \sum_{m, m'=0}^n \sum_{k \in I(m), k' \in I(m')} S_k^{(m)}(s) S_{k'}^{(m')}(t) \delta_{k,k'} \delta_{m,m'}. \quad (2.24)$$

Indicando per comodità con $\langle f, g \rangle := \int_0^1 f(u)g(u)du$ il prodotto scalare in $L^2([0, 1])$, per la relazione (2.23) possiamo scrivere $S_k^{(m)}(s) = \langle 1_{[0, s]}, H_k^{(m)} \rangle$ e analogamente $S_{k'}^{(m')}(t) = \langle 1_{[0, t]}, H_{k'}^{(m')} \rangle$. Dato che il set $\{H_k^{(m)}\}_{m \in \mathbb{N}_0, k \in I(m)}$ è ortonormale in $L^2([0, 1])$, cioè $\langle H_k^{(m)}, H_{k'}^{(m')} \rangle = \delta_{k, k'} \delta_{m, m'}$, da (2.24) otteniamo

$$\begin{aligned} \text{Cov}(B_s^{(n)}, B_t^{(n)}) &= \sum_{m, m'=0}^n \sum_{k \in I(m), k' \in I(m')} \langle 1_{[0, s]}, H_k^{(m)} \rangle \langle 1_{[0, t]}, H_{k'}^{(m')} \rangle \langle H_k^{(m)}, H_{k'}^{(m')} \rangle \\ &= \left\langle \sum_{m=0}^n \sum_{k \in I(m)} \langle 1_{[0, s]}, H_k^{(m)} \rangle H_k^{(m)}, \sum_{m'=0}^n \sum_{k' \in I(m')} \langle 1_{[0, t]}, H_{k'}^{(m')} \rangle H_{k'}^{(m')} \right\rangle. \end{aligned}$$

Dato che il set ortonormale $\{H_k^{(m)}\}_{m \in \mathbb{N}_0, k \in I(m)}$ è completo in $L^2([0, 1])$, per ogni $f \in L^2([0, 1])$ la successione delle proiezioni $\sum_{m=0}^n \sum_{k \in I(m)} \langle f, H_k^{(m)} \rangle H_k^{(m)}$ converge in $L^2([0, 1])$ per $n \rightarrow \infty$ verso f . Di conseguenza otteniamo

$$\lim_{n \rightarrow \infty} \text{Cov}(B_s^{(n)}, B_t^{(n)}) = \langle 1_{[0, s]}, 1_{[0, t]} \rangle = \int_0^1 1_{[0, s]}(u) 1_{[0, t]}(u) du = \min\{s, t\},$$

da cui segue che $\text{Cov}(B_s, B_t) = \min\{s, t\}$, grazie alla Proposizione 1.10. \square

2.4. (IR)REGOLARITÀ DELLE TRAIETTORIE

In questo paragrafo vedremo che le traiettorie del moto browniano, sebbene continue, sono piuttosto irregolari.

2.4.1. INTEGRALE DI STIELTJES. Ricordiamo che una funzione $f : [0, \infty) \rightarrow \mathbb{R}$ crescente e continua a destra determina una misura μ su $((0, \infty), \mathcal{B}((0, \infty)))$, definita da $\mu((a, b]) := f(b) - f(a)$ per intervalli $(a, b] \subseteq (0, \infty)$ e poi estesa a tutti i boreliani. Scriveremo spesso $\mu = df$ e indicheremo l'integrale con

$$\int_0^\infty h(s) df(s) := \int_0^\infty h(s) \mu(ds),$$

per ogni $h \in L^1([0, \infty), df)$. Più in generale, date due funzioni $v_1, v_2 : [0, \infty) \rightarrow \mathbb{R}$ crescenti e continue a destra, possiamo considerarne la differenza $f(t) := v_1(t) - v_2(t)$ e definire l'integrale di Stieltjes

$$\int_0^\infty h(s) df(s) := \int_0^\infty h(s) dv_1(s) - \int_0^\infty h(s) dv_2(s), \quad (2.25)$$

per ogni $h \in L^1([0, \infty), dv_1) \cap L^1([0, \infty), dv_2)$.

L'integrale $\int_0^\infty h(s) df(s)$ è dunque definito in modo naturale per ogni funzione $f : [0, \infty) \rightarrow \mathbb{R}$ continua a destra che si possa esprimere come differenza di due funzioni crescenti. Diamo ora una caratterizzazione più esplicita di tali funzioni.

2.4.2. FUNZIONI A VARIAZIONE FINITA. Data una funzione $f : [a, b] \rightarrow \mathbb{R}$, si definisce *variazione (prima) di f sull'intervallo $[a, b]$* la quantità

$$V_{[a, b]}(f) := \sup_{n \in \mathbb{N}, a=t_0 < t_1 < \dots < t_n=b} \sum_{i=1}^n |f(t_i) - f(t_{i-1})| \in [0, +\infty]. \quad (2.26)$$

Se $V_{[a,b]}(f) < \infty$, la funzione f è detta a *variazione finita* (sull'intervallo $[a, b]$). Usando la disuguaglianza triangolare, si verifica facilmente che $V_{[a,b]}(f + g) \leq V_{[a,b]}(f) + V_{[a,b]}(g)$ e $V_{[a,b]}(cf) = |c|V_{[a,b]}(f)$, per ogni $f, g : [a, b] \rightarrow \mathbb{R}$ e per ogni $c \in \mathbb{R}$. Le funzioni a variazione finita su $[a, b]$ formano dunque uno spazio vettoriale.

Una classe semplice di funzioni a variazione finita è costituito dalle funzioni monotone: in questo caso le somme in (2.26) sono telescopiche e si ha $V_{[a,b]}(f) = |f(b) - f(a)| < \infty$. Di conseguenza, ogni funzione che si può scrivere come differenza di due funzioni crescenti è a variazione finita. Mostriamo ora che anche il viceversa è vero.

Sostituendo $|f(t_i) - f(t_{i-1})|$ con $(f(t_i) - f(t_{i-1}))^+$ o $(f(t_i) - f(t_{i-1}))^-$ nella relazione (2.26), si definiscono rispettivamente le quantità $V_{[a,b]}^+(f)$ e $V_{[a,b]}^-(f)$. È facile verificare che $V_{[a,b]}(f) < \infty$ se e solo se $V_{[a,b]}^+(f) < \infty$ e $V_{[a,b]}^-(f) < \infty$, nel qual caso si ha

$$V_{[a,b]}(f) = V_{[a,b]}^+(f) + V_{[a,b]}^-(f), \quad f(b) - f(a) = V_{[a,b]}^+(f) - V_{[a,b]}^-(f).$$

Se $V_{[a,b]}(f) < \infty$, chiaramente $V_{[a,x]}(f) < \infty$ per ogni $x \in [a, b]$. Definendo $v^\pm(x) := V_{[a,x]}^\pm(f)$, dalle relazioni precedenti possiamo allora scrivere

$$f(x) = f(a) + v^+(x) - v^-(x), \quad \forall x \in [a, b]. \quad (2.27)$$

È immediato verificare che v^+ e v^- sono funzioni crescenti da $[a, b]$ in \mathbb{R} . Abbiamo dunque mostrato che *una funzione $f : [a, b] \rightarrow \mathbb{R}$ è a variazione finita su $[a, b]$ se e soltanto se si può scrivere come differenza di due funzioni crescenti.*

Una scrittura della forma (2.27) con v^+ e v^- funzioni crescenti è tutt'altro che unica: si può infatti aggiungere a v^+ e v^- una stessa arbitraria funzione crescente e la relazione (2.27) resta valida. Tuttavia, le funzioni $v^\pm(x) := V_{[a,x]}^\pm(f)$ sono *minimali*, nel senso che se $f(x) = f(a) + w_1(x) - w_2(x)$ con w_1 e w_2 crescenti, allora necessariamente $v^+(x) \leq w_1(x)$ e $v^-(x) \leq w_2(x)$, per ogni $x \in [a, b]$ (si veda la Proposizione 11.3 in [Baldi, 2000] per una dimostrazione).

È possibile mostrare che una funzione f a variazione finita è continua (risp. continua a destra) se e soltanto se lo sono sia v^+ sia v^- (si veda la Proposizione 11.4 in [Baldi, 2000]). Di conseguenza l'integrale di Stieltjes può essere definito come in (2.25) rispetto a ogni funzione $f : [a, b] \rightarrow \mathbb{R}$ continua a destra e a variazione finita, con $v_1 = v^+$ e $v_2 = v^-$.

OSSERVAZIONE 2.16. Una classe molto importante di funzioni $f : [0, \infty) \rightarrow \mathbb{R}$ è costituita dalle funzioni della forma $f(x) = f(0) + \int_0^x g(s) ds$ per un'opportuna funzione $g \in L_{loc}^1([0, \infty))$. In questo caso si dice che f *ammette derivata debole in L_{loc}^1* e, con un piccolo abuso di notazione, si scrive $g = f'.$ [†] Si noti che le funzioni di questa forma sono continue.

Scrivendo $f(x) = f(0) + \int_0^x g^+(s) ds - \int_0^x g^-(s) ds$, è chiaro che f è differenza di due funzioni crescenti, dunque è a variazione finita su ogni intervallo compatto: $V_{[0,T]}(f) < \infty$ per ogni $T \in (0, \infty).$ [‡] In questo caso l'integrale di Stieltjes (2.25) si riduce a un integrale di Lebesgue con “densità” f' , cioè $\int_0^\infty h(s) df(s) = \int_0^\infty h(s) f'(s) ds$. \square

[†]Se g è continua in $t_0 > 0$, per il teorema fondamentale del calcolo f è derivabile in t_0 e $g(t_0) = f'(t_0)$ è l'ordinaria derivata di f .

[‡]In alternativa, si può verificare direttamente che $V_{[a,b]}(f) \leq \int_a^b |f'(s)| ds < \infty$.

2.4.3. VARIAZIONE QUADRATICA DEL MOTO BROWNIANO. Sia $\{B_t\}_{t \in [0, \infty)}$ un moto browniano. Data una partizione $\pi = \{s = t_0 < t_1 < \dots < t_n = t\}$ dell'intervallo $[s, t]$, chiameremo *passo* $|\pi| = \max_{1 \leq i \leq n} (t_i - t_{i-1})$ l'ampiezza massima degli intervalli che la compongono. Introduciamo la *variazione quadratica* di B relativa alla partizione π , ponendo

$$S_\pi := \sum_{i=1}^{n-1} (B_{t_i} - B_{t_{i-1}})^2. \quad (2.28)$$

Si noti che S_π è una variabile aleatoria reale, definita sullo stesso spazio di probabilità su cui è definito il moto browniano. Il comportamento di S_π quando il passo $|\pi|$ tende a zero, cioè quando π diventa densa in $[s, t]$, è descritto dalla seguente basilare

PROPOSIZIONE 2.17. Per ogni $0 \leq s < t < \infty$ si ha che

$$\lim_{|\pi| \rightarrow 0} S_\pi = t - s \quad \text{in } L^2(\Omega, \mathcal{F}, \mathbb{P}).$$

DIMOSTRAZIONE. Possiamo scrivere

$$S_\pi - (t - s) = \sum_{i=1}^n Y_i, \quad \text{dove } Y_i := (B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1}).$$

Dato che le variabili Y_1, \dots, Y_n sono indipendenti (perché?) e a media nulla, segue che

$$\|S_\pi - (t - s)\|_2^2 = \mathbb{E}[(S_\pi - (t - s))^2] = \sum_{i=1}^n \mathbb{E}[(Y_i)^2] = \sum_{i=1}^n (t_i - t_{i-1})^2 c_i,$$

dove abbiamo posto

$$c_i := \mathbb{E} \left[\left(\left(\frac{B_{t_i} - B_{t_{i-1}}}{\sqrt{t_i - t_{i-1}}} \right)^2 - 1 \right)^2 \right] = \mathbb{E} [(Z^2 - 1)^2], \quad \text{con } Z \sim \mathcal{N}(0, 1).$$

La seconda uguaglianza segue dal fatto che $B_{t_i} - B_{t_{i-1}} \sim \mathcal{N}(0, t_i - t_{i-1})$ e mostra che in realtà $c_i = c \in (0, \infty)$ non dipende da i . Di conseguenza, essendo $|\pi| = \max_{1 \leq i \leq n} (t_i - t_{i-1})$, otteniamo la stima

$$\mathbb{E}[(S_\pi - (t - s))^2] = c \sum_{i=1}^n (t_i - t_{i-1})^2 = c |\pi| \sum_{i=1}^n (t_i - t_{i-1}) = c |\pi| (t - s),$$

da cui segue il risultato. \square

Si noti che nella definizione di S_π compare la somma dei *quadrati* degli incrementi di B calcolati sulla partizione, invece dei valori assoluti che appaiono nella definizione (2.26) di variazione di una funzione. Quando $|\pi|$ è piccolo, anche gli incrementi $B_{t_i} - B_{t_{i-1}}$ sono piccoli (perché le traiettorie di B sono continue) e di conseguenza $|B_{t_i} - B_{t_{i-1}}| \gg (B_{t_i} - B_{t_{i-1}})^2$. Alla luce di queste considerazioni, avendo appena dimostrato che per il

moto browniano la somma dei quadrati degli incrementi converge verso un limite positivo quando $|\pi| \rightarrow 0$, non è sorprendente che la variazione delle traiettorie sia q.c. *infinita*, come mostra il seguente risultato.

COROLLARIO 2.18. Quasi certamente, le traiettorie del moto browniano hanno variazione infinita su ogni intervallo. Esiste cioè $A \in \mathcal{F}$ con $P(A) = 1$ tale che per ogni $\omega \in A$ si ha $V_{[s,t]}(B(\omega)) = +\infty$, per ogni $0 \leq s < t < \infty$.

DIMOSTRAZIONE. Per definizione di moto browniano, esiste un evento $C \in \mathcal{F}$ con $P(C) = 1$ tale che, per ogni $\omega \in C$, la funzione $t \mapsto B_t(\omega)$ è continua.

Fissiamo ora $0 \leq a < b < \infty$. Data una partizione $\pi = \{a = t_0 < t_1 < \dots < t_n = b\}$ di $[a, b]$, poniamo

$$\Delta_\pi := \max_{1 \leq i \leq n} |B_{t_i} - B_{t_{i-1}}|.$$

Se $\{\pi^{(k)}\}_{k \in \mathbb{N}}$ è una arbitraria successione di partizioni di $[s, t]$ con passo $|\pi^{(k)}|$ che tende a zero, per ogni $\omega \in C$ si ha che $\lim_{k \rightarrow \infty} \Delta_{\pi^{(k)}}(\omega) = 0$, perché la funzione $u \mapsto B_u(\omega)$, essendo continua sull'intervallo chiuso e limitato $[a, b]$, è ivi uniformemente continua.

Per la Proposizione 2.17, quando $|\pi| \rightarrow 0$ si ha $S_\pi \rightarrow (b-a)$ in L^2 , quindi in probabilità, quindi q.c. lungo un'opportuna successione. Esistono dunque un evento $D_{a,b} \in \mathcal{F}$ con $P(D_{a,b}) = 1$ e una successione di partizioni $\{\pi^{(k)}\}_{k \in \mathbb{N}}$ di $[a, b]$, con passo $|\pi^{(k)}|$ che tende a zero, tali che $\lim_{k \rightarrow \infty} S_{\pi^{(k)}}(\omega) = (b-a)$ per ogni $\omega \in D_{a,b}$.

Data una partizione π di $[a, b]$, per definizione di S_π e Δ_π possiamo scrivere

$$S_\pi = \sum_{i=1}^n (B_{t_i} - B_{t_{i-1}})^2 \leq \Delta_\pi \sum_{i=1}^n |B_{t_i} - B_{t_{i-1}}| \leq \Delta_\pi V_{[a,b]}(B).$$

In particolare, per $\omega \in C \cap D_{a,b}$ otteniamo che

$$V_{[a,b]}(B(\omega)) \geq \frac{S_{\pi^{(k)}}(\omega)}{\Delta_{\pi^{(k)}}(\omega)} \longrightarrow +\infty \quad \text{per } k \rightarrow \infty,$$

poiché $S_{\pi^{(k)}}(\omega) \rightarrow (b-a) > 0$ e $\Delta_{\pi^{(k)}}(\omega) \rightarrow 0$. Quindi $V_{[a,b]}(B(\omega)) = +\infty$ per $\omega \in C \cap D_{a,b}$.

Definiamo infine

$$A := \bigcap_{0 \leq a < b < \infty, a, b \in \mathbb{Q}} C \cap D_{a,b}.$$

Chiaramente $P(A) = 1$, perché A è intersezione numerabile di eventi quasi certi. Inoltre se $\omega \in A$ si ha $V_{[s,t]}(B(\omega)) = +\infty$ per ogni $0 \leq s < t < \infty$: basta infatti considerare $a, b \in \mathbb{Q}$ con $s \leq a < b \leq t$ e di conseguenza $V_{[s,t]}(B(\omega)) \geq V_{[a,b]}(B(\omega)) = +\infty$. \square

Essendo le traiettorie del moto browniano q.c. a variazione infinita, non è possibile definire integrali del tipo $\int h(s, \omega) dB_s(\omega)$ nel senso di Stieltjes, come descritto in §2.4.1-§2.4.2. Come vedremo nel capitolo 5, imponendo opportune restrizioni sulla classe di integrandi $h(s, \omega)$, è possibile definire integrali di questo tipo.

OSSERVAZIONE 2.19. Ci si può chiedere se il limite $S_\pi \rightarrow (t-s)$ per $|\pi| \rightarrow 0$, dimostrato nella Proposizione 2.17, valga q.c. e non solo in L^2 . *La risposta è negativa*: è infatti possibile costruire una successione $\{\pi_k\}_{k \in \mathbb{N}}$ di partizioni di $[s, t]$, con $|\pi_k| \rightarrow 0$, tale che q.c. si abbia $\limsup_{k \rightarrow \infty} S_{\pi_k} \rightarrow +\infty$ (si veda l'esercizio 1.15 in [Mörters e Peres, 2009]); in particolare, segue che q.c. $\sup_\pi S_\pi = +\infty$. È comunque possibile dare condizioni sufficienti per la convergenza quasi certa: per esempio, se le partizioni π_k sono *crescenti*, nel senso che $\pi_k \subseteq \pi_{k+1}$ (a ogni passo vengono aggiunti punti alla partizione precedente), oppure se $\sum_{k \in \mathbb{N}} |\pi_k| < \infty$, allora $S_{\pi_k} \rightarrow (t-s)$ q.c. per $k \rightarrow \infty$. \square

2.4.4. RISULTATI FINI PER LE TRAIETTORIE. Esistono diversi risultati che descrivono precisamente il comportamento locale delle traiettorie del moto browniano. Cominciamo enunciando la celebre *legge del logaritmo iterato*.

TEOREMA 2.20 (LEGGE DEL LOGARITMO ITERATO). Se $B = \{B_t\}_{t \in [0, \infty)}$ è un moto browniano reale, valgono le seguenti relazioni:

$$\text{q.c.} \quad \limsup_{t \rightarrow +\infty} \frac{B_t}{\sqrt{t} \sqrt{2 \log \log t}} = 1 \quad \text{e} \quad \liminf_{t \rightarrow +\infty} \frac{B_t}{\sqrt{t} \sqrt{2 \log \log t}} = -1. \quad (2.29)$$

La dimostrazione non è difficile, ma la omettiamo per brevità (si veda il Teorema 5.1 in [Mörters e Peres, 2009], oppure il Teorema 2.11 in [Baldi, 2000] per la dimostrazione della relazione equivalente (2.31)). Dato che $-B$ è un moto browniano se B lo è, grazie alla Proposizione 2.10, è sufficiente dimostrare una sola delle due relazioni in (2.29).

Il Teorema 2.20 implica in particolare che, per ogni $\varepsilon > 0$ fissato, per q.o. $\omega \in \Omega$ esistono due successioni (aleatorie) $\{t_n = t_n(\omega)\}_{n \in \mathbb{N}}$ e $\{s_n = s_n(\omega)\}_{n \in \mathbb{N}}$, entrambe tendenti all'infinito, tali che

$$B_{t_n}(\omega) \geq (1 - \varepsilon) \sqrt{2 t_n \log \log t_n}, \quad B_{s_n}(\omega) \leq -(1 - \varepsilon) \sqrt{2 s_n \log \log s_n}. \quad (2.30)$$

Dato che q.c. la funzione $t \mapsto B_t$ è continua, queste disuguaglianze dimostrano che, per q.o. $\omega \in \Omega$, $B_t(\omega)$ visita ogni numero reale infinite volte (in particolare cambia segno infinite volte) in ogni intorno $[M, \infty)$ di infinito.

Sfruttando le proprietà di invarianza del moto browniano, è possibile trasformare le relazioni in (2.29), che riguardano proprietà asintotiche, in analoghe relazioni sulle *proprietà locali* delle traiettorie. Più precisamente, per ogni $t_0 \geq 0$ fissato, applicando le relazioni in (2.29) al moto browniano $\{t(B_{t_0+1/t} - B_{t_0})\}_{t \geq 0}$ (cf. le proprietà (b) ed (e) della Proposizione 2.10) si ottiene

$$\text{q.c.} \quad \limsup_{h \downarrow 0} \frac{B_{t_0+h} - B_{t_0}}{\sqrt{h} \sqrt{2 \log \log \frac{1}{h}}} = 1 \quad \text{e} \quad \liminf_{h \downarrow 0} \frac{B_{t_0+h} - B_{t_0}}{\sqrt{h} \sqrt{2 \log \log \frac{1}{h}}} = -1. \quad (2.31)$$

Analogamente a quanto visto in (2.30), da queste relazioni per $t_0 = 0$ segue in particolare che, per ogni $\varepsilon > 0$ fissato, per q.o. $\omega \in \Omega$ esistono due successioni (aleatorie) positive $\{h_n = h_n(\omega)\}_{n \in \mathbb{N}}$ e $\{u_n = u_n(\omega)\}_{n \in \mathbb{N}}$, entrambe tendenti a zero, tali che

$$B_{h_n}(\omega) \geq (1 - \varepsilon) \sqrt{2 h_n \log \log \frac{1}{h_n}}, \quad B_{u_n}(\omega) \leq -(1 - \varepsilon) \sqrt{2 u_n \log \log \frac{1}{u_n}}.$$

Sfruttando la continuità delle traiettorie, si ha che q.c. B_t cambia segno infinite volte in ogni intorno destro $[0, \delta)$ di 0.

Concludiamo il paragrafo con un interessante corollario del Teorema 2.20.

COROLLARIO 2.21. Sia $B = \{B_t\}_{t \in [0, \infty)}$ un moto browniano. Per qualunque punto $t_0 \geq 0$ fissato, q.c. B_t non è derivabile in $t = t_0$.

DIMOSTRAZIONE. Dalle relazioni in (2.31) segue direttamente che q.c. $\limsup_{h \downarrow 0} (B_{t_0+h} - B_{t_0})/h = +\infty$ e $\liminf_{h \downarrow 0} (B_{t_0+h} - B_{t_0})/h = -\infty$. \square

OSSERVAZIONE 2.22. Si può rafforzare il Corollario 2.21, mostrando che q.c. la funzione $t \mapsto B_t$ non è derivabile in nessun punto $t_0 \in [0, \infty)$ (si veda per esempio il Teorema 1.30 in [Mörters e Peres, 2009]). Un ipotetico analogo rafforzamento delle relazioni in (2.31) è invece falso: per esempio, se $t_0 = t_0(\omega)$ è un punto di massimo locale per il moto browniano,[†] si ha $B_{t_0+h} \leq B_{t_0}$ per h in un intorno destro di t_0 e quindi la prima relazione in (2.31) non vale. \square

2.5. PROCESSI E σ -ALGEBRE

In questo paragrafo l'insieme di indici I è arbitrario, ma nei casi concreti sarà quasi sempre un sottoinsieme di \mathbb{R} ; analogamente, lo spazio misurabile (E, \mathcal{E}) è tipicamente \mathbb{R}^d .

Ricordiamo che, data una funzione $Y : \Omega \rightarrow (G, \mathcal{G})$, dove Ω è un insieme e (G, \mathcal{G}) è uno spazio misurabile, si indica con $\sigma(Y)$ la σ -algebra generata da Y , definita come la più piccola σ -algebra su Ω che renda misurabile l'applicazione Y . Essa consiste di tutti e soli gli eventi della forma $\{Y \in A\}$, al variare di $A \in \mathcal{G}$.

Sia ora $X = \{X_t\}_{t \in I}$ un processo stocastico definito su (Ω, \mathcal{F}, P) a valori in (E, \mathcal{E}) . Come abbiamo visto in §2.1.2, il processo X può essere visto come una variabile aleatoria a valori nello spazio delle traiettorie (E^I, \mathcal{E}^I) . Di conseguenza è ben definita $\sigma(X)$, la σ -algebra su Ω generata dal processo X , costituita dagli eventi della forma $\{X \in A\} = \{\omega \in \Omega : X(\omega) \in A\}$ al variare di $A \in \mathcal{E}^I$. Per semplicità, preferiamo dare una definizione alternativa di $\sigma(X)$, che non menziona esplicitamente lo spazio delle traiettorie.

DEFINIZIONE 2.23. Dato un processo stocastico $X = \{X_t\}_{t \in I}$, definito su uno spazio di probabilità (Ω, \mathcal{F}, P) a valori in (E, \mathcal{E}) , indichiamo con $\sigma(X) := \sigma(\{X_t\}_{t \in I})$ la più piccola σ -algebra su Ω che rende misurabili tutte le componenti X_t del processo.

Secondo questa definizione, $\sigma(X)$ è la più piccola σ -algebra che contiene $\sigma(X_t)$ per ogni $t \in I$. Una base di $\sigma(X)$ è quindi data dalla famiglia di eventi J^X , definita da

$$J^X := \{\{X_{s_1} \in A_1, \dots, X_{s_k} \in A_k\}, k \in \mathbb{N}, s_i \in I, A_i \in \mathcal{E}\}. \quad (2.32)$$

Dal fatto che gli insiemi cilindrici formano una base della σ -algebra \mathcal{E}^I , cf. §2.1.1, segue che la Definizione 2.23 determina la stessa σ -algebra $\sigma(X)$ introdotta in precedenza.

[†]Per l'esistenza di tali punti, basta considerare un punto di massimo globale per il moto browniano su un intervallo compatto, per esempio sia $t_0 = t_0(\omega) \in [0, 1]$ tale che $B_{t_0(\omega)} = \max_{t \in [0, 1]} B_t(\omega)$. Un tale punto esiste q.c., per la continuità delle traiettorie, e si mostra facilmente che q.c. $0 < t_0 < 1$, dunque t_0 è un punto di massimo locale.

La σ -algebra $\sigma(X)$ consiste di tutti e soli *gli eventi che possono essere espressi in termini di X* (più precisamente, in termini di una quantità numerabile di sue componenti). Intuitivamente, $\sigma(X)$ contiene *le informazioni sul processo X* : essa consiste infatti degli eventi per i quali si può stabilire se si siano verificati o no conoscendo il processo X .

La nozione di σ -algebra generata da un processo permette di definire l'indipendenza di processi stocastici, come nel caso di variabili aleatorie.[†]

DEFINIZIONE 2.24. I processi stocastici $X^{(1)} = \{X_t^{(1)}\}_{t \in I_1}, \dots, X^{(n)} = \{X_t^{(n)}\}_{t \in I_n}$, definiti sullo stesso spazio di probabilità (Ω, \mathcal{F}, P) , si dicono *indipendenti* se lo sono le σ -algebre da loro generate $\sigma(X^{(1)}), \dots, \sigma(X^{(n)})$.

Questa definizione piuttosto astratta ha una traduzione molto esplicita per i processi gaussiani, analoga al Lemma 1.14. Dati i processi stocastici (gaussiani) $X^{(1)} = \{X_t^{(1)}\}_{t \in I_1}, \dots, X^{(n)} = \{X_t^{(n)}\}_{t \in I_n}$, essi si dicono *congiuntamente gaussiani* se il processo congiunto $\{X_t^{(i)}\}_{i \in \{1, \dots, n\}, t \in I_i}$ è gaussiano, ossia se qualsiasi combinazione lineare finita $\sum_{i=1}^n \sum_{j=1}^k c_{ij} X_{t_j}^{(i)}$ è una variabile aleatoria reale normale. Si ha allora la seguente

PROPOSIZIONE 2.25. Se i processi $X^{(1)} = \{X_t^{(1)}\}_{t \in I_1}, \dots, X^{(n)} = \{X_t^{(n)}\}_{t \in I_n}$ sono congiuntamente gaussiani, essi sono indipendenti se e solo se sono *scorrelati*, nel senso seguente: $\text{Cov}(X_s^{(i)}, X_t^{(j)}) = 0$ per ogni $i \neq j$ e per ogni s, t .

DIMOSTRAZIONE. L'indipendenza delle σ -algebre $\sigma(X^{(1)}), \dots, \sigma(X^{(n)})$ può essere verificata su una base: basta dunque mostrare che, estratto qualunque vettore finito-dimensionale $(X_{t_1, (1)}^{(1)}, \dots, X_{t_{k_1, (1)}}^{(1)})$ da $X^{(1)}$, qualunque $(X_{t_1, (2)}^{(2)}, \dots, X_{t_{k_2, (2)}}^{(2)})$ da $X^{(2)}$, ecc., questi vettori aleatori sono tra loro indipendenti. Essendo congiuntamente normali per ipotesi, sappiamo che l'indipendenza è equivalente alla scorrelazione delle rispettive componenti, cioè $\text{Cov}(X_s^{(i)}, X_t^{(j)}) = 0$ per $i \neq j$ e per ogni s, t (si veda l'Osservazione 1.15). Questa condizione è dunque sufficiente, oltre che ovviamente necessaria, per l'indipendenza dei processi $X^{(1)}, \dots, X^{(n)}$. \square

2.5.1. FILTRAZIONE NATURALE DI UN PROCESSO. Consideriamo ora il caso di un processo $X = \{X_t\}_{t \in [0, \infty)}$ indicizzato dalla semiretta reale positiva, definito su uno spazio di probabilità (Ω, \mathcal{F}, P) a valori in (E, \mathcal{E}) . Ricordando la Definizione 2.23, per ogni $s \geq 0$ indichiamo con \mathcal{F}_s^X la σ -algebra generata dal processo con insieme dei tempi ristretto a $[0, s]$, ossia $\mathcal{F}_s^X := \sigma(\{X_u\}_{u \in [0, s]})$. In altri termini, \mathcal{F}_s^X è la più piccola σ -algebra che renda misurabili tutte le applicazioni X_u , per $0 \leq u \leq s$.

Intuitivamente, la σ -algebra $\mathcal{F}_s^X = \sigma(\{X_u\}_{0 \leq u \leq s})$ contiene *le informazioni sul processo X nell'intervallo di tempo $[0, s]$* : in effetti, essa consiste di eventi per i quali si può stabilire se si siano verificati o no osservando il processo X nell'intervallo di tempo $[0, s]$. La famiglia $\{\mathcal{F}_s^X\}_{s \in [0, \infty)}$ è detta *filtrazione naturale del processo X* . Si tratta di una *famiglia crescente di σ -algebre*: $\mathcal{F}_s^X \subseteq \mathcal{F}_t^X \subseteq \mathcal{F}$ per ogni $0 \leq s < t < \infty$.

[†]In effetti, pensando un processo come una variabile aleatoria a valori nello spazio delle traiettorie, la Definizione 2.24 coincide esattamente con la definizione di indipendenza per variabili aleatorie.

Ritorniamo in dettaglio su questi concetti nel prossimo capitolo. Per il momento, forniamo un'ulteriore utile caratterizzazione alternativa del moto browniano, modificando l'ipotesi di indipendenza degli incrementi in un modo che sarà molto rilevante nel seguito.

PROPOSIZIONE 2.26. Un processo stocastico reale $B = \{B_t\}_{t \in [0, \infty)}$ è un *moto browniano* se e soltanto se soddisfa le proprietà (a), (c), (d) della Definizione 2.3 e vale inoltre la seguente

(b') per ogni $0 \leq s < t$, la variabile aleatoria $(B_t - B_s)$ è indipendente dalla σ -algebra $\mathcal{F}_s^B = \sigma(\{B_u\}_{0 \leq u \leq s})$.

DIMOSTRAZIONE. Dimostriamo un risultato più forte: *qualunque* processo stocastico $B = \{B_t\}_{t \in [0, \infty)}$ (non necessariamente un moto browniano) ha incrementi indipendenti, ossia vale la relazione (b) della Definizione 2.3, se e soltanto se vale la relazione (b').

Mostriamo innanzitutto che (b) \Rightarrow (b'). Ricordiamo che è sufficiente verificare l'indipendenza su una base e che una base di \mathcal{F}_s^B è del tipo J^X , definita in (2.32) (sostituendo X con B e restringendo gli indici s_i all'insieme $[0, s]$). Basta dunque mostrare che per ogni scelta di $0 \leq s_1 < \dots < s_k \leq s$ e $C, A_1, \dots, A_k \in \mathcal{B}(\mathbb{R})$ si ha

$$\begin{aligned} & P(\{B_{s_1} \in A_1, \dots, B_{s_k} \in A_k\} \cap \{B_t - B_s \in C\}) \\ &= P(B_{s_1} \in A_1, \dots, B_{s_k} \in A_k) \cdot P(B_t - B_s \in C). \end{aligned} \quad (2.33)$$

Introducendo il vettore $W := (B_{s_1}, \dots, B_{s_k})$ e il sottoinsieme $A := A_1 \times \dots \times A_k \subseteq \mathbb{R}^k$, possiamo scrivere $\{B_{s_1} \in A_1, \dots, B_{s_k} \in A_k\} = \{W \in A\}$. Se introduciamo il vettore $Y = (B_{s_1}, B_{s_2} - B_{s_1}, \dots, B_{s_k} - B_{s_{k-1}})$, ottenuto mediante una trasformazione lineare $Y = LW$, con L invertibile, possiamo scrivere

$$\{W \in A\} = \{Y \in LA\} = \{(B_{s_1}, B_{s_2} - B_{s_1}, \dots, B_{s_k} - B_{s_{k-1}}) \in LA\}.$$

Ricordiamo il fatto che, data una famiglia di variabili indipendenti, due sottofamiglie disgiunte sono tra loro indipendenti. Segue allora che la variabile $B_t - B_s$ è indipendente da Y , grazie alla proprietà (b), e dunque

$$P(\{Y \in LA\} \cap \{B_t - B_s \in C\}) = P(Y \in LA) \cdot P(B_t - B_s \in C).$$

Mettendo insieme le precedenti relazioni, segue che (2.33) è verificata.

Mostriamo ora che (b') \Rightarrow (b). Dobbiamo mostrare che, per ogni $k \geq 2$, fissati comunque $0 \leq t_0 < t_1 < \dots < t_k < \infty$ e $A_1, \dots, A_k \in \mathcal{B}(\mathbb{R})$, vale la relazione

$$P\left(\bigcap_{i=1}^k \{B_{t_i} - B_{t_{i-1}} \in A_i\}\right) = \prod_{i=1}^k P(B_{t_i} - B_{t_{i-1}} \in A_i).$$

Si noti che $\bigcap_{i=1}^{k-1} \{B_{t_i} - B_{t_{i-1}} \in A_i\} \in \mathcal{F}_{t_{k-1}}^B$. Infatti le variabili B_{t_i} con $1 \leq i \leq k-1$ sono $\mathcal{F}_{t_{k-1}}^B$ -misurabili, per cui lo sono anche $B_{t_i} - B_{t_{i-1}}$ (differenza di funzioni misurabili).

Scrivendo $\bigcap_{i=1}^k \{B_{t_i} - B_{t_{i-1}} \in A_i\} = \left(\bigcap_{i=1}^{k-1} \{B_{t_i} - B_{t_{i-1}} \in A_i\}\right) \cap \{B_{t_k} - B_{t_{k-1}} \in A_k\}$ e notando che $B_{t_k} - B_{t_{k-1}}$ è per ipotesi indipendente da $\mathcal{F}_{t_{k-1}}^B$, si ha che

$$\mathbb{P} \left(\bigcap_{i=1}^k \{B_{t_i} - B_{t_{i-1}} \in A_i\} \right) = \mathbb{P} \left(\bigcap_{i=1}^{k-1} \{B_{t_i} - B_{t_{i-1}} \in A_i\} \right) \cdot \mathbb{P}(B_{t_k} - B_{t_{k-1}} \in A_k).$$

Un facile argomento induttivo conclude la dimostrazione. \square

2.6. MOTO BROWNIANO MULTIDIMENSIONALE

Generalizziamo ora la definizione di moto browniano al caso multidimensionale.

DEFINIZIONE 2.27. Si dice *moto browniano d -dimensionale* qualunque processo stocastico $B = \{B_t = (B_t^{(1)}, \dots, B_t^{(d)})\}_{t \in [0, \infty)}$ a valori in \mathbb{R}^d che soddisfa le seguenti proprietà:

- (a) $B_0 = 0$ q.c.;
- (b) B ha *incrementi indipendenti*, cioè per ogni scelta di $k \geq 2$ e $0 \leq t_0 < t_1 < \dots < t_k < \infty$ i vettori aleatori $\{B_{t_i} - B_{t_{i-1}}\}_{1 \leq i \leq k}$ sono indipendenti;
- (c) B ha *incrementi stazionari gaussiani centrati*: $B_t - B_s \sim \mathcal{N}(0, (t-s)I_d)$ per ogni $0 \leq s < t$, dove I_d indica la matrice identica $d \times d$, cioè $(I_d)_{ij} = \delta_{ij}$;
- (d) q.c. B ha *traiettorie continue*, cioè q.c. la funzione $t \mapsto B_t$ è continua.

Molte proprietà del moto browniano multidimensionale sono analoghe al caso reale. Ad esempio, valgono le seguenti generalizzazioni delle Proposizioni 2.9 e 2.26.

PROPOSIZIONE 2.28. Un processo stocastico $B = \{B_t\}_{t \in [0, \infty)}$ a valori in \mathbb{R}^d è un moto browniano d -dimensionale se e soltanto se è un processo gaussiano di media nulla e di covarianza $\text{Cov}(B_s^{(i)}, B_t^{(j)}) = \delta_{ij} \min\{s, t\}$, con traiettorie q.c. continue.

PROPOSIZIONE 2.29. Un processo stocastico $B = \{B_t\}_{t \in [0, \infty)}$ a valori in \mathbb{R}^d è un moto browniano d -dimensionale se e soltanto se valgono le proprietà (a), (c), (d) della Definizione 2.27 e vale inoltre la seguente proprietà:

- (b') per ogni $0 \leq s < t$, il vettore aleatorio $(B_t - B_s)$ è indipendente dalla σ -algebra $\mathcal{F}_s^B = \sigma(\{B_u\}_{0 \leq u \leq s}) = \sigma(\{B_u^{(i)}\}_{0 \leq u \leq s, 1 \leq i \leq d})$.

Anche la Proposizione 2.10 si estende al moto browniano multidimensionale, senza bisogno di alcuna modifica nell'enunciato.

Omettiamo per brevità le dimostrazioni, analoghe al caso unidimensionale. Mostriamo solo come calcolare $\text{Cov}(B_s^{(i)}, B_t^{(j)})$ a partire dalla Definizione 2.27: per $s \leq t$ si ha

$$\text{Cov}(B_s^{(i)}, B_t^{(j)}) = \text{Cov}(B_s^{(i)}, B_t^{(j)} - B_s^{(j)}) + \text{Cov}(B_s^{(i)}, B_s^{(j)}) = s \delta_{ij}, \quad (2.34)$$

grazie all'indipendenza dei vettori aleatori $B_t - B_s$ e B_s (proprietà (b)), da cui segue quella delle componenti $B_t^{(j)} - B_s^{(j)}$ e $B_s^{(i)}$, e grazie al fatto che $B_s \sim \mathcal{N}(0, sI_n)$ (proprietà (c)).

Concludiamo la sezione con una proprietà importante, che fornisce una costruzione esplicita del moto browniano d -dimensionale a partire da d moti browniani reali indipendenti.

PROPOSIZIONE 2.30. Un processo stocastico $B = \{B_t\}_{t \in [0, \infty)}$ a valori in \mathbb{R}^d è un moto browniano d -dimensionale se e soltanto se le sue componenti $B^{(1)} = \{B_t^{(1)}\}_{t \in [0, \infty)}$, \dots , $B^{(d)} = \{B_t^{(d)}\}_{t \in [0, \infty)}$ sono moti browniani reali indipendenti.

DIMOSTRAZIONE. Se B è un moto browniano d -dimensionale, per la Proposizione 2.28 ogni componente $B^{(i)}$ è un processo gaussiano con media nulla, covarianza $\text{Cov}(B_s^{(i)}, B_t^{(i)}) = \min\{s, t\}$ e traiettorie q.c. continue. Segue che $B^{(i)}$ è un moto browniano reale, per la Proposizione 2.9. Sempre per la Proposizione 2.28, le componenti $B^{(1)}, \dots, B^{(d)}$ sono processi congiuntamente gaussiani e scorrelati, poiché $\text{Cov}(B_s^{(i)}, B_t^{(j)}) = 0$ per $i \neq j$; sono dunque indipendenti, grazie alla Proposizione 2.25.

Viceversa, se le componenti $B^{(1)}, \dots, B^{(d)}$ sono moti browniani reali indipendenti, il processo congiunto $B = \{B_t^{(i)}\}_{1 \leq i \leq d, t \geq 0}$ è gaussiano perché ogni combinazione lineare di sue componenti può essere scritta come $\sum_{i=1}^d \sum_{\ell=1}^k c_{i\ell} B_{t_\ell}^{(i)} = \sum_{i=1}^d Y_i$ e le variabili reali $Y_i := \sum_{\ell=1}^k c_{i\ell} B_{t_\ell}^{(i)}$ sono normali (perché $B^{(i)}$ è un processo gaussiano) indipendenti (perché i processi $B^{(1)}, \dots, B^{(d)}$ sono indipendenti). Grazie alla Proposizione 2.9, ciascuna componente $B^{(i)}$ ha media nulla e $\text{Cov}(B_s^{(i)}, B_t^{(i)}) = \min\{s, t\}$, mentre per $i \neq j$ si ha $\text{Cov}(B_s^{(i)}, B_t^{(j)}) = 0$ per ogni $s, t \geq 0$, poiché i processi $B^{(i)}$ e $B^{(j)}$ sono indipendenti. In definitiva, per ogni $1 \leq i, j \leq d$ e $s, t \geq 0$ si ha $\text{Cov}(B_s^{(i)}, B_t^{(j)}) = \delta_{ij} \min\{s, t\}$. Infine, q.c. le traiettorie $t \mapsto B_t = (B_t^{(1)}, \dots, B_t^{(d)})$ sono continue perché q.c. ciascuna componente $t \mapsto B_t^{(i)}$ lo è. Possiamo dunque concludere che B è un moto browniano d -dimensionale grazie alla Proposizione 2.28. \square

2.7. LA MISURA DI WIENER

Indichiamo con $C := C([0, \infty), \mathbb{R}^d)$ lo spazio delle funzioni *continue* definite su $[0, \infty)$ a valori in \mathbb{R}^d . Rendiamo C uno spazio misurabile, munendolo della σ -algebra \mathcal{B} generata dagli insiemi cilindrici, ossia quelli della forma $\{f \in C : f_{t_1} \in A_1, \dots, f_{t_k} \in A_k\}$, al variare di $k \in \mathbb{N}$, $t_1, \dots, t_k \in [0, \infty)$ e $A_1, \dots, A_k \in \mathcal{B}(\mathbb{R}^d)$. Essendo chiusa per intersezioni finite, questa classe di insiemi è una base di \mathcal{B} . Se introduciamo le *proiezioni coordinate* $\pi_t : C \rightarrow \mathbb{R}^d$ definite da $\pi_t(f) := f_t$, si verifica facilmente che \mathcal{B} è la σ -algebra generata dalle funzioni π_t , cioè $\mathcal{B} = \sigma(\{\pi_t\}_{t \geq 0})$.

Dato un moto browniano d -dimensionale $B = \{B_s\}_{s \geq 0}$, definito su uno spazio di probabilità (Ω, \mathcal{F}, P) , sappiamo che esiste $A \in \mathcal{F}$ con $P(A) = 1$ tale che la funzione $t \mapsto B_t(\omega)$ è continua per ogni $\omega \in A$. Se ridefiniamo $B_t(\omega) \equiv 0$ per $\omega \notin A$, otteniamo un

moto browniano le cui traiettorie sono continue *per ogni* $\omega \in \Omega$ e non solo q.c.. Possiamo allora vedere B come una applicazione da Ω in C :

$$\omega \longmapsto B(\omega) := \{B_s(\omega)\}_{s \in [0, \infty)} \in C.$$

Usando i sottoinsiemi cilindrici, che generano \mathcal{B} , è immediato vedere che questa applicazione è misurabile: si ha infatti per $D = \{f \in C : f_{t_1} \in A_1, \dots, f_{t_k} \in A_k\}$

$$\{B \in D\} = \{B_{t_1} \in A_1, \dots, B_{t_k} \in A_k\} \in \mathcal{F}.$$

Il moto browniano B può dunque essere visto come una *variabile aleatoria a valori in* (C, \mathcal{B}) . Di conseguenza, B induce sullo spazio d'arrivo la sua legge, una probabilità su (C, \mathcal{B}) indicata con \mathcal{W} e nota come *misura di Wiener*. Più esplicitamente, per ogni sottoinsieme $A \subseteq C$ misurabile, cioè per ogni $A \in \mathcal{B}$, si ha $\mathcal{W}(A) := P(B \in A)$.

La misura di Wiener permette una *costruzione canonica del moto browniano*. Infatti, prendendo come spazio di probabilità $(C, \mathcal{B}, \mathcal{W})$, è facile vedere che il processo stocastico delle proiezioni coordinate $\{\pi_t\}_{t \geq 0}$ è un moto browniano d -dimensionale.

OSSERVAZIONE 2.31. Se $B' = \{B'_t\}_{t \in [0, \infty)}$ è un altro moto browniano d -dimensionale, la legge \mathcal{W}' che esso induce sullo spazio (C, \mathcal{B}) è sempre la misura di Wiener \mathcal{W} . Infatti, dato un insieme cilindrico $D = \{f \in C : f_{t_1} \in A_1, \dots, f_{t_k} \in A_k\}$, si ha per definizione $\mathcal{W}(D) = P(B \in D) = P(B_{t_1} \in A_1, \dots, B_{t_k} \in A_k)$ e questa probabilità non cambia sostituendo B con B' , perché è determinata dalle leggi finito-dimensionali del moto browniano. Di conseguenza $\mathcal{W}(D) = \mathcal{W}'(D)$, e dato che gli insiemi cilindrici sono una base della σ -algebra \mathcal{B} segue che $\mathcal{W} = \mathcal{W}'$. \square

OSSERVAZIONE 2.32. Per quanto detto in §2.1.2, un moto browniano d -dimensionale $B = \{B_t\}_{t \geq 0}$, come ogni processo stocastico a valori in \mathbb{R}^d , può essere visto come variabile aleatoria a valori nello spazio delle traiettorie $((\mathbb{R}^d)^{[0, \infty)}, \mathcal{B}(\mathbb{R}^d)^{[0, \infty)})$, su cui induce la sua legge μ_B . Dato che spazio delle funzioni continue $C = C([0, \infty), \mathbb{R}^d)$ è un sottoinsieme di $(\mathbb{R}^d)^{[0, \infty)}$, si potrebbe essere tentati di definire la misura di Wiener \mathcal{W} come la restrizione della probabilità μ_B al sottoinsieme C . *Ciò non è possibile, perché C non è un sottoinsieme misurabile di $(\mathbb{R}^d)^{[0, \infty)}$, ossia $C \notin \mathcal{B}(\mathbb{R}^d)^{[0, \infty)}$* . Abbiamo infatti visto nel paragrafo 2.2.3 (cf. anche l'Osservazione 3.3 nel capitolo 3) che è possibile definire un processo $B' = \{B'_t\}_{t \geq 0}$ con le stesse leggi finito-dimensionali di B , e dunque *con la stessa legge $\mu_{B'} = \mu_B$ sullo spazio $(\mathbb{R}^d)^{[0, \infty)}$, ma con traiettorie $t \mapsto B'_t$ q.c. discontinue*. Se C fosse misurabile, si dovrebbe allora avere $1 = \mu_B(C) = \mu_{B'}(C) = 0$, il che è assurdo. † \square

OSSERVAZIONE 2.33. Sullo spazio C c'è una topologia naturale, quella della convergenza uniforme sui compatti, che è *metrizzabile*. Definendo $\|h\|_{[a, b]}^\infty := \sup_{x \in [a, b]} |h(x)|$, una distanza che induce questa topologia è data per esempio da

$$d(f, g) := \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{\|f(x) - g(x)\|_{[0, n]}^\infty}{1 + \|f(x) - g(x)\|_{[0, n]}^\infty}.$$

† Per una dimostrazione più diretta del fatto che $C \notin \mathcal{B}(\mathbb{R}^d)^{[0, \infty)}$, si veda ad esempio il Teorema 36.3 in Billingsley [1995] e le considerazioni che lo seguono, in cui si mostra che i sottoinsiemi misurabili di $(\mathbb{R}^d)^{[0, \infty)}$ sono determinati da una quantità al più numerabile di componenti.

Su C è quindi definita la corrispondente σ -algebra boreliana $\mathcal{B}(C)$, generata dagli insiemi aperti. Non è difficile mostrare che questa σ -algebra coincide con la σ -algebra \mathcal{B} generata dalle proiezioni, che abbiamo definito sopra.[†] In particolare, ogni funzione definita su C che sia continua rispetto a $d(\cdot, \cdot)$ è \mathcal{B} -misurabile.

Si noti che la *topologia* generata dalle proiezioni $\{\pi_t\}_{t \geq 0}$ è la topologia della convergenza puntuale, che è strettamente più piccola (ossia più debole) della topologia della convergenza uniforme sui compatti. Ciononostante, le σ -algebre generate da queste due topologie su C coincidono, come menzionato. \square

2.7.1. IL PRINCIPIO DI INVARIANZA DI DONSKER. Sia $\{X_n\}_{n \in \mathbb{N}}$ una successione di variabili reali i.i.d. in L^2 , definite su uno spazio di probabilità (Ω, \mathcal{F}, P) , tali che $E(X_1) = 0$ e $\text{Var}(X_1) = \sigma^2 < \infty$. Definiamo la *passeggiata aleatoria* $\{S_n\}_{n \in \mathbb{N}}$ ponendo

$$S_0 := 0, \quad S_n := \sum_{i=1}^n X_i.$$

Il celebre *teorema limite centrale* afferma che per ogni $x \in \mathbb{R}$ vale la seguente relazione:

$$\lim_{n \rightarrow \infty} P\left(\frac{S_n}{\sigma\sqrt{n}} \leq x\right) = P(Z \leq x),$$

dove Z indica una variabile aleatoria reale normale standard. Si può mostrare che ciò è equivalente al fatto che $S_n/(\sigma\sqrt{n}) \rightarrow W$ in legge per $n \rightarrow \infty$.

È possibile rafforzare notevolmente questo risultato. Definiamo la variabile \tilde{S}_t per $t \in [0, \infty)$ come l'interpolazione lineare della traiettoria $\{S_n\}_{n \in \mathbb{N}}$: poniamo cioè

$$\tilde{S}_t := (\lfloor t \rfloor + 1 - t)S_{\lfloor t \rfloor} + (t - \lfloor t \rfloor)S_{\lfloor t \rfloor + 1},$$

dove $\lfloor x \rfloor := \max\{n \in \mathbb{Z} : n \leq x\}$ indica la parte intera di un numero reale x . Introduciamo quindi, per ogni $k \in \mathbb{N}$, un processo stocastico $Y^{(k)} = \{Y_t^{(k)}\}_{t \in [0, \infty)}$ definito come il *riscaldamento diffusivo* di $\{\tilde{S}_t\}_{t \geq 0}$ di fattore k :

$$Y_t^{(k)} := \frac{\tilde{S}_{kt}}{\sigma\sqrt{k}}, \quad \forall t \geq 0.$$

Si noti che per $t = 1$ ritroviamo $Y_1^{(n)} = S_n/(\sigma\sqrt{n})$.

Dato che per costruzione il processo $Y^{(k)}$ ha traiettorie continue, per ogni $k \in \mathbb{N}$, possiamo vedere $Y^{(k)}$ come una applicazione da Ω in $C = C([0, \infty), \mathbb{R})$. Non è difficile mostrare che tale applicazione è misurabile, cioè $Y^{(k)}$ è una variabile aleatoria a valori in (C, \mathcal{B}) . È dunque ben definita la sua legge: si tratta di una probabilità sullo spazio (C, \mathcal{B}) , che indichiamo con $\mathcal{Y}^{(n)}$. Ricordiamo che C è uno spazio metrico (rispetto alla convergenza uniforme sui compatti) e che \mathcal{B} è la corrispondente σ -algebra boreliana (si veda l'Osservazione 2.33). Vale allora il seguente risultato fondamentale.

[†]L'inclusione $\mathcal{B} \subseteq \mathcal{B}(C)$ segue dal fatto che ogni proiezione π_t , essendo continua rispetto alla topologia della convergenza uniforme sui compatti, è $\mathcal{B}(C)$ -misurabile; l'inclusione inversa $\mathcal{B}(C) \subseteq \mathcal{B}$ si ottiene notando che $\|h\|_{[a, b]}^\infty = \sup_{t \in [a, b] \cap \mathbb{Q}} |h(t)| = \sup_{t \in [a, b] \cap \mathbb{Q}} |\pi_t(h)|$, da cui segue che l'applicazione $h \mapsto \|h\|_{[a, b]}^\infty$ è \mathcal{B} -misurabile (per maggiori dettagli, si veda per esempio l'esercizio 1.4 in [Baldi, 2000]).

TEOREMA 2.34 (PRINCIPIO DI INVARIANZA DI DONSKER). Per $k \rightarrow \infty$ la successione di processi $Y^{(k)}$ converge in legge verso il moto browniano. Equivalentemente, la successione di leggi $\mathcal{Y}^{(k)}$ converge debolmente verso la misura di Wiener \mathcal{W} .

Questo risultato si può formulare grossolanamente dicendo che, su larga scala, le traiettorie di una passeggiata aleatoria di media nulla e varianza finita, riscalate diffusivamente, “assomigliano” alle traiettorie del moto browniano. Per esempio, le traiettorie simulate del moto browniano illustrate nella Figura 2.1 a pagina 28 sono state ottenute a partire da una passeggiata aleatoria con incrementi gaussiani.

Uno degli aspetti più importanti del Teorema 2.34 è la sua *universalità*: qualunque sia la legge degli incrementi X_i , purché di media zero e varianza finita, la distribuzione $\mathcal{Y}^{(k)}$ delle traiettorie riscalate della passeggiata aleatoria converge per $k \rightarrow \infty$ verso lo stesso limite, cioè la legge \mathcal{W} del moto browniano. In questo senso, i dettagli “microscopici” della passeggiata aleatoria diventano irrilevanti nel limite di larga scala. Questo risultato mostra anche come il moto browniano sia un oggetto molto naturale.

Infine, il Teorema 2.34 è molto importante anche come strumento di calcolo. Infatti, per definizione di convergenza debole di misure di probabilità (si veda il paragrafo 1.5), possiamo riformulare il Teorema 2.34 nel modo seguente: per ogni funzionale $\Phi : C \rightarrow \mathbb{R}$ continuo e limitato si ha $\lim_{k \rightarrow \infty} \int_C \Phi(\zeta) \mathcal{Y}^{(k)}(d\zeta) = \int_C \Phi(\zeta) \mathcal{W}(d\zeta)$, ovvero, usando la formula del cambio di variabili (Teorema 1.6),

$$\lim_{k \rightarrow \infty} E(\Phi(Y^{(k)})) = E(\Phi(B)).$$

Questo significa che, se si conosce il valore di $E(\Phi(B))$, si conosce anche il limite della successione $E(\Phi(Y^{(k)}))$ per ogni passeggiata aleatoria di media zero e varianza finita. È naturalmente possibile anche il percorso inverso: se si sa calcolare $\lim_{k \rightarrow \infty} E(\Phi(Y^{(k)}))$ per un esempio specifico di passeggiata aleatoria con incrementi di media nulla e varianza finita, si è determinato il valore di $E(\Phi(B))$.

3. PROCESSI DI LÉVY E PROPRIETÀ DI MARKOV

Introduciamo in questo capitolo alcune nozioni basilari della teoria dei processi stocastici, illustrandone qualche interessante applicazione alla classe dei *processi di Lévy*, che include il moto browniano come caso particolare. Per la validità di diversi risultati, è richiesto che lo spazio di probabilità (Ω, \mathcal{F}, P) sia *completo*, cioè che \mathcal{F} contenga tutti gli insiemi *P-trascurabili*, ovvero tutti i sottoinsiemi degli eventi di probabilità nulla. Questa è la ragione per cui nei prossimi capitoli la completezza dello spazio di probabilità su cui lavoriamo sarà un'ipotesi frequente. Ricordiamo che è sempre possibile completare uno spazio di probabilità, come descritto nel paragrafo 1.2.3 del capitolo 1.

Prima di proseguire, ricordiamo che una probabilità su \mathbb{R}^d (o più in generale su uno spazio metrico) è determinata dagli integrali delle funzioni continue. Più precisamente, se μ, ν sono due probabilità su \mathbb{R}^d tali che $\int \Phi d\mu = \int \Phi d\nu$ per ogni funzione $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$ continua e limitata, segue che $\mu = \nu$. La dimostrazione è semplice: scegliendo le funzioni $\Phi(x) = \cos(\langle \vartheta, x \rangle)$ e $\Phi(x) = \sin(\langle \vartheta, x \rangle)$, per $\vartheta \in \mathbb{R}^d$, segue che le leggi μ e ν hanno la stessa funzione caratteristica, dunque esse coincidono.[†]

In particolare, dalla formula del cambio di variabili (Teorema 1.6) segue che due vettori aleatori $X : (\Omega, \mathcal{F}, P) \rightarrow \mathbb{R}^d$, $Y : (\Omega^*, \mathcal{F}^*, P^*) \rightarrow \mathbb{R}^d$ hanno la stessa legge se $E(\Phi(X)) = E^*(\Phi(Y))$ per ogni funzione $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$ continua e limitata.

L'uso delle funzioni continue risulta spesso utile per dimostrare l'indipendenza di un vettore aleatorio X a valori in \mathbb{R}^d , definito su uno spazio di probabilità (Ω, \mathcal{F}, P) , da una σ -algebra $\mathcal{G} \subseteq \mathcal{F}$. Per definizione, occorre mostrare che

$$P(G, X \in A) = P(G)P(X \in A),$$

per ogni $G \in \mathcal{G}$ e $A \in \mathcal{B}(\mathbb{R}^d)$. Se $P(G) = 0$ quest'uguaglianza è banalmente vera, mentre se $P(G) > 0$ la relazione si può riscrivere come $P(X \in A | G) = P(X \in A)$, per ogni $A \in \mathcal{B}(\mathbb{R}^d)$. Ciò significa che il vettore aleatorio X è indipendente dalla σ -algebra \mathcal{G} se e solo se, per ogni evento $G \in \mathcal{G}$ di probabilità positiva, la legge di X rispetto alla probabilità condizionata $P^* := P(\cdot | G)$ coincide con la legge di X (rispetto a P). Per quanto detto sopra, X è indipendente da \mathcal{G} se e solo se $E^*(\Phi(X)) = E(\Phi(X))$, cioè

$$E(\Phi(X) | G) = E(\Phi(X)), \quad (3.1)$$

per ogni $G \in \mathcal{G}$ con $P(G) > 0$ e per ogni $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$ continua e limitata.

[†]Una dimostrazione alternativa si ottiene notando che, per ogni insieme chiuso $C \subseteq \mathbb{R}^d$, si può scrivere $1_C(x) = \lim_{n \rightarrow \infty} \Phi_n(x)$, dove $\Phi_n(x) := \max\{0, 1 - n d(x, C)\}$ e $d(x, C) := \inf\{|y - z|, z \in C\}$ indica la distanza da x dall'insieme C . Sappiamo per ipotesi che $\int \Phi_n d\mu = \int \Phi_n d\nu$ per ogni $n \in \mathbb{N}$, e dato che $|\Phi_n| \leq 1$ segue per convergenza dominata che $\int 1_C d\mu = \int 1_C d\nu$, cioè $\mu(C) = \nu(C)$ per ogni insieme chiuso $C \subseteq \mathbb{R}^d$. Dato che gli insiemi chiusi sono una base della σ -algebra boreliana di \mathbb{R}^d , segue che $\mu = \nu$.

3.1. PROCESSI STOCASTICI E FILTRAZIONI

Indichiamo con \mathbb{T} un sottoinsieme di \mathbb{R} , che avrà la funzione di insieme dei tempi per i processi stocastici che considereremo. I casi che ci interessano di più sono $\mathbb{T} = \mathbb{N}_0$ e soprattutto $\mathbb{T} = [0, \infty)$ o $\mathbb{T} = [a, b]$ con $0 \leq a < b < \infty$. Ricordiamo che la nozione processo stocastico è stata introdotta nella Definizione 2.1.

3.1.1. MODIFICAZIONI E INDISTINGUIBILITÀ. Dato un processo $X = \{X_t\}_{t \in \mathbb{T}}$, definito su uno spazio di probabilità (Ω, \mathcal{F}, P) a valori in uno spazio misurabile (E, \mathcal{E}) , ricordiamo che le leggi dei vettori $(X_{t_1}, \dots, X_{t_k})$ su $(E^k, \mathcal{E}^{\otimes k})$, al variare di $k \in \mathbb{N}$ e $t_1, \dots, t_k \in \mathbb{T}$, sono dette *leggi finito-dimensionali* del processo.

Definiamo ora due importanti relazioni tra processi stocastici.

DEFINIZIONE 3.1. Due processi stocastici $X = \{X_t\}_{t \in \mathbb{T}}$, $X' = \{X'_t\}_{t \in \mathbb{T}}$ aventi lo stesso insieme dei tempi, definiti sullo stesso spazio di probabilità (Ω, \mathcal{F}, P) e a valori nello stesso spazio misurabile (E, \mathcal{E}) , si dicono:

- *modificazione* (o versione) l'uno dell'altro se, per ogni $t \in \mathbb{T}$, si ha $X_t = X'_t$ q.c.;
- *indistinguibili* se, q.c., si ha $X_t = X'_t$ per ogni $t \in \mathbb{T}$.

Si noti che l'ordine in cui compaiono “q.c.” e “per ogni $t \in \mathbb{T}$ ” è fondamentale.

Con qualche ipotesi di regolarità si possono dare riformulazioni equivalenti. Per esempio, se lo spazio di probabilità (Ω, \mathcal{F}, P) è completo, possiamo dire che i processi X e X' sono

- *modificazione* l'uno dell'altro se, per ogni $t \in \mathbb{T}$, si ha $P(X_t = X'_t) = 1$;
- *indistinguibili* se $P(X_t = X'_t \text{ per ogni } t \in \mathbb{T}) = 1$.

Le osservazioni seguenti sono facilmente verificabili.

- Se due processi X, X' sono indistinguibili, allora sono modificazione l'uno dell'altro.
- Se due processi X, X' sono modificazione l'uno dell'altro, allora hanno le stesse leggi finito-dimensionali. Infatti, per ogni $t_1, \dots, t_k \in \mathbb{T}$, i vettori aleatori $(X_{t_1}, \dots, X_{t_k})$ e $(X'_{t_1}, \dots, X'_{t_k})$ sono q.c. uguali (perché?) e dunque hanno la stessa legge.
- Se due processi X, X' sono modificazione l'uno dell'altro e se l'insieme dei tempi \mathbb{T} è numerabile, allora X e X' sono indistinguibili.

Quando l'insieme dei \mathbb{T} è più che numerabile, la nozione di indistinguibilità è invece strettamente più forte della nozione di modificazione, come mostra l'esempio seguente.

ESEMPIO 3.2. Sia (Ω, \mathcal{F}, P) uno spazio di probabilità su cui è possibile definire una variabile aleatoria reale U con legge continua, ossia senza atomi (si può considerare per esempio l'intervallo unitario $([0, 1], \mathcal{B}([0, 1]), Leb)$ munito della misura di Lebesgue, con $U(\omega) := \omega$). Definire un processo $Y = \{Y_t\}_{t \in [0, \infty)}$ ponendo $Y_t(\omega) := 1_{\{(U+t) \in \mathbb{Q}\}}(\omega)$, cioè $Y_t(\omega) = 1$ se $U(\omega) + t \in \mathbb{Q}$ mentre $Y_t(\omega) = 0$ altrimenti.

Dato *qualunque* processo reale $X = \{X_t\}_{t \in [0, \infty)}$ definito su (Ω, \mathcal{F}, P) , definiamo il processo $X' = \{X'_t\}_{t \in [0, \infty)}$ ponendo $X'_t := X_t + Y_t$. Per ogni $t \geq 0$ fissato si ha $P(X_t = X'_t) = P(Y_t = 0) = P(U + t \notin \mathbb{Q}) = 1$, perché la variabile aleatoria $U + t$ ha legge continua

per ipotesi; quindi X' è una modificazione di X . Tuttavia i processi X e X' *non* sono indistinguibili, perché per ogni $\omega \in \Omega$ si ha $X'_t(\omega) - X_t(\omega) = Y_t(\omega) \neq 0$ per infiniti valori di $t \geq 0$ (infatti la traiettoria $t \mapsto Y_t(\omega)$ vale 1 sull'insieme denso $\mathbb{Q} - U(\omega)$). \square

OSSERVAZIONE 3.3. Con riferimento all'esempio precedente, notiamo che la traiettoria $t \mapsto Y_t(\omega)$ è discontinua in ogni punto $t \in [0, \infty)$, per ogni $\omega \in \Omega$.[†] Se X è un processo con traiettorie continue q.c., segue che per q.o. $\omega \in \Omega$ la traiettoria $t \mapsto X'_t(\omega) = X_t(\omega) + Y_t(\omega)$ del processo X' è discontinua in ogni punto $t \in [0, \infty)$. I processi X , X' hanno dunque le stesse leggi finito-dimensionali (infatti X' è modificazione di X) ma q.c. X ha traiettorie continue mentre X' le ha discontinue. Questo esempio mostra che *la continuità delle traiettorie di un processo non è una proprietà delle leggi finito-dimensionali*. \square

3.1.2. CONTINUITÀ E MISURABILITÀ DI PROCESSI. Per tutto questo sottoparagrafo supponiamo che $\mathbb{T} = [0, \infty)$ oppure $\mathbb{T} = [a, b]$, con $0 \leq a < b < \infty$, e indichiamo con $\mathcal{B}(\mathbb{T})$ la σ -algebra boreliana di \mathbb{T} .

Definiamo le importanti nozioni di *continuità* e *misurabilità* per un processo. Lo spazio topologico di arrivo dei processi che considereremo nel seguito sarà quasi sempre \mathbb{R}^d .

DEFINIZIONE 3.4. Un processo stocastico $X = \{X_t\}_{t \in \mathbb{T}}$, definito su uno spazio di probabilità (Ω, \mathcal{F}, P) a valori in uno spazio topologico $(E, \mathcal{B}(E))$ (munito della σ -algebra boreliana), si dice:

- *continuo* (risp. *continuo a destra*, *continuo a sinistra*) se per ogni $\omega \in \Omega$ la funzione $t \mapsto X_t(\omega)$ è continua (risp. continua a destra, continua a sinistra) da \mathbb{T} in E ;
- *q.c. continuo* (risp. *q.c. continuo a destra*, *q.c. continuo a sinistra*) se per q.o. $\omega \in \Omega$ la funzione $t \mapsto X_t(\omega)$ è continua (risp. continua a destra, continua a sinistra) da \mathbb{T} in E .

DEFINIZIONE 3.5. Un processo $X = \{X_t\}_{t \in \mathbb{T}}$, definito su uno spazio di probabilità (Ω, \mathcal{F}, P) a valori in uno spazio misurabile (E, \mathcal{E}) , si dice *misurabile* se l'applicazione $(t, \omega) \mapsto X_t(\omega)$ è misurabile da $(\mathbb{T} \times \Omega, \mathcal{B}(\mathbb{T}) \otimes \mathcal{F})$ a valori in (E, \mathcal{E}) .

Vedremo tra poco che la misurabilità di un processo è una condizione poco restrittiva, che è verificata non appena le traiettorie del processo sono continue a destra (si vedano il Lemma 3.12 e il Lemma 3.13).

Ricordiamo che, per il teorema di Fubini (paragrafo 1.4.3 del capitolo 1), se una applicazione $(x, y) \mapsto f(x, y)$ è misurabile, allora per ogni x fissato la funzione $y \mapsto f(x, y)$

[†]Le traiettorie $t \mapsto Y_t(\omega)$ del processo che abbiamo costruito, pur essendo discontinue in ogni $t \geq 0$, coincidono con una funzione continua (identicamente nulla) al di fuori di un insieme numerabile. È possibile produrre esempi molto più irregolari: per esempio, sfruttando un risultato di N. N. Lusin e W. Sierpinski [Sur une décomposition d'un intervalle en une infinité non dénombrable d'ensembles non mesurables, C. R. Acad. Sci. Paris 165 (1917), 422-424.], si può costruire un processo $Y = \{Y_t\}_{t \geq 0}$ che sia sempre una modificazione del processo identicamente nullo, cioè $P(Y_t = 0) = 1$ per ogni $t \geq 0$, ma le cui traiettorie $t \mapsto Y_t(\omega)$ non siano "Lebesgue misurabili" (ossia non siano misurabili rispetto al completamento della σ -algebra boreliana di $[0, \infty)$ rispetto alla misura di Lebesgue) per alcun $\omega \in \Omega$.

è misurabile e, analogamente, per ogni y fissato la funzione $x \mapsto f(x, y)$ è misurabile. Tuttavia non vale il viceversa: la misurabilità delle sezioni $y \mapsto f(x, y)$, $x \mapsto f(x, y)$ non garantisce la misurabilità dell'applicazione $(x, y) \mapsto f(x, y)$.

Segue allora dalla Definizione 3.5 che, se un processo $X = \{X_t\}_{t \geq 0}$ è misurabile, le sue traiettorie $t \mapsto X_t(\omega)$ sono funzioni misurabili, per ogni $\omega \in \Omega$ fissato. La misurabilità di tutte le traiettorie non è tuttavia sufficiente a garantire che un processo sia misurabile. Si noti che se un processo reale positivo (o limitato) $X = \{X_t\}_{t \geq 0}$ è misurabile, vale la relazione $E(\int_0^1 X_t dt) = \int_0^1 E(X_t) dt$, grazie al teorema di Fubini.

3.1.3. EQUIVALENZA DI PROCESSI. Definiamo un'ulteriore relazione tra processi, che apparirà nella costruzione dell'integrale stocastico nel capitolo 5. Supponiamo sempre che $\mathbb{T} = [0, \infty)$ oppure $\mathbb{T} = [a, b]$, con $0 \leq a < b < \infty$, e indichiamo con Leb la misura di Lebesgue su \mathbb{T} .

DEFINIZIONE 3.6. Due processi stocastici $X = \{X_t\}_{t \in \mathbb{T}}$, $X' = \{X'_t\}_{t \in \mathbb{T}}$, definiti sullo stesso spazio di probabilità (Ω, \mathcal{F}, P) e a valori nello stesso spazio misurabile (E, \mathcal{E}) , si dicono *equivalenti* se si ha $X_t(\omega) = X'_t(\omega)$ per $(Leb \otimes P)$ -q.o. $(t, \omega) \in \mathbb{T} \times \Omega$.

Nel caso in cui l'insieme $\{(t, \omega) \in \mathbb{T} \times \Omega : X_t(\omega) \neq X'_t(\omega)\}$ sia misurabile (per esempio, se X e X' sono processi misurabili a valori in uno spazio metrico[†]), grazie al Teorema di Fubini possiamo scrivere

$$\begin{aligned} (Leb \otimes P)((t, \omega) \in \mathbb{T} \times \Omega : X_t(\omega) \neq X'_t(\omega)) &= \int_{\mathbb{T}} P(X_t \neq X'_t) dt \\ &= E(Leb(t \in \mathbb{T} : X_t \neq X'_t)). \end{aligned} \quad (3.2)$$

Da ciò discende che X e X' sono equivalenti se e solo se vale una delle relazioni seguenti:

- per Leb -q.o. $t \in \mathbb{T}$ si ha $P(X_t \neq X'_t) = 0$, cioè $X_t(\omega) = X'_t(\omega)$ per P -q.o. $\omega \in \Omega$;
- per P -q.o. $\omega \in \Omega$ si ha $Leb(t \in \mathbb{T} : X_t(\omega) \neq X'_t(\omega)) = 0$, cioè $X_t(\omega) = X'_t(\omega)$ per Leb -q.o. $t \in \mathbb{T}$.

Ricordiamo che se X e X' sono modificazione l'uno dell'altro, per ogni $t \in \mathbb{T}$ si ha $X_t = X'_t$ q.c.. Quindi, per processi misurabili a valori in uno spazio metrico, la nozione di equivalenza è più debole della nozione di modificazione (e, a maggior ragione, della nozione di indistinguibilità).

3.1.4. FILTRAZIONI E IPOTESI STANDARD. Dato uno spazio misurabile (Ω, \mathcal{F}) , si dice *filtrazione* una famiglia *crescente* $\{\mathcal{F}_t\}_{t \in \mathbb{T}}$ di sotto- σ -algebre di \mathcal{F} , cioè tale che $\mathcal{F}_s \subseteq \mathcal{F}_t$ per ogni $s, t \in \mathbb{T}$ con $s \leq t$. Un esempio tipico è dato dalla *filtrazione naturale* $\{\mathcal{F}_t^X\}_{t \in \mathbb{T}}$ di un qualunque processo $X = \{X_t\}_{t \in \mathbb{T}}$, definita da $\mathcal{F}_t^X := \sigma(\{X_u\}_{u \in \mathbb{T}, u \leq t})$ e introdotta nel paragrafo 2.5.1 del capitolo 2.

Data una filtrazione $\{\mathcal{F}_t\}_{t \in \mathbb{T}}$, definiamo $\mathcal{F}_\infty := \sigma(\{\mathcal{F}_t\}_{t \in \mathbb{T}})$ come la più piccola σ -algebra che contiene tutte le \mathcal{F}_t .

L'interpretazione intuitiva è che la σ -algebra \mathcal{F}_t rappresenti l'*informazione disponibile fino all'istante t* : più precisamente, \mathcal{F}_t contiene gli eventi *conoscibili entro l'istante t* , ossia quelli per cui al tempo t è possibile dire se si siano verificati oppure no. Nel caso speciale della filtrazione naturale di un processo $X = \{X_t\}_{t \in \mathbb{T}}$, la σ -algebra $\mathcal{F}_t = \mathcal{F}_t^X = \sigma(\{X_u\}_{u \in \mathbb{T}, u \leq t})$ contiene intuitivamente la "storia" del processo X fino all'istante t , ossia gli eventi esprimibili come funzione (misurabile) delle variabili $\{X_u\}_{u \in [0, t]}$.

[†]Oltre a richiedere la misurabilità dei processi X e X' , perché l'insieme $\{(t, \omega) \in \mathbb{T} \times \Omega : X_t(\omega) \neq X'_t(\omega)\}$ che appare in (3.2) sia misurabile occorre fare qualche ipotesi minimale di regolarità sullo spazio di arrivo (E, \mathcal{E}) , che garantisca che la diagonale $\{(x, y) \in E \times E : x = y\}$ sia misurabile in $(E \times E, \mathcal{E} \otimes \mathcal{E})$; è sufficiente, per esempio, richiedere che E sia uno spazio metrico (con $\mathcal{E} = \mathcal{B}(E)$).

Uno spazio di probabilità (Ω, \mathcal{F}, P) munito di una filtrazione $\{\mathcal{F}_t\}_{t \in \mathbb{T}}$ è detto *spazio (di probabilità) filtrato* e sarà indicato con $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \mathbb{T}}, P)$.

DEFINIZIONE 3.7. Dato uno spazio di probabilità *completo* (Ω, \mathcal{F}, P) , una filtrazione $\{\mathcal{F}_t\}_{t \in \mathbb{T}}$ su (Ω, \mathcal{F}, P) si dice *completa* se, per ogni $t \in \mathbb{T}$, la σ -algebra \mathcal{F}_t contiene tutti gli eventi di \mathcal{F} di probabilità nulla.

Ricordiamo che in uno spazio di probabilità completo (Ω, \mathcal{F}, P) gli insiemi P -trascurabili, ossia i sottoinsiemi degli eventi di probabilità nulla, sono essi stessi eventi.

Assumiamo d'ora in avanti che $\mathbb{T} = [0, \infty)$ oppure $\mathbb{T} = [a, b]$, con $0 \leq a < b < \infty$.

Data una filtrazione $\{\mathcal{F}_t\}_{t \in \mathbb{T}}$, definiamo $\mathcal{F}_{t+} := \bigcap_{u > t} \mathcal{F}_u$, per ogni $t < \sup(\mathbb{T})$; se $\mathbb{T} = [a, b]$, poniamo $\mathcal{F}_{b+} := \mathcal{F}_b$. Intuitivamente, la σ -algebra \mathcal{F}_{t+} contiene gli eventi conoscibili immediatamente dopo l'istante t .

DEFINIZIONE 3.8. Una filtrazione $\{\mathcal{F}_t\}_{t \in \mathbb{T}}$ si dice *continua a destra* se si ha l'uguaglianza $\mathcal{F}_t = \mathcal{F}_{t+}$ per ogni $t \in \mathbb{T}$.

Osserviamo che $\mathcal{F}_t \subseteq \mathcal{F}_{t+} \subseteq \mathcal{F}_{t+\varepsilon}$, per ogni $t \in \mathbb{T}$ e $\varepsilon > 0$, come si verifica facilmente. Si noti che \mathcal{F}_{t+} può essere strettamente più grande di \mathcal{F}_t : per esempio, se $X = \{X_s\}_{s \geq 0}$ è un processo reale, l'evento $A := \{\lim_{n \rightarrow \infty} X_{t+\frac{1}{n}} = 0\} \in \mathcal{F}_{t+}^X$, ma in generale $A \notin \mathcal{F}_t^X$.[†]

DEFINIZIONE 3.9. Diciamo che una filtrazione $\{\mathcal{F}_t\}_{t \in \mathbb{T}}$ su uno spazio di probabilità *completo* (Ω, \mathcal{F}, P) soddisfa le *ipotesi standard* se è completa e continua a destra. In questo caso, $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \mathbb{T}}, P)$ è detto *spazio (di probabilità) filtrato standard*.

Data una filtrazione generica $\{\mathcal{F}_t\}_{t \in \mathbb{T}}$ su uno spazio *completo* (Ω, \mathcal{F}, P) , se ne possono considerare alcune estensioni.

- Ponendo $\overline{\mathcal{F}}_t := \sigma(\mathcal{F}_t, \mathcal{N})$, dove $\mathcal{N} := \{C \in \mathcal{F} : P(C) = 0\}$, si ottiene una *filtrazione completa*: $\{\overline{\mathcal{F}}_t\}_{t \in \mathbb{T}}$ è la più piccola filtrazione completa che estende $\{\mathcal{F}_t\}_{t \in \mathbb{T}}$;
- Considerando $\{\mathcal{F}_{t+}\}_{t \in \mathbb{T}}$, si ottiene una *filtrazione continua a destra* (esercizio): $\{\mathcal{F}_{t+}\}_{t \in \mathbb{T}}$ è la più piccola filtrazione *continua a destra* che estende $\{\mathcal{F}_t\}_{t \in \mathbb{T}}$.
- Combinando i punti precedenti, si ottiene la filtrazione $\{\overline{\mathcal{F}}_{t+}\}_{t \in \mathbb{T}} = \{\sigma(\mathcal{F}_{t+}, \mathcal{N})\}_{t \in \mathbb{T}}$, detta *ampliamento standard* di $\{\mathcal{F}_t\}_{t \in \mathbb{T}}$: si tratta della più piccola estensione di $\{\mathcal{F}_t\}_{t \in \mathbb{T}}$ che soddisfa le ipotesi standard.

La ragione per cui insistiamo su queste proprietà è che in molti casi risulta tecnicamente conveniente lavorare con uno spazio filtrato standard (si veda per esempio l'Esecizio 1.8 nel capitolo 1 in [Karatzas e Shreve, 1998], o il Lemma 3.13 più sotto).

[†]Per esempio, sullo spazio $(\Omega = \{T, C\}, \mathcal{F} = \mathcal{P}(\Omega))$ definiamo il processo $X = \{X_s\}_{s \geq 0}$ ponendo $X_s(\omega) \equiv 0$ per $s \leq t$ mentre $X_s(\omega) := 1_{\{C\}}(\omega)$ per $s > t$. Definendo la σ -algebra banale $\mathcal{B} := \{\emptyset, \Omega\}$, la filtrazione naturale del processo X è data da $\mathcal{F}_s^X = \mathcal{B}$ per $s \leq t$ mentre $\mathcal{F}_s^X = \mathcal{F}$ per $s > t$. Si ha quindi $\mathcal{F}_t^X = \mathcal{B}$ mentre $\mathcal{F}_{t+}^X = \mathcal{F}$; dato che $A := \{\lim_{n \rightarrow \infty} X_{t+\frac{1}{n}} = 0\} = \{T\}$, segue che $A \notin \mathcal{F}_t^X$.

3.1.5. PROCESSI ADATTATI E PROGRESSIVAMENTE MISURABILI. Definiamo ora alcune importanti relazioni tra processi stocastici e filtrazioni. Assumiamo sempre che $\mathbb{T} = [0, \infty)$ oppure $\mathbb{T} = [a, b]$, con $0 \leq a < b < \infty$.

DEFINIZIONE 3.10. Un processo stocastico $\{X_t\}_{t \in \mathbb{T}}$, definito su uno spazio filtrato $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \mathbb{T}}, \mathbb{P})$ a valori in uno spazio misurabile (E, \mathcal{E}) , si dice *adattato alla filtrazione* (o *adattato tout court*) se per ogni $t \in \mathbb{T}$ la variabile X_t è \mathcal{F}_t -misurabile, cioè se X_t è misurabile come applicazione da (Ω, \mathcal{F}_t) in (E, \mathcal{E}) .

Per costruzione, ogni processo X è adattato alla sua filtrazione naturale $\{\mathcal{F}_t^X\}_{t \in \mathbb{T}}$, che è la più piccola filtrazione a cui X sia adattato. Infatti, si verifica facilmente che X è *adattato a una filtrazione* $\{\mathcal{F}_t\}_{t \in \mathbb{T}}$ se e soltanto se $\mathcal{F}_t^X \subseteq \mathcal{F}_t$ per ogni $t \in \mathbb{T}$.

Definiamo ora l'importante nozione di misurabilità progressiva.

DEFINIZIONE 3.11. Un processo $X = \{X_t\}_{t \in \mathbb{T}}$, definito su uno spazio filtrato $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \mathbb{T}}, \mathbb{P})$ a valori in uno spazio misurabile (E, \mathcal{E}) , si dice *progressivamente misurabile* se, per ogni $T \in \mathbb{T}$, l'applicazione $(t, \omega) \mapsto X_t(\omega)$ da $([a, T] \times \Omega, \mathcal{B}([a, T]) \otimes \mathcal{F}_T)$ a valori in (E, \mathcal{E}) è misurabile, dove poniamo per brevità $a := \min(\mathbb{T})$.

Se $X = \{X_t\}_{t \in \mathbb{T}}$ è un processo *progressivamente misurabile*, è facile mostrare che X è *misurabile* e *adattato*. Vale un parziale viceversa: se un processo è misurabile e adattato, si può dimostrare che ne esiste sempre una modificazione progressivamente misurabile (si tratta di un risultato tutt'altro che banale).

Le nozioni di misurabilità e di progressiva misurabilità, all'apparenza piuttosto tecniche, sono automaticamente verificate per una classe molto ampia di processi, come mostrano i seguenti risultati.

LEMMA 3.12. Se un processo $X = \{X_t\}_{t \in \mathbb{T}}$ è continuo a destra, allora è misurabile. Se X è continuo a destra e adattato, allora è progressivamente misurabile.

DIMOSTRAZIONE. Dimostriamo la seconda parte nel caso in cui $\mathbb{T} = [0, \infty)$. Fissiamo $T \geq 0$ e definiamo $X_0^{(n)} := X_0$ e $X_u^{(n)} := X_{\frac{i}{2^n}}$ per $u \in (\frac{i-1}{2^n}T, \frac{i}{2^n}T]$, dove $n \in \mathbb{N}$ e $1 \leq i \leq 2^n$. Verifichiamo che la funzione $(u, \omega) \mapsto X_u^{(n)}(\omega)$ è misurabile da $([0, T] \times \Omega, \mathcal{B}([0, T]) \otimes \mathcal{F}_T)$ a valori in (E, \mathcal{E}) : per ogni $A \in \mathcal{E}$ si ha

$$\begin{aligned} & \{(u, \omega) \in [0, T] \times \Omega : X_u^{(n)}(\omega) \in A\} \\ &= (\{0\} \times \{X_0 \in A\}) \cup \bigcup_{i=1}^{2^n} \left(\left(\frac{i-1}{2^n}T, \frac{i}{2^n}T \right] \times \{X_{\frac{i}{2^n}T} \in A\} \right) \in \mathcal{B}[0, T] \otimes \mathcal{F}_T, \end{aligned}$$

poiché per ipotesi X è adattato. Dalla continuità a destra di X si ha $X_u(\omega) = \lim_{n \rightarrow \infty} X_u^{(n)}(\omega)$, per ogni $(u, \omega) \in [0, T] \times \Omega$. La funzione $(u, \omega) \mapsto X_u(\omega)$ è dunque misurabile come limite di funzioni misurabili. \square

Le conclusioni del Lemma precedente continuano a valere anche per processi q.c. continui, a patto di lavorare con spazi di probabilità e filtrazioni complete. Più precisamente, vale la seguente estensione (omettiamo per brevità la semplice dimostrazione).

LEMMA 3.13. Se un processo $X = \{X_t\}_{t \in \mathbb{T}}$ è q.c. continuo a destra e se lo spazio di probabilità (Ω, \mathcal{F}, P) è completo, allora X è misurabile. Se X è q.c. continuo a destra e adattato a una filtrazione completa, allora X è progressivamente misurabile.

3.2. PROCESSI DI LÉVY RISPETTO A UNA FILTRAZIONE

3.2.1. MOTO BROWNIANO RISPETTO A UNA FILTRAZIONE. Avendo introdotto la nozione di spazio filtrato $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$, è utile rafforzare la definizione di moto browniano nel modo seguente.

DEFINIZIONE 3.14. Un processo stocastico $B = \{B_t\}_{t \geq 0}$ a valori in \mathbb{R}^d , definito su uno spazio filtrato $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$, è detto $\{\mathcal{F}_t\}_{t \geq 0}$ -moto browniano d -dimensionale (o moto browniano d -dimensionale rispetto alla filtrazione $\{\mathcal{F}_t\}_{t \geq 0}$) se è adattato a $\{\mathcal{F}_t\}_{t \geq 0}$ e se soddisfa le seguenti proprietà:

- (a) $B_0 = 0$ q.c.;
- (b) per ogni $0 \leq s < t$, il vettore aleatorio $B_t - B_s$ è indipendente da \mathcal{F}_s ;
- (c) per ogni $0 \leq s < t$, si ha $B_t - B_s \sim \mathcal{N}(0, (t - s)I_d)$;
- (d) B è un processo q.c. continuo.

La richiesta che B sia adattato alla filtrazione $\{\mathcal{F}_t\}_{t \geq 0}$ implica, come abbiamo già osservato, che $\{\mathcal{F}_t\}_{t \geq 0}$ deve contenere la filtrazione naturale di B : si deve cioè avere l'inclusione $\mathcal{F}_t \supseteq \mathcal{F}_t^B := \sigma(\{B_u\}_{0 \leq u \leq t})$ per ogni $t \geq 0$. Nel caso “minimale” in cui $\mathcal{F}_t = \mathcal{F}_t^B$ per ogni $t \geq 0$, ritroviamo la caratterizzazione di moto browniano fornita dalla Proposizione 2.29. In altri termini, *un moto browniano secondo la “vecchia” Definizione 2.27 non è altro che un $\{\mathcal{F}_t^B\}_{t \geq 0}$ -moto browniano secondo la “nuova” Definizione 3.14.*

In molti casi è naturale (e conveniente) considerare una filtrazione $\{\mathcal{F}_t\}_{t \geq 0}$ strettamente più grande di quella naturale del processo, come mostra l'osservazione seguente.

OSSERVAZIONE 3.15. Sia $B = \{(B_t^{(1)}, \dots, B_t^{(d)})\}_{t \geq 0}$ un moto browniano d -dimensionale e indichiamone con $\{\mathcal{G}_t := \mathcal{F}_t^B\}_{t \geq 0}$ la filtrazione naturale. Abbiamo già osservato che ogni componente $B^{(i)} = \{B_t^{(i)}\}_{t \geq 0}$ è un moto browniano reale, ma in realtà vale di più. Infatti, per il Teorema 2.29, il vettore aleatorio $B_t - B_s$ è indipendente da \mathcal{G}_s , quindi a maggior ragione ogni sua componente $B_t^{(i)} - B_s^{(i)}$ è indipendente da \mathcal{G}_s . Questo significa che il processo $B^{(i)}$ è in realtà un $\{\mathcal{G}_t\}_{t \geq 0}$ -moto browniano reale. Si osservi che $\mathcal{G}_t = \mathcal{F}_t^B$ è più ampia della filtrazione naturale $\mathcal{F}_t^{B^{(i)}} := \sigma(\{B_u^{(i)}\}_{0 \leq u \leq t})$ della componente $B^{(i)}$. \square

3.2.2. PROCESSI DI LÉVY RISPETTO A UNA FILTRAZIONE. Per i risultati che tratteremo nel resto di questo capitolo, vale la pena lavorare con una classe di processi più ampia del solo moto browniano, che andiamo ora a definire.

DEFINIZIONE 3.16 (PROCESSI DI LÉVY). Un processo stocastico $X = \{X_t\}_{t \geq 0}$ a valori in \mathbb{R}^d , definito su uno spazio filtrato $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$, è detto $\{\mathcal{F}_t\}_{t \geq 0}$ -processo di Lévy (d -dimensionale) — o processo di Lévy (d -dimensionale) rispetto alla filtrazione $\{\mathcal{F}_t\}_{t \geq 0}$ — se è adattato a $\{\mathcal{F}_t\}_{t \geq 0}$ e se soddisfa le seguenti proprietà:

- (a) $X_0 = 0$ q.c.;
- (b) per ogni $0 \leq s < t$, il vettore aleatorio $(X_t - X_s)$ è indipendente da \mathcal{F}_s ;
- (c) per ogni $0 \leq s < t$ e per ogni $h \geq 0$, i vettori $(X_{t+h} - X_{s+h})$ e $(X_t - X_s)$ hanno la stessa legge;
- (d) X è un processo q.c. continuo a destra.

Nel caso in cui $\{\mathcal{F}_t\}_{t \geq 0} = \{\mathcal{F}_t^X\}_{t \geq 0}$ sia la filtrazione naturale del processo, $X = \{X_t\}_{t \geq 0}$ è detto semplicemente processo di Lévy (d -dimensionale).

Rispetto alla Definizione 3.14 di $\{\mathcal{F}_t\}_{t \geq 0}$ -moto browniano, sono state semplicemente rilassate le condizioni (c) e (d). Di conseguenza, un $\{\mathcal{F}_t\}_{t \geq 0}$ -moto browniano d -dimensionale è un caso particolare di $\{\mathcal{F}_t\}_{t \geq 0}$ -processo di Lévy d -dimensionale, per cui tutti i risultati che seguono si applicano in particolare al moto browniano. Sottolineiamo che la classe dei processi di Lévy è ricca e contiene processi molto diversi dal moto browniano.[†]

Ricordando la dimostrazione della Proposizione 2.26, è facile vedere che la proprietà (b) della Definizione 3.16 implica l'indipendenza degli incrementi: più precisamente, se $X = \{X_t\}_{t \geq 0}$ è un $\{\mathcal{F}_t\}_{t \geq 0}$ -processo di Lévy (d -dimensionale), per ogni $k \in \mathbb{N}$ e per ogni $0 \leq t_1 < \dots < t_k$ i vettori aleatori $\{X_{t_i} - X_{t_{i-1}}\}_{1 \leq i \leq k}$ sono indipendenti.

Una conseguenza molto importante della definizione di processo di Lévy è la seguente.

PROPOSIZIONE 3.17. Sia $X = \{X_t\}_{t \geq 0}$ un $\{\mathcal{F}_t\}_{t \geq 0}$ -processo di Lévy (d -dimensionale). Allora il processo X è indipendente dalla σ -algebra \mathcal{F}_0 .

DIMOSTRAZIONE. Dobbiamo mostrare che le σ -algebre $\sigma(X)$ (si ricordi la Definizione 2.23) e \mathcal{F}_0 sono indipendenti. Ricordando che una base di $\sigma(X)$ è data dalla famiglia J^X definita nell'equazione (2.32), è sufficiente dimostrare l'indipendenza del vettore aleatorio $(X_{s_1}, \dots, X_{s_k})$ da \mathcal{F}_0 , per ogni $k \in \mathbb{N}$ e $0 \leq s_1 < \dots < s_k$ fissati.

Introduciamo gli incrementi $Y_i := X_{s_i} - X_{s_{i-1}}$, per $1 \leq i \leq k$ (con $s_0 := 0$) e mostriamo l'indipendenza delle σ -algebre $\{\mathcal{F}_0, \sigma(Y_1), \dots, \sigma(Y_k)\}$. Per ogni scelta di $D \in \mathcal{F}_0$ e di $A_1, \dots, A_k \in \mathcal{B}(\mathbb{R}^d)$, dobbiamo mostrare che

$$\mathbb{P} \left(D \cap \bigcap_{i=1}^k \{Y_i \in A_i\} \right) = \mathbb{P}(D) \prod_{i=1}^k \mathbb{P}(Y_i \in A_i). \quad (3.3)$$

Procediamo per induzione su k . Il caso $k = 1$ segue immediatamente dall'indipendenza di $Y_1 = X_{s_1} - X_0$ da \mathcal{F}_0 , per la proprietà (b) della Definizione 3.16. Sempre per la stessa

[†]Osserviamo tuttavia che, come già notato nell'Osservazione 2.5, ogni processo di Lévy X q.c. continuo (non solo q.c. continuo a destra) è dato da una semplice trasformazione del moto browniano: più precisamente, si può scrivere $X_t = aB_t + bt$, dove $a, b \in \mathbb{R}$ e $B = \{B_t\}_{t \in [0, \infty)}$ è un moto browniano.

proprietà, $Y_k = X_{s_k} - X_{s_{k-1}}$ è indipendente da $\mathcal{F}_{s_{k-1}}$. Osservando che

$$D \cap \bigcap_{i=1}^{k-1} \{Y_i \in A_i\} \in \mathcal{F}_{s_{k-1}},$$

perché i vettori $Y_i = X_{s_i} - X_{s_{i-1}}$ sono $\mathcal{F}_{s_{k-1}}$ -misurabili per $i \leq k-1$, possiamo scrivere

$$P \left(D \cap \bigcap_{i=1}^k \{Y_i \in A_i\} \right) = P \left(D \cap \bigcap_{i=1}^{k-1} \{Y_i \in A_i\} \right) \cdot P(Y_k \in A_k),$$

e la relazione (3.3) segue per l'ipotesi induttiva.

Dall'indipendenza delle σ -algebre $\{\mathcal{F}_0, \sigma(Y_1), \dots, \sigma(Y_k)\}$ segue l'indipendenza del vettore aleatorio (Y_1, \dots, Y_k) dalla σ -algebra \mathcal{F}_0 . Dato che $X_{s_i} = Y_1 + \dots + Y_i$ per ogni $1 \leq i \leq k$, il vettore aleatorio $(X_{s_1}, \dots, X_{s_k})$ è funzione misurabile di (Y_1, \dots, Y_k) e dunque è anch'esso indipendente da \mathcal{F}_0 . \square

3.2.3. AMPLIAMENTO DELLA FILTRAZIONE. Sia X un $\{\mathcal{F}_s\}_{s \geq 0}$ -processo di Lévy (d -dimensionale) e sia $\{\mathcal{F}'_s\}_{s \geq 0}$ una filtrazione ristretta che contenga la filtrazione naturale di X , ossia $\mathcal{F}_s^X \subseteq \mathcal{F}'_s \subseteq \mathcal{F}_s$ per ogni $s \geq 0$. È immediato allora verificare che X è un $\{\mathcal{F}'_s\}_{s \geq 0}$ -processo di Lévy (d -dimensionale).

Se consideriamo invece una filtrazione ampliata $\mathcal{F}''_s \supseteq \mathcal{F}_s$, non è detto che X sia un $\{\mathcal{F}''_s\}_{s \geq 0}$ -processo di Lévy (d -dimensionale), perché la proprietà (b) della Definizione 3.16 potrebbe non valere per \mathcal{F}''_s . Un caso molto importante di ampliamento che non crea problemi è dato da $\{\mathcal{F}_{s+}\}_{s \geq 0}$, come mostriamo ora.

PROPOSIZIONE 3.18. Ogni $\{\mathcal{F}_t\}_{t \geq 0}$ -processo di Lévy (d -dimensionale) $X = \{X_t\}_{t \geq 0}$ è anche un $\{\mathcal{F}_{t+}\}_{t \geq 0}$ -processo di Lévy (d -dimensionale).

DIMOSTRAZIONE. Dobbiamo solo verificare che $X_t - X_s$ è indipendente dalla σ -algebra $\mathcal{F}_{s+} := \bigcap_{\varepsilon > 0} \mathcal{F}_{s+\varepsilon}$. Per la proprietà (b) della Definizione 3.16, per ogni $\varepsilon > 0$ si ha che $X_{t+\varepsilon} - X_{s+\varepsilon}$ è indipendente da $\mathcal{F}_{s+\varepsilon}$, quindi a maggior ragione è indipendente da $\mathcal{F}_{s+} \subseteq \mathcal{F}_{s+\varepsilon}$. Di conseguenza, ricordando la relazione (3.1), per ogni $A \in \mathcal{F}_{s+}$ con $P(A) > 0$ e per ogni $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$ continua e limitata si ha

$$E(\Phi(X_{t+1/n} - X_{s+1/n}) | A) = E(\Phi(X_{t+1/n} - X_{s+1/n})), \quad \forall n \in \mathbb{N}. \quad (3.4)$$

Per definizione il processo X è q.c. continuo a destra. Dato che la funzione Φ è continua e limitata, si ha $\lim_{n \rightarrow \infty} \Phi(X_{t+1/n} - X_{s+1/n}) = \Phi(X_t - X_s)$ q.c. e, prendendo il limite $n \rightarrow \infty$ in (3.4), per convergenza dominata si ottiene

$$E(\Phi(X_t - X_s) | A) = E(\Phi(X_t - X_s)),$$

per ogni $A \in \mathcal{F}_{s+}$ e per ogni funzione $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$ continua e limitata. Ricordando ancora la relazione (3.1), l'indipendenza di $X_t - X_s$ da \mathcal{F}_{s+} è dimostrata. \square

Combinando la Proposizione 3.17 e la Proposizione 3.18 si ottengono conseguenze molto interessanti.

TEOREMA 3.19 (LEGGE 0-1 DI BLUMENTHAL). Sia $X = \{X_t\}_{t \geq 0}$ un processo di Lévy (d -dimensionale) e sia $\{\mathcal{F}_t^X\}_{t \geq 0}$ la sua filtrazione naturale. La σ -algebra \mathcal{F}_{0+}^X è banale: per ogni $A \in \mathcal{F}_{0+}^X$ si ha $P(A) = 0$ oppure $P(A) = 1$.

DIMOSTRAZIONE. Per ipotesi $X = \{X_t\}_{t \geq 0}$ è un $\{\mathcal{F}_t^X\}_{t \geq 0}$ -processo di Lévy, quindi per la Proposizione 3.18 è anche un $\{\mathcal{F}_{t+}^X\}_{t \geq 0}$ -processo di Lévy. Per la Proposizione 3.17, il processo $X = \{X_t\}_{t \geq 0}$ è indipendente da \mathcal{F}_{0+}^X , cioè $\sigma(X) := \sigma(\{X_t\}_{t \geq 0})$ è indipendente da \mathcal{F}_{0+}^X . Dato che $\mathcal{F}_{0+}^X \subseteq \sigma(X)$, segue che \mathcal{F}_{0+}^X è indipendente da sé stessa: per ogni $A \in \mathcal{F}_{0+}^X$ si ha dunque $P(A) = P(A \cap A) = P(A)^2$, per cui $P(A) = 0$ oppure $P(A) = 1$. \square

Sottolineiamo che la σ -algebra \mathcal{F}_{0+}^X non è vuota, ma contiene al contrario molti eventi interessanti. Intuitivamente, essa consiste di tutti gli eventi che si possono decidere (ossia, per i quali si può dire se si siano verificati) osservando il processo X in un intorno arbitrariamente piccolo dell'origine. Ad esempio, qualunque sia la funzione $f : [0, \infty) \rightarrow (0, \infty)$, gli eventi $\{\limsup_{h \downarrow 0} X_h/f(h) = 1\}$ e $\{\liminf_{h \downarrow 0} X_h/f(h) = -1\}$ sono in \mathcal{F}_{0+}^X : di conseguenza, per dimostrare che le relazioni in (2.31) per il moto browniano valgono q.c. (per $t_0 = 0$, senza perdita di generalità), basta mostrare che esse sono verificate su un evento di probabilità strettamente positiva.

Supponiamo ora che $\{\mathcal{F}_t\}_{t \geq 0}$ sia una filtrazione su uno spazio di probabilità (Ω, \mathcal{F}, P) completo.

PROPOSIZIONE 3.20. Ogni $\{\mathcal{F}_t\}_{t \geq 0}$ -processo di Lévy (d -dimensionale) $X = \{X_t\}_{t \geq 0}$ è anche un $\{\overline{\mathcal{F}}_{t+}\}_{t \geq 0}$ -processo di Lévy (d -dimensionale).

DIMOSTRAZIONE. Dobbiamo verificare che $X_t - X_s$ è indipendente dalla σ -algebra $\overline{\mathcal{F}}_{s+} := \sigma(\mathcal{F}_{s+}, \mathcal{N})$, dove $\mathcal{N} := \{C \in \mathcal{F} : P(C) = 0\}$. Affermiamo che vale il seguente fatto generale: se una variabile Y è indipendente da una σ -algebra $\mathcal{H} \subseteq \mathcal{F}$, lo è anche da $\overline{\mathcal{H}} := \sigma(\mathcal{H}, \mathcal{N})$. Ricordando la Proposizione 3.18 e scegliendo $Y = X_t - X_s$ e $\mathcal{H} = \mathcal{F}_{s+}$, si ha la tesi.

Resta da verificare quanto affermato. Ricordiamo che $A \in \overline{\mathcal{H}}$ se e soltanto se esistono $A' \in \mathcal{H}$ e $C \in \mathcal{N}$ tali che $A \triangle A' = C$. Dato che $P(C) = 0$, segue che $P(A) = P(A')$ e più in generale che $P(F \cap A) = P(F \cap A')$, per ogni $F \in \mathcal{F}$. Scegliendo $F = \{Y \in D\}$ e ricordando che Y è per ipotesi indipendente da \mathcal{H} , si ha dunque

$$P(Y \in D, A) = P(Y \in D, A') = P(Y \in D) P(A') = P(Y \in D) P(A),$$

da cui segue l'indipendenza di Y da $\overline{\mathcal{H}}$, per l'arbitrarietà di $D \in \mathcal{B}(\mathbb{R}^n)$ e di $A \in \overline{\mathcal{H}}$. \square

3.3. TEMPI D'ARRESTO

Ricordiamo che \mathbb{T} indica un sottoinsieme di \mathbb{R} , che ha la funzione di insieme dei tempi. Ci limitiamo per semplicità ai casi $\mathbb{T} = \mathbb{N}_0$, $\mathbb{T} = [0, \infty)$ oppure $\mathbb{T} = [a, b]$, con $0 \leq a < b < \infty$.

Introduciamo la nozione fondamentale di tempo d'arresto. Ricordiamo che, data una filtrazione $\{\mathcal{F}_t\}_{t \in \mathbb{T}}$ su uno spazio di probabilità (Ω, \mathcal{F}, P) , indichiamo con $\mathcal{F}_\infty := \sigma(\{\mathcal{F}_t\}_{t \in \mathbb{T}})$ la più piccola σ -algebra che contiene tutte le \mathcal{F}_t .

DEFINIZIONE 3.21. Dato uno spazio filtrato $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \mathbb{T}}, P)$, un *tempo d'arresto* è una variabile aleatoria $\tau : \Omega \rightarrow \mathbb{T} \cup \{+\infty\}$ tale che $\{\tau \leq t\} \in \mathcal{F}_t$, per ogni $t \in \mathbb{T}$. Si definisce la σ -algebra \mathcal{F}_τ ponendo

$$\mathcal{F}_\tau := \{A \in \mathcal{F}_\infty : A \cap \{\tau \leq t\} \in \mathcal{F}_t, \forall t \in \mathbb{T}\}. \quad (3.5)$$

Per verificare che \mathcal{F}_τ è una σ -algebra, si noti che $(A \cap \{\tau \leq t\}) \cup (A^c \cap \{\tau \leq t\}) = \{\tau \leq t\}$, da cui segue che $A^c \cap \{\tau \leq t\} = \{\tau \leq t\} \setminus (A \cap \{\tau \leq t\}) = \{\tau \leq t\} \cap (A \cap \{\tau \leq t\})^c$, quindi se $A \in \mathcal{F}_\tau$ anche $A^c \in \mathcal{F}_\tau$. Le verifiche che $\emptyset \in \mathcal{F}_\tau$ e che $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{F}_\tau$ se $A_n \in \mathcal{F}_\tau$ per ogni $n \in \mathbb{N}$ sono immediate.

Intuitivamente, un tempo d'arresto descrive un istante aleatorio deciso in base all'informazione disponibile fino al tempo presente, cioè senza guardare al futuro. Infatti la condizione $\{\tau \leq t\} \in \mathcal{F}_t$ significa che possiamo dire se l'istante τ è scoccato prima di t sulla base dell'informazione disponibile fino al tempo t .

Se la σ -algebra \mathcal{F}_t contiene intuitivamente gli eventi conoscibili entro l'istante t , la σ -algebra \mathcal{F}_τ contiene intuitivamente *gli eventi conoscibili entro l'istante (aleatorio) τ* . La condizione $A \cap \{\tau \leq t\} \in \mathcal{F}_t$ significa infatti che, quando l'istante τ è scoccato prima di t , l'evento A risulta conoscibile entro il tempo t .

OSSERVAZIONE 3.22. Nel caso discreto $\mathbb{T} = \mathbb{N}_0$, affinché τ sia un tempo d'arresto è sufficiente richiedere che $\{\tau = n\} \in \mathcal{F}_n$ per ogni $n \in \mathbb{N}$. Analogamente, gli eventi $A \in \mathcal{F}_\tau$ sono tutti e soli quelli per cui $A \cap \{\tau = n\} \in \mathcal{F}_n$ per ogni $n \in \mathbb{N}$. \square

Vediamo ora un esempio tipico di tempo d'arresto. Sia $X = \{X_t\}_{t \in \mathbb{T}}$ un processo stocastico, definito su uno spazio filtrato $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \mathbb{T}}, P)$ a valori in uno *spazio metrico* (E, \mathcal{E}) munito della σ -algebra boreliana, che sia *adattato* alla filtrazione. Dato un sottoinsieme $D \subseteq E$, definiamo il tempo d'ingresso di X in D ponendo

$$\tau_D := \inf\{t \in \mathbb{T} : X_t \in D\},$$

con la convenzione $\inf\{\emptyset\} := \infty$.

Nel caso di insieme dei tempi numerabile, $\mathbb{T} = \mathbb{N}_0$, è molto facile mostrare che τ_D è un tempo d'arresto, per ogni insieme $D \in \mathcal{E}$. Se invece l'insieme dei tempi è più che numerabile sono necessarie ipotesi aggiuntive. Per il prossimo risultato supponiamo che $\mathbb{T} = [0, \infty)$ o $\mathbb{T} = [a, b] \subseteq [0, \infty)$.

LEMMA 3.23. Se il processo $X = \{X_t\}_{t \in \mathbb{T}}$ è *adattato* e *continuo a destra*, allora:

- (a) per ogni insieme chiuso $C \subseteq E$, la variabile τ_C è un tempo d'arresto;
- (b) per ogni insieme aperto $A \subseteq E$, la variabile τ_A è un tempo d'arresto per la filtrazione $\{\mathcal{F}_{t+}\}_{t \in \mathbb{T}}$ (quindi è un tempo d'arresto se $\{\mathcal{F}_t\}_{t \in \mathbb{T}}$ è continua a destra).

Entrambe le proprietà sono vere anche nel caso in cui X sia solo *q.c. continuo a destra*, purché lo spazio di probabilità (Ω, \mathcal{F}, P) e la filtrazione $\{\mathcal{F}_t\}_{t \in \mathbb{T}}$ siano completi.

DIMOSTRAZIONE. Indicando con $d(\cdot, \cdot)$ la distanza in E , è noto (e facile da verificare) che per ogni sottoinsieme $D \subseteq E$ la funzione $x \mapsto d(x, D) := \inf_{z \in D} d(x, z)$, definita da E in \mathbb{R} , è continua; inoltre, se $C \subseteq E$ è un sottoinsieme chiuso, si ha che $d(x, C) = 0$ se e solo se $x \in C$. Di conseguenza, se X è un

processo continuo a destra, per ogni $\omega \in \Omega$ la funzione reale $u \mapsto d(X_u(\omega), C)$ è continua a destra e si annulla in tutti e soli i punti $u \geq 0$ per cui $X_u(\omega) \in C$. Possiamo dunque scrivere

$$\{\tau_C \leq t\} = \{\exists u \in \mathbb{T} \cap [0, t] : d(X_u, C) = 0\} = \left\{ \inf_{u \in (\mathbb{T} \cap [0, t] \cap \mathbb{Q}) \cup \{t\}} d(X_u, C) = 0 \right\}.$$

Se X è adattato, per ogni $u \in \mathbb{T} \cap [0, t]$ la variabile aleatoria X_u è \mathcal{F}_t -misurabile, quindi anche $d(X_u, C)$ lo è (composizione di funzioni misurabili). L'estremo inferiore di una famiglia *numerabile* di funzioni \mathcal{F}_t -misurabili è \mathcal{F}_t -misurabile, per cui l'evento in questione è in \mathcal{F}_t .

Per quanto riguarda τ_A , si noti che per ogni $s > 0$

$$\{\tau_A < s\} = \bigcup_{u \in \mathbb{T} \cap [0, s)} \{X_u \in A\} = \bigcup_{u \in \mathbb{T} \cap [0, s) \cap \mathbb{Q}} \{X_u \in A\},$$

per cui $\{\tau_A < s\} \in \mathcal{F}_s$. Si noti che $\{\tau_A \leq t\} = \bigcap_{n \geq N} \{\tau_A < t + \frac{1}{n}\}$, per ogni N fissato, da cui $\{\tau_A \leq t\} \in \mathcal{F}_{t+\frac{1}{N}}$. Dato che ciò è vero per ogni $N \in \mathbb{N}$, si ha che $\{\tau_A \leq t\} \in \bigcap_{N \in \mathbb{N}} \mathcal{F}_{t+\frac{1}{N}} = \mathcal{F}_{t+}$.

L'estensione al caso in cui X è solo q.c. continuo a destra è immediata. \square

D'ora in avanti supponiamo per semplicità che $\mathbb{T} = [0, \infty)$. Elenchiamo alcune proprietà dei tempi d'arresto, la cui verifica è lasciata come esercizio. Per ogni istante t_0 fissato, il tempo (deterministico) definito da $\tau(\omega) \equiv t_0$ per ogni $\omega \in \Omega$ è un tempo d'arresto. Inoltre, per ogni tempo d'arresto τ si ha che:

- la variabile aleatoria τ è \mathcal{F}_τ -misurabile (basta verificare che $\{\tau \leq t\} \in \mathcal{F}_\tau$ per ogni $t \in [0, \infty)$);
- $\tau + \delta$ è un tempo d'arresto, per ogni costante $\delta \geq 0$.

Dati due tempi d'arresto τ e σ , definiti sullo stesso spazio filtrato, si ha che:

- $\tau \wedge \sigma := \min\{\tau, \sigma\}$ e $\tau \vee \sigma := \max\{\tau, \sigma\}$ sono tempi d'arresto;
- se $\sigma(\omega) \leq \tau(\omega)$ per ogni $\omega \in \Omega$, allora $\mathcal{F}_\sigma \subseteq \mathcal{F}_\tau$.

LEMMA 3.24. Se $X = \{X_t\}_{t \geq 0}$ è un processo progressivamente misurabile e τ è un tempo d'arresto finito, X_τ (cioè $\omega \mapsto X_{\tau(\omega)}(\omega)$) è una variabile aleatoria \mathcal{F}_τ -misurabile.

Questo risultato vale anche nel caso in cui τ è q.c. finito, a patto di definire $X_\tau := c$ sull'evento $\{\tau = \infty\}$, dove c è un arbitrario elemento fissato di E .

DIMOSTRAZIONE. Si noti che X_τ è una funzione misurabile, in quanto composizione delle funzioni misurabili $\omega \mapsto (\omega, \tau(\omega))$ e $(\omega, t) \mapsto X_t(\omega)$. Resta da dimostrare che, per ogni $A \in \mathcal{E}$, si ha $\{X_\tau \in A\} \in \mathcal{F}_\tau$, il che equivale a dire che, per ogni $t \geq 0$, $\{\tau \leq t\} \cap \{X_\tau \in A\} \in \mathcal{F}_t$.

Poniamo $\Omega_t := \{\tau \leq t\} = \{\omega \in \Omega : \tau(\omega) \leq t\}$. Muniamo Ω_t della σ -algebra ristretta $\mathcal{F}_t| := \{A \in \mathcal{F}_t : A \subseteq \Omega_t\} = \mathcal{F}_t \cap \Omega_t$, ottenendo lo spazio misurabile $(\Omega_t, \mathcal{F}_t|)$. Si noti che l'applicazione ristretta $\tau : (\Omega_t, \mathcal{F}_t|) \rightarrow [0, \infty)$ è misurabile, in quanto

$$\{\omega \in \Omega_t : \tau(\omega) \leq s\} = \Omega_t \cap \{\tau \leq s\} = \{\tau \leq \min\{s, t\}\} \in \mathcal{F}_t|.$$

Di conseguenza, anche l'applicazione ristretta $X_\tau : (\Omega_t, \mathcal{F}_t|) \rightarrow (E, \mathcal{E})$ è misurabile, in quanto composizione delle applicazioni misurabili $\omega \mapsto (\omega, \tau(\omega))$ da $(\Omega_t, \mathcal{F}_t|)$ in $(\Omega_t \times [0, t], \mathcal{F}_t| \otimes \mathcal{B}[0, t])$ e $(\omega, t) \mapsto X_t(\omega)$ da $(\Omega_t \times [0, t], \mathcal{F}_t| \otimes \mathcal{B}[0, t])$ in (E, \mathcal{E}) . Di conseguenza, per ogni $A \in \mathcal{E}$ si ha che

$$\{\tau \leq t\} \cap \{X_\tau \in A\} = \{\omega \in \Omega_t : X_{\tau(\omega)}(\omega) \in A\} \in \mathcal{F}_t| \subseteq \mathcal{F}_t,$$

e la dimostrazione è conclusa. \square

Chiudiamo il paragrafo con un utile risultato di approssimazione: se τ è un tempo d'arresto, esiste una successione decrescente τ_n di tempi d'arresto *discreti* (che assumono cioè una quantità al più numerabile di valori) tali che, per ogni $\omega \in \Omega$, $\tau_n(\omega) \downarrow \tau(\omega)$ per $n \rightarrow \infty$ e inoltre $\tau_n(\omega) = \infty$ se e solo se $\tau(\omega) = \infty$, per ogni $n \in \mathbb{N}$. Basta infatti porre

$$\tau_n(\omega) := \sum_{k=0}^{\infty} \frac{k}{2^n} \mathbf{1}_{(\frac{k-1}{2^n}, \frac{k}{2^n}]}(\tau(\omega)) + \infty \mathbf{1}_{\{+\infty\}}(\tau(\omega)).$$

Chiaramente $\tau(\omega) \leq \tau_n(\omega) \leq \tau(\omega) + 2^{-n}$, da cui segue che $\tau_n \downarrow \tau$. Inoltre τ_n è un tempo d'arresto per ogni $n \in \mathbb{N}$, dal momento che

$$\{\tau_n \leq t\} = \bigcup_{k \leq t2^n} \left\{ \tau \in \left(\frac{k-1}{2^n}, \frac{k}{2^n} \right] \right\},$$

e $\{\tau \in (a, b]\} = \{\tau \leq b\} \cap \{\tau \leq a\}^c \in \mathcal{F}_t$ per ogni $0 \leq a < b \leq t$.

3.4. LA PROPRIETÀ DI MARKOV FORTE

Abbiamo visto nella Proposizione 2.10 (b) l'invarianza del moto browniano per traslazioni temporali. Estendiamo ora questa proprietà in modo sostanziale a tutti i processi di Lévy, considerando tempi d'arresto invece di tempi deterministici.

TEOREMA 3.25 (PROPRIETÀ DI MARKOV FORTE PER PROCESSI DI LÉVY). Sia $X = \{X_t\}_{t \geq 0}$ un $\{\mathcal{F}_t\}_{t \geq 0}$ -processo di Lévy (d -dimensionale) e sia τ un tempo d'arresto q.c. finito. Per $t \in [0, \infty)$ definiamo $Z_t := X_{\tau+t} - X_\tau$ e $\mathcal{G}_t := \mathcal{F}_{\tau+t}$. Allora il processo $Z = \{Z_t\}_{t \geq 0}$ è un $\{\mathcal{G}_t\}_{t \geq 0}$ -processo di Lévy (d -dimensionale) con la stessa legge di X .

DIMOSTRAZIONE. Notiamo innanzitutto che $\tau + t$ è un tempo d'arresto, per ogni $t \geq 0$, quindi $\mathcal{G}_t := \mathcal{F}_{\tau+t}$ è ben definita come σ -algebra, cf. (3.5). Dal fatto che $\tau + s \leq \tau + t$ per $s \leq t$ segue che $\{\mathcal{G}_t\}_{t \geq 0}$ è una filtrazione. Notiamo inoltre che, per il Lemma 3.24, $Z_t = X_{\tau+t} - X_\tau$ è $\mathcal{F}_{\tau+t}$ -misurabile, quindi il processo Z è adattato alla filtrazione $\{\mathcal{G}_t\}_{t \geq 0}$.

Mostriamo ora che il processo Z soddisfa le proprietà della Definizione 3.16. Le proprietà (a) e (d) sono immediate: infatti $Z_0 = X_\tau - X_\tau = 0$ e Z è un processo q.c. continuo a destra, perché per ipotesi X lo è. Per ottenere le proprietà (b) e (c), cominciamo a fissare due istanti $0 \leq s < t$, una funzione $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$ continua e limitata e un evento $G \in \mathcal{G}_s$ con $P(G) > 0$ e mostriamo che vale la seguente relazione:

$$E(\Phi(X_{\tau+t} - X_{\tau+s}) | G) = E(\Phi(X_{t-s})). \quad (3.6)$$

Cominciamo a supporre che il tempo d'arresto τ assuma un insieme finito o numerabile di valori $\{t_m\}_{m \in I}$. Per ogni $m \in I$ possiamo scrivere

$$E(\Phi(X_{\tau+t} - X_{\tau+s}) \mathbf{1}_G \mathbf{1}_{\{\tau=t_m\}}) = E(\Phi(X_{t_m+t} - X_{t_m+s}) \mathbf{1}_{G \cap \{\tau=t_m\}}).$$

Per ipotesi $G \in \mathcal{G}_s := \mathcal{F}_{\tau+s}$, quindi $G \cap \{\tau = t_m\} = G \cap \{\tau + s = t_m + s\} \in \mathcal{F}_{t_m+s}$. D'altro canto, X è un processo di Lévy e per la proprietà (b) della Definizione 3.16 la

variabile $X_{t_m+t} - X_{t_m+s}$ è indipendente da \mathcal{F}_{t_m+s} ; quindi le variabili $\Phi(X_{t_m+t} - X_{t_m+s})$ e $1_{G \cap \{\tau=t_m\}}$ sono indipendenti. Di conseguenza

$$\mathbb{E}(\Phi(X_{\tau+t} - X_{\tau+s})1_G 1_{\{\tau=t_m\}}) = \mathbb{E}(\Phi(X_{t_m+t} - X_{t_m+s})) \cdot \mathbb{P}(G \cap \{\tau=t_m\}).$$

Per le proprietà (a) e (b) della Definizione 3.16, $X_{t_m+t} - X_{t_m+s}$ ha la stessa legge di X_{t-s} , per cui sommando su $m \in I$ l'ultima relazione si ottiene

$$\begin{aligned} \mathbb{E}(\Phi(X_{\tau+t} - X_{\tau+s})1_G) &= \sum_{m \in I} \mathbb{E}(\Phi(X_{\tau+t} - X_{\tau+s})1_G 1_{\{\tau=t_m\}}) \\ &= \mathbb{E}(\Phi(X_{t-s})) \sum_{m \in I} \mathbb{P}(G \cap \{\tau=t_m\}) = \mathbb{E}(\Phi(X_{t-s})) \cdot \mathbb{P}(G). \end{aligned}$$

Abbiamo dunque dimostrato la relazione (3.6) se τ sia un tempo d'arresto discreto.

Per trattare il caso in cui τ è un tempo d'arresto arbitrario, sia τ_n una successione di tempi d'arresto discreti tali che $\tau_n \downarrow \tau$. Per quanto già mostrato, si ha

$$\mathbb{E}(\Phi(X_{\tau_n+t} - X_{\tau_n+s})|G) = \mathbb{E}(\Phi(X_{t-s})), \quad \forall n \in \mathbb{N}.$$

Il processo X è q.c. continuo a destra e la funzione Φ è continua, quindi $\Phi(X_{\tau_n+t} - X_{\tau_n+s}) \rightarrow \Phi(X_{\tau+t} - X_{\tau+s})$ q.c. per $n \rightarrow \infty$. Per convergenza dominata (ricordiamo che Φ è limitata) si ottiene dunque la relazione (3.6) nel caso generale.

Dato che $Z_t - Z_s = X_{\tau+t} - X_{\tau+s}$, possiamo riscrivere la relazione (3.6) come

$$\mathbb{E}(\Phi(Z_t - Z_s)|G) = \mathbb{E}(\Phi(X_{t-s})), \quad (3.7)$$

per ogni $0 \leq s < t$, per ogni funzione $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$ continua e limitata e per ogni evento $G \in \mathcal{G}_s$ con $\mathbb{P}(G) > 0$. Scegliendo $G = \Omega$, otteniamo in particolare

$$\mathbb{E}(\Phi(Z_t - Z_s)) = \mathbb{E}(\Phi(X_{t-s})). \quad (3.8)$$

Ricordando che una probabilità è determinata dagli integrali delle funzioni continue, questa relazione mostra che la legge di $Z_t - Z_s$ è una funzione di $t - s$, dunque la proprietà (c) della Definizione 3.16 è soddisfatta. Ricordando la relazione (3.1), segue da (3.7) e (3.8) che la variabile $Z_t - Z_s$ è indipendente da \mathcal{G}_s , dunque è verificata anche la proprietà (b). Abbiamo quindi provato che Z è un $\{\mathcal{G}_t\}_{t \geq 0}$ -processo di Lévy.

Resta infine da mostrare che Z e X hanno la stessa legge. Essendo entrambi processi di Lévy, è sufficiente mostrare che per ogni $t \geq 0$ le variabili Z_t e X_t hanno la stessa legge (cf. il Lemma 3.26 più sotto), e questo segue dalla relazione (3.8) con $s = 0$. \square

LEMMA 3.26. Due processi di Lévy (d -dimensionali) $X = \{X_t\}_{t \geq 0}$, $Z = \{Z_t\}_{t \geq 0}$ hanno la stessa legge se e solo se le variabili X_t e Y_t hanno la stessa legge, $\forall t \geq 0$.

DIMOSTRAZIONE. Due processi hanno la stessa legge se e solo se hanno le stesse leggi finito-dimensionali. Mostriamo dunque che, se Z_t ha la stessa legge di X_t , per ogni $t \geq 0$,

allora i vettori aleatori $(Z_{s_1}, Z_{s_2}, \dots, Z_{s_k})$ e $(X_{s_1}, X_{s_2}, \dots, X_{s_k})$ hanno la stessa legge, per ogni $k \in \mathbb{N}$ e $0 \leq s_1 < s_2 < \dots < s_k$ (il viceversa è evidente).

Convienne introdurre i vettori degli incrementi $\Delta_Z := (Z_{s_1}, Z_{s_2} - Z_{s_1}, \dots, Z_{s_k} - Z_{s_{k-1}})$ e $\Delta_X := (X_{s_1}, X_{s_2} - X_{s_1}, \dots, X_{s_k} - X_{s_{k-1}})$. Si noti che $(Z_{s_1}, Z_{s_2}, \dots, Z_{s_k}) = f(\Delta_Z)$ e analogamente $(X_{s_1}, X_{s_2}, \dots, X_{s_k}) = f(\Delta_X)$, dove f indica *la stessa funzione* misurabile.[†] Di conseguenza è sufficiente mostrare che i vettori aleatori Δ_Z e Δ_X hanno la stessa legge (perché?). Entrambi i vettori hanno componenti indipendenti (i processi di Lévy hanno incrementi indipendenti), quindi basta mostrare che le leggi delle rispettive componenti coincidono. Per la proprietà (c) della Definizione 3.16, $Z_{s_i} - Z_{s_{i-1}}$ ha la stessa legge di $Z_{s_i - s_{i-1}}$ e analogamente $X_{s_i} - X_{s_{i-1}}$ ha la stessa legge di $X_{s_i - s_{i-1}}$, quindi la conclusione segue dall'ipotesi che Z_t e X_t abbiano la stessa legge per ogni $t \geq 0$. \square

Per comodità, nei prossimi due teoremi (che sono corollari del Teorema 3.25 e della Proposizione 3.17) formuliamo esplicitamente la proprietà di Markov forte per il moto browniano, considerando innanzitutto il caso speciale di tempi d'arresto deterministici (*proprietà di Markov semplice*).

TEOREMA 3.27 (PROPRIETÀ DI MARKOV SEMPLICE PER IL MOTO BROWNIANO). Se $B = \{B_t\}_{t \geq 0}$ è un $\{\mathcal{F}_t\}_{t \geq 0}$ -moto browniano (d -dimensionale), per ogni $t_0 \geq 0$ fissato il processo $Z = \{Z_t\}_{t \geq 0}$ definito da $Z_t := B_{t_0+t} - B_{t_0}$ è un $\{\mathcal{F}_{t_0+t}\}_{t \geq 0}$ -moto browniano (d -dimensionale) indipendente dalla σ -algebra \mathcal{F}_{t_0} .

TEOREMA 3.28 (PROPRIETÀ DI MARKOV FORTE PER IL MOTO BROWNIANO). Sia $B = \{B_t\}_{t \geq 0}$ un $\{\mathcal{F}_t\}_{t \geq 0}$ -moto browniano (d -dimensionale) e sia τ un tempo d'arresto q.c. finito. Ponendo $Z_t := B_{\tau+t} - B_\tau$ per $t \geq 0$, il processo $Z = \{Z_t\}_{t \geq 0}$ è un $\{\mathcal{F}_{\tau+t}\}_{t \geq 0}$ -moto browniano (d -dimensionale) indipendente da \mathcal{F}_τ .

3.5. IL PRINCIPIO DI RIFLESSIONE

Tra le molteplici conseguenze della proprietà di Markov forte, dimostriamo il celebre *principio di riflessione* per il moto browniano reale (si veda la Figura 3.1).

TEOREMA 3.29 (PRINCIPIO DI RIFLESSIONE). Sia $B = \{B_t\}_{t \geq 0}$ un moto browniano reale e siano $S_t := \sup_{0 \leq s \leq t} B_s$ e $\tau_a := \inf\{t \geq 0 : B_t = a\}$. Per ogni $a, t > 0$ si ha

$$P(\tau_a \leq t) = P(S_t \geq a) = P(|B_t| \geq a).$$

DIMOSTRAZIONE. La prima uguaglianza è ovvia. Per la seconda, notiamo che

$$\begin{aligned} P(S_t \geq a) &= P(S_t \geq a, B_t \geq a) + P(S_t \geq a, B_t < a) \\ &= P(B_t \geq a) + P(S_t \geq a, B_t < a). \end{aligned}$$

[†] Esplicitamente: $f(y_1, \dots, y_k) = (y_1, y_1 + y_2, \dots, y_1 + \dots + y_k)$, per ogni $(y_1, \dots, y_k) \in (\mathbb{R}^d)^k$.

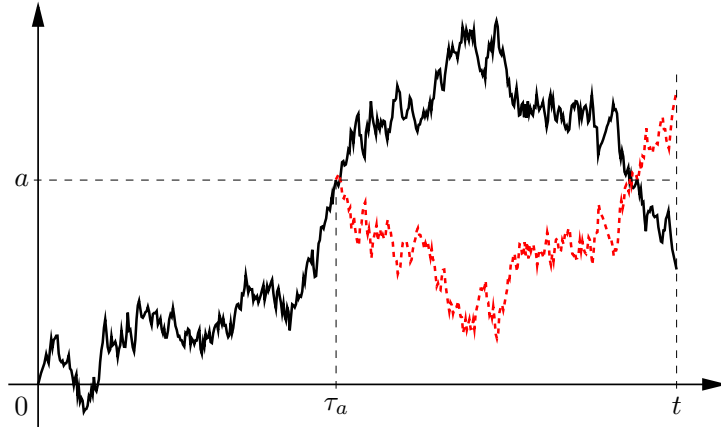


FIGURA 3.1. Rappresentazione grafica del principio di riflessione: la linea tratteggiata mostra la porzione di traiettoria $Z = \{Z_s\}_{0 \leq s \leq t - \tau_a}$ dopo la riflessione.

Introducendo il processo $Z := \{Z_s = B_{\tau_a + s} - B_{\tau_a}\}_{s \geq 0}$, possiamo scrivere

$$P(S_t \geq a, B_t < a) = P(\tau_a \leq t, B_t < a) = P(\tau_a \leq t, Z_{t - \tau_a} < 0),$$

perché $B_{\tau_a} = a$. Indicando con $C := C([0, \infty), \mathbb{R})$ lo spazio delle funzioni continue da $[0, \infty)$ in \mathbb{R} , possiamo vedere Z come una variabile aleatoria a valori in C . Definiamo

$$H_t := \{(s, f) \in [0, \infty) \times C : s \leq t \text{ e } f(t - s) < 0\}.$$

Il sottoinsieme H_t è misurabile in $[0, \infty) \times C$, perché l'applicazione $(s, f) \mapsto f(t - s)$ è continua su $[0, t] \times C$, e possiamo scrivere

$$P(S_t \geq a, B_t < a) = P((\tau_a, Z) \in H_t).$$

Per la proprietà di Markov forte (Teorema 3.28), il processo Z è un moto browniano indipendente da \mathcal{F}_{τ_a} , quindi è indipendente dalla variabile τ_a (che è \mathcal{F}_{τ_a} -misurabile e dunque $\sigma(\tau_a) \subseteq \mathcal{F}_{\tau_a}$).[†] Dato che Z ha la stessa legge di $-Z$, cioè la misura di Wiener (sono entrambi moti browniani), le variabili aleatorie (τ_a, Z) e $(\tau_a, -Z)$, a valori in $[0, \infty) \times C$, hanno *la stessa legge congiunta*, data dal prodotto della legge di τ_a con la misura di Wiener; di conseguenza $P((\tau_a, Z) \in H_t) = P((\tau_a, -Z) \in H_t)$. Si ha dunque

$$\begin{aligned} P(S_t \geq a, B_t < a) &= P((\tau_a, -Z) \in H_t) = P(\tau_a \leq t, -Z_{t - \tau_a} < 0) \\ &= P(\tau_a \leq t, B_t > a) = P(B_t > a) = P(B_t \geq a). \end{aligned}$$

Abbiamo dunque mostrato che $P(S_t \geq a) = 2P(B_t \geq a) = P(|B_t| \geq a)$. \square

Il principio di riflessione mostra che, per ogni $t \geq 0$, S_t ha la stessa legge di $|B_t|$. Sottolineiamo che l'uguaglianza in legge vale solo per un istante fissato: infatti i processi $\{S_t\}_{t \geq 0}$ e $\{|B_t|\}_{t \geq 0}$ sono molto diversi (il primo è crescente, mentre il secondo no!).

[†]Per applicare il Teorema 3.28 dovremmo teoricamente sapere che $\tau_a < \infty$ q.c., ma in realtà non ce n'è bisogno: ai fini della dimostrazione basta infatti ridefinire τ_a come $\min\{\tau_a, 2t\}$.

4. MARTINGALE

In questo capitolo richiamiamo le nozioni e i risultati fondamentali sulla speranza condizionale e la teoria delle martingale (per maggiori dettagli, si veda [Williams, 1991]). Ricordiamo la notazione $a \wedge b := \min\{a, b\}$, per $a, b \in \mathbb{R}$.

4.1. SPERANZA CONDIZIONALE

4.1.1. DEFINIZIONE. Sia (Ω, \mathcal{F}, P) uno spazio di probabilità e sia \mathcal{G} una sotto- σ -algebra di \mathcal{F} . È possibile mostrare che per ogni variabile aleatoria reale *integrabile* X definita su Ω esiste una variabile aleatoria reale Z che sia \mathcal{G} -misurabile (cioè $Z : (\Omega, \mathcal{G}) \rightarrow \mathbb{R}$ è misurabile) e tale che valga la seguente relazione:

$$\int_A X \, dP = \int_A Z \, dP, \quad \text{cioè} \quad E(X \mathbf{1}_A) = E(Z \mathbf{1}_A), \quad \forall A \in \mathcal{G}. \quad (4.1)$$

Questo è equivalente a richiedere che $E(XY) = E(ZY)$ per ogni variabile aleatoria reale Y \mathcal{G} -misurabile e limitata. La variabile Z non è unica: tuttavia, se Z_1, Z_2 sono variabili aleatorie \mathcal{G} -misurabili per cui vale (4.1), si ha che $Z_1 = Z_2$ q.c.. Risulta dunque univocamente determinata la classe di equivalenza in $L^1(\Omega, \mathcal{G}, P)$ delle variabili aleatorie Z che soddisfano la relazione (4.1), per ogni $A \in \mathcal{G}$, che è detta *speranza condizionale di X rispetto a \mathcal{G}* ed è indicata con $E(X|\mathcal{G})$. Con abuso di notazione, chiameremo “speranza condizionale” ogni specifico elemento Z di $E(X|\mathcal{G})$ e scriveremo $Z = E(X|\mathcal{G})$ q.c..

Intuitivamente, la speranza condizionale $E(X|\mathcal{G})$ è la variabile aleatoria \mathcal{G} -misurabile che meglio approssima X . Qualche esempio basilare:

- se X è \mathcal{G} -misurabile si ha $E(X|\mathcal{G}) = X$ q.c.;
- se $\mathcal{G} = \{\emptyset, \Omega\}$ si ha $E(X|\mathcal{G}) = E(X)$ q.c.;
- se $\mathcal{G} = \{\emptyset, A, A^c, \Omega\}$, per un opportuno $A \in \mathcal{F}$ con $0 < P(A) < 1$, si ha $E(X|\mathcal{G}) = a \mathbf{1}_A + b \mathbf{1}_{A^c}$ q.c., con $a = E(X|A) = \frac{1}{P(A)} \int_A X \, dP$ e analogamente $b = E(X|A^c) = \frac{1}{P(A^c)} \int_{A^c} X \, dP$.

4.1.2. PROPRIETÀ. Elenchiamo ora alcune proprietà della speranza condizionale. In tutte le relazioni che seguono, $X, Y, \{X_n\}_{n \in \mathbb{N}}$ sono variabili aleatorie reali *integrabili* definite su (Ω, \mathcal{F}, P) , \mathcal{G}, \mathcal{H} sono sotto- σ -algebre di \mathcal{F} e α, β sono numeri reali.

Cominciamo con alcune proprietà basilari:

- (*Linearità*) $E(\alpha X + \beta Y|\mathcal{G}) = \alpha E(X|\mathcal{G}) + \beta E(Y|\mathcal{G})$ q.c..
- (*Positività*) Se $X \geq 0$ q.c. allora $E(X|\mathcal{G}) \geq 0$ q.c..
- (*Jensen*) Se $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ è convessa e tale che $\varphi(X)$ sia integrabile, allora $\varphi(E(X|\mathcal{G})) \leq E(\varphi(X)|\mathcal{G})$ q.c..

Elenchiamo quindi tre proprietà squisitamente condizionali, di cui faremo uso frequente.

- (*Raffinamento*) Se $\mathcal{H} \subseteq \mathcal{G}$, allora $E(E(X|\mathcal{G})|\mathcal{H}) = E(X|\mathcal{H})$ q.c.. Segue in particolare la relazione (molto utile) $E(E(X|\mathcal{G})) = E(X)$, qualunque sia la σ -algebra \mathcal{G} .
- (*Misurabilità*) Se X è \mathcal{G} -misurabile e XY è integrabile, allora $E(XY|\mathcal{G}) = X E(Y|\mathcal{G})$ q.c.. In particolare si ritrova $E(X|\mathcal{G}) = X$ q.c. se X è \mathcal{G} -misurabile.
- (*Indipendenza*) Se X è indipendente da \mathcal{G} , $E(X|\mathcal{G}) = E(X)$ q.c..

Enunciamo infine le versioni condizionali dei classici teoremi di convergenza.

- (*Convergenza monotona*) Se $X_n \uparrow X$ q.c. per $n \rightarrow \infty$, allora $E(X_n|\mathcal{G}) \uparrow E(X|\mathcal{G})$ q.c..
- (*Lemma di Fatou*) Se $X_n \geq 0$ q.c. per ogni $n \in \mathbb{N}$ (o più in generale se $X_n \geq Y$ q.c., con $|Y|$ integrabile), allora $E(\liminf_{n \rightarrow \infty} X_n|\mathcal{G}) \leq \liminf_{n \rightarrow \infty} E(X_n|\mathcal{G})$ q.c..
- (*Convergenza dominata*) Se $|X_n| \leq Y$ q.c. per ogni $n \in \mathbb{N}$, con Y integrabile, e se $X_n \rightarrow X$ q.c. per $n \rightarrow \infty$, allora $E(X_n|\mathcal{G}) \rightarrow E(X|\mathcal{G})$ q.c..

4.1.3. ESEMPI E APPLICAZIONI. Dalla disuguaglianza di Jensen condizionale applicata alla funzione convessa $\varphi(x) = |x|^p$ (per $p \geq 1$), segue che $|E(X|\mathcal{G})|^p \leq E(|X|^p|\mathcal{G})$ q.c.. Dato che $E(E(|X|^p|\mathcal{G})) = E(|X|^p)$ per la proprietà di raffinamento, si ha

$$\|E(X|\mathcal{G})\|_p \leq \|X\|_p.$$

Questo mostra che l'applicazione che a una variabile aleatoria reale $X \in L^p(\Omega, \mathcal{F}, P)$ associa la sua speranza condizionale $E(X|\mathcal{G})$ è un operatore (lineare) *continuo*, anzi *1-lipschitziano*, da L^p in sé. In particolare, se $X_n \rightarrow X$ in L^p per $n \rightarrow \infty$ allora anche $E(X_n|\mathcal{G}) \rightarrow E(X|\mathcal{G})$ in L^p .

Vediamo ora qualche esempio di calcolo di speranze condizionali relative al moto browniano. Sia $\{\mathcal{F}_t\}_{t \geq 0}$ una filtrazione, definita sullo spazio di probabilità (Ω, \mathcal{F}, P) , e sia $B = \{B_t\}_{t \geq 0}$ un $\{\mathcal{F}_t\}_{t \geq 0}$ -moto browniano reale.

ESEMPIO 4.1. Per $s \leq t$ si ha $E(B_t|\mathcal{F}_s) = B_s$ q.c.. Infatti

$$E(B_t|\mathcal{F}_s) = E((B_t - B_s) + B_s|\mathcal{F}_s) = E(B_t - B_s|\mathcal{F}_s) + E(B_s|\mathcal{F}_s) = B_s \text{ q.c.},$$

poiché $B_t - B_s$ è indipendente da \mathcal{F}_s , mentre B_s è \mathcal{F}_s -misurabile. \square

ESEMPIO 4.2. Per $s \leq t$ si ha $E(B_t^2|\mathcal{F}_s) = B_s^2 + (t - s)$ q.c.. Infatti, scrivendo $B_t^2 = ((B_t - B_s) + B_s)^2$ e applicando le proprietà della speranza condizionale si ha

$$\begin{aligned} E(B_t^2|\mathcal{F}_s) &= E((B_t - B_s)^2|\mathcal{F}_s) + E(B_s^2|\mathcal{F}_s) + 2 E((B_t - B_s)B_s|\mathcal{F}_s) \\ &= E((B_t - B_s)^2) + B_s^2 + B_s E(B_t - B_s|\mathcal{F}_s) = (t - s) + B_s^2 \text{ q.c.}, \end{aligned}$$

dove è stato usato il fatto che $B_t - B_s \sim \mathcal{N}(0, t - s)$ è indipendente da \mathcal{F}_s . \square

ESEMPIO 4.3. Per $s \leq t$ e $\lambda \in \mathbb{R}$ (o anche $\lambda \in \mathbb{C}$) si ha $E(e^{\lambda B_t}|\mathcal{F}_s) = e^{\lambda B_s + \lambda^2(t-s)/2}$ q.c.. Infatti, scrivendo $e^{\lambda B_t} = e^{\lambda B_s} e^{\lambda(B_t - B_s)}$ si ha che

$$E(e^{\lambda B_t}|\mathcal{F}_s) = e^{\lambda B_s} E(e^{\lambda(B_t - B_s)}|\mathcal{F}_s) = e^{\lambda B_s} E(e^{\lambda(B_t - B_s)}) = e^{\lambda B_s} e^{\lambda^2(t-s)/2},$$

dove si è usato il fatto che $E(e^{\lambda Z}) = e^{\lambda^2 \sigma^2/2}$ se $Z \sim \mathcal{N}(0, \sigma^2)$. \square

4.2. MARTINGALE A TEMPO DISCRETO E CONTINUO

Per tutto il paragrafo supporremo che sia fissato uno spazio di probabilità filtrato standard $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \mathbb{T}}, \mathbb{P})$. Esempi tipici di insiemi degli indici \mathbb{T} sono \mathbb{N}_0 (o un suo sottoinsieme finito), la semiretta positiva $[0, \infty)$ oppure un intervallo $[a, b] \subseteq [0, \infty)$.

DEFINIZIONE 4.4. Un processo reale adattato $M = \{M_t\}_{t \in \mathbb{T}}$ è detto *submartingala* (risp. *supermartingala*, *martingala*) se M_t è integrabile per ogni $t \in \mathbb{T}$ e vale la seguente relazione:

$$\text{q.c.} \quad \mathbb{E}(M_t | \mathcal{F}_s) \geq M_s \quad (\text{risp.} \leq M_s, = M_s), \quad (4.2)$$

per ogni $s, t \in \mathbb{T}$ con $s \leq t$.

La condizione (4.2) può essere riespressa come $\mathbb{E}(M_t - M_s | \mathcal{F}_s) \geq 0$ (risp. ≤ 0 , $= 0$) q.c.. Si noti che M è una submartingala se e soltanto se $-M$ è una supermartingala. Analogamente, un processo è una martingala se e soltanto se è allo stesso tempo una submartingala e una supermartingala.

Si definiscono (sub,super)martingale $M = \{M_t\}_{t \geq 0}$ anche quando sullo spazio non è definita una filtrazione: in questo caso si richiede che la relazione (4.2) valga rispetto alla filtrazione naturale $\{\mathcal{F}_t^M := \sigma(\{M_u\}_{u \in [0, t] \cap \mathbb{T}})\}_{t \in \mathbb{T}}$. Quando vorremo enfatizzare la filtrazione, scriveremo che M è una $\{\mathcal{F}_t\}_{t \in \mathbb{T}}$ -(sub,super)submartingala.

Osserviamo che, se M è una submartingala, segue da (4.2) che $\mathbb{E}(M_t) \geq \mathbb{E}(M_s)$ per $t \geq s$, cioè M è *crescente* in media. Analogamente, una supermartingala è decrescente in media, mentre una martingala è costante in media.

OSSERVAZIONE 4.5. Per dimostrare che $\mathbb{E}(M_t | \mathcal{F}_s) \geq M_s$ q.c. è sufficiente mostrare che $\mathbb{E}(M_t 1_A) \geq \mathbb{E}(M_s 1_A)$ per ogni $A \in \mathcal{F}_s$. Infatti, ponendo $Z := \mathbb{E}(M_t | \mathcal{F}_s)$ per semplicità, da questa relazione segue che $\mathbb{E}(M_t 1_A) = \mathbb{E}(Z 1_A) \geq \mathbb{E}(M_s 1_A)$, dunque $\mathbb{E}((Z - M_s) 1_A) \geq 0$, per ogni $A \in \mathcal{F}_s$. Resta solo da mostrare che ciò implica che $Z - M_s \geq 0$ q.c., e avremo ottenuto la relazione desiderata $\mathbb{E}(M_t | \mathcal{F}_s) \geq M_s$ q.c..

Questo segue da un fatto generale: se Y è una variabile aleatoria integrabile e \mathcal{G} -misurabile tale che $\mathbb{E}(Y 1_A) \geq 0$ per ogni $A \in \mathcal{G}$, si ha $Y \geq 0$ q.c.. Infatti scegliendo $A = \{Y < 0\}$ si ottiene $\mathbb{E}(Y 1_{\{Y < 0\}}) \geq 0$; d'altro canto si ha chiaramente $Y 1_{\{Y < 0\}} \leq 0$, quindi $\mathbb{E}(Y 1_{\{Y < 0\}}) = 0$. Essendo $Y 1_{\{Y < 0\}} \leq 0$, ciò è possibile se e soltanto se $Y 1_{\{Y < 0\}} = 0$ q.c., che equivale a $\mathbb{P}(Y < 0) = 0$. Questo mostra che $Y \geq 0$ q.c.. \square

LEMMA 4.6. Se $M = \{M_t\}_{t \in \mathbb{T}}$ è una martingala e $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ è una funzione convessa tale che $\varphi(M_t) \in L^1$ per ogni $t \in \mathbb{T}$, il processo $\{\varphi(M_t)\}_{t \in \mathbb{T}}$ è una submartingala.

Se $M = \{M_t\}_{t \in \mathbb{T}}$ è una submartingala e φ è convessa e *crescente* tale che $\varphi(M_t) \in L^1$ per ogni $t \in \mathbb{T}$, $\{\varphi(M_t)\}_{t \in \mathbb{T}}$ è una submartingala.

DIMOSTRAZIONE. Se φ è convessa, dalla disuguaglianza di Jensen per la speranza condizionale segue che $\mathbb{E}(\varphi(M_t) | \mathcal{F}_s) \geq \varphi(\mathbb{E}(M_t | \mathcal{F}_s))$ q.c..

Se M è una martingala, si ha $E(M_t|\mathcal{F}_s) = M_s$ q.c. e dunque $E(\varphi(M_t)|\mathcal{F}_s) \geq \varphi(M_s)$ q.c., cioè la tesi. Se M è una submartingala, si ha $E(M_t|\mathcal{F}_s) \geq M_s$ q.c.; quindi, se φ è crescente, $\varphi(E(M_t|\mathcal{F}_s)) \geq \varphi(M_s)$ q.c., da cui $E(\varphi(M_t)|\mathcal{F}_s) \geq \varphi(M_s)$ q.c.. \square

In particolare, data una martingala $M = \{M_t\}_{t \in \mathbb{T}}$, i processi $\{|M_t|\}_{t \in \mathbb{T}}$ e $\{M_t^2\}_{t \in \mathbb{T}}$ sono submartingale. Attenzione che ciò non è necessariamente vero quando M è una submartingala, perché le funzioni $x \mapsto |x|$ e $x \mapsto x^2$ sono convesse ma non crescenti.

ESEMPIO 4.7. Se $\{\mathcal{F}_t\}_{t \in \mathbb{T}}$ è una filtrazione su uno spazio (Ω, \mathcal{F}, P) e $X : \Omega \rightarrow \mathbb{R}$ è una variabile aleatoria integrabile, il processo $Y = \{Y_t := E(X|\mathcal{F}_t)\}_{t \in \mathbb{T}}$ è una martingala. Infatti per $s \leq t$ si ha $\mathcal{F}_s \subseteq \mathcal{F}_t$ e dunque $E(Y_t|\mathcal{F}_s) = E(E(X|\mathcal{F}_t)|\mathcal{F}_s) = E(X|\mathcal{F}_s) = Y_s$ q.c., per la proprietà di raffinamento.

Osserviamo che, se l'insieme degli indici \mathbb{T} ha un elemento massimo, come $\mathbb{T} = \{0, \dots, T\}$ oppure $\mathbb{T} = [0, T]$, ogni martingala $M = \{M_t\}_{t \in \mathbb{T}}$ è di questa forma: infatti per la proprietà (4.2) si ha $M_s = E(M_T|\mathcal{F}_s)$ q.c. per ogni $s \in \mathbb{T}$, dove $T := \max(\mathbb{T})$. \square

ESEMPIO 4.8. Se $B = \{B_t\}_{t \geq 0}$ è un $\{\mathcal{F}_t\}_{t \geq 0}$ -moto browniano reale, i seguenti processi sono martingale:

$$\{B_t\}_{t \geq 0}, \quad \{B_t^2 - t\}_{t \geq 0}, \quad \{e^{\lambda B_t - \lambda^2 t/2}\}_{t \geq 0}, \quad \forall \lambda \in \mathbb{R},$$

come mostrano gli Esempi 4.1, 4.2 e 4.3 dello scorso paragrafo. In particolare, il moto browniano $B = \{B_t\}_{t \geq 0}$ è una martingala. \square

La teoria delle martingale è tra i capitoli più ricchi ed eleganti del calcolo delle probabilità. La nostra esposizione sarà estremamente concisa: ci limiteremo a considerare i risultati di diretto interesse per il corso, concentrandoci sui tempi d'arresto e sostanzialmente ignorando i teoremi di convergenza.

4.2.1. TEMPO DISCRETO. Consideriamo innanzitutto il caso in cui l'insieme dei tempi è discreto, $\mathbb{T} = \mathbb{N}_0$, e indichiamo con $\{\mathcal{F}_n\}_{n \in \mathbb{N}_0}$ la filtrazione. Ricordiamo che una variabile aleatoria $\tau : \Omega \rightarrow \mathbb{N}_0 \cup \{+\infty\}$ è un tempo d'arresto per $\{\mathcal{F}_n\}_{n \in \mathbb{N}_0}$ se e soltanto se $\{\tau = n\} \in \mathcal{F}_n$, per ogni $n \in \mathbb{N}_0$. Analogamente, \mathcal{F}_τ è la σ -algebra composta dagli eventi $A \in \mathcal{F}$ per cui $A \cap \{\tau = n\} \in \mathcal{F}_n$, per ogni $n \in \mathbb{N}_0$.

Osserviamo che la relazione (4.2) che definisce una submartingala può essere semplificata per processi $M = \{M_n\}_{n \in \mathbb{N}_0}$ a tempo discreto: basta richiedere che $E(M_{n+1}|\mathcal{F}_n) \geq M_n$ q.c. per ogni $n \in \mathbb{N}_0$. Infatti da questa relazione segue che

$$E(M_{n+2}|\mathcal{F}_n) = E(E(M_{n+2}|\mathcal{F}_{n+1})|\mathcal{F}_n) \geq E(M_{n+1}|\mathcal{F}_n) \geq M_n \text{ q.c.},$$

e per induzione si mostra facilmente che $E(M_k|\mathcal{F}_n) \geq M_n$ q.c. per ogni $k \geq n$. Un discorso del tutto analogo si applica alle supermartingale o martingale.

Il primo risultato che dimostriamo è che le (sub)martingale possono essere stoppate a un tempo d'arresto.

LEMMA 4.9. Se $M = \{M_n\}_{n \in \mathbb{N}_0}$ è una submartingala e τ è un tempo d'arresto, il processo arrestato $M^\tau = \{M_n^\tau\}_{n \in \mathbb{N}_0}$ definito da $M_n^\tau := M_{\tau \wedge n}$ è una submartingala. Analogamente, se M è una martingala, anche il processo M^τ è una martingala.

DIMOSTRAZIONE. Ricordiamo che $a \wedge b := \min\{a, b\}$ per $a, b \in \mathbb{R}$, cosicché $M_n^\tau(\omega) = M_{\min\{\tau(\omega), n\}}(\omega)$. Da questo segue facilmente che per ogni $n \in \mathbb{N}_0$ si ha

$$M_n^\tau = M_{\tau \wedge n} = \sum_{k=0}^n M_k \mathbf{1}_{\{\tau=k\}} + M_n \mathbf{1}_{\{\tau>n\}}. \quad (4.3)$$

Questa relazione mostra che, per ogni $n \in \mathbb{N}_0$, M_n^τ è integrabile, in quanto somma finita di variabili aleatorie integrabili, e anche che M_n^τ è \mathcal{F}_n -misurabile (si noti che $\{\tau > n\} = \{\tau \leq n\}^c \in \mathcal{F}_n$), cioè il processo M^τ è adattato.

Resta da verificare che $E(M_{n+1}^\tau | \mathcal{F}_n) \geq M_n^\tau$ q.c.. Sull'evento $\{\tau \leq n\}$ si ha $\tau \wedge n = \tau \wedge (n+1) = \tau$ e quindi $M_{n+1}^\tau = M_n^\tau$. In altri termini $(M_{n+1}^\tau - M_n^\tau) \mathbf{1}_{\{\tau \leq n\}} = 0$, per cui

$$E(M_{n+1}^\tau - M_n^\tau | \mathcal{F}_n) = E((M_{n+1}^\tau - M_n^\tau) \mathbf{1}_{\{\tau > n\}} | \mathcal{F}_n) \quad \text{q.c..}$$

D'altro canto, sull'evento $\{\tau > n\} = \{\tau \geq n+1\}$ si ha $M_n^\tau = M_n$ e $M_{n+1}^\tau = M_{n+1}$. Visto che $\{\tau > n\} \in \mathcal{F}_n$, dalle proprietà della speranza condizionale si ottiene

$$E((M_{n+1}^\tau - M_n^\tau) \mathbf{1}_{\{\tau > n\}} | \mathcal{F}_n) = \mathbf{1}_{\{\tau > n\}} E(M_{n+1} - M_n | \mathcal{F}_n) \geq 0 \quad \text{q.c.},$$

perché M è una submartingala. Abbiamo quindi mostrato che $E(M_{n+1}^\tau | \mathcal{F}_n) \geq M_n^\tau$ q.c., cioè M^τ è una submartingala. Il caso di una martingala è analogo. \square

COROLLARIO 4.10 (TEOREMA D'ARRESTO). Sia $M = \{M_n\}_{n \in \mathbb{N}_0}$ una submartingala e sia τ un tempo d'arresto, tali che una delle seguenti condizioni è verificata:

- τ è q.c. limitato, cioè esiste $N \in \mathbb{N}_0$ tale che $\tau \leq N$ q.c.;

oppure

- τ è q.c. finito e inoltre $|M_{\tau \wedge n}| \leq Y$ per ogni $n \in \mathbb{N}_0$, dove Y è una variabile aleatoria integrabile (in particolare $|M_{\tau \wedge n}| \leq K$ per una costante $K \in [0, \infty)$).

Allora la variabile M_τ è integrabile e vale la relazione

$$E(M_\tau) \geq E(M_0). \quad (4.4)$$

Se $M = \{M_n\}_{n \in \mathbb{N}_0}$ è una martingala, nelle stesse ipotesi si ha l'uguaglianza in (4.4).

DIMOSTRAZIONE. Il processo $\{M_{\tau \wedge n}\}_{n \in \mathbb{N}_0}$ è una submartingala per il Lemma 4.9, quindi

$$E(M_{\tau \wedge n}) \geq E(M_0), \quad \forall n \in \mathbb{N}_0. \quad (4.5)$$

Se $\tau \leq N$ q.c., si ha $\tau \wedge N = \tau$ q.c. e ponendo $n = N$ in (4.5) otteniamo la relazione (4.4).

Se τ è q.c. finito, per q.o. $\omega \in \Omega$ si ha $\tau(\omega) \wedge n = \tau(\omega) < \infty$ per n grande, quindi $\lim_{n \rightarrow \infty} M_{\tau(\omega) \wedge n}(\omega) = M_{\tau(\omega)}$. In altri termini, q.c. $\lim_{n \rightarrow \infty} M_{\tau \wedge n} = M_\tau$. Se $|M_{\tau \wedge n}| \leq Y$ con $Y \in L^1$, per il teorema di convergenza dominata si può passare al limite in (4.5), ottenendo $E(M_0) \leq \lim_{n \rightarrow \infty} E(M_{\tau \wedge n}) = E(M_\tau)$, cioè (4.4).

Infine, se $M = \{M_n\}_{n \in \mathbb{N}_0}$ è una martingala, entrambi i processi M e $-M$ sono submartingale. Per quanto già visto si ha dunque $E(M_0) \leq E(M_\tau)$ e $E(-M_0) \leq E(-M_\tau)$, da cui $E(M_0) = E(M_\tau)$. \square

$M = \{M_n\}_{n \in \mathbb{N}_0}$ Concludiamo con una disuguaglianza di fondamentale importanza.

TEOREMA 4.11 (DISUGUAGLIANZA MASSIMALE). Se $S = \{S_n\}_{n \in \mathbb{N}_0}$ è una submartingala, per ogni $n \in \mathbb{N}_0$ e $\lambda > 0$ vale che

$$P\left(\max_{0 \leq i \leq n} S_i \geq \lambda\right) \leq \frac{E(S_n^+)}{\lambda}, \quad P\left(\min_{0 \leq i \leq n} S_i \leq -\lambda\right) \leq \frac{E(S_n^+) - E(S_0)}{\lambda}.$$

DIMOSTRAZIONE. Fissiamo $n \in \mathbb{N}_0$ e definiamo la variabile aleatoria τ ponendo

$$\tau(\omega) := \begin{cases} \inf\{k \leq n : S_k(\omega) \geq \lambda\} & \text{se } \max_{0 \leq i \leq n} S_i(\omega) \geq \lambda \\ +\infty & \text{altrimenti} \end{cases}.$$

Si verifica facilmente che τ è un tempo d'arresto. Inoltre

$$P\left(\max_{0 \leq i \leq n} S_i \geq \lambda\right) = P(\tau \leq n) = \sum_{k=0}^n P(\tau = k) \leq \frac{1}{\lambda} \sum_{k=0}^n E(S_k \mathbf{1}_{\{\tau=k\}}),$$

dove abbiamo usato il fatto che $S_k \geq \lambda$ sull'evento $\{\tau = k\}$. Dato che $\{\tau = k\} \in \mathcal{F}_k$ e S è una submartingala, si ha $E(S_k \mathbf{1}_{\{\tau=k\}}) \leq E(S_n \mathbf{1}_{\{\tau=k\}})$ e quindi

$$P\left(\max_{0 \leq i \leq n} S_i \geq \lambda\right) \leq \frac{1}{\lambda} \sum_{k=0}^n E(S_n \mathbf{1}_{\{\tau=k\}}) = \frac{1}{\lambda} E(S_n \mathbf{1}_{\{\tau \leq n\}}) \leq \frac{1}{\lambda} E(S_n^+),$$

perché $S_n \mathbf{1}_{\{\tau \leq n\}} \leq S_n^+$. La prima disuguaglianza è dimostrata.

Per la seconda disuguaglianza, fissiamo sempre $n \in \mathbb{N}_0$ e ridefiniamo

$$\tau(\omega) := \begin{cases} \inf\{k \leq n : S_k(\omega) \leq -\lambda\} & \text{se } \min_{0 \leq i \leq n} S_i(\omega) \leq -\lambda \\ +\infty & \text{altrimenti} \end{cases}.$$

Possiamo allora scrivere

$$\begin{aligned} P\left(\min_{0 \leq i \leq n} S_i \leq -\lambda\right) &= P(\tau \leq n) \leq \frac{1}{\lambda} E(-S_\tau \mathbf{1}_{\{\tau \leq n\}}) = -\frac{1}{\lambda} E(S_{\tau \wedge n} \mathbf{1}_{\{\tau \leq n\}}) \\ &= -\frac{1}{\lambda} \left(E(S_{\tau \wedge n}) - E(S_n \mathbf{1}_{\{\tau > n\}}) \right), \end{aligned}$$

avendo usato il fatto che per definizione sull'evento $\{\tau \leq n\}$ si ha $-S_\tau \geq \lambda$ e $\tau = \tau \wedge n$. Per ottenere l'ultima uguaglianza basta scrivere $1_{\{\tau \leq n\}} = 1 - 1_{\{\tau > n\}}$ e notare che $\tau \wedge n = n$ sull'evento $\{\tau > n\}$. Analogamente a sopra, abbiamo che $E(S_n 1_{\{\tau > n\}}) \leq E(S_n^+)$, mentre $E(S_{\tau \wedge n}) \geq E(S_0)$ grazie al Corollario 4.10, perché $\tau \wedge n$ è un tempo d'arresto limitato. La seconda relazione è dunque dimostrata. \square

4.2.2. TEMPO CONTINUO. La teoria generale delle martingale a tempo continuo, in cui cioè l'insieme degli indici è più che numerabile (noi considereremo $\mathbb{T} = [0, \infty)$) è decisamente più complessa. Dal momento che è importante poter lavorare con processi continui a destra, è naturale chiedersi sotto quali condizioni una (sub,super)martingala $M = \{M_t\}_{t \geq 0}$ ammette una modificazione con tale proprietà. È possibile mostrare che, se la filtrazione $\{\mathcal{F}_t\}_{t \geq 0}$ soddisfa le ipotesi standard (come noi supponiamo), una *submartingala* M ammette una modificazione continua a destra se e soltanto se la funzione $t \mapsto E(M_t)$ è continua a destra (si veda ad esempio il Teorema 3.13 nel capitolo 1 in [Karatzas e Shreve, 1998]). In particolare, *una martingala ammette sempre una modificazione continua a destra*, perché $E(M_t) \equiv E(M_0)$ è costante. Nei fatti, avremo a che fare quasi sempre con (sub,super)martingale continue.

Elenchiamo ora le versioni a tempo continuo dei risultati dimostrati nel paragrafo precedente. Dimosteremo solo la disuguaglianza massimale (per maggiori dettagli, si veda il paragrafo 1.3 nel capitolo 1 in [Karatzas e Shreve, 1998]).

LEMMA 4.12. Se $M = \{M_t\}_{t \geq 0}$ è una submartingala continua a destra e τ è un tempo d'arresto, il *processo arrestato* $M^\tau = \{M_t^\tau\}_{t \geq 0}$ definito da $M_t^\tau := M_{\tau \wedge t}$ è una submartingala continua a destra.

Analogamente, se M è una martingala continua a destra, anche il processo M^τ è una martingala continua a destra.

COROLLARIO 4.13 (TEOREMA D'ARRESTO). Sia $M = \{M_t\}_{t \geq 0}$ una submartingala continua a destra e sia τ un tempo d'arresto, tali che una delle seguenti ipotesi è soddisfatta:

- τ è q.c. limitato, cioè esiste $T \in (0, \infty)$ tale che $\tau \leq T$ q.c.;

oppure

- τ è q.c. finito e $|M_{\tau \wedge t}| \leq Y$ per ogni $t \geq 0$, dove Y è una variabile aleatoria integrabile (in particolare $|M_{\tau \wedge n}| \leq K$ per una costante $K \in [0, \infty)$).

Allora la variabile M_τ è integrabile e vale la relazione

$$E(M_\tau) \geq E(M_0). \quad (4.6)$$

Se $M = \{M_t\}_{t \geq 0}$ è una martingala, nelle stesse ipotesi si ha l'uguaglianza in (4.6).

TEOREMA 4.14 (DISUGUAGLIANZA MASSIMALE). Per una submartingala $S = \{S_t\}_{t \geq 0}$ continua a destra vale che, per ogni $t \geq 0$ e $\lambda > 0$,

$$P \left(\sup_{u \in [0, t]} S_u \geq \lambda \right) \leq \frac{E(S_t^+)}{\lambda}, \quad P \left(\inf_{u \in [0, t]} S_u \leq -\lambda \right) \leq \frac{E(S_t^+) - E(S_0)}{\lambda}.$$

DIMOSTRAZIONE. Dimostriamo la prima relazione (la seconda è del tutto analoga). Fissiamo $0 =: t_0 < t_1 < \dots < t_k := t$ e consideriamo il processo a tempo discreto $\{S_{t_i}\}_{0 \leq i \leq k}$, che è una submartingala rispetto alla filtrazione $\{\mathcal{F}_{t_i}\}_{0 \leq i \leq k}$. La disuguaglianza massimale a tempo discreto (Teorema 4.11) dà, per ogni $\varepsilon \in (0, \lambda)$,

$$P \left(\max_{u \in \{t_0, t_1, \dots, t_k\}} S_u \geq \lambda - \varepsilon \right) \leq \frac{E(S_t^+)}{\lambda - \varepsilon}.$$

Fissiamo ora una successione crescente di partizioni $\pi^{(n)} := \{t_1^{(n)}, \dots, t_{k_n}^{(n)}\}$ per cui si abbia $\bigcup_{n \in \mathbb{N}} \pi^{(n)} = [0, t] \cap \mathbb{Q}$. Sfruttando l'inclusione naturale di eventi e la continuità dal basso della probabilità, si ottiene allora

$$\begin{aligned} P \left(\sup_{u \in [0, t] \cap \mathbb{Q}} S_u \geq \lambda \right) &\leq P \left(\sup_{u \in [0, t] \cap \mathbb{Q}} S_u > \lambda - \varepsilon \right) \\ &= \lim_{n \rightarrow \infty} P \left(\max_{u \in \{t_0^{(n)}, \dots, t_{k_n}^{(n)}\}} S_u > \lambda - \varepsilon \right) \leq \frac{E(S_t^+)}{\lambda - \varepsilon}. \end{aligned}$$

Ma $\sup_{u \in [0, t] \cap \mathbb{Q}} S_u = \sup_{u \in [0, t]} S_u$, per la continuità a destra di S , per cui

$$P \left(\sup_{u \in [0, t]} S_u \geq \lambda \right) \leq \frac{E(S_t^+)}{\lambda - \varepsilon} \quad \forall \lambda > 0, \quad \forall \varepsilon \in (0, \lambda).$$

Prendendo il limite $\varepsilon \downarrow 0$, si conclude la dimostrazione. \square

OSSERVAZIONE 4.15. Nel caso di una *supermartingala* $S = \{S_t\}_{t \geq 0}$ continua a destra, le relazioni del Teorema 4.14 si riformulano nel modo seguente:

$$P \left(\inf_{u \in [0, t]} S_u \leq -\lambda \right) \leq \frac{E(S_t^-)}{\lambda}, \quad P \left(\sup_{u \in [0, t]} S_u \geq \lambda \right) \leq \frac{E(S_t^-) + E(S_0)}{\lambda},$$

per ogni $t \geq 0$ e $\lambda > 0$. \square

ESEMPIO 4.16. Sia $B = \{B_t\}_{t \geq 0}$ un moto browniano reale e definiamo per $a, b > 0$ la variabile $\tau_{-a, b} := \inf\{s \geq 0 : B_s \notin (-a, b)\}$, che dà il primo istante in cui B esce dall'intervallo $(-a, b)$. Già sappiamo che $\tau_{-a, b}$ è un tempo d'arresto q.c. finito, come conseguenza del principio di riflessione (infatti $\tau_{-a, b} = \min\{\tau_{-a}, \tau_b\}$). Mostriamo ora che la legge della variabile $B_{\tau_{-a, b}}$ è data da

$$P(B_{\tau_{-a, b}} = -a) = \frac{b}{a+b}, \quad P(B_{\tau_{-a, b}} = b) = \frac{a}{a+b}. \quad (4.7)$$

Sappiamo che B è una martingala continua. Dato che $|B_{\tau_{-a,b} \wedge t}| \leq \max\{a, b\}$ per ogni $t \geq 0$, il Corollario 4.13 dà $E(B_{\tau_{-a,b}}) = E(B_0) = 0$. Per la continuità delle traiettorie di B , la variabile $B_{\tau_{-a,b}}$ può assumere solo i due valori $-a, b$, per cui

$$0 = E(B_{\tau_{-a,b}}) = -a P(B_{\tau_{-a,b}} = -a) + b P(B_{\tau_{-a,b}} = b).$$

Dato che $P(B_{\tau_{-a,b}} = -a) + P(B_{\tau_{-a,b}} = b) = 1$, si ottengono le relazioni in (4.7). \square

4.3. VARIAZIONE QUADRATICA DI UNA MARTINGALA CONTINUA

In questo paragrafo esaminiamo più da vicino le martingale $M = \{M_t\}_{t \geq 0}$ continue di quadrato integrabile, per cui $E(M_t^2) < \infty$ (cioè $M_t \in L^2$) per ogni $t \geq 0$.

Abbiamo visto che il moto browniano ha traiettorie continue ma piuttosto irregolari, in quanto di variazione infinita. Questo fenomeno non è una strana peculiarità del moto browniano, ma è una conseguenza diretta del fatto che il moto browniano è una martingala continua. In effetti, per ogni martingala $M = \{M_t\}_{t \geq 0}$ di quadrato integrabile, vale la seguente relazione:

$$E(M_t^2 - M_s^2 | \mathcal{F}_s) = E((M_t - M_s)^2 | \mathcal{F}_s), \quad (4.8)$$

come si verifica facilmente osservando che $E(M_t M_s | \mathcal{F}_s) = M_s E(M_t | \mathcal{F}_s) = M_s^2$ q.c., per le proprietà della speranza condizionale e per la definizione di martingala. Prendendo il valore atteso di entrambi i membri in (4.8), segue facilmente che

$$E(M_t^2) - E(M_0^2) = E\left(\sum_{i=1}^n (M_{t_i} - M_{t_{i-1}})^2\right), \quad (4.9)$$

per ogni partizione $\pi = \{0 =: t_0 < t_1 < \dots < t_n := t\}$ dell'intervallo $[0, t]$. Notiamo ora che, per il Lemma 4.6, il processo $\{M_t^2\}_{t \geq 0}$ è una submartingala e quindi $E(M_t^2) - E(M_0^2) \geq 0$.

- Il caso in cui $E(M_t^2) = E(M_0^2)$ è poco interessante: dalla relazione (4.9) segue infatti che, per ogni partizione $\pi = \{0 =: t_0 < t_1 < \dots < t_n := t\}$, si deve avere q.c. $M_{t_0} = M_{t_1} = \dots = M_{t_n}$. Se M ha traiettorie continue, ciò implica che $M_s = M_0$ per ogni $s \in [0, t]$, ossia le traiettorie di M sono costanti nell'intervallo $[0, t]$.
- Viceversa, se $E(M_t^2) > E(M_0^2)$, la relazione (4.9) suggerisce che M possa avere variazione quadratica positiva, o più precisamente che $\sum_{i=1}^n (M_{t_i} - M_{t_{i-1}})^2$ non tenda a zero q.c. per $n \rightarrow \infty$ (si noti che il membro di sinistra in (4.9) è positivo e non dipende dalla partizione π). Se ciò avviene, è facile mostrare che con probabilità positiva le traiettorie di M hanno variazione *infinita*, come per il moto browniano.

Veniamo ora a una trattazione più precisa. Partiamo da una considerazione che a prima vista sembra avere poco a che vedere con la variazione quadratica. Dato un moto browniano reale $\{B_t\}_{t \geq 0}$, sappiamo che il processo $\{B_t^2 - t\}_{t \geq 0}$ è una martingala. Possiamo riformulare questo fatto dicendo che si può *compensare* la submartingala $\{B_t^2\}_{t \geq 0}$ rendendola una martingala, sottraendo un opportuno processo crescente (in questo caso pari a t).

È naturale chiedersi se un risultato analogo valga più in generale: data una martingala continua di quadrato integrabile $M = \{M_t\}_{t \geq 0}$, è possibile *compensare* la submartingala

$M^2 = \{M_t^2\}_{t \geq 0}$, ovvero trovare un processo continuo e crescente $A = \{A_t\}_{t \geq 0}$ tale che $M^2 - A = \{M_t^2 - A_t\}_{t \geq 0}$ sia una *martingala*? La risposta è affermativa: non solo un tale processo A esiste, ma coincide con la variazione quadratica di M , analogamente al caso del moto browniano. Più precisamente, vale il seguente risultato.

TEOREMA 4.17. Sia $M = \{M_t\}_{t \geq 0}$ una martingala continua di quadrato integrabile rispetto a una filtrazione completa $\{\mathcal{F}_t\}_{t \geq 0}$. Allora esiste un unico processo $A = \{A_t\}_{t \geq 0}$ crescente, continuo, adattato e nullo al tempo zero tale che $M^2 - A$ sia una martingala. Per ogni $t \geq 0$, la variabile A_t è la variazione quadratica di M sull'intervallo $[0, t]$:

$$A_t = \lim_{|\pi| \rightarrow 0} \sum_{i=1}^n (M_{t_i} - M_{t_{i-1}})^2 \quad \text{in probabilità,} \quad (4.10)$$

dove indichiamo con $\pi = \{0 =: t_0 < t_1 < \dots < t_n := t\}$ le partizioni di $[0, t]$.

Il processo $A = \{A_t\}_{t \geq 0}$ è detto *variazione quadratica* o *processo crescente* della martingala M ed è indicato con $A = \langle M \rangle$ (scriveremo dunque $A_t = \langle M \rangle_t$).

OSSERVAZIONE 4.18. La ragione per cui il processo A che *compensa* la submartingala M^2 è dato dalla variazione quadratica di M , definita in (4.10), si può intuire considerando martingale $M = \{M_n\}_{n \in \mathbb{N}}$ a tempo discreto. Ponendo infatti $A_n := \sum_{i=1}^n (M_i - M_{i-1})^2$, è immediato verificare che il processo $L = \{L_n := M_n^2 - A_n\}_{n \in \mathbb{N}}$ è una martingala:

$$\mathbb{E}(L_n - L_{n-1} | \mathcal{F}_{n-1}) = \mathbb{E}(M_n^2 - M_{n-1}^2 | \mathcal{F}_{n-1}) - \mathbb{E}((M_n - M_{n-1})^2 | \mathcal{F}_{n-1}) = 0,$$

avendo usato la relazione (4.8). □

Omettiamo la dimostrazione del Teorema 4.17, che è piuttosto tecnica (si veda il paragrafo 2.3 in [Durrett, 1996]). Per le martingale M di interesse che incontreremo nel seguito del corso, saremo in gradi di costruire esplicitamente (“con le mani”) un processo A crescente, continuo, adattato e nullo al tempo zero tale che $M^2 - A$ sia una martingala. Il Teorema 4.17 assicura che tale processo A è proprio la variazione quadratica $\langle M \rangle$ di M , definita in (4.10), ma non avremo bisogno di questo fatto.

OSSERVAZIONE 4.19. Applicando il Teorema 4.17, non è difficile dimostrare che le traiettorie di una martingala continua di quadrato integrabile $M = \{M_t\}_{t \geq 0}$ su qualunque intervallo sono a variazione infinita, oppure sono costanti. Più precisamente, q.c. sull'evento $\{A_t - A_s > 0\}$ si ha $V_{[s,t]}(u \mapsto M_u) = +\infty$, mentre sull'evento $\{A_t - A_s = 0\}$ si ha che $u \mapsto M_u$ è costante sull'intervallo $[s, t]$. □

Definiamo infine la *covariazione quadratica* di due martingale.

DEFINIZIONE 4.20. Siano $M = \{M_t\}_{t \geq 0}$ e $N = \{N_t\}_{t \geq 0}$ due martingale continue di quadrato integrabile, definite sullo stesso spazio filtrato $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ con filtrazione completa $\{\mathcal{F}_t\}_{t \geq 0}$. Definiamo il processo $\langle M, N \rangle = \{\langle M, N \rangle_t\}_{t \geq 0}$, detto *covariazione quadratica di M e N* , ponendo

$$\langle M, N \rangle_t := \frac{1}{4} (\langle M + N \rangle_t - \langle M - N \rangle_t). \quad (4.11)$$

Si ha allora che $MN - \langle M, N \rangle = \{M_t N_t - \langle M, N \rangle_t\}_{t \geq 0}$ è una martingala. Inoltre $\langle M, N \rangle$ è l'unico processo $A = \{A_t\}_{t \geq 0}$ a *variazione finita, continuo, adattato e nullo al tempo zero* tale che $MN - A$ sia una martingala.

La verifica che $MN - \langle M, N \rangle$ è una martingala è immediata: basta osservare che $MN = \frac{1}{4}((M+N)^2 - (M-N)^2)$ e applicare il Teorema 4.17. Inoltre, usando le relazioni (4.10) e (4.11) è facile mostrare che

$$\langle M, N \rangle_t = \lim_{|\pi| \rightarrow 0} \sum_{i=1}^n (M_{t_i} - M_{t_{i-1}})(N_{t_i} - N_{t_{i-1}}) \quad \text{in probabilità,}$$

dove indichiamo con $\pi = \{0 =: t_0 < t_1 < \dots < t_n := t\}$ le partizioni di $[0, t]$.

Sottolineiamo che il processo $\langle M, N \rangle$ in generale *non è crescente*, ma le sue traiettorie sono a *variazione finita*: infatti è chiaro dalla definizione (4.11) che $\langle M, N \rangle$ è la differenza di due processi crescenti.

5. INTEGRALE STOCASTICO

Per tutto il capitolo fissiamo uno spazio filtrato standard $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$, su cui è definito un $\{\mathcal{F}_t\}_{t \geq 0}$ -moto browniano reale $B = \{B_t\}_{t \geq 0}$ e su cui saranno definiti tutti i processi che introdurremo.

In questo capitolo definiamo l'integrazione rispetto al moto browniano per un'ampia classe di processi. La strategia è di definire l'integrale innanzitutto per una classe di processi "elementari" e di estenderlo poi a processi più generali per continuità. Cominciamo pertanto a richiamare alcuni risultati standard sull'estensione di operatori, formulati in un contesto leggermente più generale del solito.

5.1. PROLUNGAMENTO DI ISOMETRIE

Si dice *spazio pseudometrico* un insieme E munito di una *pseudodistanza* $d(\cdot, \cdot)$, ossia di una funzione $d : E \times E \rightarrow \mathbb{R}$ tale che per ogni $x, y, z \in E$ valgano le seguenti proprietà:

$$d(x, x) = 0, \quad d(x, y) = d(y, x), \quad d(x, y) \leq d(x, z) + d(z, y).$$

Come è ben noto, se si rafforza la prima condizione richiedendo che $d(x, y) = 0$ se e solo se $x = y$, la funzione $d(\cdot, \cdot)$ si dice *distanza* e lo spazio E si dice *spazio metrico*. Ciò che differenzia uno spazio pseudometrico rispetto a uno spazio metrico è che ci possono essere punti $x, y \in E$ distinti (cioè $x \neq y$) tali che $d(x, y) = 0$.

Un caso tipico è dato dagli spazi di *variabili aleatorie* $L^p(\Omega, \mathcal{F}, P)$: date infatti due variabili aleatorie $X, Y \in L^p(\Omega, \mathcal{F}, P)$ tali che $d(X, Y) := \|X - Y\|_p = (E(|X - Y|^p))^{1/p} = 0$, non si ha necessariamente $X = Y$ (ma solo $X = Y$ q.c.). Come abbiamo già ricordato, se si identificano le variabili aleatorie q.c. uguali, il relativo spazio delle *classi di equivalenza* (che, con abuso di notazione, si indica ancora con $L^p(\Omega, \mathcal{F}, P)$) diventa uno spazio metrico.

Data una successione di punti $\{x_n\}_{n \in \mathbb{N}}$ in uno spazio pseudometrico E e un punto $x \in E$, si dice che x_n converge verso x (e si scrive $x_n \rightarrow x$) se si ha $\lim_{n \rightarrow \infty} d(x_n, x) = 0$. A differenza di quanto accade per gli spazi metrici, il limite in generale non è unico: in effetti, se $x_n \rightarrow x$, allora $x_n \rightarrow y$ per ogni $y \in E$ con $d(x, y) = 0$.

Sottolineiamo che la pseudodistanza è una funzione continua, nel senso che se $x_n \rightarrow x$ e $y_n \rightarrow y$ si ha $d(x_n, y_n) \rightarrow d(x, y)$. Infatti, per la disuguaglianza triangolare,

$$d(x_n, y_n) \leq d(x_n, x) + d(x, y) + d(y, y_n), \quad d(x, y) \leq d(x, x_n) + d(x_n, y_n) + d(y_n, y),$$

dunque $|d(x_n, y_n) - d(x, y)| \leq d(x_n, x) + d(y_n, y) \rightarrow 0$.

Un sottoinsieme S di uno spazio pseudometrico E si dice *denso* se per ogni $x \in E$ esiste una successione di punti $x_n \in S$ tali che $x_n \rightarrow x$. Una successione $\{x_n\}_{n \in \mathbb{N}}$ in uno spazio pseudometrico E si dice di Cauchy se $\forall \varepsilon > 0$ esiste $n_0 < \infty$ tale che $d(x_n, x_m) < \varepsilon$ per

ogni $n, m \geq n_0$. È facile vedere che in qualunque spazio pseudometrico ogni successione convergente è di Cauchy. Se vale anche il viceversa, ossia se per ogni successione $\{x_n\}_{n \in \mathbb{N}}$ di Cauchy in E esiste $x \in E$ tale che $x_n \rightarrow x$, lo spazio pseudometrico E si dice *completo*.

Come è noto, $L^p(\Omega, \mathcal{F}, P)$ come spazio di *classi di equivalenza* è uno spazio metrico completo; come spazio di variabili aleatorie, è invece uno spazio pseudometrico completo. In effetti, se $X_n \rightarrow X$ in $L^p(\Omega, \mathcal{F}, P)$, allora $X_n \rightarrow X'$ per ogni altra variabile aleatoria $X' \in L^p(\Omega, \mathcal{F}, P)$ tale che $X = X'$ q.c..

Possiamo finalmente enunciare e dimostrare il risultato principale sull'estensione di isometrie densamente definite.

TEOREMA 5.1. Siano E uno spazio pseudometrico, F uno spazio pseudometrico completo, S un sottoinsieme denso di E e $J : S \rightarrow F$ un'isometria:

$$d(J(x), J(y)) = d(x, y), \quad \forall x, y \in S. \quad (5.1)$$

Allora esiste un'isometria $\bar{J} : E \rightarrow F$ che estende J a tutto E , ossia:

$$\bar{J}(x) = J(x), \quad \forall x \in S, \quad d(\bar{J}(x), \bar{J}(y)) = d(x, y), \quad \forall x, y \in E. \quad (5.2)$$

Se $\hat{J} : E \rightarrow F$ è un'altra isometria che estende J , si ha $d(\hat{J}(x), \bar{J}(x)) = 0$ per ogni $x \in E$. In particolare, F è uno spazio *metrico* (e non solo pseudometrico), l'estensione isometrica $\bar{J} : E \rightarrow F$ di J è unica.

DIMOSTRAZIONE. Cominciamo a verificare l'ultima affermazione. Sia $x \in E$ e consideriamo una successione $\{x_n\}_{n \in \mathbb{N}}$ in S convergente a x , cioè $d(x_n, x) \rightarrow 0$ (una tale successione esiste perché per ipotesi S è denso in E). Per la disuguaglianza triangolare

$$d(\hat{J}(x), \bar{J}(x)) \leq d(\hat{J}(x), \hat{J}(x_n)) + d(\hat{J}(x_n), \bar{J}(x_n)) + d(\bar{J}(x_n), \bar{J}(x)).$$

Per ipotesi \bar{J} e \hat{J} coincidono su S , essendo estensioni di J , quindi $\hat{J}(x_n) = \bar{J}(x_n)$ e $d(\hat{J}(x_n), \bar{J}(x_n)) = 0$. Inoltre $d(\bar{J}(x_n), \bar{J}(x)) = d(x_n, x)$ e $d(\hat{J}(x_n), \hat{J}(x)) = d(x_n, x)$, per l'isometria di \bar{J} e \hat{J} . Dato che $d(x_n, x) \rightarrow 0$, segue che $d(\hat{J}(x), \bar{J}(x)) = 0$ per ogni $x \in E$. Se lo spazio di arrivo F è metrico, ciò implica che $\hat{J}(x) = \bar{J}(x)$ per ogni $x \in E$.

Mostriamo ora l'esistenza di un operatore \bar{J} che soddisfa (5.2). Per $x \in S$ poniamo $\bar{J}(x) := J(x)$. Per $x \in E \setminus S$, fissiamo un'arbitraria successione $\{\tilde{x}_n\}_{n \in \mathbb{N}}$ in S che converge verso x . Essendo convergente, $\{\tilde{x}_n\}_{n \in \mathbb{N}}$ è di Cauchy in E e quindi la successione delle immagini $\{J(\tilde{x}_n)\}_{n \in \mathbb{N}}$ è di Cauchy in F , poiché $d(J(\tilde{x}_n), J(\tilde{x}_m)) = d(\tilde{x}_n, \tilde{x}_m)$, grazie a (5.1). Essendo per ipotesi F completo, esiste almeno un punto limite per la successione $\{J(\tilde{x}_n)\}_{n \in \mathbb{N}}$: indicheremo con $\bar{J}(x)$ uno di tali punti, scelto arbitrariamente ma fissato una volta per tutte, per cui si ha $J(\tilde{x}_n) \rightarrow \bar{J}(x)$.

Abbiamo quindi definito un operatore $\bar{J} : E \rightarrow F$ che per costruzione soddisfa la prima relazione in (5.2). Per verificare la seconda relazione, siano $x, y \in E$. Se $x \in E \setminus S$, sia $\tilde{x}_n \in S$ la successione fissata sopra nella definizione di \bar{J} , mentre se $x \in S$ poniamo $\tilde{x}_n \equiv x$; analogamente definiamo la successione \tilde{y}_n . Per costruzione $\tilde{x}_n \rightarrow x$, $J(\tilde{x}_n) \rightarrow \bar{J}(x)$

e analogamente $\tilde{y}_n \rightarrow y$, $J(\tilde{y}_n) \rightarrow \bar{J}(y)$, quindi per continuità della pseudodistanza si ha

$$d(J(\tilde{x}_n), J(\tilde{y}_n)) \rightarrow d(\bar{J}(x), \bar{J}(y)), \quad d(\tilde{x}_n, \tilde{y}_n) \rightarrow d(x, y).$$

Dato che $d(J(\tilde{x}_n), J(\tilde{y}_n)) = d(\tilde{x}_n, \tilde{y}_n)$ per ogni $n \in \mathbb{N}$, grazie alla relazione (5.1), passando al limite $n \rightarrow \infty$ otteniamo $d(\bar{J}(x), \bar{J}(y)) = d(x, y)$. \square

OSSERVAZIONE 5.2. Abbiamo enunciato il Teorema 5.1 per operatori isometrici perché è il caso che ci interessa per l'integrale stocastico. Sottolineiamo tuttavia che la dimostrazione si estende quasi senza modifiche ad operatori *lipschitziani* $J : S \rightarrow F$: più precisamente, se $d(J(x), J(y)) \leq C d(x, y)$ per ogni $x, y \in S$, con $C \geq 0$, allora esiste un operatore $\bar{J} : E \rightarrow F$ che estende J e tale che $d(\bar{J}(x), \bar{J}(y)) \leq C d(x, y)$ per ogni $x, y \in E$; inoltre, tale operatore è unico se lo spazio d'arrivo F è metrico e completo. \square

Specializziamo ora il Teorema 5.1 al caso in cui E ed F sono spazi vettoriali e l'operatore J è lineare. L'analogo vettoriale di uno spazio pseudo metrico è dato da uno spazio *seminormato*: si tratta di uno spazio vettoriale reale E munito di una *seminorma*, cioè di una funzione $\|\cdot\| : E \rightarrow \mathbb{R}$ tale che per ogni $x, y \in E$ e per ogni $\lambda \in \mathbb{R}$ si abbia

$$\|0\| = 0, \quad \|\lambda x\| = |\lambda| \|x\|, \quad \|x + y\| \leq \|x\| + \|y\|.$$

Se si impone la condizione più forte che $\|x\| = 0$ se e solo se $x = 0$, la funzione $\|\cdot\|$ si dice *norma* e lo spazio vettoriale E si dice *spazio normato*. Ogni spazio seminormato (risp. normato) E è in particolare uno spazio pseudometrico (risp. metrico), in cui la pseudodistanza (risp. distanza) è definita da $d(x, y) := \|x - y\|$, per cui si applicano tutti i concetti definiti in precedenza: convergenza di successioni, densità di un sottoinsieme, completezza dello spazio, ... Notiamo che le operazioni di somma e di prodotto per scalari sono funzioni continue: più precisamente, se $x_n \rightarrow x$ e $y_n \rightarrow y$ si ha $\alpha x_n + \beta y_n \rightarrow \alpha x + \beta y$ per ogni $\alpha, \beta \in \mathbb{R}$, come si verifica facilmente usando le proprietà della seminorma.

Per quanto ci riguarda, l'esempio tipico di spazio seminormato (risp. normato) completo è dato dallo spazio di variabili aleatorie (risp. di classi di equivalenza) $L^p(\Omega, \mathcal{F}, P)$, in cui $X \mapsto \|X\|_p := (E(|X|^p))^{1/p}$ è una seminorma.

Veniamo ora al risultato annunciato.

COROLLARIO 5.3. Siano E uno spazio seminormato, F uno spazio seminormato completo, $S \subseteq E$ un *sottospazio vettoriale* denso e $J : S \rightarrow F$ un'*isometria lineare*:

$$J(\alpha x + \beta y) = \alpha J(x) + \beta J(y), \quad \forall \alpha, \beta \in \mathbb{R}, \forall x, y \in S, \quad (5.3)$$

$$\|J(x)\| = \|x\|, \quad \forall x \in S. \quad (5.4)$$

Allora esiste un'*isometria lineare* $\bar{J} : E \rightarrow F$ che estende J a tutto E :

$$\begin{aligned} \bar{J}(x) &= J(x), \quad \forall x \in S, \\ \bar{J}(\alpha x + \beta y) &= \alpha \bar{J}(x) + \beta \bar{J}(y), \quad \forall \alpha, \beta \in \mathbb{R}, \forall x, y \in E, \\ \|\bar{J}(x)\| &= \|x\|, \quad \forall x \in E. \end{aligned} \quad (5.5)$$

Se $\hat{J} : E \rightarrow F$ è un'altra isometria lineare che estende J , si ha $\|\hat{J}(x) - \bar{J}(x)\| = 0$ per ogni $x \in E$. In particolare, F è uno spazio *normato* (e non solo seminormato), esiste un'unica isometria lineare $\bar{J} : E \rightarrow F$ che estende J .

DIMOSTRAZIONE. Dalle proprietà (5.3), (5.4) segue che $d(J(x), J(y)) = \|J(x) - J(y)\| = \|J(x - y)\| = \|x - y\| = d(x, y)$ per ogni $x, y \in E$, dunque l'ipotesi (5.1) è soddisfatta. Il Teorema 5.1 garantisce l'esistenza di un'isometria $\bar{J} : E \rightarrow F$ che estende J , cioè tale che $\bar{J}(x) = J(x)$ per $x \in S$ e $\|\bar{J}(y) - \bar{J}(x)\| = \|y - x\|$ per ogni $x, y \in E$. Inoltre, se lo spazio F è normato, dunque metrico, tale operatore \bar{J} è unico.

Resta solo da mostrare che \bar{J} è lineare, ossia che soddisfa la relazione (5.5). Osserviamo che, essendo isometrico, l'operatore \bar{J} è continuo: più precisamente, se $z_n \rightarrow z$ si ha $\bar{J}(z_n) \rightarrow \bar{J}(z)$, perché per costruzione $\|\bar{J}(z_n) - \bar{J}(z)\| = \|z_n - z\| \rightarrow 0$.

Dati $x, y \in E$, siano $\{x_n\}_{n \in \mathbb{N}}$ e $\{y_n\}_{n \in \mathbb{N}}$ due successioni in S che convergono verso x e y rispettivamente. Per ogni $\alpha, \beta \in \mathbb{R}$ si ha $\alpha x_n + \beta y_n \rightarrow \alpha x + \beta y$, per continuità della somma e del prodotto per scalari, e per la continuità di \bar{J} si ottiene

$$\alpha \bar{J}(x_n) + \beta \bar{J}(y_n) \rightarrow \alpha \bar{J}(x) + \beta \bar{J}(y), \quad \bar{J}(\alpha x_n + \beta y_n) \rightarrow \bar{J}(\alpha x + \beta y).$$

Per la validità della relazione (5.5) è dunque sufficiente mostrare che $\bar{J}(\alpha x_n + \beta y_n) = \alpha \bar{J}(x_n) + \beta \bar{J}(y_n)$ per ogni $n \in \mathbb{N}$. Per costruzione $x_n, y_n \in S$ e quindi anche $\alpha x_n + \beta y_n \in S$, perché S è un sottospazio vettoriale. Dato che per costruzione \bar{J} è un'estensione di J , si ha $\bar{J}(x_n) = J(x_n)$, $\bar{J}(y_n) = J(y_n)$ e $J(\alpha x_n + \beta y_n) = J(\alpha x + \beta y)$. La conclusione segue allora dall'ipotesi (5.3). \square

OSSERVAZIONE 5.4. Supponiamo che, nelle stesse ipotesi del Teorema 5.3, esista una forma bilineare $\langle \cdot, \cdot \rangle$ su E tale che $\|x\| = \sqrt{\langle x, x \rangle}$ per ogni $x \in E$, e analogamente per F . Allora l'operatore lineare \bar{J} preserva, oltre alla seminorma, anche la forma bilineare, cioè

$$\langle \bar{J}(x), \bar{J}(y) \rangle = \langle x, y \rangle, \quad \forall x, y \in E.$$

Basta infatti notare che la forma bilineare si può ricostruire dalla seminorma grazie alla relazione $\langle a, b \rangle = \frac{1}{4}(\|a + b\|^2 - \|a - b\|^2)$, nota come *identità di polarizzazione*. \square

5.2. L'INTEGRALE STOCASTICO IN $M^2[a, b]$

Ricordiamo che è fissato uno spazio filtrato standard $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$, su cui è definito un $\{\mathcal{F}_t\}_{t \geq 0}$ -moto browniano reale $B = \{B_t\}_{t \geq 0}$ e su cui saranno definiti tutti i processi che introdurremo.

5.2.1. SPAZI DI PROCESSI. Per $0 \leq a < b < \infty$ fissati, introduciamo lo spazio $M^2[a, b]$ dei processi per i quali definiremo inizialmente l'integrale stocastico.

DEFINIZIONE 5.5. Indichiamo con $M^2[a, b]$ lo spazio vettoriale dei processi reali $X = \{X_t\}_{t \in [a, b]}$ *progressivamente misurabili* e tali che

$$\|X\|_{M^2}^2 := E \left(\int_a^b X_t^2 dt \right) < \infty. \quad (5.6)$$

Ricordando che un processo progressivamente misurabile è misurabile, per il Teorema di Fubini si ha $\|X\|_{M^2}^2 = \int_{\Omega \times [a, b]} X_t(\omega)^2 (P \otimes Leb)(d\omega, dt)$. Questo mostra che

$$M^2[a, b] \subseteq L^2(\Omega \times [a, b], \mathcal{F} \otimes \mathcal{B}([a, b]), P \otimes Leb),$$

e $\|\cdot\|_{M^2}$ non è altro che l'ordinaria seminorma dello spazio $L^2(\Omega \times [a, b])$. Lo spazio $M^2[a, b]$ è dunque *seminormato*.[†] Notiamo anche che la seminorma è indotta dalla forma bilineare $\langle X, Y \rangle_{M^2} := E(\int_a^b X_t Y_t dt)$, si ha cioè $\|X\|_{M^2}^2 = \langle X, X \rangle_{M^2}$ per ogni $X \in M^2[a, b]$.

OSSERVAZIONE 5.6. Se $X = \{X_t\}_{t \in [a, b]}$ è un processo reale misurabile — in particolare, se X è progressivamente misurabile — per il teorema di Fubini la quantità $N_X(\omega) := \int_a^b X_t(\omega)^2 dt \in [0, \infty]$, ossia il quadrato della norma della traiettoria $t \mapsto X_t(\omega)$ in $L^2([a, b], \mathcal{B}([a, b]), Leb)$, è ben definita per ogni $\omega \in \Omega$ ed è una variabile aleatoria. Possiamo allora vedere $M^2[a, b]$ come lo spazio dei processi $X = \{X_t\}_{t \in [a, b]}$ progressivamente misurabili tali che $N_X \in L^1(\Omega, \mathcal{F}, P)$. \square

Non è difficile verificare che i seguenti processi $X = \{X_t\}_{t \in [a, b]}$ sono in $M^2[a, b]$:

$$X_t = B_t, \quad X_t = \int_a^t |B_s|^p ds, \quad \forall p \geq 0, \quad X_t = e^{\lambda B_t}, \quad \forall \lambda \in \mathbb{R}, \quad X_t = \sup_{a \leq s \leq t} B_s,$$

mentre ad esempio il processo $X_t := e^{(B_t)^3}$ non lo è, perché $\|X\|_{M^2} = \infty$. Vedremo più avanti come sia possibile trattare anche questo genere di processi.

Diremo che un processo $X = \{X_s\}_{s \in [a, b]}$ è *semplice* se le sue traiettorie sono costanti a tratti con istanti di salto deterministici:

$$X_s(\omega) = \sum_{i=0}^{k-1} C_i(\omega) 1_{[t_i, t_{i+1})}(s), \quad \forall s \in [a, b], \quad (5.7)$$

con $a = t_0 < t_1 < \dots < t_k = b$ e $C_i : (\Omega, \mathcal{F}, P) \rightarrow \mathbb{R}$ variabili aleatorie.

Definiremo inizialmente l'integrale stocastico su un opportuno spazio di processi semplici $\mathcal{S}[a, b]$, che ora introduciamo, che avranno la funzione di approssimanti.

DEFINIZIONE 5.7. Indichiamo con $\mathcal{S}[a, b]$ lo spazio vettoriale dei processi $X = \{X_s\}_{s \in [a, b]}$ semplici, cioè della forma (5.7), tali che per ogni $i = 0, \dots, k-1$ la variabile aleatoria C_i è \mathcal{F}_{t_i} -misurabile e di quadrato integrabile (cioè $C_i \in L^2(\Omega, \mathcal{F}_{t_i}, P)$).

[†]Si noti che $\|X\|_{M^2} = 0$ non implica che il processo X sia identicamente nullo, ma soltanto che $X_t(\omega) = 0$ per $(Leb \otimes P)$ -q.o. $(t, \omega) \in [a, b] \times \Omega$.

Una prima osservazione elementare è che lo spazio $\mathcal{S}[a, b]$ è contenuto in $M^2[a, b]$. Sia infatti $X \in \mathcal{S}[a, b]$, della forma (5.7). Dato $t \in [a, b]$, se $i \in \{0, \dots, k-1\}$ è tale che $t \in [t_i, t_{i+1})$, si ha che $X_t = C_i$ è \mathcal{F}_{t_i} -misurabile per ipotesi, quindi a maggior ragione è \mathcal{F}_t -misurabile (infatti $\mathcal{F}_{t_i} \subseteq \mathcal{F}_t$, essendo $t_i \leq t$). Ciò mostra che il processo X è adattato, quindi progressivamente misurabile perché continuo a destra, grazie al Lemma 3.12. Resta da verificare che $\|X\|_{M^2} < \infty$. Si osservi che $X_s^2 = \sum_{i=0}^{k-1} C_i^2 \mathbf{1}_{[t_i, t_{i+1})}(s)$, perché $\mathbf{1}_{[t_i, t_{i+1})}(s) \mathbf{1}_{[t_j, t_{j+1})}(s) \equiv 0$ per $i \neq j$. Essendo per ipotesi $C_i \in L^2(\Omega)$, si ottiene

$$\|X\|_{M^2}^2 = \mathbb{E} \left(\int_a^b X_t^2 dt \right) = \mathbb{E} \left(\sum_{i=0}^{k-1} \int_a^b C_i^2 \mathbf{1}_{[t_i, t_{i+1})}(t) dt \right) = \sum_{i=0}^{k-1} \mathbb{E}(C_i^2) (t_{i+1} - t_i) < \infty,$$

dunque $X \in M^2[a, b]$.

Un risultato meno evidente è che i processi in $\mathcal{S}[a, b]$ possono essere effettivamente usati come approssimanti per i processi in $M^2[a, b]$, come mostra la seguente proposizione (la cui dimostrazione è posposta al sottoparagrafo 5.2.5).

PROPOSIZIONE 5.8 (DENSITÀ DI $\mathcal{S}[a, b]$ IN $M^2[a, b]$). Per ogni processo $X \in M^2[a, b]$ esiste una successione di processi $X^{(n)} \in \mathcal{S}[a, b]$ tale che $\|X^{(n)} - X\|_{M^2} \rightarrow 0$ per $n \rightarrow \infty$.

5.2.2. L'INTEGRALE STOCASTICO PER PROCESSI SEMPLICI. Cominciamo a definire l'integrale stocastico quando l'integrando è un processo in $\mathcal{S}[a, b]$.

DEFINIZIONE 5.9 (INTEGRALE STOCASTICO DI PROCESSI SEMPLICI). Dato un processo semplice $X \in \mathcal{S}[a, b]$, della forma (5.7), si dice *integrale stocastico* di X rispetto al moto browniano B la variabile aleatoria

$$J_{a,b}(X) := \int_a^b X_t dB_t := \sum_{i=0}^{k-1} C_i (B_{t_{i+1}} - B_{t_i}). \quad (5.8)$$

Sottolineiamo che l'integrale stocastico $\int_a^b X_t dB_t$ è una variabile aleatoria definita sullo stesso spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ su cui è definito il moto browniano B . Volendo essere più espliciti, possiamo scrivere

$$J_{a,b}(X)(\omega) := \left(\int_a^b X_t dB_t \right)(\omega) := \sum_{i=0}^{k-1} C_i(\omega) (B_{t_{i+1}}(\omega) - B_{t_i}(\omega)), \quad \forall \omega \in \Omega. \quad (5.9)$$

Si noti che questa definizione è piuttosto naturale: se B_t e X_t descrivono rispettivamente il valore di una azione e il numero di azioni in mio possesso all'istante t , l'integrale $\int_a^b X_t dB_t$ rappresenta il guadagno (con segno!) delle mie azioni tra gli istanti a e b . Alla luce di questa interpretazione, la richiesta che C_i sia \mathcal{F}_{t_i} -misurabile, e dunque che il processo X sia adattato, appare molto ragionevole: infatti il numero di azioni che possiedo all'istante t_i può essere deciso solo in funzione dell'informazione \mathcal{F}_{t_i} disponibile fino a quell'istante, senza guardare al futuro.

OSSERVAZIONE 5.10. Non è evidente che la Definizione 5.9 sia ben posta, poiché la rappresentazione (5.7) di un processo semplice non è unica: per esempio, se X è dato da (5.7), si può aggiungere “fittiziamente” un punto $\bar{t} \in (t_i, t_{i+1})$ mantenendo la stessa variabile C_i nei due sottointervalli $[t_i, \bar{t})$ e $[\bar{t}, t_{i+1})$. Tuttavia, non è difficile verificare che, se si usa una diversa rappresentazione $X_t(\omega) = \sum_{i=0}^{k'-1} C'_i(\omega) 1_{[t'_i, t'_{i+1})}(t)$ per il processo X dato da (5.7), l'integrale $\int_a^b X_t dB_t$ definito in (5.8) non cambia. \square

Le proprietà fondamentali dell'integrale stocastico di processi semplici sono date nella seguente proposizione, la cui dimostrazione è posposta al sottoparagrafo 5.2.6.

PROPOSIZIONE 5.11. L'integrale stocastico di processi semplici $X \mapsto J_{a,b}(X)$ definisce un operatore lineare: per ogni scelta di $\alpha, \beta \in \mathbb{R}$ e $X, Y \in \mathcal{S}[a, b]$

$$J_{a,b}(\alpha X + \beta Y)(\omega) = \alpha J_{a,b}(X)(\omega) + \beta J_{a,b}(Y)(\omega), \quad \forall \omega \in \Omega.$$

Per ogni $X \in \mathcal{S}[a, b]$ si ha $J_{a,b}(X) \in L^2(\Omega)$ e valgono le seguenti relazioni:

$$\mathbb{E}(J_{a,b}(X) | \mathcal{F}_a) = 0, \quad \mathbb{E}(J_{a,b}(X)^2 | \mathcal{F}_a) = \mathbb{E}\left(\int_a^b X_t^2 dt \middle| \mathcal{F}_a\right), \quad \text{q.c.} \quad (5.10)$$

Di conseguenza, per ogni $X \in \mathcal{S}[a, b]$ si ha che

$$\mathbb{E}(J_{a,b}(X)) = 0, \quad \|J_{a,b}(X)\|_{L^2(\Omega)} = \|X\|_{M^2}. \quad (5.11)$$

5.2.3. COSTRUZIONE DELL'INTEGRALE STOCASTICO. Abbiamo finalmente tutti gli elementi per estendere la definizione dell'integrale stocastico $\int_a^b X_t dB_t$ a ogni processo $X \in M^2[a, b]$, sfruttando il Corollario 5.3. Si noti infatti che:

- $M^2[a, b]$ è uno spazio seminormato;
- $L^2(\Omega, \mathcal{F}, \mathbb{P})$ come spazio di variabili aleatorie è uno spazio *seminormato* completo, mentre come spazio di classi di equivalenza è uno spazio *normato* completo (spazio di Banach), come abbiamo ricordato in precedenza;
- $\mathcal{S}[a, b]$ è un sottospazio vettoriale denso di $M^2[a, b]$, grazie alla Proposizione 5.8;
- $J_{a,b} : \mathcal{S}[a, b] \rightarrow L^2(\Omega, \mathcal{F}, \mathbb{P})$ (dove $L^2(\Omega, \mathcal{F}, \mathbb{P})$ indica lo spazio di funzioni o di classi di equivalenza, a nostra scelta) è un operatore lineare e isometrico, per la Proposizione 5.11.

Possiamo allora applicare il Corollario 5.3 con

$$E = M^2[a, b], \quad S = \mathcal{S}[a, b], \quad F = L^2(\Omega, \mathcal{F}, \mathbb{P}), \quad J = J_{a,b}, \quad (5.12)$$

che garantisce l'esistenza di un operatore $\bar{J} : M^2[a, b] \rightarrow L^2(\Omega, \mathcal{F}, \mathbb{P})$ lineare e isometrico che estende $J = J_{a,b}$. Tale operatore è detto *integrale stocastico* e, con un piccolo abuso di notazione, sarà ancora indicato con $J_{a,b}(X) = \int_a^b X_t dB_t$.

È importante sottolineare che l'estensione (lineare e isometrica) di $J_{a,b}$ a tutto lo spazio $M^2[a, b]$ è *unica* se pensiamo lo spazio d'arrivo $L^2(\Omega, \mathcal{F}, P)$ come spazio di classi di equivalenza di variabili aleatorie. In altri termini, per ogni $X \in M^2[a, b]$ l'integrale stocastico $J_{a,b}(X) = \int_a^b X_t dB_t$ determina univocamente una classe di equivalenza di variabili aleatorie q.c. uguali, tuttavia il valore $J_{a,b}(X)(\omega)$ per un fissato $\omega \in \Omega$ (cioè la variabile aleatoria specifica all'interno della classe di equivalenza) non è definito canonicamente.[†] Ciononostante, si usa talvolta la notazione $\int_a^b X_t(\omega) dB_t(\omega)$, come se l'integrale fosse definito traiettoria per traiettoria. Nel seguito sottintenderemo spesso questa ambiguità nell'interpretazione di $L^2(\Omega, \mathcal{F}, P)$.

Riassumendo, possiamo dare la seguente definizione.

DEFINIZIONE 5.12 (INTEGRALE STOCASTICO). Si definisce *integrale stocastico* l'unico operatore $X \mapsto J_{a,b}(X) = \int_a^b X_t dB_t$, definito per $X \in M^2[a, b]$ a valori in $L^2(\Omega, \mathcal{F}, P)$, con le seguenti proprietà:

- $J_{a,b}$ è un operatore *lineare*, cioè $J_{a,b}(\alpha X + \beta Y) = \alpha J_{a,b}(X) + \beta J_{a,b}(Y)$ q.c. per ogni $\alpha, \beta \in \mathbb{R}$ e $X, Y \in M^2[a, b]$, e *isometrico*:

$$\|J_{a,b}(X)\|_{L^2(\Omega)} = \|X\|_{M^2}, \quad \forall X \in M^2[a, b]; \quad (5.13)$$

- se $X \in \mathcal{S}[a, b]$ è un processo semplice, $J_{a,b}(X)$ è dato dalla Definizione 5.8.

OSSERVAZIONE 5.13. Nonostante abbiamo usato un risultato piuttosto astratto, quale il Corollario 5.3, è utile tenere a mente che la definizione dell'integrale stocastico è in realtà abbastanza concreta. In effetti, come è chiaro dalla dimostrazione del Teorema 5.1, per costruire l'integrale stocastico $\int_a^b X_t dB_t$ per un processo $X \in M^2$ si considera una qualunque successione $\{X^{(n)}\}_{n \in \mathbb{N}} \in \mathcal{S}[a, b]$ di processi semplici che converge verso X in $M^2[a, b]$ e si definisce $\int_a^b X_t dB_t$ come il limite in $L^2(\Omega, \mathcal{F}, P)$ delle variabili aleatorie $\int_a^b X_t^{(n)} dB_t$, che sono definite esplicitamente in (5.8). \square

OSSERVAZIONE 5.14. Nel caso di processi integrandi deterministici, cioè se $X_t(\omega) = f(t)$ non dipende da $\omega \in \Omega$, la condizione $X \in M^2[a, b]$ equivale alla richiesta che $f : [a, b] \rightarrow \mathbb{R}$ sia misurabile e $\int_a^b |f(t)|^2 dt < \infty$, cioè $f \in L^2([a, b])$. In questo caso speciale, l'integrale stocastico $\int_a^b f(t) dB_t$ è detto *integrale di Wiener* e può essere costruito in modo più diretto ed elementare. Inoltre la variabile aleatoria $\int_a^b f(t) dB_t$, che sappiamo avere media nulla e varianza $\sigma^2 = \|f\|_{L^2([a, b])}^2 = \int_a^b f(t)^2 dt$, è in questo caso *una variabile normale* $\mathcal{N}(0, \sigma^2)$. Questo in generale non è vero per $X \in M^2[a, b]$. \square

Elenchiamo alcune semplici proprietà dell'integrale stocastico.

[†]Vedremo più avanti come sia possibile definire appropriatamente il valore puntuale $J_{a,b}(X)(\omega)$ in modo che l'integrale stocastico sia una funzione regolare dell'estremo di integrazione (cf. il paragrafo 5.3).

PROPOSIZIONE 5.15. Per ogni scelta di $X, Y \in M^2$ valgono le seguenti proprietà:

$$\mathbb{E} \left(\int_a^b X_t dB_t \right) = 0; \quad (5.14)$$

$$\text{Var} \left(\int_a^b X_t dB_t \right) = \mathbb{E} \left(\int_a^b X_t^2 dt \right); \quad (5.15)$$

$$\text{Cov} \left(\int_a^b X_t dB_t, \int_a^b Y_t dB_t \right) = \mathbb{E} \left(\int_a^b X_t Y_t dt \right); \quad (5.16)$$

$$\int_a^b X_t dB_t = \int_a^b Y_t dB_t \text{ q.c.} \iff \mathbb{E} \left(\int_a^b (X_t - Y_t)^2 dt \right) = 0. \quad (5.17)$$

DIMOSTRAZIONE. Usando per brevità la notazione compatta $J_{a,b}(X) = \int_a^b X_t dB_t$ e sfruttando la relazione (5.14), possiamo riscrivere (5.15) come $\|J_{a,b}(X)\|_{L^2(\Omega)} = \|X\|_{M^2}$, che non è altro che la proprietà di isometria (5.13); analogamente, la relazione (5.16) può essere riscritta nella forma $\langle J_{a,b}(X), J_{a,b}(Y) \rangle_{L^2(\Omega)} = \langle X, Y \rangle_{M^2}$ e la sua validità segue facilmente dall'isometria (5.13) per polarizzazione, cf. l'Osservazione 5.4.

La relazione (5.17) si dimostra notando che $\|X - Y\|_{M^2} = \|J_{a,b}(X) - J_{a,b}(Y)\|_{L^2(\Omega)}$, per l'isometria (5.13) e la linearità dell'integrale stocastico, quindi si ha $\|X - Y\|_{M^2} = 0$ (cioè la condizione a destra in (5.17)) se e solo se $\|J_{a,b}(X) - J_{a,b}(Y)\|_{L^2(\Omega)} = 0$, il che è equivalente a $J_{a,b}(X) = J_{a,b}(Y)$ q.c..

Resta solo da dimostrare (5.14), cioè $\mathbb{E}(J_{a,b}(X)) = 0$ per ogni $X \in M^2[a, b]$. Per costruzione, data un'arbitraria successione $\{X^{(n)}\}_{n \in \mathbb{N}}$ di processi in $\mathcal{S}[a, b]$ che converge verso X in $M^2[a, b]$, si ha $J_{a,b}(X) = \lim_{n \rightarrow \infty} J_{a,b}(X^{(n)})$ in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Dato che la convergenza in L^2 implica la convergenza dei valori attesi e visto che $\mathbb{E}(J_{a,b}(X^{(n)})) = 0$ per ogni $n \in \mathbb{N}$, per la Proposizione 5.11, segue che $\mathbb{E}(J_{a,b}(X)) = 0$. \square

Enunciamo infine esplicitamente la continuità dell'integrale stocastico come operatore da $M^2[a, b]$ in $L^2(\Omega, \mathcal{F}, \mathbb{P})$, che è una conseguenza diretta della proprietà di isometria.

PROPOSIZIONE 5.16. Per ogni successione $\{X^{(n)}\}_{n \in \mathbb{N}}$ di processi che converge in $M^2[a, b]$ verso un processo X , la successione dei corrispondenti integrali stocastici $\int_a^b X_t^{(n)} dB_t$ converge in $L^2(\Omega, \mathcal{F}, \mathbb{P})$ verso $\int_a^b X_t dB_t$.

DIMOSTRAZIONE. Per ipotesi $X^{(n)} \rightarrow X$ in $M^2[a, b]$, cioè $\|X^{(n)} - X\|_{M^2} \rightarrow 0$, quindi grazie alla relazione (5.13) $\|J_{a,b}(X^{(n)}) - J_{a,b}(X)\|_{L^2(\Omega)} = \|X^{(n)} - X\|_{M^2} \rightarrow 0$, cioè $J_{a,b}(X^{(n)}) \rightarrow J_{a,b}(X)$ in $L^2(\Omega)$. \square

Il resto di questo paragrafo è dedicato alla dimostrazione delle Proposizioni 5.8 (nel sottoparagrafo 5.2.5) e 5.11 (nel sottoparagrafo 5.2.6). Vale però la pena fare una piccola digressione, per capire meglio il ruolo della misurabilità progressiva.

5.2.4. DIGRESSIONE. Se si vuole sviluppare una teoria dell'integrazione stocastica, è piuttosto naturale l'idea di definire inizialmente l'integrale stocastico per processi semplici della forma (5.7), tramite l'equazione (5.8) (la quale è ben definita anche se $X \notin \mathcal{S}[a, b]$), e poi estendere la definizione a processi $X = \{X_t\}_{a \leq t \leq b}$ più generali approssimando (in un senso opportuno) X con una successione di processi semplici $X^{(n)} \in \mathcal{S}[a, b]$ e definendo $\int_a^b X_t dB_t$ come il limite (in un senso opportuno) di $\int_a^b X_t^{(n)} dB_t$.

Il punto delicato è capire quali condizioni imporre sul processo X e sui processi semplici $X^{(n)}$ perché questa procedura funzioni. Restrizioni a priori molto forti, quali ad esempio la continuità e l'uniforme limitatezza delle traiettorie di X , non risultano sufficienti. La ragione di queste difficoltà sta nel fatto più volte menzionato che le traiettorie del moto browniano hanno variazione infinita.

L'idea fondamentale di Itô [1944] è stata proprio quella di restringersi ai processi $X = \{X_t\}_{a \leq t \leq b}$ *progressivamente misurabili*, o più precisamente allo spazio $M^2[a, b]$. In questo modo, come abbiamo visto, si possono sfruttare le proprietà probabilistiche del moto browniano B per mostrare che, per un'opportuna successione di processi semplici $X^{(n)} \in \mathcal{S}[a, b]$ che approssimano X , si ha la *convergenza in* $L^2(\Omega, \mathcal{F}, P)$ degli integrali stocastici $\int_a^b X_t^{(n)} dB_t$ verso una variabile aleatoria limite, che è per definizione $\int_a^b X_t dB_t$.

Per capire meglio i problemi che sorgono e il ruolo della progressiva misurabilità, prendiamo come integrando il moto browniano stesso: $X = \{B_t\}_{t \in [a, b]}$. Data una successione di partizioni $\pi^{(n)} = \{a = t_0^{(n)} < t_1^{(n)} < \dots < t_{k_n}^{(n)} = b\}$ dell'intervallo $[a, b]$ con passo che tende a zero, due possibili scelte di approssimanti di X , a priori entrambe legittime, sono per esempio

$$X_t^{(n)} := \sum_{i=1}^{k_n} B_{t_{i-1}^{(n)}} \mathbf{1}_{[t_{i-1}^{(n)}, t_i^{(n)})}(t), \quad \tilde{X}_t^{(n)} := \sum_{i=1}^{k_n} B_{t_i^{(n)}} \mathbf{1}_{[t_{i-1}^{(n)}, t_i^{(n)})}(t).$$

Il problema è che queste due scelte portano a risultati diversi! Infatti per $n \rightarrow \infty$

$$\int_a^b \tilde{X}_t^{(n)} dB_t - \int_a^b X_t^{(n)} dB_t = \sum_{i=1}^{k_n} (B_{t_i^{(n)}} - B_{t_{i-1}^{(n)}})^2 \longrightarrow b - a \quad \text{in } L^2,$$

grazie alla Proposizione 2.17. Qui emerge chiaramente come la differenza dei risultati sia proprio dovuta al fatto che il moto browniano ha variazione quadratica positiva. Il processo approssimante “giusto” per noi è $X^{(n)}$, che è progressivamente misurabile (facile esercizio), a differenza di $\tilde{X}^{(n)}$ che non è neanche adattato.

5.2.5. DIMOSTRAZIONE DELLA PROPOSIZIONE 5.8. Cominciamo con alcuni risultati deterministici di approssimazione. Sia $L^2([a, b]) = L^2([a, b], \mathcal{B}([a, b]), dt)$ lo spazio delle funzioni misurabili da $[a, b]$ a \mathbb{R} di quadrato integrabile rispetto alla misura di Lebesgue, la cui norma sarà indicata con $\|f\|_2 := (\int_a^b f(t)^2 dt)^{1/2}$. Definiamo per $n \in \mathbb{N}$ l'operatore lineare $P_n : L^2([a, b]) \rightarrow L^2([a, b])$ mediante

$$(P_n f)(t) := \sum_{i=1}^{\lfloor n(b-a) \rfloor - 1} c_{n,i}(f) \mathbf{1}_{[a + \frac{i}{n}, a + \frac{i+1}{n})}(t), \quad \text{dove} \quad c_{n,i}(f) := \frac{1}{(\frac{1}{n})} \int_{a + \frac{i-1}{n}}^{a + \frac{i}{n}} f(s) ds.$$

In parole, l'operatore P_n approssima la funzione f con una funzione $P_n f$ costante a tratti su intervalli di ampiezza $\frac{1}{n}$: il valore di $P_n f$ in ciascuno di questi intervalli è dato dalla media di f sull'intervallo precedente. Si noti che se $f \in L^2([a, b])$ si ha $f \in L^1([a, b])$, perché $[a, b]$ è uno spazio di misura finito, quindi $c_{n,i}(f) < \infty$ per ogni i, n .

Dimostriamo ora alcune proprietà basilari dell'operatore P_n .

LEMMA 5.17. P_n è un operatore 1-lipschitziano, cioè $\|P_n f\|_2 \leq \|f\|_2$ per ogni $f \in L^2([a, b])$. Inoltre $P_n f \rightarrow f$ in $L^2([a, b])$ per $n \rightarrow \infty$, per ogni $f \in L^2([a, b])$.

DIMOSTRAZIONE. Per la prima parte, si noti che

$$(c_{n,i}(f))^2 = \left(\frac{1}{(\frac{1}{n})} \int_{a+\frac{i-1}{n}}^{a+\frac{i}{n}} f(s) ds \right)^2 \leq \frac{1}{(\frac{1}{n})} \int_{a+\frac{i-1}{n}}^{a+\frac{i}{n}} f(s)^2 ds,$$

avendo usato la disuguaglianza di Jensen, da cui si ottiene

$$\begin{aligned} \|P_n f\|_2^2 &= \int_a^b (P_n f)(t)^2 dt = \sum_{i=1}^{\lfloor n(b-a) \rfloor - 1} (c_{n,i}(f))^2 \cdot \frac{1}{n} \\ &\leq \sum_{i=1}^{\lfloor n(b-a) \rfloor - 1} \int_{a+\frac{i-1}{n}}^{a+\frac{i}{n}} f(s)^2 ds \leq \int_a^b f(t)^2 dt = \|f\|_2^2. \end{aligned}$$

Per la seconda parte, osserviamo innanzitutto che se $g : [a, b] \rightarrow \mathbb{R}$ è continua si ha $(P_n g)(t) \rightarrow g(t)$ per $n \rightarrow \infty$, per ogni $t \geq 0$ (esercizio). Essendo continua su un compatto, g è limitata: $\sup_{t \in [a, b]} |g(t)| \leq K$ con $K \in (0, \infty)$, da cui segue che anche $\sup_{t \in [a, b]} |(P_n g)(t)| \leq K$, poiché chiaramente $|c_{n,i}(f)| \leq K$. Si ha allora che $\int_a^b |(P_n g)(t) - g(t)|^2 dt = \|P_n g - g\|_2^2 \rightarrow 0$, per convergenza dominata, cioè $P_n g \rightarrow g$ in $L^2([a, b])$.

Il caso generale di $f \in L^2([a, b])$ segue facilmente per approssimazione. Dato che le funzioni continue sono dense in $L^2([a, b])$, per ogni $\varepsilon > 0$ esiste una funzione continua g tale che $\|f - g\|_2 \leq \frac{\varepsilon}{3}$, per cui $\|P_n f - P_n g\|_2 = \|P_n(f - g)\|_2 \leq \|f - g\|_2 \leq \frac{\varepsilon}{3}$. Dato che $P_n g \rightarrow g$ in $L^2([a, b])$, fissiamo n_0 in modo che $\|P_n g - g\|_2 \leq \frac{\varepsilon}{3}$ per $n \geq n_0$. Dalla disuguaglianza triangolare segue dunque che per $n \geq n_0$

$$\|P_n f - f\|_2 \leq \|P_n f - P_n g\|_2 + \|P_n g - g\|_2 + \|g - f\|_2 \leq \varepsilon.$$

Questo mostra che $P_n f \rightarrow f$ in $L^2([a, b])$ per $n \rightarrow \infty$, per ogni $f \in L^2([a, b])$. \square

Passiamo ora all'approssimazione di processi. Notiamo innanzitutto che se $X \in M^2[a, b]$ allora per q.o. $\omega \in \Omega$ la funzione $u \mapsto X_u(\omega)$, che indicheremo con $X.(\omega)$, è in $L^2([a, b])$: infatti $\|X\|_{M^2}^2 = E \left(\int_a^b X_t^2 dt \right) < \infty$ e dunque $\int_a^b X_t(\omega)^2 dt = \|X.(\omega)\|_2^2 < \infty$ q.c.. Possiamo allora definire l'operatore $\hat{P}_n : M^2[a, b] \rightarrow M^2[a, b]$ ponendo $(\hat{P}_n X)_t(\omega) := 0$ se $X.(\omega) \notin L^2([a, b])$, mentre se $X.(\omega) \in L^2([a, b])$ definiamo $(\hat{P}_n X)_t(\omega) := (P_n X.(\omega))(t)$, cioè

appliciamo l'operatore P_n sopra introdotto alla funzione $u \mapsto X_u(\omega)$. Più esplicitamente:

$$(\hat{P}_n X)_t = \sum_{i=1}^{\lfloor n(b-a) \rfloor - 1} C_{n,i} 1_{[a+\frac{i}{n}, a+\frac{i+1}{n})}(t), \quad \text{dove} \quad C_{n,i} = \frac{1}{(\frac{1}{n})} \int_{a+\frac{i-1}{n}}^{a+\frac{i}{n}} X_s ds. \quad (5.18)$$

Chiaramente $\hat{P}_n X$ è un processo con traiettorie costanti a tratti, della forma (5.7). Il fatto che $C_{n,i}$ sia $\mathcal{F}_{a+\frac{i}{n}}$ -misurabile segue dal fatto che il processo X è progressivamente misurabile, per definizione di $M^2[a, b]$ (intuitivamente, $C_{n,i}$ è funzione di $\{X_s\}_{s \leq a+\frac{i}{n}}$). Infine, applicando la disuguaglianza di Jensen si ha che $C_{n,i} \in L^2(\Omega)$:

$$E(C_{n,i}^2) \leq \frac{1}{(\frac{1}{n})} \int_{a+\frac{i-1}{n}}^{a+\frac{i}{n}} E(X_s^2) ds \leq n \|X\|_{M^2}^2 < \infty.$$

In definitiva, abbiamo mostrato che $\hat{P}_n X \in \mathcal{S}[a, b]$, per ogni $X \in M^2[a, b]$ e $n \in \mathbb{N}$. La dimostrazione della Proposizione 5.8 è allora completata con il seguente lemma. \square

LEMMA 5.18. Per ogni $X \in M^2[a, b]$ si ha che $\hat{P}_n X \rightarrow X$ in $M^2[a, b]$ per $n \rightarrow \infty$.

DIMOSTRAZIONE. L'affermazione $\hat{P}_n X \rightarrow X$ in $M^2[a, b]$ è equivalente a richiedere che

$$\lim_{n \rightarrow \infty} \|\hat{P}_n X - X\|_{M^2}^2 = \lim_{n \rightarrow \infty} E(A_n(\omega)) = 0, \quad (5.19)$$

dove abbiamo posto per comodità

$$A_n(\omega) := \int_a^b |(\hat{P}_n X)_t(\omega) - X_t(\omega)|^2 dt = \|P_n X(\omega) - X(\omega)\|_2^2,$$

dove ricordiamo che $\|\cdot\|_2$ denota la norma in $L^2([a, b])$. Dato che $X(\omega)$ in $L^2([a, b])$ per q.o. $\omega \in \Omega$, grazie al Lemma 5.17 si ha che $P_n X(\omega) \rightarrow X(\omega)$ in $L^2([a, b])$, cioè $A_n(\omega) \rightarrow 0$. La relazione (5.19) segue dunque dal teorema di convergenza dominata: infatti applicando la disuguaglianza $\|a - b\|^2 \leq (\|a\| + \|b\|)^2 \leq 2(\|a\|^2 + \|b\|^2)$ e ancora il Lemma 5.17 possiamo scrivere

$$A_n(\omega) \leq 2(\|P_n X(\omega)\|_2^2 + \|X(\omega)\|_2^2) \leq 4\|X(\omega)\|_2^2,$$

e questa variabile è integrabile: infatti

$$E(\|X(\omega)\|_2^2) = E\left(\int_a^b X_t(\omega)^2 dt\right) = \|X\|_{M^2}^2 < \infty,$$

perché per ipotesi $X \in M^2[a, b]$. \square

5.2.6. DIMOSTRAZIONE DELLA PROPOSIZIONE 5.11. Mostriamo che per ogni processo semplice $X \in \mathcal{S}[a, b]$ si ha $J_{a,b}(X) \in L^2(\Omega)$. Se X è della forma (5.7), per definizione (cf. (5.8)) si ha

$$J_{a,b}(X) = \sum_{i=0}^{k-1} C_i (B_{t_{i+1}} - B_{t_i}). \quad (5.20)$$

Basta dunque mostrare che $C_i (B_{t_{i+1}} - B_{t_i}) \in L^2(\Omega)$, per ogni $i = 0, \dots, k-1$. Per definizione di $\{\mathcal{F}_t\}_{t \geq 0}$ -moto browniano, $(B_{t_{i+1}} - B_{t_i})$ è indipendente dalla σ -algebra \mathcal{F}_{t_i} , mentre C_i è \mathcal{F}_{t_i} -misurabile per definizione di $\mathcal{S}[a, b]$. In particolare, le variabili aleatorie $(B_{t_{i+1}} - B_{t_i})^2$ e C_i^2 sono indipendenti, e anche integrabili, perché $(B_{t_{i+1}} - B_{t_i})$ è normale e $C_i \in L^2(\Omega)$ per ipotesi. Di conseguenza il loro prodotto è integrabile: $E(C_i^2 (B_{t_{i+1}} - B_{t_i})^2) = E(C_i^2) E((B_{t_{i+1}} - B_{t_i})^2) < \infty$, cioè $C_i (B_{t_{i+1}} - B_{t_i}) \in L^2(\Omega)$.

Per quanto riguarda la linearità, dati due processi semplici $X, Y \in \mathcal{S}[a, b]$, conviene esprimerli nella forma (5.7) usando *lo stesso insieme di tempi* t_1, \dots, t_k (eventualmente aggiungendo istanti di salto fittizi) e dalla definizione (5.8) segue immediatamente che $J_{a,b}(\alpha X + \beta Y) = \alpha J_{a,b}(X) + \beta J_{a,b}(Y)$.

Verifichiamo ora le relazioni in (5.10). La prima è immediata: ricordando (5.20) si ha

$$\begin{aligned} E(J_{a,b}(X) | \mathcal{F}_a) &= \sum_{i=0}^{k-1} E(C_i (B_{t_{i+1}} - B_{t_i}) | \mathcal{F}_a) = \sum_{i=0}^{k-1} E(E(C_i (B_{t_{i+1}} - B_{t_i}) | \mathcal{F}_{t_i}) | \mathcal{F}_a) \\ &= \sum_{i=0}^{k-1} E(C_i | \mathcal{F}_a) E(B_{t_{i+1}} - B_{t_i}) = 0, \end{aligned}$$

avendo usato ancora il fatto che $(B_{t_{i+1}} - B_{t_i})$ è indipendente da \mathcal{F}_{t_i} mentre C_i è \mathcal{F}_{t_i} -misurabile. Allo stesso modo possiamo scrivere

$$\begin{aligned} E(J_{a,b}(X)^2 | \mathcal{F}_a) &= \sum_{i=0}^{k-1} E(C_i^2 (B_{t_{i+1}} - B_{t_i})^2 | \mathcal{F}_a) \\ &\quad + 2 \sum_{0 \leq i < j \leq k-1} E(C_i (B_{t_{i+1}} - B_{t_i}) C_j (B_{t_{j+1}} - B_{t_j}) | \mathcal{F}_a). \end{aligned}$$

Per $i < j$ si ha $t_{i+1} \leq t_j$. Prendendo la speranza condizionale rispetto a \mathcal{F}_{t_j} e osservando che C_i, C_j e $(B_{t_{i+1}} - B_{t_i})$ sono \mathcal{F}_{t_j} -misurabili, si ottiene

$$\begin{aligned} &E(C_i (B_{t_{i+1}} - B_{t_i}) C_j (B_{t_{j+1}} - B_{t_j}) | \mathcal{F}_a) \\ &= E(E(C_i (B_{t_{i+1}} - B_{t_i}) C_j (B_{t_{j+1}} - B_{t_j}) | \mathcal{F}_{t_j}) | \mathcal{F}_a) \\ &= E(C_i (B_{t_{i+1}} - B_{t_i}) C_j E((B_{t_{j+1}} - B_{t_j}) | \mathcal{F}_{t_j}) | \mathcal{F}_a) = 0, \end{aligned}$$

essendo $E((B_{t_{j+1}} - B_{t_j}) | \mathcal{F}_{t_j}) = E((B_{t_{j+1}} - B_{t_j})) = 0$, perché $(B_{t_{j+1}} - B_{t_j})$ è indipendente da \mathcal{F}_{t_j} . Analogamente, $E((B_{t_{i+1}} - B_{t_i})^2 | \mathcal{F}_{t_i}) = E(B_{t_{i+1}} - B_{t_i})^2 = (t_{i+1} - t_i)$, da cui

$$\begin{aligned} E(C_i^2 (B_{t_{i+1}} - B_{t_i})^2 | \mathcal{F}_a) &= E(E(C_i^2 (B_{t_{i+1}} - B_{t_i})^2 | \mathcal{F}_{t_i}) | \mathcal{F}_a) \\ &= E(C_i^2 E((B_{t_{i+1}} - B_{t_i})^2 | \mathcal{F}_{t_i}) | \mathcal{F}_a) = E(C_i^2 | \mathcal{F}_a) (t_{i+1} - t_i). \end{aligned}$$

Questo mostra che

$$\mathbb{E} \left(J_{a,b}(X)^2 \mid \mathcal{F}_a \right) := \sum_{i=0}^{k-1} \mathbb{E}(C_i^2 \mid \mathcal{F}_a) (t_{i+1} - t_i) = \mathbb{E} \left(\int_a^b X_t^2 dt \mid \mathcal{F}_a \right),$$

che coincide con la seconda relazione in (5.10).

Infine, le relazioni in (5.11) sono una conseguenza immediata di quelle in (5.10): basta prendere il valore atteso di entrambi i membri. \square

5.2.7. IL RUMORE BIANCO. In ingegneria e in fisica si fa spesso uso del *rumore bianco* (*white noise* nella letteratura anglofona), intendendo con ciò un processo gaussiano reale $N = \{N_s\}_{s \geq 0}$ con media nulla e covarianza $\text{Cov}(N_s, N_t) = \delta(s - t)$, dove $\delta(\cdot)$ indica la delta di Dirac, cioè la “funzione” reale tale che $\int_A \delta(s) ds = 1$ e $\int_{A^c} \delta(s) ds = 0$, per ogni insieme aperto A che contiene 0.

È ben noto che una funzione siffatta non può esistere, di conseguenza la definizione di $\{N_s\}_{s \geq 0}$ come processo gaussiano reale è mal posta.[†] Vogliamo però mostrare informalmente che l'integrale $B_t := \int_0^t N_s ds$ non è altro che il moto browniano. In effetti, se N è un processo gaussiano di media nulla, anche $B = \{B_t := \int_0^t N_s ds\}_{t \geq 0}$ lo è, in quanto funzione lineare di N . Resta solo da calcolarne la covarianza: per $s < t$

$$\begin{aligned} \text{Cov}(B_s, B_t) &= \mathbb{E} \left[\left(\int_0^s N_u du \right) \left(\int_0^t N_v dv \right) \right] = \int_0^s \int_0^t \mathbb{E}(N_u N_v) dv du \\ &= \int_0^s \int_0^t \delta(v - u) dv du = \int_0^s \mathbf{1}_{[0,t)}(u) du = \min\{s, t\}. \end{aligned}$$

Dalla formula $B_t = \int_0^t N_s ds$ segue dunque che il rumore bianco può essere pensato come la derivata prima del moto browniano (che in effetti sappiamo non esistere!): $N_s = \frac{dB_s}{ds}$, o se si preferisce $dB_s = N_s ds$.

Grazie a queste proprietà, è possibile fornire una “dimostrazione” euristica della proprietà di isometria dell'integrale stocastico. Consideriamo innanzitutto il caso di integrandi deterministici (integrale di Wiener):

$$\begin{aligned} \mathbb{E} \left[\left(\int_a^b f(s) dB_s \right) \left(\int_a^b g(t) dB_t \right) \right] &= \int_a^b \int_a^b f(s) g(t) \mathbb{E}[N_s N_t] ds dt \\ &= \int_a^b g(t) \left(\int_a^b f(s) \delta(s - t) ds \right) dt = \int_a^b f(t) g(t) dt = \langle f, g \rangle_{L^2(\mathbb{R}^+)}, \end{aligned}$$

avendo usato la proprietà $\int_a^b f(s) \delta(s - t) ds = f(t)$. Con argomenti analoghi è possibile trattare anche il caso di integrandi stocastici $X = \{X_s\}_{s \geq 0}$, $Y = \{Y_t\}_{t \geq 0}$, usando la relazione $\mathbb{E}(X_s N_s Y_t N_t) = X_s Y_t \delta(s - t)$, da cui si ricava (5.16).

[†]Si può definire N come processo a valori nelle distribuzioni, ma non esploreremo questa strada.

5.3. L'INTEGRALE STOCASTICO COME PROCESSO

Ricordiamo che è fissato uno spazio filtrato standard $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$, su cui è definito un $\{\mathcal{F}_t\}_{t \geq 0}$ -moto browniano reale $B = \{B_t\}_{t \geq 0}$. Per tutto questo paragrafo fissiamo $T > 0$ e lavoriamo con processi $X \in M^2[0, T]$.

Dato $X \in M^2[0, T]$, per ogni intervallo $[a, b] \subseteq [0, T]$ il processo ristretto $\{X_t\}_{t \in [a, b]}$ è chiaramente in $M^2[a, b]$, per cui è ben definito l'integrale $\int_a^b X_t dB_t$. In alternativa, si può considerare l'integrale $\int_0^T X_t 1_{[a, b]}(t) dB_t$ del processo $\{X_t 1_{[a, b]}(t)\}_{t \in [0, T]} \in M^2[0, T]$. Approssimando gli integrandi con processi semplici, è facile mostrare che questi due integrali in effetti coincidono (come classi di equivalenza in $L^2(\Omega, \mathcal{F}, P)$), ossia

$$\int_a^b X_t dB_t = \int_0^T X_t 1_{[a, b]}(t) dB_t, \quad \text{q.c.} \quad (5.21)$$

Notiamo che, grazie alla proprietà (5.17), è indifferente usare $1_{[a, b]}(t)$ o $1_{[a, b]}(t)$ all'interno dell'integrale. Dalla relazione (5.21) e dalla linearità dell'integrale stocastico in $M^2[0, T]$, si ricava l'abituale relazione di additività dell'integrale rispetto agli estremi di integrazione: per ogni scelta di $0 \leq a < b < c \leq T$

$$\int_a^c X_t dB_t = \int_a^b X_t dB_t + \int_b^c X_t dB_t, \quad \text{q.c.}$$

5.3.1. L'INTEGRALE STOCASTICO COME MARTINGALA CONTINUA. Dato $X \in M^2[0, T]$, definiamo il processo $I = \{I_t = I_t(X)\}_{t \in [0, T]}$ ponendo

$$I_t := J_{0, t}(X) = \int_0^t X_s dB_s. \quad (5.22)$$

Il processo I descrive l'integrale stocastico in funzione dell'estremo di integrazione. Si noti che $I_t - I_s = J_{s, t}(X) = \int_s^t X_u dB_u$.

Abbiamo già osservato che c'è una certa arbitrarietà nella definizione di $I_t(\omega)$ per ogni $\omega \in \Omega$, dal momento che l'integrale stocastico identifica una classe di equivalenza di variabili aleatorie e non una variabile aleatoria precisa. Dimostriamo ora che il processo $I = \{I_t\}_{t \in [0, T]}$ è una martingala di quadrato integrabile, con variazione quadratica esplicita; mostriamo inoltre che è possibile fissare le versioni di $I_t(\omega)$ per diversi valori di t in modo "canonico" che fa sì che le traiettorie $t \mapsto I_t(\omega)$ siano continue.

D'ora in avanti, quando avremo a che fare con il processo $I = \{I_t = \int_0^t X_u dB_u\}_{t \geq 0}$, supporremo sempre di averne fissato una versione continua.

TEOREMA 5.19. Se $X \in M^2[0, T]$, il processo $I = \{I_t\}_{t \in [0, T]}$ definito in (5.22) è una martingala di quadrato integrabile, la cui variazione quadratica è data da

$$\langle I \rangle_t = \int_0^t X_u^2 du. \quad (5.23)$$

Esiste inoltre una modificazione di I con traiettorie continue.

DIMOSTRAZIONE. Cominciamo a mostrare che il processo $I = \{I_t\}_{t \in [0, T]}$ è adattato, cioè che I_t è \mathcal{F}_t -misurabile, per ogni $t \in [0, T]$. A tal fine, per $t \in [0, T]$ fissato, sia $\{X^{(n)}\}_{n \in \mathbb{N}}$ una successione di processi semplici in $\mathcal{S}[0, t]$ che convergono in $M^2[0, t]$ verso il processo ristretto $\{X_s\}_{s \in [0, t]}$. Definendo $I_t^{(n)} := \int_0^t X_u^{(n)} dB_u$, per costruzione dell'integrale stocastico (o per la Proposizione 5.16) si ha $I_t^{(n)} \rightarrow I_t$ in $L^2(\Omega)$. Dal fatto che la σ -algebra \mathcal{F}_t contiene tutti gli eventi di misura nulla segue che il limite in $L^2(\Omega)$ di variabili aleatorie \mathcal{F}_t -misurabili è \mathcal{F}_t -misurabile.[†] Per dimostrare la \mathcal{F}_t -misurabilità di I_t basta allora mostrare che $I_t^{(n)}$ è \mathcal{F}_t -misurabile, per ogni $n \in \mathbb{N}$. Per la Definizione 5.9 di integrale stocastico di processi semplici, se $X^{(n)}$ è della forma (5.7) (con $[a, b] = [0, t]$) si ha $I_t^{(n)} = \int_0^t X_s^{(n)} dB_s = \sum_{i=0}^{k-1} C_i (B_{t_{i+1}} - B_{t_i})$, dove $0 = t_0 < t_1 < \dots < t_k = t$. Per costruzione le variabili C_j e B_{t_j} sono \mathcal{F}_{t_j} misurabili e quindi \mathcal{F}_t -misurabili, perché $t_j \in [0, t]$ per ogni $j = 0, \dots, k$, quindi anche $I_t^{(n)}$ è \mathcal{F}_t -misurabile.

Mostriamo ora che vale la relazione di martingala: $E(I_t | \mathcal{F}_s) = I_s$ q.c., per ogni $s < t$. Dato che $I_t - I_s = J_{s,t}(X) = \int_s^t X_u dB_u$, basta mostrare che

$$E(J_{s,t}(X) | \mathcal{F}_s) = 0, \quad \text{q.c.} \quad (5.24)$$

Per $0 \leq s < t \leq T$ fissati, sia $\{X^{(n)}\}_{n \in \mathbb{N}}$ una successione di processi semplici in $\mathcal{S}[s, t]$ che convergono in $M^2[s, t]$ verso il processo ristretto $\{X_u\}_{u \in [s, t]}$, in modo che $J_{s,t}(X^{(n)}) \rightarrow J_{s,t}(X)$ in $L^2(\Omega)$ (si ricordi la Proposizione 5.16). Grazie alla prima relazione in (5.10), valida per processi in $\mathcal{S}[s, t]$, sappiamo che $E(J_{s,t}(X^{(n)}) | \mathcal{F}_s) = 0$ per ogni $n \in \mathbb{N}$. Dato che la speranza condizionale è un operatore continuo in L^2 , possiamo passare al limite in questa relazione, ottenendo (5.24). Dato che $I_t = \int_0^t X_u dB_u \in L^2(\Omega)$ per ogni $t \in [0, T]$, per costruzione dell'integrale stocastico, abbiamo mostrato che il processo $I = \{I_t\}_{t \in [0, T]}$ è una martingala di quadrato integrabile.

Mostriamo ora che il processo $\langle I \rangle = \{\langle I \rangle_t\}_{t \in [0, T]}$ definito in (5.23) è effettivamente la variazione quadratica di I (si ricordi il Teorema 4.17). Omettiamo la verifica che $\langle I \rangle$ è un processo crescente, continuo, adattato e nullo al tempo zero (esercizio), limitandoci a mostrare che il processo $\{I_t^2 - \langle I \rangle_t\}_{t \in [0, T]} = \{I_t^2 - \int_0^t X_u^2 du\}_{t \in [0, T]}$ è una martingala: con una semplice manipolazione algebrica, basta mostrare che $E(I_t^2 - I_s^2 | \mathcal{F}_s) = E(\int_s^t X_u^2 du | \mathcal{F}_s)$ q.c., per ogni $0 \leq s < t \leq T$. Dato che I è una martingala, si verifica facilmente (si ricordi la relazione (4.8)) che $E(I_t^2 - I_s^2 | \mathcal{F}_s) = E((I_t - I_s)^2 | \mathcal{F}_s)$; dato che $I_t - I_s = J_{s,t}(X)$, resta solo da mostrare che

$$E(J_{s,t}(X)^2 | \mathcal{F}_s) = E\left(\int_s^t X_u^2 du \middle| \mathcal{F}_s\right). \quad (5.25)$$

Per $0 \leq s < t \leq T$ fissati, sia $\{X^{(n)}\}_{n \in \mathbb{N}}$ una successione di processi semplici in $\mathcal{S}[s, t]$ che convergono in $M^2[s, t]$ verso il processo ristretto $\{X_u\}_{u \in [s, t]}$, in modo che $J_{s,t}(X^{(n)}) \rightarrow J_{s,t}(X)$ in $L^2(\Omega)$. Grazie alla seconda relazione in (5.10), valida per processi semplici

[†]Una successione convergente in $L^2(\Omega)$ converge in probabilità, quindi ammette una sottosuccessione che converge q.c.; di conseguenza, la variabile aleatoria limite può essere scritta come limite puntuale della sottosuccessione, al di fuori di un evento di probabilità nulla.

in $\mathcal{S}[s, t]$, la relazione (5.25) è verificata rimpiazzando X con $X^{(n)}$, per ogni $n \in \mathbb{N}$. La validità della relazione (5.25) per ogni $X \in M^2[0, T]$ segue allora passando al limite, notando che per $n \rightarrow \infty$

$$J_{s,t}(X^{(n)})^2 \longrightarrow J_{s,t}(X)^2 \quad \text{e} \quad \int_s^t (X_u^{(n)})^2 du \longrightarrow \int_s^t X_u^2 du \quad \text{in } L^1(\Omega) \quad (5.26)$$

e sfruttando la continuità della speranza condizionale in L^1 .

Entrambe le relazioni in (5.26) seguono dal seguente fatto generale: se $(E, \mathcal{E}, \mathbb{P})$ è uno spazio di probabilità e $Y_n \rightarrow Y$ in $L^2(E, \mathcal{E}, \mathbb{P})$, allora $Y_n^2 \rightarrow Y^2$ in $L^1(E, \mathcal{E}, \mathbb{P})$. Infatti

$$\|Y_n^2 - Y^2\|_{L^1} = \mathbb{E}(|Y_n^2 - Y^2|) = \mathbb{E}(|Y_n - Y| |Y_n + Y|) \leq \|Y_n - Y\|_{L^2} \|Y_n + Y\|_{L^2}$$

per la disuguaglianza di Cauchy-Schwarz, e visto che $Y_n \rightarrow Y$ in L^2 si ha che $\|Y_n - Y\|_{L^2} \rightarrow 0$ e $\|Y_n + Y\|_{L^2}$ è limitato, quindi $Y_n^2 \rightarrow Y^2$ in L^1 . Scegliendo $Y_n = J_{s,t}(X^{(n)})$ e $Y = J_{s,t}(X)$ si ottiene la prima relazione in (5.26), perché per costruzione $J_{s,t}(X^{(n)}) \rightarrow J_{s,t}(X)$ in $L^2(\Omega)$. Scegliendo invece $Y_n = X^{(n)}$ e $Y = X$, viste come variabili aleatorie definite sullo spazio $(E, \mathcal{E}, \mathbb{P}) = ([s, t] \times \Omega, \mathcal{B}([s, t]) \otimes \mathcal{F}, dt \otimes P)$, si ottiene la seconda relazione, perché per costruzione si ha $X^{(n)} \rightarrow X$ in $L^2([s, t] \times \Omega)$ e

$$\begin{aligned} \left\| \int_s^t (X_u^{(n)})^2 du - \int_s^t X_u^2 du \right\|_{L^1(\Omega)} &= \mathbb{E} \left(\left| \int_s^t (X_u^{(n)})^2 du - \int_s^t X_u^2 du \right| \right) \\ &\leq \mathbb{E} \left(\int_s^t |(X_u^{(n)})^2 - X_u^2| du \right) = \|(X^{(n)})^2 - X^2\|_{L^1([s, t] \times \Omega)}. \end{aligned}$$

Mostriamo infine che esiste una modificazione del processo $I = \{I_t\}_{t \in [0, T]}$ con traiettorie continue. Sia $\{X^{(n)}\}_{n \in \mathbb{N}}$ una successione di processi semplici in $\mathcal{S}[0, T]$ che converge in $M^2[0, T]$ verso X , cioè $\|X^{(n)} - X\|_{M^2[0, T]} \rightarrow 0$ per $n \rightarrow \infty$, a meno di estrarre una sottosuccessione, possiamo supporre che $\|X^{(n)} - X\|_{M^2[0, T]} \leq \frac{1}{2} \frac{1}{n^3}$. Di conseguenza, si ha $\|X^{(n)} - X^{(n+1)}\|_{M^2[0, T]} \leq \frac{1}{n^3}$ per ogni $n \in \mathbb{N}$, per la disuguaglianza triangolare.

Definendo $I_t^{(n)} := \int_0^t X_u^{(n)} dB_u$, il processo $I^{(n)} = \{I_t^{(n)}\}_{t \in [0, T]}$ è una martingala di quadrato integrabile, per quanto visto nella prima parte della dimostrazione. Mostriamo ora che il processo $I^{(n)}$ è q.c. continuo. È facile verificare che per ogni $0 \leq c < d \leq T$ e $0 \leq t \leq T$ si ha $\int_0^t \mathbf{1}_{[c, d)}(s) dB_s = B_{d \wedge t} - B_{c \wedge t}$ (basta distinguere i tre casi $t < c$, $t \in [c, d)$ e $t > d$ e ricordare che $x \wedge y := \min\{x, y\}$). Essendo il processo $X^{(n)} \in \mathcal{S}[0, T]$ della forma (5.7) (con $[a, b] = [0, T]$), per la linearità dell'integrale stocastico si ha

$$(I_t^{(n)})(\omega) = \left(\int_0^t X_s^{(n)} dB_s \right) (\omega) = \sum_{i=0}^{k-1} C_i(\omega) (B_{t_{i+1} \wedge t}(\omega) - B_{t_i \wedge t}(\omega)).$$

Per definizione di moto browniano, esiste $C \in \mathcal{F}$ con $P(C) = 1$ tale che la traiettoria $t \mapsto B_t(\omega)$ è continua per ogni $\omega \in C$. Dato che la funzione $t \mapsto t_i \wedge t$ è continua, segue che per ogni $\omega \in C$ la traiettoria $t \mapsto (I_t^{(n)})(\omega)$ è continua. Questo mostra che il processo $I^{(n)}$ è q.c. continuo.

Dato che $I^{(n)}$ è una martingala di quadrato integrabile, il processo $(I^{(n)} - I^{(n+1)})^2$ è una submartingala, per il Lemma 4.6. Dato che tale submartingala è q.c. continua,

possiamo applicare la disuguaglianza massimale, cf. il Teorema 4.14, ottenendo

$$\begin{aligned} \mathbb{P} \left(\|I^{(n)} - I^{(n+1)}\|_\infty \geq \frac{1}{n^2} \right) &= \mathbb{P} \left(\sup_{0 \leq u \leq T} |I_u^{(n)} - I_u^{(n+1)}| \geq \frac{1}{n^2} \right) \\ &= \mathbb{P} \left(\sup_{0 \leq u \leq T} (I_u^{(n)} - I_u^{(n+1)})^2 \geq \frac{1}{n^4} \right) \leq n^4 \mathbb{E} ((I_T^{(n)} - I_T^{(n+1)})^2) \\ &= n^4 \mathbb{E} \left[\left(\int_0^T (X_s^{(n)} - X_s^{(n+1)}) dB_s \right)^2 \right] = n^4 \|X^{(n)} - X^{(n+1)}\|_{M^2[0,T]}^2 \leq \frac{1}{n^2}, \end{aligned}$$

dove l'ultima uguaglianza segue dall'isometria (5.13) dell'integrale stocastico, l'ultima disuguaglianza segue dall'ipotesi $\|X^{(n)} - X^{(n+1)}\|_{M^2[0,T]} \leq \frac{1}{n^3}$ e dove abbiamo posto come al solito $\|f\|_\infty := \sup_{0 \leq u \leq T} |f(u)|$. Dato che $\sum_{n \in \mathbb{N}} \frac{1}{n^2} < \infty$, per il lemma di Borel-Cantelli esiste un evento A con $\mathbb{P}(A) = 1$ tale che per ogni $\omega \in A$ si ha $\|I^{(n)}(\omega) - I^{(n+1)}(\omega)\|_\infty \leq \frac{1}{n^2}$ per n grande, o più precisamente per $n \geq n_0(\omega)$, con $n_0(\omega) < \infty$. Applicando la disuguaglianza triangolare, per $m \geq n \geq n_0(\omega)$ si ottiene

$$\|I^{(m)}(\omega) - I^{(n)}(\omega)\|_\infty \leq \sum_{k=n}^{m-1} \|I^{(k+1)}(\omega) - I^{(k)}(\omega)\|_\infty \leq \sum_{k=n}^{\infty} \frac{1}{k^2} \leq \frac{(cost.)}{n}.$$

Ciò significa che per ogni $\omega \in A \cap C$ la successione di funzioni continue $\{u \mapsto I_u^{(n)}(\omega)\}_{n \in \mathbb{N}}$ è di Cauchy per $\|\cdot\|_\infty$, quindi converge uniformemente per $n \rightarrow \infty$ verso una funzione continua, che indichiamo con $u \mapsto I_u^{(\infty)}(\omega)$. Ponendo $I_u^{(\infty)}(\omega) \equiv 0$ per $\omega \notin A \cap C$, abbiamo definito un processo $I^{(\infty)} = \{I_u^{(\infty)}\}_{u \in [0,T]}$ con traiettorie continue.

Resta solo da verificare che $I^{(\infty)}$ è una modificazione di I . Fissiamo $t \in [0, T]$. Per costruzione di $I^{(\infty)}$, si ha $I_t^{(n)} \rightarrow I_t^{(\infty)}$ q.c.. D'altro canto, per la Proposizione 5.16 si ha $I_t^{(n)} := J_{0,t}(X^{(n)}) \rightarrow J_{0,t}(X) =: I_t$ in $L^2(\Omega)$, perché per costruzione $X^{(n)} \rightarrow X$ in $M^2[0, T]$. Dato che una successione convergente in $L^2(\Omega)$ ha una sottosuccessione che converge q.c., i due limiti sono q.c. uguali: si ha dunque $I_t^{(\infty)} = I_t$ q.c. \square

OSSERVAZIONE 5.20. Vale la pena sottolineare un aspetto importante, messo in luce nell'ultima parte della dimostrazione del Teorema 5.19. Dato $X \in M^2[0, T]$, esiste una successione $\{X^{(n)}\}_{n \in \mathbb{N}}$ di processi semplici $M^2[0, T]$ tali che q.c.

$$\sup_{t \in [0, T]} \left| \int_0^t X_u^{(n)} dB_u - \int_0^t X_u dB_u \right| \rightarrow 0 \quad \text{per } n \rightarrow \infty,$$

dove fissiamo una versione continua del processo $\int_0^t X_u dB_u$. Più precisamente, qualunque successione di processi $\{X^{(n)}\}_{n \in \mathbb{N}}$ tale che $\|X^{(n)} - X\|_{M^2[0, T]} \leq \frac{1}{2} \frac{1}{n^3}$ ha questa proprietà. \square

5.3.2. TEMPI D'ARRESTO E LOCALITÀ.

Vediamo ora due risultati molto utili.

Mostriamo innanzitutto che la relazione $\int_0^t X_u dB_u = \int_0^T X_u 1_{[0,t)}(u) dB_u$ vale anche quando il tempo deterministico t è sostituito da un tempo d'arresto.

PROPOSIZIONE 5.21. Sia $X \in M^2[0, T]$ e sia $\{I_t = \int_0^t X_u dB_u\}_{t \geq 0}$ una versione continua dell'integrale stocastico. Per ogni tempo d'arresto τ tale che $\tau \leq T$ q.c., vale la relazione:

$$I_\tau = \int_0^\tau X_u dB_u = \int_0^T X_u 1_{[0, \tau)}(u) dB_u \quad \text{q.c.}, \quad (5.27)$$

dove $(I_\tau)(\omega) := I_{\tau(\omega)}(\omega)$, per ogni $\omega \in \Omega$.

In altri termini, considerare l'integrale stocastico I_t del processo X fino al tempo t e poi porre $t = \tau$ è la stessa cosa che fare l'integrale stocastico del processo $\{X_u 1_{[0, \tau)}(u)\}_{u \in [0, T]}$. (Si noti che la prima uguaglianza in (5.27) è solo una questione di notazioni.)

DIMOSTRAZIONE. Il processo $\{1_{[0, \tau)}(u)\}_{u \in [0, T]}$ è progressivamente misurabile, perché è continuo a destra e adattato: infatti si ha che $1_{[0, \tau)}(u) = 1_{\{\tau > u\}}(\omega)$ è \mathcal{F}_u -misurabile, perché $\{\tau > u\} = \{\tau \leq u\}^c \in \mathcal{F}_u$. Si verifica immediatamente che il prodotto di processi progressivamente misurabili è progressivamente misurabile, e dato che $1_{[0, \tau)}(u) \leq 1$ segue che $\{X_u 1_{[0, \tau)}(u)\}_{u \in [0, T]} \in M^2[0, T]$. Questo mostra che il membro di destra nella relazione (5.27) è ben definito. Chiaramente anche il membro di sinistra è ben definito q.c.: $(I_\tau)(\omega) := I_{\tau(\omega)}(\omega)$ per ogni $\omega \in \Omega$ tale che $\tau(\omega) \leq T$. Resta solo da mostrare che queste due variabili sono q.c. uguali.

Sia $\{\tau_n\}_{n \in \mathbb{N}}$ una successione di tempi d'arresto che assumono valori discreti, tali che $\tau_n \downarrow \tau$ q.c. per $n \rightarrow \infty$ (ridefinendo τ_n come $\tau_n \wedge T$, possiamo assumere che $\tau_n \leq T$). Supponiamo di aver dimostrato la relazione (5.27) per τ_n , cioè $I_{\tau_n} = \int_0^{\tau_n} X_u 1_{[0, \tau_n)}(u) dB_u$. Per $n \rightarrow \infty$ si ha $I_{\tau_n} \rightarrow I_\tau$ q.c., poiché abbiamo fissato per ipotesi una versione di I con traiettorie continue. Dato che $X_u(\omega) 1_{[0, \tau_n(\omega))}(u) \rightarrow X_u(\omega) 1_{[0, \tau(\omega))}(u)$ per q.o. (u, ω) , per convergenza dominata ($|X_u(\omega) 1_{[0, \tau_n(\omega))}(u)| \leq |X_u(\omega)|$) si ha $X 1_{[0, \tau_n)} \rightarrow X 1_{[0, \tau)}$ in $M^2[0, T]$, quindi $\int_0^{\tau_n} X_u 1_{[0, \tau_n)}(u) dB_u \rightarrow \int_0^{\tau} X_u 1_{[0, \tau)}(u) dB_u$ in $L^2(\Omega)$. Dato che una successione convergente in L^2 ha una sottosuccessione convergente q.c., le due variabili aleatorie limite I_τ e $\int_0^{\tau} X_u 1_{[0, \tau)}(u) dB_u$ sono q.c. uguali e la relazione (5.27) è dimostrata.

Sia $\{X^{(n)}\}_{n \in \mathbb{N}}$ una successione di processi semplici che converge verso X in $M^2[0, T]$. Supponiamo di aver dimostrato la relazione (5.27) con $X^{(n)}$ al posto di X , cioè $I_\tau^{(n)} = \int_0^{\tau} X_u^{(n)} 1_{[0, \tau)}(u) dB_u$, dove $I_t^{(n)} := \int_0^t X_u^{(n)} dB_u$. Se scegliamo $X^{(n)}$ in modo che $\|X^{(n)} - X\|_{M^2[0, T]} \leq \frac{1}{2} \frac{1}{n^3}$, abbiamo visto nella dimostrazione del Teorema 5.19 che per q.o. $\omega \in \Omega$ si ha la convergenza di $I_t^{(n)}(\omega)$ per $n \rightarrow \infty$ verso $I_t(\omega)$, (uniformemente) per ogni $t \in [0, T]$; scegliendo $t = \tau(\omega)$ si ha che $I_{\tau(\omega)}^{(n)}(\omega) \rightarrow I_{\tau(\omega)}(\omega)$, cioè $I_\tau^{(n)} \rightarrow I_\tau$ q.c.. Analogamente, $\|X^{(n)} 1_{[0, \tau)} - X 1_{[0, \tau)}\|_{M^2[0, T]} \leq \|X^{(n)} - X\|_{M^2[0, T]} \leq \frac{1}{2} \frac{1}{n^3}$, quindi anche $\int_0^{\tau} X_u^{(n)} 1_{[0, \tau)}(u) dB_u \rightarrow \int_0^{\tau} X_u 1_{[0, \tau)}(u) dB_u$ q.c.. Questo mostra che $I_\tau = \int_0^{\tau} X_u 1_{[0, \tau)}(u) dB_u$ q.c., cioè la relazione (5.27).

Resta infine da dimostrare che la relazione (5.27) è verificata quando X è un processo semplice e τ assume un insieme discreto di valori: in questo caso l'integrale stocastico è dato dalla formula elementare (5.8) e la validità di (5.27) si verifica facilmente con un calcolo diretto. \square

Mostriamo infine che l'integrale stocastico, pur non essendo definito puntualmente per ogni $\omega \in \Omega$, è tuttavia un operatore che agisce localmente. Dato un evento $A \in \mathcal{F}$, diciamo che una proprietà vale "per q.o. $\omega \in A$ " intendendo che esiste $N \in \mathcal{F}$ con $P(N) = 0$ tale che la proprietà vale per ogni $\omega \in A \setminus N$.

PROPOSIZIONE 5.22 (LOCALITÀ DELL'INTEGRALE STOCASTICO). Sia $A \in \mathcal{F}$ un evento e siano $X, Y \in M^2[0, T]$ processi tali che per q.o. $\omega \in A$ si abbia $X_u(\omega) = Y_u(\omega)$

per *Leb*-q.o. $u \in [0, T]$. Allora gli integrali stocastici di X e Y coincidono q.c. su A :

$$\text{per q.o. } \omega \in A \text{ si ha } \left(\int_0^t X_u dB_u \right) (\omega) = \left(\int_0^t Y_u dB_u \right) (\omega), \quad \forall t \in [0, T]. \quad (5.28)$$

Sottolineiamo che affinché l'uguaglianza (5.28) valga *per ogni* $t \in [0, T]$ occorre scegliere le versioni continue dei processi $\{\int_0^t X_u dB_u\}_{t \geq 0}$ e $\{\int_0^t Y_u dB_u\}_{t \geq 0}$ (in caso contrario, l'uguaglianza vale solo per *Leb*-q.o. $t \in [0, T]$).

DIMOSTRAZIONE. Introduciamo i processi semplici $X^{(n)} := \hat{P}_n X$ e $Y^{(n)} := \hat{P}_n Y$ come in (5.18). Essendo definiti puntualmente per ogni ω , segue dalle ipotesi che per q.o. $\omega \in A$ si ha $X_u^{(n)}(\omega) = Y_u^{(n)}(\omega)$ per ogni $u \in [0, T]$ e quindi $\int_0^t X_u^{(n)}(\omega) dB_u(\omega) = \int_0^t Y_u^{(n)}(\omega) dB_u(\omega)$ per ogni $t \in [0, T]$, direttamente dalla definizione (5.8) di integrale stocastico di processi semplici.

Per il Lemma 5.18, $X^{(n)} \rightarrow X$ e $Y^{(n)} \rightarrow Y$ in $M^2[0, T]$, per cui possiamo estrarre sottosuccessioni $\tilde{X}^{(n)}$ e $\tilde{Y}^{(n)}$ tali che $\|\tilde{X}^{(n)} - X\|_{M^2[0, T]} \leq \frac{1}{n^3}$ e $\|\tilde{Y}^{(n)} - Y\|_{M^2[0, T]} \leq \frac{1}{n^3}$. Segue allora dall'Osservazione 5.20 che per q.o. $\omega \in \Omega$ si ha

$$\left(\int_0^t \tilde{X}_u^{(n)} dB_u \right) (\omega) \rightarrow \left(\int_0^t X_u dB_u \right) (\omega), \quad \left(\int_0^t \tilde{Y}_u^{(n)} dB_u \right) (\omega) \rightarrow \left(\int_0^t Y_u dB_u \right) (\omega),$$

uniformemente in $t \in [0, T]$. In particolare, per q.o. $\omega \in A$ si ha $(\int_0^t X_u dB_u)(\omega) = (\int_0^t Y_u dB_u)(\omega)$ per ogni $t \in [0, T]$. \square

5.4. L'INTEGRALE STOCASTICO IN $M_{\text{loc}}^2[a, b]$

Ricordiamo che è fissato uno spazio filtrato standard $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$, su cui è definito un $\{\mathcal{F}_t\}_{t \geq 0}$ -moto browniano reale $B = \{B_t\}_{t \geq 0}$.

Abbiamo definito l'integrale stocastico $\int_0^t X_u dB_u$ per processi progressivamente misurabili tali che $E(\int_0^t X_u^2 du) < \infty$. Mostriamo che questa ipotesi si può rilassare, richiedendo solo che $N_X(\omega) = \int_0^t X_u(\omega)^2 du < \infty$ per q.o. $\omega \in \Omega^\dagger$ (si ricordi l'Osservazione 5.6).

5.4.1. LO SPAZIO $M_{\text{loc}}^2[a, b]$. Siano $0 \leq a < b < \infty$ numeri reali fissati.

DEFINIZIONE 5.23. Indichiamo con $M_{\text{loc}}^2[a, b]$ lo spazio vettoriale dei processi $X = \{X_t\}_{t \in [a, b]}$ progressivamente misurabili tali che $\int_a^b X_t^2(\omega) dt < \infty$ per q.o. $\omega \in \Omega$.

OSSERVAZIONE 5.24. Ogni processo $X = \{X_t\}_{t \in [a, b]}$ adattato e q.c. continuo appartiene allo spazio $M_{\text{loc}}^2[a, b]$. Infatti il processo X è progressivamente misurabile per il Lemma 3.13 (ricordiamo che lo spazio di probabilità e la filtrazione sono completi per ipotesi). Inoltre, per q.o. $\omega \in \Omega$ la funzione $t \mapsto X_t(\omega)$ è continua sull'intervallo compatto $[a, b]$, dunque è limitata; in particolare, $\int_a^b X_t^2(\omega) dt < \infty$. \square

[†]Quest'ultima è sostanzialmente la condizione più debole per poter definire l'integrale stocastico $\int_0^t X_u dB_u$: si veda ad esempio il Problema 4.11 nel capitolo 3 in [Karatzas e Shreve, 1998].

Fissiamo ora $[0, T] \subseteq [0, \infty)$ ed estendiamo la definizione dell'integrale stocastico a processi in $M_{\text{loc}}^2[0, T]$, mediante una procedura detta *localizzazione*. Risulta conveniente costruire da subito l'intero processo $I_t = \int_0^t X_u dB_u$, per ogni $t \in [0, T]$.

Fissato un processo $X = \{X_t\}_{t \in [0, T]}$ in $M_{\text{loc}}^2[0, T]$, introduciamo per $n \in \mathbb{N}$ la variabile aleatoria τ_n definita da

$$\tau_n := \inf \left\{ t \in [0, T] : \int_0^t X_u^2 du > n \right\}, \quad (5.29)$$

con la convenzione $\inf \emptyset := \infty$. La variabile τ_n è un tempo d'arresto per il Lemma 3.23: infatti il processo $\{\int_0^t X_u^2 du\}_{t \in [0, T]}$ è progressivamente misurabile, in quanto q.c. continuo e adattato (perché?). Osserviamo che $\{\tau_n\}_{n \in \mathbb{N}}$ è una successione crescente di variabili aleatorie a valori in $[0, T] \cup \{\infty\}$ e vale l'uguaglianza

$$A_n := \{\tau_n = \infty\} = \left\{ \int_0^T X_u^2 du \leq n \right\}. \quad (5.30)$$

Questo mostra che la successione di eventi $\{A_n\}_{n \in \mathbb{N}}$ è crescente e si ha

$$\mathbb{P} \left(\bigcup_{n \in \mathbb{N}} A_n \right) = \mathbb{P} \left(\int_0^T X_u^2 du < \infty \right) = 1, \quad (5.31)$$

perché per ipotesi $X \in M_{\text{loc}}^2[0, T]$.

Definiamo per ogni $n \in \mathbb{N}$ un processo $X^{(n)} = \{X_u^{(n)}\}_{u \in [0, T]}$ ponendo

$$X_u^{(n)} := X_u \mathbf{1}_{[0, \tau_n)}(u), \quad \text{ossia} \quad X_u^{(n)}(\omega) := X_u(\omega) \mathbf{1}_{[0, \tau_n(\omega))}(u).$$

Dato che q.c. la funzione $t \mapsto \int_0^t X_u^2 du$ è continua, segue dalla definizione di τ_n che q.c.

$$\int_0^T (X_u^{(n)})^2 du = \int_0^T X_u^2 \mathbf{1}_{[0, \tau_n)}(u) du = \int_0^{\tau_n \wedge T} X_u^2 du \leq n.$$

Di conseguenza $\mathbb{E}(\int_0^T (X_u^{(n)})^2 du) \leq n < \infty$. Essendo $X^{(n)}$ progressivamente misurabile (esercizio), si ha dunque $X^{(n)} \in M^2[0, T]$, per ogni $n \in \mathbb{N}$. È dunque ben definito il processo $I^{(n)} = \{I_t^{(n)}\}_{t \in [0, T]}$ dato da

$$I_t^{(n)} := \int_0^t X_u^{(n)} dB_u = \int_0^t X_u \mathbf{1}_{[0, \tau_n)}(u) dB_u, \quad (5.32)$$

e inoltre $I^{(n)}$ è una martingala di quadrato integrabile. Grazie al Teorema 5.19, possiamo scegliere una versione continua di $I^{(n)}$, e così facciamo. Vogliamo ora mostrare che per $n \rightarrow \infty$ il processo $I^{(n)}$ converge verso un processo limite, che sarà per definizione l'integrale stocastico di X .

Fissiamo $m \in \mathbb{N}$. L'osservazione fondamentale è che sull'evento $A_m := \{\tau_m = \infty\}$ si ha q.c. $I_t^{(n)} = I_t^{(m)}$ per ogni $n \geq m$ e per ogni $t \in [0, T]$. La spiegazione intuitiva è

molto semplice: se $\tau_m(\omega) = \infty$, a maggior ragione $\tau_n(\omega) = \infty$ per $n \geq m$; nella relazione (5.32) si ha allora $1_{[0, \tau_n(\omega))}(u) = 1_{[0, \infty)}(u) \equiv 1$ per ogni $u \in [0, t]$, dunque $I_t^{(n)}$ non dipende da $n \geq m$. Formalizziamo questo argomento: i processi $\{X_u 1_{[0, \tau_n)}(u)\}_{u \in [0, T]}$ e $\{X_u 1_{[0, \tau_m)}(u)\}_{u \in [0, T]}$ coincidono sull'evento A_m , perché per $\omega \in A_m$ si ha $\tau_n(\omega) = \tau_m(\omega) = \infty$, di conseguenza per la Proposizione 5.22 esiste $N_m \in \mathcal{F}$ con $P(N_m) = 0$ tale che per ogni $\omega \in A_m \setminus N_m$ si ha $I_t^{(n)}(\omega) = I_t^{(m)}(\omega)$ per ogni $t \in [0, T]$ e per ogni $n \geq m$. In altri termini, per $\omega \in A_m \setminus N_m$ la funzione $\{I_t^{(n)}(\omega)\}_{t \in [0, T]}$ non dipende da $n \geq m$, quindi esiste il limite

$$I_t(\omega) := \lim_{n \rightarrow \infty} I_t^{(n)}(\omega), \quad \forall t \in [0, T]. \quad (5.33)$$

Dato che $m \in \mathbb{N}$ è arbitrario, questo limite esiste per ogni $\omega \in A := \bigcup_{m \in \mathbb{N}} (A_m \setminus N_m)$. Osservando che $A \supseteq (\bigcup_{m \in \mathbb{N}} A_m) \setminus (\bigcup_{m \in \mathbb{N}} N_m)$, possiamo scrivere

$$P(A) \geq P\left(\bigcup_{m \in \mathbb{N}} A_m\right) - P\left(\bigcup_{m \in \mathbb{N}} N_m\right) = P\left(\bigcup_{m \in \mathbb{N}} A_m\right) = 1,$$

grazie alla relazione (5.31). In definitiva, possiamo dare la seguente

DEFINIZIONE 5.25. Per ogni $X \in M_{\text{loc}}^2[0, T]$, si definisce *integrale stocastico di X rispetto a B* il processo $I = \{I_t =: \int_0^t X_u dB_u\}_{t \in [0, T]}$ definito dal limite in (5.33) per ogni $\omega \in A$ (dove $P(A) = 1$); per $\omega \notin A$ poniamo $I_t(\omega) \equiv 0$.

OSSERVAZIONE 5.26. Dato $X \in M_{\text{loc}}^2[0, T]$ e $n > 0$, dalla costruzione data segue che per q.o. $\omega \in \{\tau_n = \infty\} = \{\int_0^T X_s^2 ds \leq n\}$ si ha $\{I_t(\omega)\}_{t \in [0, T]} = \{I_t^{(n)}(\omega)\}_{t \in [0, T]}$, ossia

$$\left(\int_0^t X_u dB_u\right)(\omega) = \left(\int_0^t X_u 1_{[0, \tau_n)}(u) dB_u\right)(\omega), \quad \forall t \in [0, T].$$

In particolare, per q.o. $\omega \in \Omega$ esiste $n_0 = n_0(\omega) < \infty$ tale che $I_t(\omega) = I_t^{(n_0)}(\omega)$ per ogni $t \in [0, T]$. Di conseguenza, anche per $X \in M_{\text{loc}}^2[0, T]$ l'integrale stocastico $I = \{I_t\}_{t \in [0, T]}$ ha traiettorie continue, perché $\{X_t 1_{[0, \tau_n)}(t)\}_{t \in [0, T]} \in M^2[0, T]$ e dunque abbiamo potuto scegliere versioni continue dei processi $\{I_t^{(n)}(\omega)\}_{t \in [0, T]}$. \square

OSSERVAZIONE 5.27. Si noti che la definizione (5.33) è effettivamente un'estensione dell'integrale stocastico, cioè se $X \in M^2[0, T]$ la variabile I_t definita in questo modo coincide con l'integrale stocastico definito in precedenza. Abbiamo già notato che se $X \in M_{\text{loc}}^2[0, T]$ (a maggior ragione se $X \in M^2[0, T]$) si ha $\lim_{n \rightarrow \infty} X_u(\omega) 1_{[0, \tau_n(\omega))}(u) = X_u(\omega)$ per q.o. $\omega \in \Omega$ e per ogni $u \in [0, T]$. Di conseguenza, se $X \in M^2[0, T]$, per convergenza dominata ($|X_u(\omega) 1_{[0, \tau_n(\omega))}(u)| \leq |X_u(\omega)|$) segue che $X 1_{[0, \tau_n)} \rightarrow X$ in $M^2[0, T]$. Per la Proposizione 5.16, la variabile $I_t^{(n)} = \int_0^t X_u^{(n)} 1_{[0, \tau_n)}(u) du$ converge in $L^2(\Omega)$ per $n \rightarrow \infty$ verso l'integrale stocastico $\int_0^t X_u du$, che risulta dunque q.c. uguale alla variabile I_t definita in (5.33). \square

5.4.2. PRIME PROPRIETÀ. Dato $X \in M_{\text{loc}}^2[0, T]$, abbiamo definito l'integrale stocastico $I = \{I_t = \int_0^t X_u dB_u\}_{t \in [0, T]}$ e abbiamo visto che è un processo con traiettorie continue. È immediato verificare che l'integrale stocastico è un operatore lineare su $M_{\text{loc}}^2[0, T]$.

Purtroppo molte proprietà possedute dall'integrale stocastico per processi in $M^2[0, T]$ vengono perse. Per esempio, la variabile aleatoria $\int_0^t X_u dB_u$ in generale *non è integrabile*, a maggior ragione non è in $L^2(\Omega)$.

Non ha molto senso parlare di isometria per l'integrale stocastico in $M_{\text{loc}}^2[0, T]$, in quanto su $M_{\text{loc}}^2[0, T]$ non c'è una struttura naturale di spazio metrico. Ci si può tuttavia chiedere se valga qualche forma di continuità analoga alla Proposizione 5.16. La risposta è affermativa a patto di sostituire la convergenza in $M^2[0, T]$ e $L^2(\Omega)$ con la convergenza in probabilità, come mostra il seguente risultato.

PROPOSIZIONE 5.28. Siano $\{X^{(n)}\}_{n \in \mathbb{N}}$, X processi in $M_{\text{loc}}^2[0, T]$ con la proprietà che $\int_0^T |X_u^{(n)} - X_u|^2 du \rightarrow 0$ in probabilità. Allora anche $\int_0^T X_u^{(n)} dB_u \rightarrow \int_0^T X_u dB_u$ in probabilità.

Prima della dimostrazione, enunciamo un'utile disuguaglianza.

PROPOSIZIONE 5.29. Per ogni $X \in M_{\text{loc}}^2[0, T]$ e per ogni $\varepsilon, \delta > 0$ si ha che

$$\mathbb{P} \left(\sup_{t \in [0, T]} \left| \int_0^t X_s dB_s \right| > \varepsilon \right) \leq \frac{\delta}{\varepsilon^2} + \mathbb{P} \left(\int_0^T X_s^2 ds > \delta \right).$$

DIMOSTRAZIONE. Introduciamo il tempo d'arresto $\tau := \inf\{t \in [0, T] : \int_0^t X_s^2 ds > \delta\}$ e poniamo $I_t := \int_0^t X_s dB_s$ e $\tilde{I}_t := \int_0^t X_s \mathbf{1}_{[0, \tau)}(s) dB_s$. Grazie all'Osservazione 5.26, per q.o. $\omega \in \{\tau = \infty\} = \{\int_0^T X_s^2 ds \leq \delta\}$ si ha $I_t(\omega) = \tilde{I}_t(\omega)$ per ogni $t \in [0, T]$, per cui

$$\begin{aligned} \mathbb{P} \left(\sup_{t \in [0, T]} |I_t| > \varepsilon \right) &\leq \mathbb{P} \left(\sup_{t \in [0, T]} |I_t| > \varepsilon, \tau = \infty \right) + \mathbb{P}(\tau < \infty) \\ &\leq \mathbb{P} \left(\sup_{t \in [0, T]} |\tilde{I}_t| > \varepsilon \right) + \mathbb{P} \left(\int_0^T X_s^2 ds > \delta \right). \end{aligned}$$

Si noti che $\tilde{I}_t = \int_0^t X_s \mathbf{1}_{[0, \tau)}(s) dB_s$ è una martingala continua, perché $\{X_s \mathbf{1}_{[0, \tau)}(s)\}_{s \in [0, T]} \in M^2[0, T]$ (infatti, per definizione di τ , $\int_0^T (X_s \mathbf{1}_{[0, \tau)}(s))^2 ds = \int_0^{\tau \wedge T} X_s^2 ds \leq \delta$). Applicando la disuguaglianza massimale alla submartingala continua \tilde{I}_t^2 e l'isometria dell'integrale stocastico in $M^2[0, T]$, si ottiene

$$\mathbb{P} \left(\sup_{t \in [0, T]} |\tilde{I}_t| > \varepsilon \right) \leq \frac{1}{\varepsilon^2} \mathbb{E}((\tilde{I}_T)^2) = \frac{1}{\varepsilon^2} \mathbb{E} \left(\int_0^T X_s^2 \mathbf{1}_{[0, \tau)}(s) ds \right) = \frac{1}{\varepsilon^2} \mathbb{E} \left(\int_0^{\tau \wedge T} X_s^2 ds \right) \leq \frac{\delta}{\varepsilon^2}. \quad \square$$

DIMOSTRAZIONE DELLA PROPOSIZIONE 5.28. Per ogni $\varepsilon > 0$, applicando la Proposizione 5.29 al processo $X^{(n)} - X$ con $\delta = \varepsilon^3/2$, si ha

$$\mathbb{P} \left(\left| \int_0^T (X_s^{(n)} - X_s) dB_s \right| > \varepsilon \right) \leq \frac{\varepsilon}{2} + \mathbb{P} \left(\int_0^T (X_s^{(n)} - X_s)^2 ds > \frac{\varepsilon^3}{2} \right).$$

Per ipotesi l'ultimo termine tende a zero per $n \rightarrow \infty$, quindi è minore di $\varepsilon/2$ per n grande. \square

Restano anche validi i risultati descritti in §5.3.2, che riformuliamo per $X \in M_{\text{loc}}^2[0, T]$ nelle proposizioni seguenti (che non dimostriamo).

PROPOSIZIONE 5.30. Se $X \in M_{\text{loc}}^2[0, T]$ e τ è un tempo d'arresto tale che $\tau \leq T$ q.c., vale la relazione

$$I_\tau =: \int_0^\tau X_u dB_u = \int_0^T X_u \mathbf{1}_{[0, \tau)}(u) dB_u \quad \text{q.c.}$$

DIMOSTRAZIONE. Poniamo $\tau_n := \inf\{t \in [0, T] : \int_0^t X_u^2 du > n\}$, $X_t^{(n)} := X_t \mathbf{1}_{[0, \tau_n)}(t)$ e $I_t^{(n)} := \int_0^t X_u^{(n)} dB_u$. Ricordando la costruzione dell'integrale stocastico per $X \in M_{\text{loc}}^2[0, T]$, per q.o. $\omega \in \Omega$ si ha $\sup_{t \in [0, T]} |I_t^{(n)}(\omega) - I_t(\omega)| \rightarrow 0$ per $n \rightarrow \infty$; in particolare, $I_{\tau(\omega)}^{(n)}(\omega) \rightarrow I_{\tau(\omega)}(\omega)$. Ricordando che $X^{(n)} \in M^2[0, T]$ per ogni $n \in \mathbb{N}$, grazie alla Proposizione 5.21 applicata a $I^{(n)}$ possiamo scrivere

$$I_\tau = \lim_{n \rightarrow \infty} I_\tau^{(n)} = \lim_{n \rightarrow \infty} \int_0^T X_u^{(n)} \mathbf{1}_{[0, \tau)}(u) dB_u \quad \text{q.c.}$$

Se mostriamo che $\int_0^T |X_u^{(n)} \mathbf{1}_{[0, \tau)}(u) - X_u \mathbf{1}_{[0, \tau)}(u)|^2 du \rightarrow 0$ in probabilità per $n \rightarrow \infty$, applicando la Proposizione 5.28 otteniamo che

$$I_\tau = \lim_{n \rightarrow \infty} \int_0^T X_u^{(n)} \mathbf{1}_{[0, \tau)}(u) dB_u = \int_0^T X_u \mathbf{1}_{[0, \tau)}(u) dB_u \quad \text{q.c.},$$

completando la dimostrazione. Se fissiamo $\omega \in \Omega$ tale che $\tau_n(\omega) \rightarrow +\infty$ e $\int_0^T |X_u(\omega)|^2 du < \infty$, per tali ω si ha $X_u^{(n)} \mathbf{1}_{[0, \tau)}(u) = X_u \mathbf{1}_{[0, \tau_n)}(u) \mathbf{1}_{[0, \tau)}(u) \rightarrow X_u \mathbf{1}_{[0, \tau)}(u)$ per ogni $u \in [0, T]$; dato che

$$|X_u^{(n)} \mathbf{1}_{[0, \tau)}(u) - X_u \mathbf{1}_{[0, \tau)}(u)| = |X_u| \mathbf{1}_{[0, \tau)}(u) |\mathbf{1}_{[0, \tau_n)}(u) - 1| \leq |X_u|,$$

per convergenza dominata si ottiene $\int_0^T |X_u^{(n)} \mathbf{1}_{[0, \tau)}(u) - X_u \mathbf{1}_{[0, \tau)}(u)|^2 du \rightarrow 0$. Dato che $\tau_n \rightarrow \infty$ e $\int_0^T |X_u|^2 du < \infty$ q.c., abbiamo mostrato che $\int_0^T |X_u^{(n)} \mathbf{1}_{[0, \tau)}(u) - X_u \mathbf{1}_{[0, \tau)}(u)|^2 du \rightarrow 0$ q.c. \square

PROPOSIZIONE 5.31 (LOCALITÀ DELL'INTEGRALE STOCASTICO). Sia $A \in \mathcal{F}$ un evento e siano $X, Y \in M_{\text{loc}}^2[0, T]$ processi tali che per q.o. $\omega \in A$ si abbia $X_u(\omega) = Y_u(\omega)$ per *Leb*-q.o. $u \in [0, T]$. Allora i rispettivi integrali stocastici coincidono q.c. su A :

$$\text{per q.o. } \omega \in A \text{ si ha } \left(\int_0^t X_u dB_u \right)(\omega) = \left(\int_0^t Y_u dB_u \right)(\omega), \quad \forall t \in [0, T]. \quad (5.34)$$

DIMOSTRAZIONE. Per costruzione dell'integrale stocastico in $M_{\text{loc}}^2[0, T]$, per q.o. $\omega \in \Omega$ si ha

$$\int_0^t X_u dB_u = \lim_{n \rightarrow \infty} \int_0^t X_u^{(n)} dB_u, \quad \text{e} \quad \int_0^t Y_u dB_u = \lim_{n \rightarrow \infty} \int_0^t Y_u^{(n)} dB_u \quad \text{uniformemente in } t \in [0, T], \quad (5.35)$$

dove $X_u^{(n)} := X_u \mathbf{1}_{[0, \tau_n^X)}(u)$, $\tau_n^X := \inf\{t \in [0, T] : \int_0^t X_u^2 du > n\}$ e analogamente $Y_u^{(n)} := Y_u \mathbf{1}_{[0, \tau_n^Y)}(u)$, $\tau_n^Y := \inf\{t \in [0, T] : \int_0^t Y_u^2 du > n\}$. Per ipotesi, per q.o. $\omega \in A$ si ha $X_u(\omega) = Y_u(\omega)$ per *Leb*-q.o. $u \in [0, T]$, quindi $\tau_n^X(\omega) = \tau_n^Y(\omega)$ e di conseguenza $X_u^{(n)}(\omega) = Y_u^{(n)}(\omega)$ per *Leb*-q.o. $u \in [0, T]$. Segue allora dalla Proposizione 5.22 che per q.o. $\omega \in A$ si ha $(\int_0^t X_u^{(n)} dB_u)(\omega) = (\int_0^t Y_u^{(n)} dB_u)(\omega)$ per ogni $t \in [0, T]$. Ricordando la relazione (5.35), la dimostrazione è conclusa. \square

Mostriamo infine che per integrandi continui l'integrale stocastico si ottiene come limite delle somme di Riemann.

PROPOSIZIONE 5.32. Sia $X = \{X_t\}_{t \in [0, T]}$ un processo adattato e q.c. continuo (di conseguenza $X \in M_{\text{loc}}^2[0, T]$). Per ogni successione di partizioni $\pi^{(n)} = \{0 =: t_0^{(n)} < t_1^{(n)} < \dots < t_{k_n}^{(n)} := T\}$ di passo tendente a zero si ha:

$$\sum_{i=0}^{k_n-1} X_{t_i^{(n)}} (B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}}) \xrightarrow{n \rightarrow \infty} \int_0^T X_u dB_u \quad \text{in probabilità.} \quad (5.36)$$

DIMOSTRAZIONE. Notiamo che il membro sinistro in (5.36) coincide con $\int_0^T X_u^{(n)} dB_u$, dove poniamo $X_u^{(n)} := \sum_{i=0}^{k_n-1} X_{t_i^{(n)}} 1_{[t_i^{(n)}, t_{i+1}^{(n)}]}(u)$. La funzione $u \mapsto X_u$ è q.c. continua su $[0, T]$, quindi uniformemente continua, quindi $\sup_{u \in [0, T]} |X_u - X_u^{(n)}| \rightarrow 0$ q.c. per $n \rightarrow \infty$. Di conseguenza anche $\int_0^T |X_u^{(n)} - X_u|^2 du \rightarrow 0$ q.c. e dunque per la Proposizione 5.28 $\int_0^T X_u^{(n)} dB_u \rightarrow \int_0^T X_u dB_u$ in probabilità per $n \rightarrow \infty$. \square

5.4.3. MARTINGALE LOCALI. Abbiamo già osservato che in generale l'integrale stocastico $\int_0^t X_u dB_u$ non è una variabile integrabile per $X \in M_{\text{loc}}^2[0, T]$. Di conseguenza, il processo $I = \{I_t = \int_0^t X_u dB_u\}_{t \in [0, T]}$ in generale può non essere una martingala. Tuttavia esso è una *martingala locale*, nel senso della definizione seguente.

DEFINIZIONE 5.33. Un processo stocastico reale $M = \{M_t\}_{t \in \mathbb{T}}$, definito su uno spazio filtrato $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \mathbb{T}}, \mathbb{P})$, è detto *martingala locale* se esiste una successione di tempi d'arresto $\{\tau_n\}_{n \in \mathbb{N}}$ per cui $\lim_{n \rightarrow \infty} \tau_n = \infty$ q.c. e tali che per ogni $n \in \mathbb{N}$ il processo arrestato $M^{\tau_n} = \{M_t^{\tau_n} := M_{t \wedge \tau_n}\}_{t \in \mathbb{T}}$ sia una martingala.

PROPOSIZIONE 5.34. Per ogni processo $X \in M_{\text{loc}}^2[0, T]$, l'integrale stocastico $I = \{I_t = \int_0^t X_u dB_u\}_{t \in [0, T]}$ è una martingala locale.

DIMOSTRAZIONE. Basta scegliere i tempi d'arresto $\{\tau_n\}_{n \in \mathbb{N}}$ definiti in (5.29), per i quali si ha $\tau_n \rightarrow \infty$ q.c. per $n \rightarrow \infty$: in effetti, dalla relazione (5.30) e dal fatto che $\int_0^T X_u^2 du < \infty$ q.c. segue che per q.o. $\omega \in \Omega$ esiste $n_0(\omega) < \infty$ tale che $\tau_n(\omega) = \infty$ per ogni $n \geq n_0(\omega)$.

Notiamo che $1_{[0, t \wedge \tau_n)}(u) = 1_{[0, \tau_n)}(u) 1_{[0, t)}(u)$. Applicando la Proposizione 5.30 al tempo d'arresto $t \wedge \tau_n \leq T$ e la relazione (5.21), si ottiene

$$\begin{aligned} I_{t \wedge \tau_n} &:= \int_0^{t \wedge \tau_n} X_u dB_u = \int_0^T X_u 1_{[0, t \wedge \tau_n)}(u) dB_u \\ &= \int_0^T (X_u 1_{[0, \tau_n)}(u)) 1_{[0, t)}(u) dB_u = \int_0^t X_u 1_{[0, \tau_n)}(u) dB_u. \end{aligned}$$

Come abbiamo notato in precedenza, il processo $\{X_u 1_{[0, \tau_n)}(u)\}_{u \in [0, T]}$ è in $M^2[0, T]$, quindi $\{I_t^{\tau_n} = I_{t \wedge \tau_n}\}_{t \in [0, T]}$ è una martingala per il Teorema 5.19. \square

Segue immediatamente dalla Definizione 5.33 che una martingala è una martingala locale (basta scegliere $\tau_n \equiv \infty$). Il viceversa non è vero, per esempio perché una martingala locale non è necessariamente integrabile. Questa non è tuttavia la sola mancanza: esistono

infatti martingale locali integrabili (o anche uniformemente integrabili) che non sono martingale. Sono pertanto utili condizioni sufficienti per concludere che una martingala locale è una vera martingala, come quelle descritte nel lemma seguente.

LEMMA 5.35. Sia $M = \{M_t\}_{t \in \mathbb{T}}$ una martingala locale.

- Se esiste una variabile aleatoria integrabile Y tale che $|M_t| \leq Y$ q.c., per ogni $t \geq 0$ (in particolare se M è limitata), allora M è una martingala.
- Se $M_t \geq 0$ q.c., per ogni $t \in \mathbb{T}$, allora M è una *supermartingala*.

DIMOSTRAZIONE. Per ipotesi, esiste una successione di tempi d'arresto $\{\tau_n\}_{n \in \mathbb{N}}$, tali che $\tau_n \rightarrow \infty$ per $n \rightarrow \infty$, tali che $\{M_{t \wedge \tau_n}\}_{t \in \mathbb{T}}$ è una martingala. In particolare, per ogni $s, t \in \mathbb{T}$ con $s < t$ si ha q.c.

$$\mathbb{E}(M_{t \wedge \tau_n} | \mathcal{F}_s) = M_{s \wedge \tau_n}, \quad (5.37)$$

e inoltre $M_{t \wedge \tau_n}$ è una variabile aleatoria (integrabile) \mathcal{F}_t -misurabile, per ogni $t \in \mathbb{T}$.

Dato che q.c. $\tau_n \rightarrow \infty$ per $n \rightarrow \infty$, si ha q.c. $M_t = \lim_{n \rightarrow \infty} M_{t \wedge \tau_n}$, per ogni $t \in \mathbb{T}$. Di conseguenza, M_t è \mathcal{F}_t -misurabile per ogni $t \geq 0$ e dunque il processo M è adattato.

Se $|M_t| \leq Y$ per ogni $t \geq 0$, con Y integrabile, segue che $M_t \in L^1$ per ogni $t \geq 0$. Applicando il teorema di convergenza dominata per la speranza condizionale in (5.37), si ottiene $\mathbb{E}(M_t | \mathcal{F}_s) = M_s$ q.c., cioè M è una martingala.

Supponiamo ora che $M_t \geq 0$ q.c., per ogni $t \geq 0$. Dalla relazione (5.37) segue che $\mathbb{E}(M_{t \wedge \tau_n}) = \mathbb{E}(M_0)$, per ogni $n \in \mathbb{N}$ e $t \geq 0$, quindi per il lemma di Fatou

$$\mathbb{E}(M_t) = \mathbb{E} \left(\lim_{n \rightarrow \infty} M_{t \wedge \tau_n} \right) \leq \lim_{n \rightarrow \infty} \mathbb{E}(M_{t \wedge \tau_n}) \leq \mathbb{E}(M_0) < \infty.$$

Questo mostra che $M_t \in L^1$ per ogni $t \geq 0$. Applicando il Lemma di Fatou per la speranza condizionale in (5.37), si ha infine q.c.

$$\mathbb{E}(M_t | \mathcal{F}_s) = \mathbb{E} \left(\liminf_{n \rightarrow \infty} M_{t \wedge \tau_n} \mid \mathcal{F}_s \right) \leq \liminf_{n \rightarrow \infty} \mathbb{E}(M_{t \wedge \tau_n} | \mathcal{F}_s) = \lim_{n \rightarrow \infty} M_{s \wedge \tau_n} = M_s,$$

cioè M è una supermartingala. □

5.5. GLI SPAZI \mathcal{M}^2 E $\mathcal{M}_{\text{loc}}^2$

Finora abbiamo considerato processi $X = \{X_t\}_{t \in [0, T]}$ indicizzati da un intervallo limitato $[0, T]$. Talora risulta però utile lavorare con processi il cui insieme dei tempi è l'intera semiretta positiva. Diamo quindi le seguenti definizioni.

DEFINIZIONE 5.36. Indichiamo con \mathcal{M}^2 (risp. $\mathcal{M}_{\text{loc}}^2$) lo spazio vettoriale dei processi progressivamente misurabili $X = \{X_t\}_{t \in [0, \infty)}$ tali che per ogni $T > 0$ si ha $X = \{X_t\}_{t \in [0, T]} \in M^2[0, T]$ (risp. $X = \{X_t\}_{t \in [0, T]} \in M_{\text{loc}}^2[0, T]$).

Un processo $X = \{X_t\}_{t \geq 0}$ è dunque in \mathcal{M}^2 (risp. in $\mathcal{M}_{\text{loc}}^2$) se e solo se è progressivamente misurabile e per ogni $T > 0$ si ha $E(\int_0^T X_t^2 dt) < \infty$ (risp. $\int_0^T X_t^2 dt < \infty$ q.c.). Chiaramente si ha l'inclusione $\mathcal{M}^2 \subseteq \mathcal{M}_{\text{loc}}^2$.

Sottolineiamo che per $X \in \mathcal{M}^2$ si può avere $E(\int_0^\infty X_t^2 dt) = \infty$; analogamente, per $X \in \mathcal{M}_{\text{loc}}^2$ si può avere $\int_0^\infty X_t^2 dt = \infty$ q.c..

Se $X \in \mathcal{M}_{\text{loc}}^2$ (in particolare, se $X \in \mathcal{M}^2$), l'integrale stocastico $I_t = \int_0^t X_u dB_u$ è ben definito per ogni $t \in [0, \infty)$ e valgono i risultati visti nei paragrafi precedenti. Sottolineiamo in particolare che:

- se $X \in \mathcal{M}^2$, il processo $I = \{I_t\}_{t \geq 0}$ è una martingala di quadrato integrabile che ammette una versione continua, la cui variazione quadratica è $\langle I \rangle_t = \int_0^t X_u^2 du$;
- se $X \in \mathcal{M}_{\text{loc}}^2$, il processo $I = \{I_t\}_{t \geq 0}$ è una *martingala locale* che ammette una versione continua; il processo $\langle I \rangle_t = \int_0^t X_u^2 du$, ben definito per ogni $t \in [0, \infty)$, sarà ancora detto *variazione quadratica della martingala locale* I .

6. CALCOLO STOCASTICO E APPLICAZIONI

In questo capitolo dimostriamo la *formula di Itô*, il cuore del calcolo stocastico, e ne discutiamo alcune applicazioni. Per tutto il capitolo, supporremo di avere fissato uno spazio filtrato standard $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$, su cui è definito un $\{\mathcal{F}_t\}_{t \geq 0}$ -moto browniano reale (o vettoriale, quando specificato) $B = \{B_t\}_{t \geq 0}$.

6.1. FORMULA DI ITÔ PER IL MOTO BROWNIANO

Sia $s \mapsto b_s$ una funzione reale definita su $[0, t]$ di classe C^1 , cioè che ammette derivata prima b' continua. In particolare, b ha variazione finita e la corrispondente misura è $db_s = b'_s ds$, cioè per ogni $f : [0, t] \rightarrow \mathbb{R}$ misurabile e limitata si ha

$$\int_0^t f(s) db_s = \int_0^t f(s) b'_s ds.$$

Un caso particolare in cui questo integrale si calcola esplicitamente si ha quando l'integrando è della forma $f(s) = F'(b_s)$, con $F : \mathbb{R} \rightarrow \mathbb{R}$ funzione di classe C^1 . Infatti, grazie alla *formula di derivazione delle funzioni composte* (*chain rule* nella letteratura anglofona) si ha $F'(b_s) b'_s = \frac{d}{ds} F(b_s)$ e applicando il *teorema fondamentale del calcolo* si ottiene

$$\int_0^t F'(b_s) db_s = F(b_t) - F(b_0).$$

In particolare, scegliendo $F(x) = x^2$ si ottiene $2 \int_0^t b_s db_s = b_t^2$.

Ci si può chiedere se valga un'analoga formula per l'integrale stocastico. Consideriamo il caso semplice dell'integrale $2 \int_0^t B_s dB_s$. Secondo le regole dell'integrale ordinario dovrebbe dare B_t^2 , ma questo non è possibile: infatti sappiamo che il processo $\{\int_0^t B_s dB_s\}_{t \geq 0}$ è una martingala (nulla al tempo zero), poiché $B \in \mathcal{M}^2$, quindi si deve avere $E(2 \int_0^t B_s dB_s) = 0$, mentre invece $E(B_t^2) = t$. Per calcolare l'integrale, sia $\pi = \{0 = t_0 < t_1 < \dots < t_k = t\}$ una partizione di $[0, t]$. Osservando che $x^2 - y^2 = 2y(x - y) + (x - y)^2$, scriviamo

$$B_t^2 = \sum_{i=0}^{k-1} B_{t_{i+1}}^2 - B_{t_i}^2 = 2 \sum_{i=0}^{k-1} B_{t_i} (B_{t_{i+1}} - B_{t_i}) + \sum_{i=0}^{k-1} (B_{t_{i+1}} - B_{t_i})^2.$$

Se ora prendiamo una successione di partizioni $\pi^{(n)}$ di passo tendente a zero, il secondo termine converge in L^2 (quindi in probabilità) verso t mentre il primo termine converge

in probabilità verso l'integrale stocastico $2 \int_0^t B_s dB_s$ per la Proposizione 5.32, perché B ha traiettorie continue. Otteniamo dunque la formula

$$2 \int_0^t B_s dB_s = B_t^2 - t,$$

che contiene un termine extra rispetto all'integrale ordinario. Si noti che il valore atteso del membro destro di questa relazione è correttamente nullo.

Il caso di un integrale $\int_0^t F'(B_s) dB_s$ generale, per F di classe C^2 (cioè che ammette derivate prima e seconda continue), porta alla celebre *formula di Itô*.

TEOREMA 6.1 (FORMULA DI ITÔ). Se $F : \mathbb{R} \rightarrow \mathbb{R}$ è di classe C^2 , si ha q.c.

$$F(B_t) - F(B_0) = \int_0^t F'(B_s) dB_s + \frac{1}{2} \int_0^t F''(B_s) ds, \quad \forall t \geq 0. \quad (6.1)$$

Prima di procedere alla dimostrazione, si noti che l'integrale stocastico in (6.1) è ben posto, perché il processo $\{F'(B_s)\}_{s \geq 0}$ è (adattato e) q.c. continuo, quindi in $\mathcal{M}_{\text{loc}}^2$.

DIMOSTRAZIONE DEL TEOREMA 6.1. Cominciamo a considerare il caso in cui F'' è limitata: $C := \sup_{x \in \mathbb{R}} |F''(x)| < \infty$. Sia $\pi = \{0 = t_0 < t_1 < \dots < t_k = t\}$ una partizione di $[0, t]$. Lo sviluppo di Taylor al secondo ordine con resto di Lagrange dà $F(y) - F(x) = F'(x)(y - x) + \frac{1}{2} F''(z)(y - x)^2$, per un opportuno $z \in [x, y]$ (se $x \leq y$, altrimenti $z \in [y, x]$), per cui possiamo scrivere

$$\begin{aligned} F(B_t) - F(B_0) &= \sum_{i=0}^{k-1} (F(B_{t_{i+1}}) - F(B_{t_i})) \\ &= \sum_{i=0}^{k-1} F'(B_{t_i})(B_{t_{i+1}} - B_{t_i}) + \frac{1}{2} \sum_{i=0}^{k-1} F''(B_{s_i})(B_{t_{i+1}} - B_{t_i})^2, \end{aligned} \quad (6.2)$$

dove $s_i \in [t_i, t_{i+1}]$ (è stata usata la continuità delle traiettorie del moto browniano).

Fissiamo $t > 0$ e sia $\pi = \pi^{(n)} = \{0 = t_0^{(n)} < t_1^{(n)} < \dots < t_{k_n}^{(n)} = t\}$ una successione di partizioni con passo tendente a zero. Dato che il processo $\{F'(B_s)\}_{s \geq 0}$ è adattato e q.c. continuo, il primo termine in (6.2) converge in probabilità per $n \rightarrow \infty$ verso $\int_0^t F'(B_s) dB_s$, grazie all'approssimazione di Riemann dell'integrale stocastico in $\mathcal{M}_{\text{loc}}^2$ (cf. Proposizione 5.32).[†] Mostriamo che anche il secondo termine[‡] $X_n := \sum_{i=0}^{k-1} F''(B_{s_i})(B_{t_{i+1}} - B_{t_i})^2$ converge in probabilità per $n \rightarrow \infty$, verso la variabile aleatoria $\int_0^t F''(s) ds$. Segue allora dalla relazione (6.2) che $F(B_t) - F(B_0)$ converge in probabilità per $n \rightarrow \infty$ verso il membro destro in (6.1). Ma la variabile aleatoria $F(B_t) - F(B_0)$ non dipende da $n \in \mathbb{N}$, per cui può convergere in probabilità solo se è q.c. uguale al suo limite. Questo mostra che,

[†]Si può mostrare che la convergenza ha luogo anche in L^2 , per l'isometria dell'integrale stocastico in $M^2[0, t]$ (si usa il fatto che F'' è limitata, da cui segue che $|F'(x)| \leq a + b|x|$, con $a, b \in (0, \infty)$).

[‡]Per alleggerire la notazione, omettiamo d'ora in avanti la dipendenza da n in $k = k_n$, nei punti $t_i = t_i^{(n)}$ della partizione $\pi^{(n)}$ e in $s_i = s_i^{(n)}$.

per ogni $t > 0$ fissato, l'uguaglianza in (6.1) vale q.c.. Dato che l'intersezione numerabile di eventi quasi certi è un evento quasi certo, si ha che q.c. la relazione (6.1) vale per ogni $t \in \mathbb{Q} \cap [0, \infty)$. Infine, poiché entrambi i membri in (6.1) sono q.c. continui in t , segue che q.c. la relazione vale per ogni $t \in [0, \infty)$.

Resta solo da mostrare che, per ogni $t > 0$ fissato, $X_n := \sum_{i=0}^{k-1} F''(B_{s_i})(B_{t_{i+1}} - B_{t_i})^2$ converge in probabilità per $n \rightarrow \infty$ verso $A := \int_0^t F''(B_s) ds$. Nel seguito, per dimostrare la convergenza in probabilità di una successione di variabili aleatorie, sfrutteremo il Corollario 1.12: ci basterà dunque mostrare che per ogni sottosuccessione esiste una sotto-sottosuccessione che converge in probabilità.

Introduciamo un primo processo approssimante $Y_n := \sum_{i=0}^{k-1} F''(B_{t_i})(B_{t_{i+1}} - B_{t_i})^2$, per il quale possiamo scrivere

$$\begin{aligned} |X_n - Y_n| &\leq \sum_{i=0}^{k-1} |F''(B_{t_i}) - F''(B_{s_i})| (B_{t_{i+1}} - B_{t_i})^2 \\ &\leq \left(\sup_{r, s \in [0, t], |r-s| \leq |\pi^{(n)}|} |F''(B_r) - F''(B_s)| \right) \sum_{i=0}^{k-1} (B_{t_{i+1}} - B_{t_i})^2. \end{aligned} \quad (6.3)$$

Dato che q.c. la funzione $s \mapsto F''(B_s)$ è continua, essa è uniformemente continua su $[0, t]$, dunque il sup in (6.3) tende a zero q.c. per $n \rightarrow \infty$. Per quanto riguarda la somma, sappiamo che essa converge verso t in L^2 , quindi in probabilità. Da ciò segue facilmente che $|X_n - Y_n| \rightarrow 0$ in probabilità: infatti, per ogni sottosuccessione si può estrarre una sotto-sottosuccessione tale che la somma in (6.3) converge q.c. verso t ; lungo tale sotto-sottosuccessione si ha allora $|X_n - Y_n| \rightarrow 0$ q.c., dunque in probabilità.

Introduciamo il secondo processo approssimante $Z_n := \sum_{i=0}^{k-1} F''(B_{t_i})(t_{i+1} - t_i)$. Definendo $\Delta_i := (B_{t_{i+1}} - B_{t_i})^2 - (t_{i+1} - t_i)$ possiamo scrivere

$$(Y_n - Z_n)^2 = \left(\sum_{i=0}^{k-1} F''(B_{t_i}) \Delta_i \right)^2 = \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} F''(B_{t_i}) F''(B_{t_j}) \Delta_i \Delta_j.$$

Per $i < j$ si ha $E(F''(B_{t_i}) F''(B_{t_j}) \Delta_i \Delta_j) = E(F''(B_{t_i}) F''(B_{t_j}) \Delta_i E(\Delta_j | \mathcal{F}_{t_j})) = 0$, perché $(F''(B_{t_i}) F''(B_{t_j}) \Delta_i)$ è \mathcal{F}_{t_j} -misurabile, mentre Δ_j è indipendente da \mathcal{F}_{t_j} e ha media nulla. Un analogo discorso vale per $i > j$. Dato che $|F''(x)| \leq C < \infty$, si ha dunque

$$E[(Y_n - Z_n)^2] = \sum_{i=0}^{k-1} E[F''(B_{t_i})^2] E[\Delta_i^2] \leq C^2 \sum_{i=0}^{k-1} E[\Delta_i^2].$$

Per l'invarianza di scala $E[(B_t - B_s)^2 - (t - s)] = c(t - s)^2$, dove abbiamo posto $c := E[(Z^2 - 1)^2] \in (0, \infty)$ con $Z \sim \mathcal{N}(0, 1)$, per cui

$$E[(Y_n - Z_n)^2] \leq C^2 c \sum_{i=0}^{k-1} (t_{i+1} - t_i)^2 \leq C^2 c |\pi^{(n)}| \sum_{i=0}^{k-1} (t_{i+1} - t_i) = C^2 c t |\pi^{(n)}|.$$

Questo mostra che $|Y_n - Z_n| \rightarrow 0$ in L^2 , quindi in probabilità.

Infine, ricordando che $A = \int_0^t F''(B_s) ds$, è chiaro che per $n \rightarrow \infty$ $|Z_n - A| \rightarrow 0$ q.c., quindi in probabilità. Infatti q.c. la funzione $s \mapsto F''(B_s)$ è continua su $[0, t]$ e di conseguenza le somme di Riemann convergono verso il corrispondente integrale.

Infine, per la disuguaglianza triangolare possiamo scrivere

$$|X_n - A| \leq |X_n - Y_n| + |Y_n - Z_n| + |Z_n - A|.$$

Avendo mostrato che i tre termini nel membro destro di questa relazione convergono a zero in probabilità, segue facilmente che $|X_n - A| \rightarrow 0$ in probabilità: infatti, per ogni sottosuccessione basta scegliere una sotto-sottosuccessione $\{n_k\}_{k \in \mathbb{N}}$ lungo cui i tre termini tendono a zero q.c. e si ha che anche $|X_{n_k} - A| \rightarrow 0$ q.c., dunque in probabilità. Questo conclude la dimostrazione nell'ipotesi in cui F'' è limitata.

Il caso in cui F'' non è limitata si ottiene per approssimazione. Sia infatti $\{F_n\}_{n \in \mathbb{N}}$ una successione di funzioni di classe C^2 , con F_n'' limitata, tali che per $n \rightarrow \infty$ si abbia la convergenza di F_n , F_n' e F_n'' verso rispettivamente F , F' , F'' , uniformemente su ogni compatto.[†] Per ogni $L > 0$ e $\varepsilon > 0$ esiste dunque $n_0 = n_0(\varepsilon, L) < \infty$ tale che per ogni $n \geq n_0$ e $x \in [-L, L]$

$$|F_n(x) - F(x)| \leq \varepsilon, \quad |F_n'(x) - F'(x)| \leq \varepsilon, \quad |F_n''(x) - F''(x)| \leq \varepsilon. \quad (6.4)$$

Per ogni $t \geq 0$ fissato, la relazione (6.1) con F sostituito da F_n vale q.c., per ogni $n \in \mathbb{N}$. Ora mostriamo che esiste una sottosuccessione lungo cui ciascun termine in (6.1) contenente F_n converge q.c. verso lo stesso termine contenente F : da ciò segue che per ogni $t \geq 0$ fissato la relazione (6.1) vale q.c., senza vincoli su F'' .

Per il membro di sinistra non ci sono problemi: sappiamo che $F_n(x) \rightarrow F(x)$ per ogni $x \in \mathbb{R}$ e ponendo $x = B_t(\omega)$ si ha la convergenza q.c.. Anche il secondo termine nel membro destra di (6.1) è facile: per q.o. $\omega \in \Omega$ la funzione $s \mapsto B_s(\omega)$ è continua, quindi limitata su $[0, t]$. Esiste dunque $L = L(\omega)$ tale che $B_s(\omega) \in [-L, L]$ per ogni $s \in [0, t]$, quindi grazie a (6.4) per $n \geq n_0$ si ha $\int_0^t |F_n''(B_s(\omega)) - F''(B_s(\omega))| ds \leq \varepsilon t$. Questo mostra che $\int_0^t |F_n''(B_s(\omega)) - F''(B_s(\omega))| ds \rightarrow 0$ per $n \rightarrow \infty$. Resta infine il primo termine nel membro destra di (6.1): con analoghi argomenti, è immediato vedere che per q.o. $\omega \in \Omega$ si ha $\int_0^t |F_n'(B_s(\omega)) - F'(B_s(\omega))|^2 ds \rightarrow 0$ per $n \rightarrow \infty$, quindi $\int_0^t F_n'(B_s) dB_s$ converge verso $\int_0^t F'(B_s) dB_s$ in probabilità, per le proprietà dell'integrale stocastico in $\mathcal{M}_{\text{loc}}^2$, e dunque si ha la convergenza q.c. per una sottosuccessione. \square

OSSERVAZIONE 6.2. La formula di Itô (6.1) si scrive spesso in forma differenziale:

$$dF(B_t) = F'(B_t) dB_t + \frac{1}{2} F''(B_t) dt. \quad (6.5)$$

[†]È facile costruire una tale successione, “tagliando” i valori di $F''(x)$ più grandi di n o più piccoli di $-n$. Più precisamente, poniamo $g_n(x) := \max\{F''(x) \wedge n, -n\}$ e definiamo

$$F_n''(x) := g_n(x), \quad F_n'(x) := F'(0) + \int_0^x F_n''(y) dy, \quad F_n(x) := F(0) + \int_0^x F_n'(y) dy,$$

dove $\int_0^x \dots := -\int_x^0 \dots$ per $x < 0$. Queste definizioni sono consistenti, cioè F_n' e F_n'' sono effettivamente le derivate prima e seconda di F_n . Dato che F'' è una funzione continua, è limitata su ogni compatto: per ogni $L > 0$ esiste n_0 tale che $\max_{x \in [-L, L]} |F''(x)| \leq n_0$, da cui segue che per ogni $n \geq n_0$ si ha $g_n(x) \equiv F''(x)$ per ogni $x \in [-L, L]$, per definizione di g_n . Di conseguenza anche $F_n(x) \equiv F(x)$ per ogni $x \in [-L, L]$: quindi su ogni compatto non solo F_n converge uniformemente a F , ma addirittura coincide con F per n grande (e analogamente per F_n' e F_n'').

Sottolineiamo che si tratta solo di una notazione compatta, il cui significato è precisamente la formula di Itô (6.1). Questa relazione può essere vista come la *chain rule* (regola di derivazione di funzioni composte) per l'integrale stocastico. \square

La formula di Itô può essere vista come la versione stocastica del *teorema fondamentale del calcolo* (o anche della *chain rule*, per l'Osservazione 6.2). In effetti, essa permette di “calcolare” — o meglio, di esprimere in forma più semplice — una classe particolare di integrali stocastici. Dato che $B_0 = 0$, possiamo infatti riscrivere (6.1) come

$$\int_0^t F'(B_s) dB_s = F(B_t) - F(0) - \frac{1}{2} \int_0^t F''(B_s) ds, \quad (6.6)$$

e si noti che quello che compare nel membro di destra è un *integrale ordinario*, rispetto alla misura di Lebesgue.

Osserviamo che il membro destro in (6.6) è ben definito come integrale ordinario per ogni $\omega \in \Omega$ per cui la funzione $s \mapsto B_s(\omega)$ è continua (o anche solo misurabile). Abbiamo dunque un insieme di ω “universale” su cui sono definiti canonicamente gli integrali stocastici della forma $\int_0^t f(B_s) dB_s$, per ogni funzione f di classe C^1 e per ogni $t \geq 0$ (basta porre $F(x) := \int_0^x f(z) dz$).

In realtà, come avremo modo di apprezzare nel seguito, l'applicazione fondamentale della formula di Itô non consiste tanto nel “calcolare” gli integrali stocastici della forma $\int_0^t F'(B_s) dB_s$, cf. (6.6), quanto piuttosto nell'esprimere ogni processo $\{F(B_t)\}_{t \geq 0}$, con $F(\cdot)$ di classe C^2 , come somma di una *martingala locale* (l'integrale stocastico $\int_0^t F'(B_s) dB_s$) e di un *processo a variazione finita*[†] (l'integrale ordinario $\frac{1}{2} \int_0^t F''(B_s) ds$), cf. (6.11).

6.2. PROCESSI DI ITÔ E FORMULA DI ITÔ GENERALE

6.2.1. PROCESSI DI ITÔ. Sappiamo che per definire l'integrale stocastico $\int_0^t X_s dB_s$ per ogni $t \geq 0$ è necessario che il processo $X = \{X_s\}_{s \geq 0}$ sia in $\mathcal{M}_{\text{loc}}^2$. Per definire l'integrale ordinario $\int_0^t X_s ds$ per ogni $t \geq 0$ è sufficiente richiedere che X sia nello spazio $\mathcal{M}_{\text{loc}}^1$, lo spazio dei processi progressivamente misurabili con traiettorie localmente integrabili.

DEFINIZIONE 6.3. Indichiamo con $M_{\text{loc}}^1[0, T]$ lo spazio vettoriale dei processi $X = \{X_t\}_{t \in [a, b]}$ progressivamente misurabili tali che $\int_0^T |X_t| dt < \infty$ q.c.. Indichiamo con $\mathcal{M}_{\text{loc}}^1$ lo spazio vettoriale dei processi progressivamente misurabili $X = \{X_t\}_{t \in [0, \infty)}$ tali che per ogni $T > 0$ si ha $\{X_t\}_{t \in [0, T]} \in M_{\text{loc}}^1[0, T]$.

La formula di Itô mostra che, per ogni $F : \mathbb{R} \rightarrow \mathbb{R}$ di classe C^2 , il processo $\{F(B_t)\}_{t \geq 0}$ si scrive come somma di due processi: l'integrale stocastico $\int_0^t F'(B_s) dB_s$ e l'integrale ordinario $\frac{1}{2} \int_0^t F''(B_s) ds$. Questo motiva la prossima importante definizione.

[†]Ricordiamo che, se $g : [0, T] \rightarrow \mathbb{R}$ è una funzione integrabile, l'integrale ordinario $t \mapsto \int_0^t g(s) ds$ è una funzione a variazione finita, cf. il paragrafo 2.4.

DEFINIZIONE 6.4. Un processo stocastico reale *q.c. continuo* $X = \{X_t\}_{t \geq 0}$ è detto *processo di Itô* se esistono $\varphi = \{\varphi_t\}_{t \geq 0} \in \mathcal{M}_{\text{loc}}^2$ e $\psi = \{\psi_t\}_{t \geq 0} \in \mathcal{M}_{\text{loc}}^1$ tali che q.c.

$$X_t - X_0 = \int_0^t \varphi_s dB_s + \int_0^t \psi_s ds, \quad \forall t \geq 0. \quad (6.7)$$

Indicheremo questo fatto con la notazione differenziale $dX_t = \varphi_t dB_t + \psi_t dt$.

Come abbiamo già osservato, un'ampia classe di processi di Itô è data dai processi della forma $\{F(B_t)\}_{t \geq 0}$, qualunque sia $F: \mathbb{R} \rightarrow \mathbb{R}$ di classe C^2 .

Si noti che nella Definizione 6.4 richiediamo che X sia un processo q.c. continuo. Questa non è una restrizione: infatti se X deve soddisfare la relazione (6.7), esso ammette una versione continua, per le proprietà dell'integrale stocastico e dell'integrale ordinario.

Notiamo che se $\psi \equiv 0$ si ha $dX_t = \varphi_t dB_t$, dunque X è una *martingala locale*.[†] Questa osservazione sarà molto utile nel seguito.

OSSERVAZIONE 6.5. Un processo di Itô è per definizione un processo dato dalla somma di un integrale stocastico $I_t := \int_0^t \varphi_s dB_s$ e di un integrale ordinario $R_t := \int_0^t \psi_s ds$. È importante sottolineare che questi due processi hanno proprietà radicalmente differenti. Infatti, q.c. le traiettorie del processo R_t hanno *variazione finita* su ogni intervallo $[0, T]$. D'altro canto, sappiamo che il processo I_t è una *martingala locale*: analogamente alle martingale di quadrato integrabile, si può mostrare che q.c. le sue traiettorie hanno *variazione infinita* su ogni intervallo (escludendo il caso banale in cui siano costanti). Sfruttando queste proprietà, è possibile mostrare che la *decomposizione di un processo di Itô* X nella forma (6.7) è *unica*, nel senso che i processi $I_t := \int_0^t \varphi_s dB_s$ e $R_t := \int_0^t \psi_s ds$ sono univocamente determinati da X , a meno di indistinguibilità. Da ciò segue che i processi integrandi $\varphi = \{\varphi_s(\omega)\}_{s \geq 0}$ e $\psi = \{\psi_s(\omega)\}_{s \geq 0}$ sono univocamente determinati per P-q.o. $\omega \in \Omega$ e per Leb-q.o. $s \geq 0$. \square

6.2.2. FORMULA DI ITÔ GENERALE. Se X è un processo di Itô, $dX_s = \varphi_s dB_s + \psi_s ds$, possiamo definire l'integrale rispetto a X ponendo semplicemente

$$\int_0^t Y_s dX_s := \int_0^t Y_s \varphi_s dB_s + \int_0^t Y_s \psi_s ds, \quad (6.8)$$

per ogni processo $Y = \{Y_s\}_{s \geq 0}$ progressivamente misurabile per cui gli integrali abbiano senso, cioè tale che $\{Y_s \varphi_s\}_{s \geq 0} \in \mathcal{M}_{\text{loc}}^2$ e $\{Y_s \psi_s\}_{s \geq 0} \in \mathcal{M}_{\text{loc}}^1$. Per esempio, oltre a essere progressivamente misurabile, basta che Y abbia q.c. traiettorie localmente limitate (in particolare, basta che sia q.c. continuo).

Dato il processo di Itô X con decomposizione $dX_s = \varphi_s dB_s + \psi_s ds$, definiamo *variazione quadratica* $\langle X \rangle_t$ di X la variazione quadratica dell'integrale stocastico che

[†]Vale anche il viceversa: un processo di Itô X con $dX_t = \varphi_t dB_t + \psi_t dt$ è una martingala locale rispetto alla filtrazione $\{\mathcal{F}_t\}_{t \geq 0}$ fissata sullo spazio soltanto se $\psi_t \equiv 0$. L'enfasi sulla filtrazione è di fondamentale importanza! Si può infatti verificare che il processo $Y_t := B_t - \int_0^t \frac{B_s}{s} ds$, che ha differenziale stocastico $dY_t = dB_t - \frac{B_t}{t} dt$ (in particolare $\psi_t \neq 0$) è un moto browniano. Come ogni moto browniano, il processo Y è una martingala rispetto alla sua filtrazione naturale $\{\mathcal{G}_t = \sigma(\{Y_s\}_{s \leq t})\}_{t \geq 0}$. Il punto è che il moto browniano originale B non è un $\{\mathcal{G}_t\}_{t \geq 0}$ -moto browniano, quindi $\{\mathcal{G}_t\}_{t \geq 0}$ non può essere presa come filtrazione sullo spazio (Ω, \mathcal{F}, P) .

compare nella sua decomposizione, ossia

$$\langle X \rangle_t := \int_0^t \varphi_s^2 ds. \quad (6.9)$$

Si può dimostrare che $\langle X \rangle_t$ è il limite in probabilità della somma $\sum_{i=0}^{k-1} (X_{t_{i+1}} - X_{t_i})^2$ lungo una partizione $\pi = \{0 = t_0 < t_1 < \dots < t_k = t\}$ di $[0, t]$, quando il passo della partizione tende verso zero, ma non avremo bisogno di questo fatto. Si noti che per definizione $\langle X \rangle_t$ è un processo di Itô, il cui differenziale stocastico è dato da

$$d\langle X \rangle_t = \varphi_t^2 dt.$$

Possiamo quindi definire l'integrale rispetto a $\langle X \rangle$ ponendo

$$\int_0^t Y_s d\langle X \rangle_s := \int_0^t Y_s \varphi_s^2 ds, \quad (6.10)$$

per ogni processo Y per cui ciò abbia senso. Si noti che se X è un moto browniano, si ha $\langle X \rangle_t = t$ e di conseguenza $d\langle X \rangle_t = dt$.

Diremo che una funzione $F = F(t, x) : \mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}$ è di classe $C^{1,2}$ se è derivabile con continuità una volta in t e due volte in x , ossia se le derivate parziali $\frac{\partial F}{\partial t}(t, x)$, $\frac{\partial F}{\partial x}(t, x)$ e $\frac{\partial^2 F}{\partial x^2}(t, x)$ esistono e sono funzioni continue di $(t, x) \in \mathbb{R}^+ \times \mathbb{R}$. È prassi indicare la derivata temporale con un punto e le derivate spaziali con gli apici, ossia

$$\dot{F}(t, x) := \frac{\partial F}{\partial t}(t, x), \quad F'(t, x) := \frac{\partial F}{\partial x}(t, x), \quad F''(t, x) := \frac{\partial^2 F}{\partial x^2}(t, x).$$

Enunciamo ora (senza dimostrazione) una generalizzazione della formula di Itô.

TEOREMA 6.6 (FORMULA DI ITÔ GENERALIZZATA). Se $X = \{X_t\}_{t \geq 0}$ è un processo di Itô, con $dX_t = \varphi_t dB_t + \psi_t dt$, e $F = F(t, x) : \mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}$ è di classe $C^{1,2}$, si ha q.c. per ogni $t \geq 0$

$$\begin{aligned} F(t, X_t) - F(0, X_0) &= \int_0^t \dot{F}(s, X_s) ds + \int_0^t F'(s, X_s) dX_s \\ &\quad + \frac{1}{2} \int_0^t F''(s, X_s) d\langle X \rangle_s. \end{aligned} \quad (6.11)$$

In notazione differenziale:

$$dF(t, X_t) = \dot{F}(t, X_t) dt + F'(t, X_t) dX_t + \frac{1}{2} F''(t, X_t) d\langle X \rangle_t \quad (6.12)$$

Ricordando le relazioni (6.8) e (6.10), possiamo riscrivere il membro destro in (6.11) nel modo seguente:

$$\int_0^t F'(s, X_s) \varphi_s dB_s + \int_0^t \left(\dot{F}(s, X_s) + F'(s, X_s) \psi_s + \frac{1}{2} F''(s, X_s) \varphi_s^2 \right) ds.$$

Questo mostra che, per ogni processo di Itô $X = \{X_t\}_{t \geq 0}$ e per ogni funzione $F = F(t, x) : \mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}$ di classe $C^{1,2}$, il processo $\{F(t, X_t)\}_{t \geq 0}$ è un processo di Itô, il cui differenziale stocastico è dato da

$$dF(t, X_t) = F'(t, X_t) \varphi_t dB_t + \left(\dot{F}(t, X_t) + F'(t, X_t) \psi_t + \frac{1}{2} F''(t, X_t) \varphi_t^2 \right) dt.$$

6.3. QUALCHE ESEMPIO

6.3.1. MOTO BROWNIANO GEOMETRICO. Vogliamo ora determinare il processo di Itô $X = \{X_t\}_{t \geq 0}$ che risolve la seguente equazione differenziale stocastica:

$$\begin{cases} dX_t = b X_t dt + \sigma X_t dB_t \\ X_0 = x \end{cases}, \quad (6.13)$$

dove $b \in \mathbb{R}$, $\sigma > 0$ e $x > 0$. Procediamo euristicamente per “indovinare” la soluzione: se assumiamo che $X_t \neq 0$ per ogni t , possiamo dividere per X_t , ottenendo

$$\frac{dX_t}{X_t} = b dt + \sigma dB_t. \quad (6.14)$$

Il membro di sinistra fa pensare al differenziale di $\log X_t$. In effetti, se assumiamo che $X_t > 0$ per ogni t , dalla formula di Itô si ha $d(\log X_t) = \frac{1}{X_t} dX_t - \frac{1}{2} \frac{1}{X_t^2} d\langle X \rangle_t$.[†] Dall'equazione (6.13) è chiaro che $d\langle X \rangle_t = \sigma^2 X_t^2 dt$, per cui da (6.14) si ottiene

$$d(\log X_t) = \left(b - \frac{1}{2} \sigma^2 \right) dt + \sigma dB_t,$$

e integrando da 0 a t si ha $\log X_t - \log X_0 = (b - \frac{1}{2} \sigma^2) t + \sigma B_t$, ovvero

$$X_t = x \exp \left(\left(b - \frac{1}{2} \sigma^2 \right) t + \sigma B_t \right). \quad (6.15)$$

Questo processo è detto *moto browniano geometrico*.

Questa derivazione euristica suggerisce che, se esiste un processo positivo soluzione dell'equazione (6.13), esso è necessariamente un moto browniano geometrico. La dimostrazione rigorosa di questo fatto sarà una conseguenza dei teoremi di esistenza e unicità per equazioni differenziali stocastiche, che vedremo nel prossimo capitolo.

Mostriamo ora che effettivamente il processo X definito da (6.15) risolve l'equazione (6.13). Chiaramente $X_0 = x$, inoltre scrivendo $X_t = x e^{Y_t}$, dove $dY_t = (b - \frac{1}{2} \sigma^2) dt + \frac{1}{2} \sigma dB_t$,

[†]A priori l'applicazione della formula di Itô non è giustificata, perché il logaritmo non è definito su tutto \mathbb{R} . Tuttavia, se $X_t > 0$ per ogni t , è possibile mostrare che la formula di Itô è effettivamente valida, usando opportuni tempi d'arresto (nello spirito della dimostrazione del Lemma 6.13). In ogni caso, questa derivazione serve soltanto a “indovinare” la soluzione (6.15) dell'equazione differenziale stocastica (6.13), che verifichiamo poi essere effettivamente soluzione.

possiamo applicare la formula di Itô (6.12), ottenendo

$$\begin{aligned} dX_t &= d(x e^{Y_t}) = x e^{Y_t} dY_t + \frac{1}{2} x e^{Y_t} d\langle Y \rangle_t \\ &= X_t \left[\left(b - \frac{1}{2} \sigma^2 \right) dt + \frac{1}{2} \sigma dB_t \right] + \frac{1}{2} X_t \sigma^2 dt = b X_t dt + \sigma X_t dB_t, \end{aligned}$$

cioè l'equazione (6.13) è verificata.

6.3.2. SUPERMARTINGALA ESPONENZIALE. Dato un processo $\varphi = \{\varphi_t\}_{t \in [0, T]} \in M_{\text{loc}}^2[0, T]$, definiamo il processo $Z = \{Z_t\}_{t \in [0, T]}$ ponendo

$$Z_t := \exp \left(\int_0^t \varphi_s dB_s - \frac{1}{2} \int_0^t \varphi_s^2 ds \right). \quad (6.16)$$

Si noti che possiamo scrivere

$$Z_t = \exp(X_t), \quad \text{dove} \quad dX_t := \varphi_t dB_t - \frac{1}{2} \varphi_t^2 dt.$$

Applicando la formula di Itô (6.12) si ricava

$$dZ_t = e^{X_t} dX_t + \frac{1}{2} e^{X_t} d\langle X \rangle_t = e^{X_t} \left(\varphi_t dB_t - \frac{1}{2} \varphi_t^2 dt \right) + \frac{1}{2} e^{X_t} \varphi_t^2 dt,$$

quindi i termini a variazione finita si cancellano e si ottiene

$$dZ_t = Z_t \varphi_t dB_t. \quad (6.17)$$

Questa relazione mostra che Z è una martingala locale. Dato che $Z_t > 0$, segue dal Lemma 5.35 che Z è una supermartingala. Il processo Z è detto *supermartingala esponenziale*.

Essendo Z una supermartingala, si ha $E(Z_t) \leq E(Z_0) = 1$, per ogni $t \geq 0$. È di fondamentale importanza dare condizioni che garantiscano che $Z = \{Z_t\}_{t \in [0, T]}$ sia una vera martingala, come vedremo a proposito del Teorema di Girsanov.

Una condizione necessaria e sufficiente, benché implicita, è che $E(Z_T) = 1$ (che implica $E(Z_t) = 1$ per ogni $t \in [0, T]$: infatti $1 = E(Z_0) \geq E(Z_t) \geq E(Z_T)$ per la proprietà di supermartingala). Questo segue dal fatto generale che una supermartingala costante in media è una martingala. Infatti per la proprietà di supermartingala vale che $Z_s - E(Z_t | \mathcal{F}_s) \geq 0$ e se Z è costante in media si ha $E[Z_s - E(Z_t | \mathcal{F}_s)] = E(Z_s) - E(Z_t) = 0$, per cui la variabile $Z_s - E(Z_t | \mathcal{F}_s)$ deve essere q.c. nulla: $Z_s = E(Z_t | \mathcal{F}_s)$ q.c. e dunque Z è una martingala.

Nel Lemma 6.17 più sotto dimostreremo una condizione esplicita, ma molto restrittiva, che garantisce che $E(Z_T) = 1$. Due condizioni più deboli sono descritte nella seguente proposizione, che non dimostreremo (si vedano la Proposizione 7.23 in [Baldi, 2000] e la Proposizione 5.12 nel capitolo 3 in [Karatzas e Shreve, 1998]).

PROPOSIZIONE 6.7. Sia Z la supermartingala esponenziale definita in (6.16).

- Se $E[\exp(\frac{1}{2} \int_0^T \varphi_s^2 ds)] < \infty$, allora $E(Z_T) = 1$ (*criterio di Novikov*).
- Se esiste $a > 0$ tale che $E[\exp(a\varphi_s^2)] < \infty$, $\forall s \in [0, T]$, allora $E(Z_T) = 1$.

6.4. IL CASO MULTIDIMENSIONALE

Per quanto non ci siano novità sostanziali, è molto importante per le applicazioni estendere la teoria dell'integrazione stocastica al caso di processi vettoriali. *Supporremo dunque in questo paragrafo che $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ sia uno spazio filtrato standard, su cui è definito un $\{\mathcal{F}_t\}_{t \geq 0}$ -moto browniano $B = \{B_t = (B_t^{(1)}, \dots, B_t^{(d)})\}_{t \geq 0}$ a valori in \mathbb{R}^d .*

Definiamo $\mathcal{M}_{\text{loc}}^2(n \times d)$ come lo spazio dei processi $\varphi = \{(\varphi_t)_{ij}\}_{t \geq 0, 1 \leq i \leq n, 1 \leq j \leq d}$ tali che per ogni $1 \leq i \leq n$, $1 \leq j \leq d$ il processo $\varphi_{ij} = \{(\varphi_t)_{ij}\}_{t \geq 0}$ sia in $\mathcal{M}_{\text{loc}}^2$. Si noti che $\varphi_t = \{(\varphi_t)_{ij}\}_{1 \leq i \leq n, 1 \leq j \leq d}$ può essere vista come una matrice $n \times d$, per ogni $t \geq 0$. In modo analogo si definiscono gli spazi $\mathcal{M}^2(n \times d)$ e $\mathcal{M}_{\text{loc}}^1(n \times d)$.

Dato un processo $\varphi \in \mathcal{M}_{\text{loc}}^2(n \times d)$, è possibile definire l'integrale stocastico $\int_0^t \varphi_s \cdot dB_s$ per ogni $t \geq 0$ come il processo a valori in \mathbb{R}^n definito da

$$\left(\int_0^t \varphi_s \cdot dB_s \right)_i := \sum_{j=1}^d \int_0^t (\varphi_s)_{ij} dB_s^{(j)}, \quad \forall 1 \leq i \leq n.$$

In altre parole, il termine $\varphi_s \cdot dB_s$ va interpretato come il prodotto tra la matrice $n \times d$ -dimensionale φ_s e il vettore d -dimensionale dB_s . In analogia col caso unidimensionale, se $\varphi \in \mathcal{M}^2(n \times d)$ le componenti del processo $\int_0^t \varphi_s \cdot dB_s$ sono martingale di quadrato integrabile, mentre se $\varphi \in \mathcal{M}_{\text{loc}}^2(n \times d)$ esse sono in generale solo martingale locali.

Un processo n -dimensionale $X = \{X_t\}_{t \geq 0}$ è detto processo di Itô se esistono $\varphi \in \mathcal{M}_{\text{loc}}^2(n \times d)$ e $\psi \in \mathcal{M}_{\text{loc}}^1(n \times 1)$ tali che

$$X_t - X_0 = \int_0^t \varphi_s \cdot dB_s + \int_0^t \psi_s ds, \quad \text{cioè} \quad dX_t = \varphi_t \cdot dB_t + \psi_t dt,$$

in perfetta analogia col caso unidimensionale. In particolare, per ogni $t \geq 0$ possiamo definire l'integrale rispetto a X . Limitandoci per semplicità al caso di processi integrandi $Y = \{(Y_t)_i\}_{t \geq 0, 1 \leq i \leq n}$ a valori nelle matrici $1 \times n$, poniamo

$$\int_0^t Y_s \cdot dX_s := \int_0^t Y_s \cdot \varphi_s \cdot dB_s + \int_0^t Y_s \cdot \psi_s ds, \quad (6.18)$$

sotto l'ipotesi che $\{(Y_s \cdot \varphi_s)_i\}_{s \geq 0, 1 \leq i \leq d} \in \mathcal{M}_{\text{loc}}^2(1 \times d)$ e $\{Y_s \cdot \psi_s\}_{s \geq 0} \in \mathcal{M}_{\text{loc}}^1$, dove naturalmente $Y_s \cdot \varphi_s$ e $Y_s \cdot \psi_s$ vanno intesi come prodotti di matrici, ossia

$$(Y_s \cdot \varphi_s)_i := \sum_{k=1}^n (Y_s)_k (\varphi_s)_{ki}, \quad \text{per ogni } 1 \leq i \leq d, \quad Y_s \cdot \psi_s := \sum_{k=1}^n (Y_s)_k (\psi_s)_k.$$

Al solito, è sufficiente che ogni componente di Y sia progressivamente misurabile e abbia q.c. traiettorie localmente limitate (in particolare basta che Y sia q.c. continuo).

Infine, definiamo la *covarianza quadratica* $\langle X, X \rangle = \{\langle X^{(i)}, X^{(j)} \rangle_t\}_{t \geq 0, 1 \leq i, j \leq n}$ di un processo di Itô n -dimensionale X , con decomposizione $dX_t = \varphi_t \cdot dB_t + \psi_t dt$, come il processo a valori nelle matrici $n \times n$ definito da

$$\langle X^{(i)}, X^{(j)} \rangle_t := \int_0^t (\varphi_s \cdot \varphi_s^*)_{ij} ds = \int_0^t \sum_{k=1}^d (\varphi_s)_{ik} (\varphi_s^*)_{kj} ds.$$

Si noti che $\langle X^{(i)}, X^{(j)} \rangle_t$ è un processo di Itô, con differenziale stocastico dato da

$$d\langle X^{(i)}, X^{(j)} \rangle_t = (\varphi_t \cdot \varphi_t^*)_{ij} dt = \left(\sum_{k=1}^d (\varphi_t)_{ik} (\varphi_t^*)_{kj} \right) dt. \quad (6.19)$$

Possiamo quindi definire l'integrale rispetto al processo $\langle X^{(i)}, X^{(j)} \rangle_t$ nel modo già visto.

OSSERVAZIONE 6.8. Una regola pratica molto utile per “calcolare” $d\langle X^{(i)}, X^{(j)} \rangle_t$ senza dover ricordare la formula (6.19) è la seguente: si scrive $d\langle X^{(i)}, X^{(j)} \rangle_t = \langle dX_t^{(i)}, dX_t^{(j)} \rangle$, si scrivono le componenti $X^{(i)}$ e $X^{(j)}$ usando la decomposizione $dX_t = \varphi_t \cdot dB_t + \psi_t dt$, si sviluppa per bilinearità e si semplifica l'espressione risultante usando le regole

$$\langle dB_t^{(i)}, dB_t^{(j)} \rangle = \delta_{ij} dt, \quad \langle dB_t^{(i)}, dt \rangle = 0, \quad \langle dt, dt \rangle = 0.$$

In questo modo si ricava

$$\begin{aligned} \langle dX_t^{(i)}, dX_t^{(j)} \rangle &= \left\langle \sum_{k=1}^d (\varphi_t)_{ik} dB_t^{(k)} + (\psi_t)_i dt, \sum_{l=1}^d (\varphi_t)_{jl} dB_t^{(l)} + (\psi_t)_j dt \right\rangle \\ &= \sum_{k,l=1}^d (\varphi_t)_{ik} (\varphi_t)_{jl} \langle dB_t^{(k)}, dB_t^{(l)} \rangle = \sum_{k,l=1}^d (\varphi_t)_{ik} (\varphi_t)_{jl} \delta_{kl} dt = (\varphi_t \varphi_t^*)_{ij} dt, \end{aligned}$$

in accordo con la definizione (6.19). □

Possiamo ora formulare la versione multidimensionale della formula di Itô. Una funzione $F = F(t, x) : \mathbb{R}^+ \times \mathbb{R}^n \rightarrow \mathbb{R}$ è detta di classe $C^{1,2}$ se le sue derivate parziali di ordine uno in t e di ordine due in x esistono e sono funzioni continue:

$$\dot{F}(t, x) := \frac{\partial F}{\partial t}(t, x), \quad F'_i(t, x) := \frac{\partial F}{\partial x_i}(t, x), \quad F''_{ij}(t, x) := \frac{\partial^2 F}{\partial x_i \partial x_j}(t, x),$$

per ogni $(t, x) \in \mathbb{R}^+ \times \mathbb{R}^n$ e per ogni $1 \leq i, j \leq n$.

TEOREMA 6.9 (FORMULA DI ITÔ MULTIDIMENSIONALE). Se $X = \{X_t\}_{t \geq 0}$ un processo di Itô n -dimensionale, $dX_t = \varphi_t dB_t + \psi_t dt$, e $F = F(t, x) : \mathbb{R}^+ \times \mathbb{R}^n \rightarrow \mathbb{R}$ è una funzione di classe $C^{1,2}$, si ha q.c. per ogni $t \geq 0$

$$\begin{aligned} F(t, X_t) - F(0, X_0) &= \int_0^t \dot{F}(s, X_s) ds + \sum_{i=1}^n \int_0^t F'_i(s, X_s) dX_s^{(i)} \\ &\quad + \frac{1}{2} \sum_{i,j=1}^n \int_0^t F''_{ij}(s, X_s) d\langle X^{(i)}, X^{(j)} \rangle_s. \end{aligned} \quad (6.20)$$

In notazione differenziale:

$$\begin{aligned} dF(t, X_t) &= \dot{F}(t, X_t) dt + \sum_{i=1}^n F'_i(t, X_t) dX_t^{(i)} \\ &\quad + \frac{1}{2} \sum_{i,j=1}^n F''_{ij}(t, X_t) d\langle X^{(i)}, X^{(j)} \rangle_t. \end{aligned} \quad (6.21)$$

OSSERVAZIONE 6.10. Si può riscrivere l'equazione (6.21) in forma più compatta:

$$dF(t, X_t) = \dot{F}(t, X_t) dt + F'(t, X_t) \cdot dX_t + \frac{1}{2} \text{Tr}(F''(t, X_t) \cdot d\langle X, X \rangle_t), \quad (6.22)$$

dove ricordiamo che $\text{Tr}(CD) := \sum_{i,j=1}^n C_{ij} D_{ji}$ per ogni coppia di matrici $n \times n$ C, D . \square

La formula di Itô multidimensionale (6.20) è un po' involuta, ma si semplifica in alcuni casi particolari interessanti. L'esempio più importante si ha quando X è il moto browniano B : infatti in questo caso $\psi_s \equiv 0$ e $(\varphi_t)_{ij} = \delta_{ij}$ è la matrice identica, per cui $d\langle X^{(i)}, X^{(j)} \rangle_t = \delta_{ij} dt$. Vale la pena enunciare esplicitamente la formula di Itô in questo caso speciale. Per ragioni estetiche, indichiamo questa volta il *gradiente* di F rispetto a x con $\nabla F := F' = \{\frac{\partial F}{\partial x_i}\}_{1 \leq i \leq n}$, e introduciamo il *laplaciano* $\Delta F := \text{Tr}(F'') = \sum_{i=1}^n \frac{\partial^2 F}{\partial x_i^2}$.

COROLLARIO 6.11 (FORMULA DI ITÔ PER IL MOTO BROWNIANO d -DIMENSIONALE). Per ogni funzione $F = F(t, x) : \mathbb{R}^+ \times \mathbb{R}^n \rightarrow \mathbb{R}$ di classe $C^{1,2}$, si ha q.c. per ogni $t \geq 0$

$$F(t, B_t) - F(0, B_0) = \int_0^t \nabla F(s, B_s) \cdot dB_s + \int_0^t \left[\dot{F}(s, B_s) + \frac{1}{2} \Delta F(s, B_s) \right] ds. \quad (6.23)$$

In notazione differenziale:

$$dF(t, B_t) = \nabla F(t, B_t) \cdot dB_t + \left(\dot{F}(t, B_t) + \frac{1}{2} \Delta F(t, B_t) \right) dt. \quad (6.24)$$

Queste formule sono alla base di alcune fondamentali applicazioni del moto browniano alle funzioni armoniche e al problema di Dirichlet, che discutiamo nella prossimo paragrafo.

Un altro caso speciale particolarmente interessante della formula di Itô multidimensionale è il seguente. Supponiamo ora che $B = \{B_t\}_{t \geq 0}$ sia un moto browniano reale ($d = 1$) e che $X = \{X_t\}_{t \geq 0}$, $Y = \{Y_t\}_{t \geq 0}$ siano due processi di Itô reali, con differenziali stocastici

$$dX_t = \varphi_t^X dB_t + \psi_t^X dt, \quad dY_t = \varphi_t^Y dB_t + \psi_t^Y dt.$$

Introduciamo la *covarianza quadratica* $\langle X, Y \rangle_t := \int_0^t \varphi_s^X \varphi_s^Y ds$, di modo che

$$d\langle X, Y \rangle_t = \varphi_t^X \varphi_t^Y dt.$$

(Si veda l'Osservazione 6.8 per una “motivazione” empirica.) Una semplice applicazione della formula di Itô (6.21) al processo bidimensionale (X_t, Y_t) con la funzione $F(x, y) := xy$ conduce all'importante corollario seguente.

COROLLARIO 6.12 (FORMULA DI INTEGRAZIONE PER PARTI STOCASTICA). Per ogni coppia di processi di Itô reali $X = \{X_t\}_{t \geq 0}$, $Y = \{Y_t\}_{t \geq 0}$, si ha q.c. per ogni $t \geq 0$

$$X_t Y_t - X_0 Y_0 = \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle X, Y \rangle_t. \quad (6.25)$$

In notazione differenziale:

$$d(X_t Y_t) = X_t dY_t + Y_t dX_t + d\langle X, Y \rangle_t. \quad (6.26)$$

6.5. MOTO BROWNIANO E LAPLACIANO

Fissiamo $x \in \mathbb{R}^d$ e indichiamo con $B = \{B_t\}_{t \geq 0}$ un moto browniano d -dimensionale che parte da x . Questo significa semplicemente che $B_t = x + \beta_t$, dove $\beta = \{\beta_t\}_{t \geq 0}$ è un moto browniano d -dimensionale standard. Indicheremo per chiarezza con P_x e E_x la probabilità e il valore atteso.

Riscriviamo la formula di Itô per B e per funzioni $F(t, x) = F(x)$ non dipendenti dal tempo: dalle relazioni (6.23) e (6.24) segue che per ogni funzione $F : \mathbb{R}^d \rightarrow \mathbb{R}$ di classe C^2 (ricordando che $B_0 = x$)

$$F(B_t) - F(x) = \int_0^t \nabla F(B_s) \cdot dB_s + \frac{1}{2} \int_0^t \Delta F(B_s) ds,$$

ossia in notazione differenziale

$$dF(B_t) = \nabla F(B_t) \cdot dB_t + \frac{1}{2} \Delta F(B_t) dt.$$

Una conseguenza fondamentale di queste formule è che se F è una funzione armonica, cioè se $\Delta F = 0$, allora il processo $\{F(B_t)\}_{t \geq 0}$ è una martingala locale.

Conseguenze molto interessanti si ottengono per funzioni armoniche F definite su un sottoinsieme di \mathbb{R}^d , nel qual caso occorre essere più precisi. Dato un sottoinsieme $A \subseteq \mathbb{R}^d$, indichiamo con

$$\tau_A := \inf\{t \geq 0 : B_t \in A\}$$

il tempo d'ingresso in A del moto browniano B . Ricordiamo che se A è chiuso (o aperto), τ_A è un tempo d'arresto e, se $\tau_A < \infty$ q.c., B_{τ_A} è una variabile aleatoria.

LEMMA 6.13. Sia $D \subseteq \mathbb{R}^d$ un insieme aperto e connesso, sia $F : D \rightarrow \mathbb{R}$ una funzione armonica (cioè di classe C^2 e tale che $\Delta F(x) = 0$ per ogni $x \in D$) e sia G un insieme aperto limitato tale che $\overline{G} \subseteq D$. Per ogni $x \in G$ si ha $\tau_{G^c} < \infty$, P_x -q.c., e vale la relazione

$$F(x) = E_x(F(B_{\tau_{G^c}})). \quad (6.27)$$

DIMOSTRAZIONE. Cominciamo a mostrare che $P_x(\tau_{G^c} < \infty) = 1$. Per ipotesi G è limitato, dunque $G \subseteq [-L, L]^d$ per qualche $L > 0$. Dato che $\{B_t^{(1)} - x_1\}_{t \geq 0}$ è un moto browniano reale standard, sappiamo che P_x -q.c. $\limsup_{t \rightarrow \infty} B_t^{(1)} = +\infty$. Di conseguenza, per P_x -q.o. $\omega \in \Omega$ esiste $t_0(\omega) < \infty$ tale che $B_{t_0(\omega)}^{(1)}(\omega) > L$. Quindi $B_{t_0(\omega)}(\omega) \notin G$, ovvero $\tau_{G^c}(\omega) \leq t_0(\omega) < \infty$.

Consideriamo ora il processo $M = \{M_t\}_{t \geq 0}$ definito da

$$M_t := F(B_{t \wedge \tau_{G^c}})$$

e mostriamo che è una martingala. Dato che \overline{G} è chiuso e limitato, la funzione F è limitata su \overline{G} . Sia $\Psi : \mathbb{R}^d \rightarrow \mathbb{R}$ una funzione di classe C^2 su tutto \mathbb{R}^d che coincida con F su \overline{G} .[†] Applicando la formula di Itô e ricordando che $B_0 = x$ si ha

$$\Psi(B_t) - \Psi(x) = \int_0^t \nabla \Psi(B_s) \cdot dB_s + \frac{1}{2} \int_0^t \Delta \Psi(B_s) ds.$$

Sostituendo t con $t \wedge \tau_{G^c}$, per le proprietà dell'integrale stocastico (e di quello ordinario) possiamo scrivere

$$\Psi(B_{t \wedge \tau_{G^c}}) - \Psi(x) = \int_0^t 1_{[0, \tau_{G^c})}(s) \nabla \Psi(B_s) \cdot dB_s + \frac{1}{2} \int_0^t 1_{[0, \tau_{G^c})}(s) \Delta \Psi(B_s) ds.$$

Dato che Ψ coincide con F su \overline{G} e dato che $B_{t \wedge \tau_{G^c}} \in \overline{G}$, si ha $\Psi(B_{t \wedge \tau_{G^c}}) = F(B_{t \wedge \tau_{G^c}})$. Analogamente, per $s \leq \tau_{G^c}$ si ha $B_s \in \overline{G}$ e quindi $\nabla \Psi(B_s) = \nabla F(B_s)$, mentre $\Delta \Psi(B_s) = \Delta F(B_s) = 0$ perché per ipotesi F è armonica su $D \supseteq \overline{G}$. Otteniamo dunque la relazione

$$M_t := F(B_{t \wedge \tau_{G^c}}) = F(x) + \int_0^t 1_{[0, \tau_{G^c})}(s) \nabla F(B_s) \cdot dB_s,$$

che mostra che M è una martingala locale. Dato che $B_s \in \overline{G}$ per $s \leq \tau_{G^c}$, segue che $|\nabla F(B_s) 1_{[0, \tau_{G^c})}(s)| \leq \sup_{x \in \overline{G}} |\nabla F(x)| < \infty$, perché la funzione $x \mapsto \nabla F(x)$ è continua

[†]Per esempio basta definire $\Psi(x) := F(x) I(x)$, dove $I : \mathbb{R}^d \rightarrow [0, 1]$ è una funzione di classe C^∞ tale che $I(x) = 1$ per ogni $x \in \overline{G}$ e $I(x) = 0$ per $x \notin D$. Una tale I si ottiene per esempio ponendo $I(x) := 1_{G_\varepsilon} * \varrho$, dove $*$ indica la convoluzione, $G_\varepsilon := \{x \in \mathbb{R}^d : \text{dist}(x, G) < \varepsilon\}$, ϱ è una funzione C^∞ e di integrale uno con supporto in $\{x \in \mathbb{R}^d : |x| < \varepsilon\}$ e $0 < \varepsilon < \frac{1}{2} \text{dist}(G, D^c)$.

su \overline{G} e dunque limitata. Questo mostra che l'integrando $\{\nabla F(B_s) \mathbf{1}_{[0, \tau_{G^c})}(s)\}_{s \geq 0}$ è in $\mathcal{M}^2(1 \times d)$, quindi M è una vera martingala (di quadrato integrabile).

Dato che una martingala è costante in media, si ha

$$F(x) = E_x(M_0) = E_x(M_t) = E_x(F(B_{t \wedge \tau_{G^c}})),$$

per ogni $t \geq 0$. Per $t \rightarrow \infty$ si ha $t \wedge \tau_{G^c} \rightarrow \tau_{G^c}$ q.c., perché $P_x(\tau_{G^c} < \infty) = 1$, quindi anche $F(B_{t \wedge \tau_{G^c}}) \rightarrow F(B_{\tau_{G^c}})$ q.c. per la continuità di F . Dato che $|F(B_{t \wedge \tau_{G^c}})| \leq \sup_{x \in \overline{G}} |F(x)| < \infty$, per convergenza dominata si ottiene $E_x(F(B_{\tau_{G^c}})) = F(x)$, cioè la relazione (6.27). \square

6.5.1. IL PROBLEMA DI DIRICHLET. Un problema classico in elettrostatica consiste nel determinare il potenziale elettrico in una regione dello spazio D , quando se ne conosce il valore sulla frontiera ∂D .

Più precisamente, dato un insieme aperto e limitato $D \subseteq \mathbb{R}^d$ e assegnata una funzione $f : \partial D \rightarrow \mathbb{R}$ continua, il *problema di Dirichlet* consiste nel determinare (se esiste) una funzione $F : \overline{D} \rightarrow \mathbb{R}$ che soddisfi le seguenti relazioni:

$$F \text{ è continua su } \overline{D}, \text{ di classe } C^2 \text{ su } D \text{ e } \begin{cases} \Delta F(x) = 0 & \forall x \in D \\ F(x) = f(x) & \forall x \in \partial D \end{cases}. \quad (6.28)$$

Vale allora il seguente risultato.

PROPOSIZIONE 6.14. Se esiste una funzione F soluzione del problema (6.28), essa è unica ed è data da

$$F(x) = E_x(f(B_{\tau_{D^c}})), \quad \forall x \in \overline{D}. \quad (6.29)$$

DIMOSTRAZIONE. Per $n \in \mathbb{N}$ poniamo $D_n := \{x \in \mathbb{R}^d : \text{dist}(x, D^c) > \frac{1}{n}\}$. Se F è soluzione di (6.28), possiamo applicare il Lemma 6.13 con $G = D_n$, ottenendo

$$F(x) = E_x(F(B_{\tau_{D_n^c}})), \quad \forall x \in D_n. \quad (6.30)$$

Mostriamo ora che $\tau_{D_n^c} \rightarrow \tau_{D^c}$ q.c. per $n \rightarrow \infty$. Si noti che $\tau_{D_n^c}$ è crescente in n e quindi q.c. esiste $\bar{\tau} := \lim_{n \rightarrow \infty} \tau_{D_n^c}$. Resta da mostrare che $\bar{\tau} = \tau_{D^c}$ q.c.. Da un lato si ha per definizione $\tau_{D_n^c} \leq \tau_{D^c}$ per ogni $n \in \mathbb{N}$, quindi $\bar{\tau} \leq \tau_{D^c}$. Dall'altro lato, per continuità delle traiettorie si ha q.c. $B_{\bar{\tau}} = \lim_{n \rightarrow \infty} B_{\tau_{D_n^c}}$, quindi $\text{dist}(B_{\bar{\tau}}, D^c) = \lim_{n \rightarrow \infty} \text{dist}(B_{\tau_{D_n^c}}, D^c) = 0$, perché $\text{dist}(B_{\tau_{D_n^c}}, D^c) = \frac{1}{n}$. Di conseguenza, q.c. $B_{\bar{\tau}} \in D^c$ e dunque $\tau_{D^c} \leq \bar{\tau}$.

Per continuità delle traiettorie di B e della funzione F su \overline{D} , segue che $F(B_{\tau_{D_n^c}}) \rightarrow F(B_{\tau_{D^c}})$ q.c. per $n \rightarrow \infty$. Dato che F è continua sull'insieme chiuso e limitato \overline{D} , essa è limitata e quindi $|F(B_{\tau_{D_n^c}})| \leq \sup_{x \in \overline{D}} |F(x)| < \infty$. Possiamo dunque applicare il teorema di convergenza dominata: per ogni $x \in D$ fissato, si ha $x \in D_n$ per n sufficientemente grande, quindi passando al limite in (6.30) si ottiene la relazione (6.29), per ogni $x \in D$. Infine, per $x \in \partial D$ si ha $\tau_{D^c} = 0$ e dunque la relazione (6.29) vale banalmente, perché $B_0 = x$ e $F(x) = f(x)$ per $x \in \partial D$, grazie a (6.28). \square

Per quanto riguarda l'esistenza della soluzione del problema (6.28), è naturale considerare la funzione F definita da (6.29) (dopotutto, se una soluzione esiste, essa deve essere data da tale relazione). È chiaro che $F(x) = f(x)$ per $x \in \partial D$, perché in questo caso $\tau_{D^c} = 0$. È anche vero che F è armonica su D , cioè di classe C^2 e tale che $\Delta F(x) = 0$ per ogni $x \in D$, come mostriamo qui sotto. Tuttavia, la funzione F in generale non è continua su \overline{D} , cioè non è detto che $F(y) \rightarrow f(x)$ per $y \rightarrow x \in \partial D$: in questo caso il problema di Dirichlet (6.28) *non ammette soluzione*.

Affinché la funzione F definita da (6.29) sia continua su \overline{D} occorrono ipotesi aggiuntive su D . Per esempio, una condizione sufficiente è che ∂D sia una varietà differenziabile di classe C^1 . Per maggiori dettagli si veda il paragrafo 4.2 in [Karatzas e Shreve, 1998].

Per mostrare che la funzione F definita dalla relazione (6.29) è armonica in D , mostreremo che F soddisfa la *proprietà del valor medio*: per ogni $x \in D$ e per ogni $r > 0$ tale che $\mathcal{B}(x, r) := \{y \in \mathbb{R}^d : |y - x| \leq r\} \subseteq D$, si ha

$$F(x) = \int_{\partial \mathcal{B}(x, r)} F(y) \mu_{x, r}(dy), \quad (6.31)$$

dove $\partial \mathcal{B}(x, r) := \{y \in \mathbb{R}^d : |y - x| = r\}$ e dove $\mu_{x, r}$ indica la misura di superficie su $\partial \mathcal{B}(x, r)$, normalizzata in modo che sia una probabilità: $\mu_{x, r}(\partial \mathcal{B}(x, r)) = 1$. È infatti un risultato classico di analisi che se una funzione soddisfa la proprietà del valor medio su un insieme aperto D , essa è armonica su D (si veda la Proposizione 2.5 nel capitolo 4 in [Karatzas e Shreve, 1998]).

La prima osservazione è che il secondo membro di (6.31) si può scrivere come $E_x(F(B_{\tau_{\mathcal{B}(x, r)}}))$, perché la legge della variabile $B_{\tau_{\mathcal{B}(x, r)}}$ è proprio $\mu_{x, r}$. Questo segue dal fatto che il moto browniano in \mathbb{R}^d è invariante per rotazioni, quindi anche la legge di $B_{\tau_{\mathcal{B}(x, r)}}$, che è una misura su $\partial \mathcal{B}(x, r)$, deve essere invariante per rotazioni (di centro x) e la misura di superficie normalizzata è l'unica probabilità su $\partial \mathcal{B}(x, r)$ con questa proprietà. Resta dunque da mostrare che $F(x) = E_x(F(B_{\tau_{\mathcal{B}(x, r)}}))$.

Per questa relazione, usiamo una proprietà della speranza condizionale che ora descriviamo. Siano X, Y variabili aleatorie, definite su (Ω, \mathcal{F}, P) a valori negli spazi misurabili (E_1, \mathcal{E}_1) e (E_2, \mathcal{E}_2) rispettivamente, e sia $g : E_1 \times E_2 \rightarrow \mathbb{R}$ una funzione misurabile e limitata. Supponiamo che \mathcal{G} sia una sotto σ -algebra di \mathcal{F} tale che X sia \mathcal{G} -misurabile, mentre Y sia indipendente da \mathcal{G} (in particolare le variabili X e Y sono indipendenti). Allora

$$E(g(X, Y) | \mathcal{G}) = \tilde{g}(X), \quad \text{dove} \quad \tilde{g}(x) := E(g(x, Y)). \quad (6.32)$$

Occorre mostrare che $E(g(X, Y) 1_G) = E(\tilde{g}(X) 1_G)$, per ogni $G \in \mathcal{G}$. Questa relazione è immediata da verificare se $g(a, b) = 1_{A \times B}(a, b) = 1_A(a) 1_B(b)$, con $A \in \mathcal{E}_1$ e $B \in \mathcal{E}_2$. Inoltre le funzioni g che soddisfano tale relazione costituiscono uno spazio vettoriale che contiene le costanti e chiuso per limiti crescenti. Dato che la famiglia $\{A \times B : A \in \mathcal{E}_1, B \in \mathcal{E}_2\}$ è una base di $\mathcal{E}_1 \otimes \mathcal{E}_2$, segue dal teorema di Classe Monotona che la relazione (6.32) vale per ogni g misurabile e limitata.

Possiamo infine applicare la relazione (6.32) con $P = P_x$, $\mathcal{G} = \mathcal{F}_{\tau_{\mathcal{B}(x, r)}}$, $X = B_{\tau_{\mathcal{B}(x, r)}}$, $Y = \{B_{\tau_{\mathcal{B}(x, r)} + t} - B_{\tau_{\mathcal{B}(x, r)}}\}_{t \geq 0}$ e $g(a, b) = f(a + b_\tau)$, dove $\tau := \inf\{s \geq 0 : a + b_s \notin D\}$. La condizione che Y sia indipendente da \mathcal{G} segue dalla proprietà di Markov forte del moto browniano. Con queste definizioni si ha $g(X, Y) = f(B_{\tau_{D^c}})$ e inoltre $\tilde{g}(a) = E_x(g(a, Y)) = E_a(f(B_{\tau_{D^c}})) =: F(a)$, perché $a + Y$ rispetto a P_x è un moto browniano che parte da a e per la definizione (6.29). Si ottiene dunque la relazione

$$E_x(f(B_{\tau_{D^c}}) | \mathcal{F}_{\tau_{\mathcal{B}(x, r)}}) = F(B_{\tau_{\mathcal{B}(x, r)}}).$$

Infine, dato che $\mathcal{F}_{\tau_{\mathcal{B}(x, r)}} \subseteq \mathcal{F}_{\tau_{D^c}}$, dalla definizione (6.29) possiamo scrivere

$$F(x) = E_x(f(B_{\tau_{D^c}})) = E_x[E_x(f(B_{\tau_{D^c}}) | \mathcal{F}_{\tau_{\mathcal{B}(x, r)}})] = E_x(F(B_{\tau_{\mathcal{B}(x, r)}})),$$

e per quanto già detto la relazione (6.31) è dimostrata.

6.5.2. TRANSIENZA E RICORRENZA DEL MOTO BROWNIANO. Introduciamo la funzione $F : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$ definita da

$$F(z) := \begin{cases} \frac{1}{|z|^{d-2}} & \text{se } d \geq 3 \\ \log |z| & \text{se } d = 2 \\ |z| & \text{se } d = 1 \end{cases}.$$

È un fatto noto (e facilmente dimostrabile) che F è una funzione armonica sul dominio $D := \mathbb{R}^d \setminus \{0\}$. Consideriamo la corona sferica

$$G := \{z \in \mathbb{R}^d : r < |z| < R\}, \quad \text{dove } 0 < r < R < \infty,$$

che soddisfa le ipotesi del Lemma 6.13. Quindi, per ogni $x \in G$, il tempo di uscita τ_{G^c} dall'insieme G del moto browniano che parte in x è q.c. finito, e si ha

$$F(x) = E_x(F(B_{\tau_{G^c}})).$$

Per continuità delle traiettorie, $B_{\tau_{G^c}} \in \partial G = \{z \in \mathbb{R}^d : |z| = r \text{ o } |z| = R\}$, e dato che $F(z) = F(|z|)$ si ottiene

$$F(|x|) = F(r) P_x(|B_{\tau_{G^c}}| = r) + F(R) P_x(|B_{\tau_{G^c}}| = R).$$

Visto che $P_x(|B_{\tau_{G^c}}| = r) + P_x(|B_{\tau_{G^c}}| = R) = 1$, si ricava facilmente che

$$P_x(|B_{\tau_{G^c}}| = r) = \frac{F(|x|) - F(R)}{F(r) - F(R)} = \begin{cases} \frac{\frac{1}{|x|^{d-2}} - \frac{1}{R^{d-2}}}{\frac{1}{r^{d-2}} - \frac{1}{R^{d-2}}} & \text{se } d \geq 3 \\ \frac{\log R - \log |x|}{\log R - \log r} & \text{se } d = 2 \\ \frac{R - |x|}{R - r} & \text{se } d = 1 \end{cases}. \quad (6.33)$$

Notiamo che $\{|B_{\tau_{G^c}}| = r\}$ è l'evento “il moto browniano che parte da $x \in G$ raggiunge la sfera interna $\{|z| = r\}$ prima di quella esterna $\{|z| = R\}$ ”, da cui si evince che tale evento è crescente in R . Consideriamo quindi l'evento limite

$$A_r = \lim_{R \rightarrow \infty} \{|B_{\tau_{G^c}}| = r\} = \bigcup_{R > 0} \{|B_{\tau_{G^c}}| = r\},$$

che si può descrivere come “per qualche $R > 0$, il moto browniano che parte da x raggiunge la sfera $\{|z| = r\}$ prima della sfera $\{|z| = R\}$ ”. Dato che q.c. le traiettorie di B sono continue, esse sono limitate su ogni intervallo di tempo limitato, per cui l'evento A_r non è altro che “il moto browniano che parte da x raggiunge in tempo finito la sfera $\{|z| = r\}$ ”.

Prendendo il limite $R \rightarrow \infty$ in (6.33) e usando la continuità dal basso della probabilità, si ottiene quindi l'importante relazione seguente, valida per ogni $x \in \mathbb{R}^d$, $r > 0$ con $r < |x|$:

$$P_x(A_r) = P_x(B \text{ raggiunge in tempo finito la sfera } \{|z| = r\}) = \begin{cases} \left(\frac{r}{|x|}\right)^{d-2} & \text{se } d \geq 3 \\ 1 & \text{se } d \leq 2 \end{cases}.$$

È chiaro che c'è una grossa differenza tra i casi $d \leq 2$ e $d \geq 3$.

- Se $d = 1$ o $d = 2$, per ogni $r < |x|$ fissato si ha $P_x(A_r) = 1$, ossia q.c. il moto browniano che parte da x visita la palla di raggio r centrata nell'origine. Prendendo una successione $r_n \downarrow 0$ si ha $P_x(\bigcap_{n \in \mathbb{N}} A_{r_n}) = 1$, ossia q.c. il moto browniano che parte da x visita ogni intorno dell'origine. Per simmetria, anche l'evento $C_x :=$ "il moto browniano che parte dall'origine visita ogni intorno di x " ha probabilità uno, quindi anche $P_0(\bigcap_{x \in \mathbb{Q}^d} C_x) = 1$, ossia q.c. il moto browniano (che parte dall'origine) visita ogni intorno di ogni punto di coordinate razionali. Ciò significa che per $d = 1$ e $d = 2$ q.c. le traiettorie del moto browniano d -dimensionale sono dense in \mathbb{R}^d . Per tale ragione, si dice che il moto browniano in dimensione 1 e 2 è un processo *ricorrente*.
- Se $d \geq 3$, per ogni $r < |x|$ fissato si ha $P_x(A_r) = (r/|x|)^{d-2} < 1$, ossia il moto browniano che parte da x ha una probabilità positiva di non visitare mai la palla di raggio r centrata nell'origine. È possibile inoltre dimostrare che si ha q.c. $\lim_{t \rightarrow \infty} |B_t| = +\infty$. Si dice in questo caso che il moto browniano in dimensione $d \geq 3$ è *transiente*.

6.6. IL TEOREMA DI GIRSANOV

6.6.1. PRELUDIO. Le leggi normali multivariate in \mathbb{R}^d con matrice delle covarianze non singolare sono assolutamente continue rispetto alla misura di Lebesgue, con densità strettamente positiva ovunque. Pertanto, traslando una tale legge si ottiene una nuova legge assolutamente continua rispetto alla legge di partenza.

Consideriamo per esempio un vettore aleatorio normale con media nulla e matrice delle covarianze K non singolare: $Z \sim N(0, K)$ con $\det(K) \neq 0$. Dato $a = (a_1, \dots, a_d) \in \mathbb{R}^d$, per ogni insieme $A \in \mathcal{B}(\mathbb{R}^d)$ possiamo scrivere

$$\begin{aligned} P(Z + a \in A) &= P(Z \in A - a) = \int_{A-a} f_Z(z) dz = \int_{A-a} \frac{e^{-\frac{1}{2}\langle z, K^{-1}z \rangle}}{(2\pi)^{d/2} \sqrt{|\det(K)|}} dz \\ &= \int_A \frac{e^{-\frac{1}{2}\langle (x-a), K^{-1}(x-a) \rangle}}{(2\pi)^{d/2} \sqrt{|\det(K)|}} dx = \int_{\mathbb{R}^d} 1_A(x) e^{\langle x, K^{-1}a \rangle - \frac{1}{2}\langle a, K^{-1}a \rangle} f_Z(x) dx. \end{aligned} \tag{6.34}$$

Questa relazione mostra che la legge del vettore aleatorio $Z + a$ è assolutamente continua rispetto alla legge del vettore Z , con densità di Radon-Nikodym data da

$$P(Z + a \in dx) = e^{\langle x, K^{-1}a \rangle - \frac{1}{2}\langle a, K^{-1}a \rangle} P(Z \in dx).$$

Grazie alla formula del cambio di variabili (Teorema 1.6), possiamo riscrivere la relazione (6.34) come

$$P(Z + a \in A) = E \left(1_A(Z) e^{\langle Z, K^{-1}a \rangle - \frac{1}{2} \langle a, K^{-1}a \rangle} \right). \quad (6.35)$$

Se introduciamo una nuova probabilità Q su (Ω, \mathcal{F}) , definita da

$$Q(d\omega) := e^{\langle Z(\omega), K^{-1}a \rangle - \frac{1}{2} \langle a, K^{-1}a \rangle} P(d\omega),$$

possiamo riscrivere la relazione (6.35) come $P(Z + a \in A) = Q(Z \in A)$. Sostituendo A con $A + a$ si ottiene dunque

$$P(Z \in A) = Q(Z - a \in A), \quad \forall A \in \mathcal{B}(\mathbb{R}^d).$$

Questo mostra che il vettore aleatorio traslato $Z - a$ rispetto alla nuova probabilità Q ha la stessa legge del vettore aleatorio Z rispetto alla probabilità originale P .

Mostriamo ora che è possibile estendere questa proprietà delle leggi normali multivariate al moto browniano.

6.6.2. IL TEOREMA DI GIRSANOV. Ricordiamo che è fissato uno spazio filtrato standard $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$, su cui è definito un $\{\mathcal{F}_t\}_{t \geq 0}$ -moto browniano reale $B = \{B_t\}_{t \geq 0}$. Per tutto questo paragrafo, lavoreremo in realtà con la filtrazione $\{\mathcal{F}_t\}_{t \in [0, T]}$ e il moto browniano $B = \{B_t\}_{t \in [0, T]}$ con insieme dei tempi ristretto a $[0, T]$.

Consideriamo la “traslazione” $B - F$ del moto browniano B mediante un processo $F = \{F_t\}_{t \in [0, T]}$. Il teorema di Girsanov, che ora dimostriamo, afferma che se il processo F è della forma $F_t = \int_0^t \varphi_s ds$, dove $\varphi = \{\varphi_s\}_{s \in [0, T]}$ in $M_{\text{loc}}^2[0, T]$, è possibile definire una nuova legge Q_T su (Ω, \mathcal{F}) tale che $B - F$ sia un moto browniano rispetto a Q_T .

Dato un processo $\varphi = \{\varphi_s\}_{s \in [0, T]}$ in $M_{\text{loc}}^2[0, T]$, ricordiamo la definizione (6.16) della supermartingala esponenziale $Z = \{Z_t\}_{t \in [0, T]}$:

$$Z_t = Z_t(\varphi) := \exp \left(\int_0^t \varphi_s dB_s - \frac{1}{2} \int_0^t \varphi_s^2 ds \right). \quad (6.36)$$

Sappiamo che Z è una supermartingala, in particolare $E(Z_T) \leq E(Z_0) = 1$. Se supponiamo che $E(Z_T) = 1$ (che è equivalente, come abbiamo visto in § 6.3.2, a richiedere che Z sia una martingala), possiamo definire una nuova legge Q_T su Ω , ponendo

$$Q_T(d\omega) := Z_T(\omega) P(d\omega), \quad \text{cioè} \quad Q_T(A) := E_P(1_A Z_T), \quad \forall A \in \mathcal{F}, \quad (6.37)$$

dove indichiamo con E_P il valore atteso rispetto a P , per distinguerlo da quello rispetto a Q_T , che indicheremo con E_{Q_T} . Abbiamo quindi il seguente fondamentale risultato.

TEOREMA 6.15 (GIRSANOV). Se $E_P(Z_T) = 1$, il processo $\tilde{B} = \{\tilde{B}_t\}_{t \in [0, T]}$ definito da

$$\tilde{B}_t = B_t - \int_0^t \varphi_s ds, \quad (6.38)$$

è un $\{\mathcal{F}_t\}_{t \in [0, T]}$ -moto browniano rispetto alla probabilità Q_T .

Ricordiamo che due condizioni sufficienti esplicite che garantiscono che $E_P(Z_T) = 1$ sono date nella Proposizione 6.7 (che non abbiamo dimostrato). Dimostriamo invece una condizione più forte nel Lemma 6.17 qui sotto.

OSSERVAZIONE 6.16. Sempre sotto l'ipotesi $E_P(Z_T) = 1$, si può definire una legge Q_t su Ω per ogni $t \in [0, T]$, mediante l'equazione (6.37). Allora, per ogni $t \in [0, T]$, la legge Q_T coincide con Q_t su \mathcal{F}_t . Infatti, dato che Z è una martingala, per $A \in \mathcal{F}_t$ si ha

$$\begin{aligned} Q_T(A) &= E_P(1_A Z_T) = E_P[E_P(1_A Z_T \mid \mathcal{F}_t)] = E_P[1_A E_P(Z_T \mid \mathcal{F}_t)] \\ &= E_P[1_A Z_t] = Q_t(A). \end{aligned} \quad (6.39)$$

Se $\varphi \in M_{\text{loc}}^2$ e se $E_P(Z_T) = 1$ per ogni $T > 0$, è possibile definire una legge Q_∞ su Ω , che coincide con Q_t su \mathcal{F}_t , per ogni $t \in [0, \infty)$, tale che il processo $\{\tilde{B}_t\}_{t \in [0, \infty)}$ definito in (6.38) sia un $\{\mathcal{F}_t\}_{t \in [0, \infty)}$ -moto browniano rispetto a Q_∞ . Tuttavia Q_∞ in generale *non* è assolutamente continua rispetto a P . Per maggiori dettagli, si veda la discussione che segue il Teorema 5.1 nel capitolo 3 in [Karatzas e Shreve, 1998]. \square

6.6.3. PREPARAZIONE. Prima di dimostrare il Teorema 6.15, abbiamo bisogno di alcuni risultati preparatori. Definiamo $M_{\text{loc}}^{2, \mathbb{C}}[0, T]$ come lo spazio dei processi $\varphi = \{\varphi_s\}_{s \in [0, T]}$ a valori in \mathbb{C} tali che le parti reale $\Re(\varphi) = \{\Re(\varphi_s)\}_{s \in [0, T]}$ e immaginaria $\Im(\varphi) = \{\Im(\varphi_s)\}_{s \in [0, T]}$ di φ sono entrambe processi in $M_{\text{loc}}^2[0, T]$. Equivalentemente, $\varphi \in M_{\text{loc}}^{2, \mathbb{C}}[0, T]$ se e solo se $\Re(\varphi)$ e $\Im(\varphi)$ sono processi progressivamente misurabili e si ha $\int_0^T |\varphi_s|^2 ds < \infty$ q.c..

Se $\varphi \in M_{\text{loc}}^{2, \mathbb{C}}[0, T]$, definiamo l'integrale stocastico

$$\int_0^t \varphi_s dB_s := \int_0^t \Re(\varphi_s) dB_s + i \int_0^t \Im(\varphi_s) dB_s,$$

così come l'integrale ordinario

$$\begin{aligned} \int_0^t \varphi_s^2 ds &:= \int_0^t \Re(\varphi_s^2) ds + i \int_0^t \Im(\varphi_s^2) ds \\ &= \int_0^t (\Re(\varphi_s)^2 - \Im(\varphi_s)^2) ds + i \int_0^t 2 \Re(\varphi_s) \Im(\varphi_s) ds. \end{aligned}$$

In particolare, si può definire senza problemi $Z_t = Z_t(\varphi)$ mediante la relazione (6.36) per ogni $\varphi \in M_{\text{loc}}^{2, \mathbb{C}}[0, T]$. Vale allora il seguente

LEMMA 6.17. Se $\varphi \in M_{\text{loc}}^{2, \mathbb{C}}[0, T]$ è tale che esiste una costante reale $C < \infty$ per cui $\int_0^T |\varphi_s|^2 ds \leq C$ q.c., si ha $E_P(Z_T(\varphi)) = 1$.

DIMOSTRAZIONE. Cominciamo a considerare il caso in cui φ è reale. Sappiamo che $Z = \{Z_t(\varphi)\}_{t \in [0, T]}$ è una martingala locale: in effetti $dZ_t = \varphi_t Z_t dB_t$. Definiamo ora la

variabile $Y := \sup_{t \in [0, T]} Z_t$ e assumiamo che $E(Y^2) < \infty$. Segue allora facilmente che il processo integrando $\{\varphi_s Z_s\}_{s \in [0, T]} \in M^2[0, T]$:

$$E_P \left(\int_0^T \varphi_s^2 Z_s^2 ds \right) \leq E_P \left(Y^2 \int_0^T \varphi_s^2 ds \right) \leq C E_P(Y^2) < \infty,$$

avendo usato l'ipotesi $\int_0^T \varphi_s^2 ds \leq C$ q.c.. Dalla relazione $dZ_t = \varphi_t Z_t dB_t$ segue allora che Z è una vera martingala: in particolare $E_P(Z_T) = E(Z_0) = 1$.

Resta da mostrare che effettivamente $Y := \sup_{t \in [0, T]} Z_t \in L^2$. Si noti che $Z_t(\varphi)^3 = Z_t(3\varphi) \cdot e^{3 \int_0^t \varphi_s^2 ds} \leq Z_t(3\varphi) e^{3C}$ q.c., per cui possiamo scrivere

$$P(Y > \lambda) = P \left(\sup_{t \in [0, T]} Z_t(\varphi)^3 > \lambda^3 \right) \leq P \left(\sup_{t \in [0, T]} Z_t(3\varphi) > \lambda^3 e^{-3C} \right).$$

Dato che $\{Z_t(3\varphi)\}_{t \geq 0}$ è una supermartingala continua, possiamo applicare la disuguaglianza massimale (Osservazione 4.15) ottenendo

$$P(Y > \lambda) \leq \frac{e^{3C}}{\lambda^3} (E_P(Z_0(3\varphi)) + E_P(Z_T(3\varphi)^-)) = \frac{e^{3C}}{\lambda^3},$$

perché $Z_0(3\varphi) = 1$ e $Z_T(3\varphi) \geq 0$. Questo mostra che $Y \in L^2$, poiché

$$E_P(Y^2) = \int_0^\infty P(Y^2 > x) dx \leq 1 + \int_1^\infty P(Y > \sqrt{x}) dx < \infty.$$

Il caso in cui φ è complessa è analogo. Dato che nella formula di Itô (6.20) la funzione F appare linearmente, la formula si può applicare senza problemi al caso di funzioni $F(x)$ di classe C^2 a valori complessi, in particolare all'esponenziale complesso. Applicando dunque la formula di Itô alla relazione (6.36), si ricava che anche per $\varphi \in M_{\text{loc}}^{2, \mathbb{C}}[0, T]$ il processo $\{Z_t = Z_t(\varphi)\}_{t \in [0, T]}$ soddisfa l'equazione $dZ_t = \varphi_t Z_t dB_t$, ossia

$$d\Re(Z_t) = \Re(\varphi_t Z_t) dB_t, \quad d\Im(Z_t) = \Im(\varphi_t Z_t) dB_t.$$

Se mostriamo che i processi $\{\Re(\varphi_t Z_t)\}_{t \in [0, T]}$ e $\{\Im(\varphi_t Z_t)\}_{t \in [0, T]}$ sono nello spazio $M^2[0, T]$, segue allora che $\Re(Z)$ e $\Im(Z)$ sono entrambe vere martingale e non soltanto martingale locali. Di conseguenza $E_P(\Re(Z_T)) = E_P(\Re(Z_0)) = 1$ e $E_P(\Im(Z_T)) = E_P(\Im(Z_0)) = 0$, da cui segue che $E_P(Z_T) = 1$, cioè quanto vogliamo dimostrare.

Resta infine da mostrare che $\{\Re(\varphi_t Z_t)\}_{t \in [0, T]} \in M^2[0, T]$ e $\{\Im(\varphi_t Z_t)\}_{t \in [0, T]} \in M^2[0, T]$. Dato che $|\Re(\varphi_t Z_t)| \leq |\varphi_t Z_t|$ e $\Im(\varphi_t Z_t) \leq |\varphi_t Z_t|$, basta osservare che

$$E_P \left(\int_0^T |\varphi_s Z_s|^2 ds \right) \leq E_P \left(Y^2 \int_0^T |\varphi_s|^2 ds \right) \leq C E_P(Y^2),$$

dove questa volta abbiamo posto $Y := \sup_{t \in [0, T]} |Z_t|$. Dobbiamo solo mostrare che $Y \in L^2$, ma questo è semplice: infatti

$$\begin{aligned} |Z_t(\varphi)| &= \left| e^{\int_0^t \varphi_s dB_s} \right| e^{-\frac{1}{2} \int_0^t \varphi_s^2 ds} = e^{\int_0^t \Re(\varphi_s) dB_s} e^{-\frac{1}{2} \int_0^t \Re(\varphi_s^2) ds} \\ &= Z_t(\Re(\varphi)) e^{\frac{1}{2} \int_0^t (\Re(\varphi_s)^2 - \Re(\varphi_s^2)) ds} = Z_t(\Re(\varphi)) e^{\frac{1}{2} \int_0^t \Im(\varphi_s)^2 ds} \leq Z_t(\Re(\varphi)) e^{\frac{1}{2} C}. \end{aligned}$$

Dato che $\int_0^T \Re(\varphi_s)^2 ds \leq \int_0^T |\varphi_s|^2 ds \leq C$ q.c. per ipotesi, segue dalla prima parte della dimostrazione che $\sup_{t \in [0, T]} Z_t(\Re(\varphi)) \in L^2$. \square

6.6.4. DIMOSTRAZIONE DEL TEOREMA 6.15. Per semplicità, dimostriamo solamente che il processo \tilde{B} rispetto a \mathbb{Q}_T è un moto browniano.[†] Dato che \tilde{B} è q.c. continuo, basta mostrare che le sue leggi finito-dimensionali rispetto a \mathbb{Q}_T coincidono con quelle del moto browniano. Dato che il moto browniano è un (il) processo gaussiano di media nulla e matrice delle covarianze $s \wedge t := \min\{s, t\}$, ci basta mostrare che per ogni $k \in \mathbb{N}$, $0 < t_1 < \dots < t_k \leq T$ e $\vartheta = (\vartheta_1, \dots, \vartheta_k) \in \mathbb{R}^k$ si ha

$$\mathbb{E}_{\mathbb{Q}_T} \left(e^{i \sum_{j=1}^k \vartheta_j \tilde{B}_{t_j}} \right) = e^{-\frac{1}{2} \sum_{j,l=1}^k (t_j \wedge t_l) \vartheta_j \vartheta_l}. \quad (6.40)$$

Ricordando la definizione (6.37) di \mathbb{Q}_T , possiamo riformulare (6.40) nel modo seguente:

$$\mathbb{E}_P \left(e^{i \sum_{j=1}^k \vartheta_j \tilde{B}_{t_j}} Z_T(\varphi) \right) = e^{-\frac{1}{2} \sum_{j,l=1}^k (t_j \wedge t_l) \vartheta_j \vartheta_l}. \quad (6.41)$$

Cominciamo a supporre che $\int_0^T \varphi_s^2 ds \leq C$ q.c., per un'opportuna costante $C < \infty$. Definendo il processo complesso $\tilde{\varphi}_s := \varphi_s + i \sum_{j=1}^k \vartheta_j \mathbf{1}_{[0, t_j)}(s)$, si deduce facilmente che $\int_0^T |\tilde{\varphi}_s|^2 ds \leq C'$ q.c., per un'opportuna costante $C' < \infty$. Possiamo quindi applicare il Lemma 6.17, che dà $\mathbb{E}_P(Z_T(\tilde{\varphi})) = 1$. Con calcoli elementari si ottiene

$$\begin{aligned} \int_0^T \tilde{\varphi}_s dB_s &= \int_0^T \varphi_s dB_s + i \sum_{j=1}^k \vartheta_j B_{t_j}, \\ \frac{1}{2} \int_0^T (\tilde{\varphi}_s)^2 ds &= \frac{1}{2} \int_0^T \varphi_s^2 ds - \frac{1}{2} \sum_{j,l=1}^k \vartheta_j \vartheta_l (t_j \wedge t_l) + i \sum_{j=1}^k \vartheta_j \int_0^{t_j} \varphi_s ds, \end{aligned}$$

da cui si ha

$$Z_T(\tilde{\varphi}) = Z_T(\varphi) e^{i \sum_{j=1}^k \vartheta_j \tilde{B}_{t_j}} e^{\frac{1}{2} \sum_{j,l=1}^k \vartheta_j \vartheta_l (t_j \wedge t_l)}.$$

La condizione $\mathbb{E}_P(Z_T(\tilde{\varphi})) = 1$ equivale dunque alla relazione (6.41).

Nel caso generale si procede per localizzazione. Definiamo

$$\tau_n := \inf \left\{ t \in [0, T] : \int_0^t \varphi_s^2 ds > n \right\},$$

con la convenzione $\inf\{\emptyset\} := +\infty$, e poniamo $\varphi_s^{(n)} := \varphi_s \mathbf{1}_{[0, \tau_n)}(s)$. Si osservi che $\int_0^T (\varphi_s^{(n)})^2 ds \leq n$ per costruzione. Per quanto appena dimostrato, la relazione (6.41) vale dunque con φ rimpiazzato da $\varphi^{(n)}$, per ogni $n \in \mathbb{N}$, cioè

$$\mathbb{E}_P \left(e^{i \sum_{j=1}^k \vartheta_j \tilde{B}_{t_j}^{(n)}} Z_T(\varphi^{(n)}) \right) = e^{-\frac{1}{2} \sum_{j,l=1}^k (t_j \wedge t_l) \vartheta_j \vartheta_l}, \quad (6.42)$$

[†]La dimostrazione che \tilde{B} è un $\{\mathcal{F}_t\}_{t \in [0, T]}$ -moto browniano non è molto più difficile: si veda il Teorema 7.22 (che richiama il Teorema 4.26 e l'Esercizio 3.11) in [Baldi, 2000].

dove abbiamo posto $\tilde{B}_t^{(n)} := B_t - \int_0^t \varphi_s^{(n)} ds$.

Per definizione di integrale stocastico in $M_{\text{loc}}^2[0, T]$, si ha $\int_0^T \varphi_s^{(n)} dB_s \rightarrow \int_0^T \varphi_s dB_s$ q.c. per $n \rightarrow \infty$. Dato che anche $\int_0^T (\varphi_s^{(n)})^2 ds \rightarrow \int_0^T \varphi_s^2 ds$ e $\int_0^T \varphi_s^{(n)} ds \rightarrow \int_0^T \varphi_s ds$ q.c. (per convergenza dominata), segue che $Z_T(\varphi^{(n)}) \rightarrow Z_T(\varphi)$ e $\tilde{B}_{t_j}^{(n)} \rightarrow \tilde{B}_{t_j}$ q.c.. Per concludere la dimostrazione ci basta dunque mostrare che si può passare al limite $n \rightarrow \infty$ in (6.42).

Per semplicità di notazioni, poniamo $Z_n := Z_T(\varphi^{(n)})$ e $\Theta_n := \sum_{j=1}^k \vartheta_j \tilde{B}_{t_j}^{(n)}$ e analogamente $Z := Z_T(\varphi)$ e $\Theta := \sum_{j=1}^k \vartheta_j \tilde{B}_{t_j}$. Possiamo scrivere

$$\begin{aligned} |\mathbb{E}_P(e^{i\Theta_n} Z_n) - \mathbb{E}_P(e^{i\Theta} Z)| &\leq \mathbb{E}_P(|e^{i\Theta_n} Z_n - e^{i\Theta} Z|) \\ &\leq \mathbb{E}_P(|e^{i\Theta_n} (Z_n - Z)|) + \mathbb{E}_P(|(e^{i\Theta_n} - e^{i\Theta}) Z|) \\ &\leq \mathbb{E}_P(|Z - Z_n|) + \mathbb{E}_P(|e^{i\Theta_n} - e^{i\Theta}| |Z|). \end{aligned}$$

Sappiamo che $\Theta_n \rightarrow \Theta$ q.c. per $n \rightarrow \infty$, quindi il secondo termine tende a zero per convergenza dominata. Per quanto riguarda il primo termine, sappiamo che $Z_n \rightarrow Z$ q.c. e inoltre $\mathbb{E}_P(Z_n) = \mathbb{E}_P(Z) = 1$ per ogni $n \in \mathbb{N}$. Applicando il Lemma 6.18 più in basso, si ottiene allora $\mathbb{E}_P(|Z - Z_n|) \rightarrow 0$ per $n \rightarrow \infty$.

Avendo mostrato che $\mathbb{E}_P(e^{i\Theta} Z) = \lim_{n \rightarrow \infty} \mathbb{E}_P(e^{i\Theta_n} Z_n) = \exp(-\frac{1}{2} \sum_{j,l=1}^k (t_j \wedge t_l) \vartheta_j \vartheta_l)$, la relazione (6.41) è verificata e questo conclude la dimostrazione. \square

LEMMA 6.18 (SCHEFFÉ). Siano $\{Z_n\}_{n \in \mathbb{N}}$, Z variabili aleatorie positive e integrabili tali che $Z_n \rightarrow Z$ q.c.. Allora $\mathbb{E}(|Z - Z_n|) \rightarrow 0$ se e solo se $\mathbb{E}(Z_n) \rightarrow \mathbb{E}(Z)$.

DIMOSTRAZIONE. Si noti che $(Z - Z_n)^+ \leq Z^+ = Z \in L^1$, perché $Z_n \geq 0$. Dato che $(Z - Z_n)^+ \rightarrow 0$ q.c. per $n \rightarrow \infty$, per convergenza dominata si ha $\mathbb{E}_P((Z - Z_n)^+) \rightarrow 0$.

Se supponiamo che $\mathbb{E}(Z_n) \rightarrow \mathbb{E}(Z)$ per $n \rightarrow \infty$, segue che $\mathbb{E}_P(Z - Z_n) \rightarrow 0$ e dato che $|x - y| = 2(x - y)^+ - (x - y)$ si ottiene

$$\mathbb{E}_P(|Z - Z_n|) = 2 \mathbb{E}_P((Z - Z_n)^+) - \mathbb{E}_P(Z - Z_n) \rightarrow 0.$$

Viceversa, per la disuguaglianza triangolare si ha $|\mathbb{E}(Z_n) - \mathbb{E}(Z)| \leq \mathbb{E}(|Z_n - Z|)$ e quindi se $\mathbb{E}_P(|Z - Z_n|) \rightarrow 0$ si ottiene $\mathbb{E}(Z_n) \rightarrow \mathbb{E}(Z)$. \square

6.6.5. LA FORMULA DI CAMERON-MARTIN. Introduciamo lo spazio $D[0, T]$ delle funzioni $f : [0, T] \rightarrow \mathbb{R}$ assolutamente continue, con $f(0) = 0$ e con derivata prima in L^2 :

$$D[0, T] := \left\{ f : [0, T] \rightarrow \mathbb{R} : \exists g \in L^2[0, T] : f(t) = \int_0^t g(s) ds, \forall t \in [0, T] \right\}$$

(scriveremo $g(s) = f'(s)$). Se $B = \{B_t\}_{t \in [0, T]}$ è un moto browniano reale e $f \in D[0, T]$, definiamo il processo (deterministico) $\varphi_s := f'(s)$ e poniamo

$$Z_T := e^{\int_0^T f'(s) dB_s - \frac{1}{2} \int_0^T f'(s)^2 ds}.$$

Grazie al Lemma 6.17 si ha $E(Z_T) = 1$, quindi possiamo applicare il Teorema 6.15: definita la probabilità $Q(d\omega) := Z_T(\omega) P(d\omega)$, il processo $\tilde{B} := B - f$, è un moto browniano rispetto a Q (si noti infatti che $\tilde{B}_t := B_t - \int_0^t \varphi_s ds = B_t - f(t)$).

Dato un qualunque sottoinsieme misurabile A dello spazio $C([0, T], \mathbb{R})$ delle funzioni continue da $[0, T]$ in \mathbb{R} , possiamo dunque scrivere

$$P(B \in A) = Q(\tilde{B} \in A) = E_P(1_{\{\tilde{B} \in A\}} Z_T).$$

Sostituendo A con $A - f$ e applicando la formula del cambio di variabili (Teorema 1.6), si ottiene la celebre *formula di Cameron-Martin*:

$$\begin{aligned} P(B + f \in A) &= E_P \left(1_{\{B \in A\}} e^{\int_0^T f'(s) dB_s - \frac{1}{2} \int_0^T f'(s)^2 ds} \right) \\ &= \int_{C([0, T], \mathbb{R})} 1_A(x) e^{\int_0^T f'(s) dx_s - \frac{1}{2} \int_0^T f'(s)^2 ds} \mathcal{W}(dx), \end{aligned} \quad (6.43)$$

dove \mathcal{W} indica la misura di Wiener, cioè la legge del moto browniano, su $C([0, T], \mathbb{R})$. Si noti che $\int_0^T f'(s) dx_s$ indica proprio l'integrale di Itô, ben definito per ν -q.o. $x \in C([0, T], \mathbb{R})$

Dato che l'esponenziale nel membro destro in (6.43) è q.c. strettamente positivo, se $P(B \in A) > 0$ allora anche $P(B \pm f \in A) > 0$, per ogni $f \in D[0, T]$. Ad esempio, posto $S_t := \sup_{s \in [0, t]} B_s$, per ogni $c > 0$ fissato si ha

$$P(B_t < c \text{ per ogni } t \leq T) = P(S_T < c) = P(|B_T| < c) > 0,$$

per il principio di riflessione. Di conseguenza, per ogni $f \in D[0, T]$ si ha

$$P(B_t < c + f(t) \text{ per ogni } t \leq T) > 0.$$

La formula di Cameron-Martin (6.43) mostra che, per ogni $T > 0$, la legge $\tilde{\mathcal{W}}$ del processo $\tilde{B} = B + f$ è assolutamente continua rispetto alla misura di Wiener \mathcal{W} , con densità di Radon-Nikodym data da

$$\frac{d\tilde{\mathcal{W}}}{d\mathcal{W}}(x) = \exp \left(\int_0^T f'(s) dx_s - \frac{1}{2} \int_0^T f'(s)^2 ds \right),$$

per \mathcal{W} -q.o. $x \in C([0, T], \mathbb{R})$. In altre parole, traslando la misura di Wiener \mathcal{W} mediante una funzione $f \in D[0, T]$, si ottiene una legge $\tilde{\mathcal{W}}$ assolutamente continua rispetto a \mathcal{W} . Si dice pertanto che la misura di Wiener è *quasi-invariante* per traslazioni in $D[0, T]$.

È interessante notare che invece la legge di $\sigma B = \{\sigma B_t\}_{t \in [0, T]}$, per $\sigma > 0$, $\sigma \neq 1$, non è assolutamente continua rispetto alla legge di B . Infatti, definendo il sottoinsieme

$$C_\sigma := \left\{ h \in C([0, T], \mathbb{R}) : \limsup_{t \downarrow 0} \frac{h(t)}{\sqrt{t} \sqrt{2 \log \log(1/t)}} = \sigma \right\},$$

per la legge del logaritmo iterato si ha $P(\sigma B \in C_\sigma) = 1$, mentre $P(B \in C_\sigma) = 0$.

7. EQUAZIONI DIFFERENZIALI STOCASTICHE

In questo capitolo ci concentriamo sulle equazioni differenziali stocastiche, dimostrando l'esistenza e l'unicità di soluzioni sotto ipotesi standard e discutendo brevemente alcune applicazioni alla teoria delle equazioni differenziali alle derivate parziali.

7.1. DEFINIZIONI

Siamo interessati alle equazioni differenziali stocastiche della forma

$$\begin{cases} dX_t = \sigma(t, X_t) dB_t + b(t, X_t) dt \\ X_0 = x \end{cases}, \quad (7.1)$$

dove sono assegnati il punto $x \in \mathbb{R}^n$ e le funzioni misurabili $b : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, $\sigma : [0, T] \times \mathbb{R}^n \rightarrow \mathcal{M}(n \times d, \mathbb{R})$. Indicheremo le componenti di b e σ con $b_i(t, x)$ e $\sigma_{ij}(t, x)$, per $1 \leq i \leq n$, $1 \leq j \leq d$.

Definiamo innanzitutto precisamente la nozione di soluzione.

DEFINIZIONE 7.1. Una soluzione dell'equazione (7.1) è il dato di:

- uno spazio filtrato standard $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0, T]}, P)$;
- un $\{\mathcal{F}_t\}_{t \in [0, T]}$ -moto browniano $B = \{B_t\}_{t \in [0, T]}$, definito su Ω a valori in \mathbb{R}^d ;
- un processo $X = \{X_t\}_{t \in [0, T]}$ definito su Ω a valori in \mathbb{R}^n , *q.c. continuo* e adattato alla filtrazione $\{\mathcal{F}_t\}_{t \in [0, T]}$, che soddisfa (7.1), cioè tale che

$$\{\sigma_{ij}(s, X_s)\}_{s \in [0, T]} \in M_{loc}^2[0, T], \quad \{b_i(s, X_s)\}_{s \in [0, T]} \in M_{loc}^1[0, T], \quad (7.2)$$

e per ogni $t \in [0, T]$

$$X_t = x + \int_0^t \sigma(s, X_s) dB_s + \int_0^t b(s, X_s) ds. \quad (7.3)$$

Una soluzione è detta *forte* se il processo X è adattato all'ampliamento standard $\{\mathcal{G}_{t+}\}_{t \in [0, T]}$ della filtrazione naturale del moto browniano.

Con abuso di notazione, scriveremo che un processo X è soluzione dell'equazione (7.1), sottintendendo lo spazio di probabilità, la filtrazione e il moto browniano. Scriviamo

esplicitamente l'equazione (7.3) in componenti: per ogni $i \in \{1, \dots, n\}$

$$X_t^{(i)} = x_i + \int_0^t \sum_{j=1}^d \sigma_{ij}(s, X_s) dB_s^{(j)} + \int_0^t b_i(s, X_s) ds.$$

Osserviamo che la condizione (7.2), necessaria affinché gli integrali in (7.3) abbiano senso, è automaticamente soddisfatta se le funzioni b e σ sono continue.

OSSERVAZIONE 7.2. L'esempio più semplice di equazione differenziale della forma (7.1) si ha quando $\sigma(t, x) \equiv \sigma \in \mathcal{M}(n \times d, \mathbb{R})$, e $b(t, x) \equiv b \in \mathbb{R}^n$. L'equazione diventa dunque $dX_t = \sigma dB_t + b dt$, che ha come soluzione il processo $X_t = x + \sigma B_t + b t$, detto moto browniano con *deriva* (o *drift*) $b \in \mathbb{R}^n$ e *matrice di dispersione* $\sigma \in \mathcal{M}(n \times d, \mathbb{R})$.

Più in generale, quando $\sigma(t, x) = \sigma(x)$ e $b(t, x) = b(x)$ non dipendono dal tempo, un processo X soluzione dell'equazione (7.1) è detto *diffusione*. Intuitivamente, una diffusione è un processo che si comporta localmente come un moto browniano con drift e matrice di dispersione: $X_{t+\varepsilon} - X_t \approx \sigma(X_t)(B_{t+\varepsilon} - B_t) + b(X_t)\varepsilon + o(\varepsilon)$. \square

Noi ci concentreremo sull'*esistenza (e unicità) di soluzioni forti*. Più precisamente, per ogni spazio filtrato standard $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0, T]}, \mathbb{P})$ su cui è definito un $\{\mathcal{F}_t\}_{t \in [0, T]}$ -moto browniano $B = \{B_t\}_{t \in [0, T]}$ a valori in \mathbb{R}^d , mostreremo che — sotto opportune ipotesi su $b(t, x)$ e $\sigma(t, x)$ — esiste un processo $X = \{X_t\}_{t \in [0, T]}$, definito su Ω e adattato all'ampliamento standard della filtrazione naturale di B , che è soluzione (forte) dell'equazione (7.1), e che tale processo è unico a meno di indistinguibilità.

Notiamo tuttavia che esistono equazioni differenziali stocastiche per cui esistono soluzioni ma che non ammettono soluzioni forti: in altre parole, una soluzione X deve necessariamente “contenere più aleatorietà” di quella del moto browniano. Un esempio celebre è dato dall'*equazione di Tanaka*: $dX_t = \text{sign}(X_t) dB_t$, dove $\text{sign}(x) := 1$ se $x > 0$ e $\text{sign}(x) = -1$ se $x \leq 0$ (si veda l'Esempio 3.5 nel capitolo 5 in Karatzas e Shreve [1998]).

Introduciamo infine alcune nozioni di unicità per l'equazione (7.1).

DEFINIZIONE 7.3. Diremo che per l'equazione (7.1) c'è *unicità in legge* se, quali che siano le soluzioni $\{(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0, T]}, \mathbb{P}), B, X\}$, $\{(\Omega', \mathcal{F}', \{\mathcal{F}'_t\}_{t \in [0, T]}, \mathbb{P}'), B', X'\}$, i processi X e X' hanno la stessa legge su $C([0, T], \mathbb{R}^n)$ (equivalentemente, se hanno le stesse leggi finito-dimensionali).

Diremo che per l'equazione (7.1) c'è *unicità per traiettorie* se, quali che siano le soluzioni X e X' , definite sullo stesso spazio filtrato standard $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0, T]}, \mathbb{P})$ e con lo stesso moto browniano B , i processi X e X' sono indistinguibili, cioè si ha $\mathbb{P}(X_t = X'_t \forall t \in [0, T]) = 1$.

È possibile mostrare che l'unicità per traiettorie implica l'unicità in legge: si veda la Proposizione 3.20 nel capitolo 5 in Karatzas e Shreve [1998].

7.2. ESISTENZA E UNICITÀ DI SOLUZIONI FORTI

Dimostreremo ora l'esistenza di soluzioni forti e l'unicità per traiettorie per l'equazione differenziale stocastica (7.1), sotto opportune condizioni.

IPOTESI 7.4. Le funzioni $b : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ e $\sigma : [0, T] \times \mathbb{R}^n \rightarrow \mathcal{M}(n \times d, \mathbb{R})$ sono misurabili ed esistono costanti $L, M < \infty$ tali che per ogni $t \in [0, T]$ e $x, y \in \mathbb{R}^n$

$$|b(t, x)|^2 \leq M(1 + |x|^2), \quad |\sigma(t, x)|^2 \leq M(1 + |x|^2), \quad (7.4)$$

$$|b(t, x) - b(t, y)| \leq L|x - y|, \quad |\sigma(t, x) - \sigma(t, y)| \leq L|x - y|. \quad (7.5)$$

Naturalmente per $x \in \mathbb{R}^n$ indichiamo con $|x|$ la norma euclidea, e analogamente per $x \in \mathcal{M}(n \times d, \mathbb{R})$ poniamo $|x|^2 := \sum_{i=1}^n \sum_{j=1}^d (x_{ij})^2$.

TEOREMA 7.5. Siano b, σ funzioni che soddisfano l'Ipotesi 7.4 e sia $x \in \mathbb{R}^n$. Allora:

- per l'equazione (7.1) c'è *unicità per traiettorie*;
- per l'equazione (7.1) c'è *esistenza di soluzioni forti*: più precisamente, per ogni spazio filtrato standard $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0, T]}, \mathbb{P})$, su cui è definito un $\{\mathcal{F}_t\}_{t \in [0, T]}$ -moto browniano d -dimensionale B , esiste un processo $X = \{X_t\}_{t \in [0, T]}$ definito su Ω a valori in \mathbb{R}^n che è soluzione forte dell'equazione (7.1).
- ogni soluzione $X = \{X_t\}_{t \in [0, T]}$ dell'equazione (7.1) è in $M^2[0, T]$.

Si noti che, fissato lo spazio $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0, T]}, \mathbb{P})$, la soluzione forte dell'equazione (7.1) costruita nel Teorema 7.5 è necessariamente unica (a meno di indistinguibilità), grazie all'unicità per traiettorie. Per questa ragione si dice che, sotto l'Ipotesi 7.4, per l'equazione (7.1) c'è *esistenza e unicà di soluzioni forti*.

Un'altra conseguenza immediata del Teorema 7.5 è che, sotto l'Ipotesi 7.4, *ogni soluzione dell'equazione (7.1) è forte*. Consideriamo infatti una soluzione generica, cioè un processo X' , definito su uno spazio filtrato standard $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0, T]}, \mathbb{P})$ su cui è definito un $\{\mathcal{F}_t\}_{t \in [0, T]}$ -moto browniano d -dimensionale B , che risolve l'equazione (7.1). Grazie al Teorema 7.5, esiste un processo X definito su Ω che è soluzione *forte* dell'equazione, cioè che è adattato all'ampliamento standard $\{\bar{\mathcal{G}}_{t+}\}_{t \in [0, T]}$ della filtrazione naturale di B . Per l'unicità per traiettorie, X è indistinguibile da X' , pertanto anche X' è adattato alla filtrazione $\{\bar{\mathcal{G}}_{t+}\}_{t \in [0, T]}$ e dunque anche X' è soluzione forte.

OSSERVAZIONE 7.6. Il dato iniziale $x \in \mathbb{R}^n$ dell'equazione (7.1) può essere sostituito da una variabile aleatoria $X \in L^2$ che sia *indipendente da* B . La dimostrazione dell'esistenza di una soluzione (adattata alla filtrazione $\{\sigma(X, \bar{\mathcal{G}}_{t+})\}_{t \in [0, T]}$) e dell'unicità per traiettorie in questo caso può essere ottenuta apportando minime variazioni alla dimostrazione che presentiamo. \square

Per la dimostrazione sarà utile il classico Lemma di Gronwall.

LEMMA 7.7 (GRONWALL). Sia $g : [0, T] \rightarrow \mathbb{R}$ una funzione *misurabile e limitata*, che soddisfa la seguente relazione: esistono $a, b \geq 0$ tali che

$$g(t) \leq a + b \int_0^t g(s) ds, \quad \forall t \in [0, T]. \quad (7.6)$$

Allora per ogni $t \in [0, T]$ vale che $g(t) \leq a e^{bt}$.

DIMOSTRAZIONE. Mostriamo che per ogni $n \in \mathbb{N}$ vale la seguente relazione:

$$g(t) \leq a \sum_{k=0}^n \frac{(bt)^k}{k!} + b^{n+1} \int_0^t ds_1 \int_0^{s_1} ds_2 \cdots \int_0^{s_n} ds_{n+1} g(s_{n+1}). \quad (7.7)$$

Il caso $n = 0$ non è altro che la relazione (7.6). Per il passo induttivo, basta notare che $g(s_{n+1}) \leq a + b \int_0^{s_{n+1}} ds_{n+2} g(s_{n+2})$, sempre per la relazione (7.6), ottenendo

$$\begin{aligned} & b^{n+1} \int_0^t ds_1 \int_0^{s_1} ds_2 \cdots \int_0^{s_n} ds_{n+1} g(s_{n+1}) \\ & \leq b^{n+1} a \frac{t^{n+1}}{(n+1)!} + b^{n+2} \int_0^t ds_1 \int_0^{s_1} ds_2 \cdots \int_0^{s_n} ds_{n+1} \int_0^{s_{n+1}} ds_{n+2} g(s_{n+2}), \end{aligned}$$

dove abbiamo usato l'integrale elementare $\int_0^t ds_1 \int_0^{s_1} ds_2 \cdots \int_0^{s_n} ds_{n+1} = \frac{t^{n+1}}{(n+1)!}$. Questo mostra che la relazione (7.7) vale per ogni $n \in \mathbb{N}$.

Dato che per ipotesi $|g(x)| \leq M$ con $M < \infty$, l'ultimo termine in (7.7) è limitato in valore assoluto da $M b^{n+1} \frac{t^{n+1}}{(n+1)!}$ e quindi tende a zero per $n \rightarrow \infty$. Prendendo dunque il limite $n \rightarrow \infty$ in (7.7) si ottiene $g(t) \leq a e^{bt}$. \square

Passiamo ora alla dimostrazione del Teorema 7.5. Per semplificare le notazioni, considereremo solo il caso unidimensionale $d = n = 1$, in particolare b e σ sono funzioni da $[0, T] \times \mathbb{R}$ in \mathbb{R} . Trattiamo separatamente l'unicità e l'esistenza.

7.2.1. TEOREMA 7.5: UNICITÀ. Siano fissati $x \in \mathbb{R}$ e uno spazio di probabilità (Ω, \mathcal{F}, P) , munito di filtrazione $\{\mathcal{F}_t\}_{t \in [0, T]}$ che soddisfa le ipotesi standard, su cui è definito un $\{\mathcal{F}_t\}_{t \in [0, T]}$ -moto browniano reale B . Faremo uso frequente della relazione $(x_1 + \dots + x_n)^2 \leq n(x_1^2 + \dots + x_n^2)$, che si dimostra facilmente:

$$(x_1 + \dots + x_n)^2 = \sum_{i,j=1}^n x_i x_j \leq \sum_{i,j=1}^n \frac{x_i^2 + x_j^2}{2} = n(x_1^2 + \dots + x_n^2),$$

avendo usato la disuguaglianza $ab \leq \frac{a^2+b^2}{2}$, che segue da $(a-b)^2 \geq 0$.

Supponiamo che $X = \{X_t\}_{t \in [0, T]}$ e $X' = \{X'_t\}_{t \in [0, T]}$ siano due processi reali definiti su Ω , entrambi soluzioni dell'equazione (7.1), cioè per ogni $t \in [0, T]$

$$X_t = x + \int_0^t \sigma(s, X_s) dB_s + \int_0^t b(s, X_s) ds,$$

e analogamente per X' . Se introduciamo il tempo d'arresto

$$\tau_n := \inf \{t \geq 0 : |X_t| \geq n \text{ o } |X'_t| \geq n\}, \quad (7.8)$$

con la convenzione $\inf \emptyset := +\infty$, possiamo dunque scrivere per ogni $t \in [0, T]$

$$\begin{aligned} X_{t \wedge \tau_n} &= x + \int_0^{t \wedge \tau_n} \sigma(s, X_s) dB_s + \int_0^{t \wedge \tau_n} b(s, X_s) ds \\ &= x + \int_0^t \sigma(s, X_s) 1_{[0, \tau_n)}(s) dB_s + \int_0^t b(s, X_s) 1_{[0, \tau_n)}(s) ds, \end{aligned}$$

e analogamente per X' . Sottraendo le relazioni per X e X' si ottiene dunque

$$\begin{aligned} E[(X_{t \wedge \tau_n} - X'_{t \wedge \tau_n})^2] &\leq 2 \left(E \left[\left(\int_0^t (\sigma(s, X_s) - \sigma(s, X'_s)) 1_{[0, \tau_n)}(s) dB_s \right)^2 \right] \right. \\ &\quad \left. + E \left[\left(\int_0^t (b(s, X_s) - b(s, X'_s)) 1_{[0, \tau_n)}(s) ds \right)^2 \right] \right) \\ &\leq 2 \left(E \left[\int_0^t (\sigma(s, X_s) - \sigma(s, X'_s))^2 1_{[0, \tau_n)}(s) ds \right] \right. \\ &\quad \left. + T E \left[\int_0^t (b(s, X_s) - b(s, X'_s))^2 1_{[0, \tau_n)}(s) ds \right] \right), \end{aligned}$$

dove abbiamo usato la disuguaglianza di Cauchy-Schwarz per l'integrale ordinario e la seguente relazione per l'integrale stocastico:

$$E \left[\left(\int_0^t Z_s dB_s \right)^2 \right] \leq E \left[\int_0^t Z_s^2 ds \right], \quad (7.9)$$

valida per ogni processo $Z = \{Z_s\}_{s \in [0, t]} \in M_{\text{loc}}^2[0, t]$. In effetti, se $Z \in M^2[0, t]$ questa relazione è un'uguaglianza, per l'isometria dell'integrale stocastico, mentre se $Z \in M_{\text{loc}}^2[0, t] \setminus M^2[0, t]$ il membro destro vale $+\infty$ e la disuguaglianza è banalmente verificata. Usando l'Ipotesi 7.4 abbiamo dunque

$$\begin{aligned} E[(X_{t \wedge \tau_n} - X'_{t \wedge \tau_n})^2] &\leq 2L^2(1+T) \int_0^t E[(X_s - X'_s)^2 1_{[0, \tau_n)}(s)] ds \\ &\leq 2L^2(1+T) \int_0^t E[(X_{s \wedge \tau_n} - X'_{s \wedge \tau_n})^2] ds. \end{aligned}$$

Abbiamo dunque mostrato che la funzione $g(t) := E[(X_{t \wedge \tau_n} - X'_{t \wedge \tau_n})^2]$ soddisfa la relazione $g(t) \leq a + b \int_0^t g(s) ds$, con $a = 0$ e $b = 2L^2(1+T)$. Si osservi che g è misurabile, perché i processi X e X' , essendo continui e adattati, sono (progressivamente) misurabili. Inoltre g è limitata: $g(t) \leq 2(E[X_{t \wedge \tau_n}^2] + E[X'_{t \wedge \tau_n}^2]) \leq 4n^2$, grazie alla definizione (7.8) di τ_n . Applicando il Lemma 7.7 si ottiene dunque $g(t) \leq 0$ per ogni $t \in [0, T]$. Dato che chiaramente $g(t) = E[(X_{t \wedge \tau_n} - X'_{t \wedge \tau_n})^2] \geq 0$, segue che $g(t) = 0$ per ogni $t \in [0, T]$.

Questo mostra che, per ogni $t \in [0, T]$ fissato, si ha $X_{t \wedge \tau_n}(\omega) = X'_{t \wedge \tau_n}(\omega)$ per ogni $n \in \mathbb{N}$ e per q.o. $\omega \in \Omega$. Dato che $\lim_{n \rightarrow \infty} \tau_n = \infty$ q.c., per q.o. ω si ha $t \wedge \tau_n(\omega) = t$ per n grande e dunque $X_t(\omega) = X'_t(\omega)$. Si ha quindi q.c. $X_t = X'_t$ per ogni $t \in [0, T] \cap \mathbb{Q}$ e per la continuità delle traiettorie di X e X' segue che q.c. $X_t = X'_t$ per ogni $t \in [0, T]$, cioè i processi X e X' sono indistinguibili. L'unicità per traiettorie è dunque dimostrata. \square

7.2.2. TEOREMA 7.5: ESISTENZA. Siano fissati $x \in \mathbb{R}$ e uno spazio di probabilità (Ω, \mathcal{F}, P) , su cui è definito un moto browniano reale $B = \{B_t\}_{t \geq 0}$. Scegliamo come filtrazione l'ampliamento standard $\{\bar{\mathcal{G}}_{t+}\}_{t \in [0, T]}$ della filtrazione naturale del moto browniano. Costruiremo su questo spazio un processo X continuo e adattato che è soluzione (forte) dell'equazione differenziale stocastica (7.1) e mostreremo che $X \in M^2[0, T]$. Il processo X sarà ottenuto con un procedimento di iterazione, analogo a quanto avviene per le equazioni differenziali ordinarie.

Se sullo spazio (Ω, \mathcal{F}, P) è già definita una filtrazione $\{\mathcal{F}_t\}_{t \in [0, T]}$ che soddisfa le ipotesi standard e B è un $\{\mathcal{F}_t\}_{t \in [0, T]}$ -moto browniano reale, il nostro procedimento produce una soluzione X adattata a $\{\bar{\mathcal{G}}_{t+}\}_{t \in [0, T]}$, dunque a maggior ragione adattata a $\{\mathcal{F}_t\}_{t \in [0, T]}$, dal momento che $\bar{\mathcal{G}}_{t+} \subseteq \mathcal{F}_t$. Avendo già dimostrato l'unicità per traiettorie, segue che ogni altra soluzione Y definita su Ω , a priori adattata a $\{\mathcal{F}_t\}_{t \in [0, T]}$, è indistinguibile da X , e dunque è anch'essa adattata a $\{\bar{\mathcal{G}}_{t+}\}_{t \in [0, T]}$. Questo mostra che, sotto l'Ipotesi 7.4, qualunque soluzione dell'equazione differenziale stocastica (7.1) è una soluzione forte.

Per $Y \in M^2[0, T]$, definiamo il processo $J(Y) = \{J_t(Y)\}_{t \in [0, T]}$ ponendo

$$J_t(Y) := x + \int_0^t \sigma(s, Y_s) dB_s + \int_0^t b(s, Y_s) ds. \quad (7.10)$$

Grazie all'Ipotesi 7.4 è immediato verificare che entrambi i processi $\{\sigma(s, Y_s)\}_{s \in [0, T]}$ e $\{b(s, Y_s)\}_{s \in [0, T]}$ sono in $M^2[0, T]$, per cui gli integrali in (7.10) sono ben definiti. In effetti, applicando la relazione (7.9), la disuguaglianza di Cauchy-Schwarz e l'Ipotesi 7.4, per $t \in [0, T]$ possiamo scrivere

$$\begin{aligned} E[J_t(Y)^2] &\leq 3 \left(x^2 + E \left[\left(\int_0^t \sigma(s, Y_s) dB_s \right)^2 \right] + E \left[\left(\int_0^t b(s, Y_s) ds \right)^2 \right] \right) \\ &\leq 3 \left(x^2 + E \left[\int_0^t \sigma(s, Y_s)^2 ds \right] + T E \left[\int_0^t b(s, Y_s)^2 ds \right] \right) \\ &\leq 3 \left(x^2 + M(1+T) \int_0^t (1 + E[Y_s^2]) ds \right) \\ &\leq 3 \left(x^2 + M(1+T)T + M(1+T) \|Y\|_{M^2[0, T]}^2 \right). \end{aligned}$$

Dato che l'ultimo membro è finito e non dipende da t , integrando questa relazione si ottiene che $\|J(Y)\|_{M^2[0, T]}^2 = \int_0^T E[J_t(Y)^2] dt < \infty$. Questo mostra che $J(Y) \in M^2[0, T]$, per ogni $Y \in M^2[0, T]$, ovvero J è un operatore da $M^2[0, T]$ in sé.

Usando ancora la relazione (7.9), la disuguaglianza di Cauchy-Schwarz e l'Ipotesi 7.4, per $Y, Y' \in M^2[0, T]$ e per $t \in [0, T]$ abbiamo

$$\begin{aligned} & \mathbb{E} [(J_t(Y) - J_t(Y'))^2] \\ & \leq 2 \left(\mathbb{E} \left[\left(\int_0^t (\sigma(s, Y_s) - \sigma(s, Y'_s)) dB_s \right)^2 \right] + \mathbb{E} \left[\left(\int_0^t (b(s, Y_s) - b(s, Y'_s)) ds \right)^2 \right] \right) \\ & \leq 2 \left(\mathbb{E} \left[\int_0^t (\sigma(s, Y_s) - \sigma(s, Y'_s))^2 ds \right] + T \mathbb{E} \left[\int_0^t (b(s, Y_s) - b(s, Y'_s))^2 ds \right] \right) \\ & \leq 2(1+T) L^2 \mathbb{E} \left[\int_0^t (Y_s - Y'_s)^2 ds \right]. \end{aligned}$$

Ponendo $C := 2(1+T) L^2$, abbiamo mostrato che per ogni $t \in [0, T]$

$$\mathbb{E} [(J_t(Y) - J_t(Y'))^2] \leq C \int_0^t \mathbb{E} [(Y_s - Y'_s)^2] ds. \quad (7.11)$$

Questa relazione ci tornerà utile tra poco. Per il momento, notiamo che il membro destro di (7.11) è limitato da $C \|Y - Y'\|_{M^2[0, T]}^2$, per cui integrando la relazione si ottiene $\|J(Y) - J(Y')\|_{M^2[0, T]}^2 \leq C T \|Y - Y'\|_{M^2[0, T]}^2$. Questo mostra che $J : M^2[0, T] \rightarrow M^2[0, T]$ è un operatore continuo.

Definiamo ora ricorsivamente una successione di processi $X^{(n)} = \{X_t^{(n)}\}_{t \in [0, T]} \in M^2[0, T]$, ponendo $X^{(1)} \equiv x$ e $X^{(n+1)} := J(X^{(n)})$, per ogni $n \in \mathbb{N}$. Più esplicitamente, per $t \in [0, T]$ e $n \in \mathbb{N}$

$$X_t^{(1)} \equiv x, \quad X_t^{(n+1)} := J_t(X^{(n)}) = x + \int_0^t \sigma(s, X_s^{(n)}) dB_s + \int_0^t b(s, X_s^{(n)}) ds.$$

Applicando la relazione (7.11), per ogni $n \geq 2$ e $t \in [0, T]$ si ha

$$\mathbb{E} [(X_t^{(n+1)} - X_t^{(n)})^2] \leq C \int_0^t \mathbb{E} [(X_s^{(n)} - X_s^{(n-1)})^2] ds \quad (7.12)$$

Per il caso $n = 1$, usando ancora la relazione (7.9), la disuguaglianza di Cauchy-Schwarz e l'Ipotesi 7.4, si ha

$$\begin{aligned} \mathbb{E} [(X_t^{(2)} - X_t^{(1)})^2] & \leq 2 \left(\mathbb{E} \left[\left(\int_0^t \sigma(s, x) dB_s \right)^2 \right] + \left(\int_0^t b(s, x) ds \right)^2 \right) \\ & \leq 2 \left(\int_0^t \sigma(s, x)^2 ds + T \int_0^t b(s, x)^2 ds \right) \leq 2M(1+T)T(1+|x|^2) =: c, \end{aligned} \quad (7.13)$$

e chiaramente $c < \infty$. Usando le relazioni (7.12) e (7.13), è immediato verificare per induzione che vale la seguente stima, per ogni $t \in [0, T]$ e $n \in \mathbb{N}$:

$$\mathbb{E} [(X_t^{(n+1)} - X_t^{(n)})^2] \leq c C^{n-1} \frac{t^{n-1}}{(n-1)!},$$

che integrata in $t \in [0, T]$ dà

$$\|X^{(n+1)} - X^{(n)}\|_{M^2[0,T]}^2 \leq c C^{n-1} \frac{T^n}{n!}.$$

Questa relazione mostra che la successione di processi $\{X^{(n)}\}_{n \in \mathbb{N}}$ è di Cauchy in $M^2[0, T]$: infatti per $m > n$ possiamo scrivere

$$\|X^{(m)} - X^{(n)}\|_{M^2[0,T]} \leq \sum_{k=n}^{m-1} \|X^{(k+1)} - X^{(k)}\|_{M^2[0,T]} \leq \sum_{k=n}^{\infty} \left(c C^{k-1} \frac{T^k}{k!} \right)^{1/2},$$

e la serie può essere resa piccola a piacere scegliendo n grande, perché è convergente (esercizio). Per la completezza di $M^2[0, T]$, esiste dunque un processo $X \in M^2[0, T]$ tale che $X^{(n)} \rightarrow X$ in $M^2[0, T]$.

Mostriamo infine che X è soluzione di (7.1). Per costruzione $X^{(n+1)} := J(X^{(n)})$, per ogni $n \in \mathbb{N}$. Dato che $X^{(n)} \rightarrow X$ in $M^2[0, T]$, prendendo il limite $n \rightarrow \infty$ e usando il fatto che $J : M^2[0, T] \rightarrow M^2[0, T]$ è un operatore continuo si ottiene la relazione $X = J(X)$, ovvero

$$X_t = x + \int_0^t \sigma(s, X_s) dB_s + \int_0^t b(s, X_s) ds, \quad (7.14)$$

che non è altro che la forma integrale dell'equazione (7.1).

Per costruzione, X è stato determinato come elemento di $M^2[0, T]$, cioè come classe di equivalenza di processi. Occorre dimostrare che è possibile scegliere un vero processo X , cioè un rappresentante della classe di equivalenza, che sia continuo (come prescritto dalla Definizione 7.1). Ma la relazione (7.14) mostra X è somma di un integrale stocastico e di un integrale ordinario, dunque ne esiste sempre una versione continua: questa versione di X costituisce dunque una soluzione dell'equazione (7.1).

Dato che fin dall'inizio abbiamo scelto come filtrazione il completamento $\{\bar{\mathcal{G}}_{t+}\}_{t \in [0, T]}$ della filtrazione naturale del moto browniano, il processo X ottenuto è una soluzione forte dell'equazione (7.1). \square

7.3. LA FORMULA DI FEYNMAN-KAC

Siano assegnate le funzioni $b : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, $\sigma : [0, T] \times \mathbb{R}^n \rightarrow \mathcal{M}(n \times d, \mathbb{R})$, che supporremo continue e soddisfacenti la condizione di crescita lineare (7.4). Facciamo l'ipotesi che, per ogni $s \in [0, T]$ e per ogni $x \in \mathbb{R}^n$, esista una soluzione $X = \{X_t\}_{t \in [s, T]}$ dell'equazione differenziale stocastica

$$\begin{cases} dX_t = \sigma(t, X_t) dB_t + b(t, X_t) dt \\ X_s = x \end{cases}, \quad (7.15)$$

definita su uno spazio filtrato standard $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [s, T]}, \mathbb{P})$ su cui è definito un $\{\mathcal{F}_t\}_{t \in [s, T]}$ -moto browniano $B = \{B_t\}_{t \in [s, T]}$ a valori in \mathbb{R}^d . Più esplicitamente,

$$X_t = x + \int_s^t \sigma(u, X_u) dB_u + \int_s^t b(u, X_u) du, \quad \forall t \in [s, T]. \quad (7.16)$$

Supponiamo inoltre che ci sia unicità in legge per questa equazione. Per quanto visto, *una condizione sufficiente è che le funzioni b, σ soddisfino l'Ipotesi 7.4* (nel qual caso c'è unicità per traiettorie e la soluzione X dell'equazione (7.15) è in $M^2[s, T]$).

Naturalmente la soluzione X dipenderà dal tempo iniziale s e dal dato iniziale x . Con un piccolo abuso di notazione, è prassi indicare questa dipendenza nella probabilità: scriveremo cioè $P_{s,x}(X_t \in A)$ e $E_{s,x}(F(X_t))$ per calcolare probabilità e valori attesi relativi al processo $X = \{X_t\}_{t \in [s, T]}$ soluzione dell'equazione (7.15), (7.16).

Introduciamo ora, per ogni $t \in [0, T]$, un operatore differenziale L_t del secondo ordine, associato all'equazione (7.15). L'operatore L_t agisce sulle funzioni $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ di classe C^2 nel modo seguente:

$$\begin{aligned} L_t \varphi(x) &:= \frac{1}{2} \text{Tr}(\varphi''(x) \sigma(t, x) \sigma(t, x)^*) + b(t, x) \cdot \varphi'(x) \\ &= \frac{1}{2} \sum_{i,j=1}^n (\sigma \sigma^*)_{ij}(t, x) \frac{\partial^2}{\partial x_i \partial x_j} \varphi(x) + \sum_{i=1}^n b_i(t, x) \frac{\partial}{\partial x_i} \varphi(x), \end{aligned}$$

dove naturalmente $(\sigma \sigma^*)_{ij}(t, x) := \sum_{k=1}^d \sigma_{ik}(t, x) \sigma_{jk}(t, x)$.

Consideriamo infine la seguente equazione differenziale alle derivate parziali, per una funzione $F = F(t, x) : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$ di classe C^1 in t e C^2 in x :

$$\begin{cases} -\frac{\partial}{\partial t} F(t, x) = L_t F(t, x) - k(t, x) F(t, x) + g(t, x) & \forall t \in [0, T], x \in \mathbb{R}^n \\ F(T, x) = f(x) & \forall x \in \mathbb{R}^n \end{cases}, \quad (7.17)$$

dove $g(t, x), f(x) \in \mathbb{R}$ e $k(t, x) \geq 0$ sono funzioni *continue* assegnate. Si noti che si specifica il valore di $F(t, x)$ all'istante *finale* $t = T$.

Facciamo ora l'ipotesi che le funzioni $g(t, x)$ e $f(x)$ sono a crescita polinomiale in x , uniformemente in t . Si può allora dimostrare che se esiste una soluzione $F(t, x)$ dell'equazione (7.17) a crescita polinomiale in x uniformemente in t , essa è unica e ammette la seguente formula di rappresentazione, nota come *formula di Feynman-Kac*:

$$F(t, x) = E_{t,x} \left[f(X_T) e^{-\int_t^T k(u, X_u) du} + \int_t^T g(s, X_s) e^{-\int_t^s k(u, X_u) du} ds \right]. \quad (7.18)$$

È anche possibile dare condizioni sotto le quali la funzione definita da (7.18) è effettivamente soluzione dell'equazione (7.17): per maggiori dettagli, si vedano il Teorema 7.6 e il Remark 7.8 nel capitolo 5 in [Karatzas e Shreve, 1998] (in cui sono anche descritte precisamente le ipotesi di crescita polinomiale su g, f e F).

Per semplicità, ci limiteremo a fornire una dimostrazione della relazione (7.18) sotto ipotesi più forti: *supporremo che le derivate $\frac{\partial}{\partial x_i} F(t, x)$ siano limitate su $[0, T] \times \mathbb{R}^n$ e che la soluzione X dell'equazione (7.15) sia in $M^2[s, T]$* . Il caso generale si ottiene con un procedimento di localizzazione: per maggiori dettagli, si veda la dimostrazione del Teorema 7.6 nel capitolo 5 in [Karatzas e Shreve, 1998].

DIMOSTRAZIONE DELLA FORMULA (7.18). Supponiamo che F sia una soluzione dell'equazione (7.17) e fissiamo $t \in [0, T]$. Per $s \in [t, T]$, la formula di Itô dà

$$dF(s, X_s) = (\dot{F}(s, X_s) + L_s F(s, X_s)) ds + F'(s, X_s) \sigma(s, X_s) dB_s.$$

Dato che $d(e^{-\int_t^s k(u, X_u) du}) = -k(s, X_s) e^{-\int_t^s k(u, X_u) du} ds$, per la formula di integrazione per parti stocastica (Corollario 6.12) si ha

$$\begin{aligned} & d\left(F(s, X_s) e^{-\int_t^s k(u, X_u) du}\right) \\ &= F(s, X_s) d\left(e^{-\int_t^s k(u, X_u) du}\right) + e^{-\int_t^s k(u, X_u) du} dF(s, X_s) \\ &= e^{-\int_t^s k(u, X_u) du} \left\{ (-k F + \dot{F} + L_s F)(s, X_s) ds + F'(s, X_s) \sigma(s, X_s) dB_s \right\} \\ &= -e^{-\int_t^s k(u, X_u) du} g(s, X_s) ds + e^{-\int_t^s k(u, X_u) du} F'(s, X_s) \sigma(s, X_s) dB_s, \end{aligned} \quad (7.19)$$

avendo usato il fatto che F è soluzione dell'equazione (7.17).

Per ipotesi $k(u, x) \geq 0$ e $|F'(u, x)| \leq C$ per ogni $(u, x) \in [0, T] \times \mathbb{R}^n$, per un'opportuna costante $C < \infty$, e inoltre $|\sigma(s, x)| \leq M(1 + |x|)$, quindi

$$\left| e^{-\int_t^s k(u, X_u) du} F'(s, X_s) \sigma(s, X_s) \right| \leq (cost.)(1 + |X_s|).$$

Questo mostra che il processo $\{e^{-\int_t^s k(u, X_u) du} F'(s, X_s) \sigma(s, X_s)\}_{s \in [t, T]}$ è in $M^2[s, T]$, perché per ipotesi $X \in M^2[s, T]$. Quindi l'integrale stocastico dato dall'ultimo termine in (7.19) è una vera martingala (di quadrato integrabile). Integrando la relazione (7.19) tra $s = t$ e $s = T$ e prendendo il valore atteso $E_{t,x}$, dato che la martingala ha valore atteso nullo si ottiene

$$\begin{aligned} & E_{t,x} \left(F(T, X_T) e^{-\int_t^T k(u, X_u) du} \right) - E_{t,x}(F(t, X_t)) \\ &= -E_{t,x} \left(\int_t^T g(s, X_s) e^{-\int_t^s k(u, X_u) du} ds \right). \end{aligned}$$

La formula (7.18) segue osservando che $E_{t,x}(F(t, X_t)) = F(t, x)$, per definizione di $E_{t,x}$, e $F(T, X_T) = f(X_T)$, grazie all'equazione (7.17). \square

7.3.1. IL CASO INDIPENDENTE DAL TEMPO. Nel caso particolare in cui le funzioni $\sigma(t, x) = \sigma(x)$, $b(t, x) = b(x)$, $g(t, x) = g(x)$ e $k(t, x) = k(x)$ non dipendano esplicitamente dal tempo, anche l'operatore $L_t = L$ non dipende dal tempo ed è dato da

$$L\varphi(x) = \frac{1}{2} \sum_{i,j=1}^n (\sigma\sigma^*)_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} \varphi(x) + \sum_{i=1}^n b_i(x) \frac{\partial}{\partial x_i} \varphi(x).$$

Se $F(t, x)$ è soluzione dell'equazione (7.17), la funzione $\Psi(t, x) := F(T - t, x)$, definita per $(t, x) \in [0, T] \times \mathbb{R}^n$, è soluzione dell'equazione

$$\begin{cases} \frac{\partial}{\partial t} \Psi(t, x) = L\Psi(t, x) - k(x) \Psi(t, x) + g(x) & \forall t \in [0, T], x \in \mathbb{R}^n \\ \Psi(0, x) = f(x) & \forall x \in \mathbb{R}^n \end{cases}. \quad (7.20)$$

Se siamo nelle condizioni in cui la formula di Feynman-Kac (7.18) definisce una soluzione dell'equazione (7.17), la soluzione $\Psi(t, x)$ dell'equazione (7.20) si ottiene rimpiazzando t con $T - t$ nella formula (7.18). Possiamo dunque scrivere

$$\Psi(t, x) = E_{T-t, x} \left[f(X_T) e^{-\int_{T-t}^T k(X_u) du} + \int_{T-t}^T g(X_s) e^{-\int_{T-t}^s k(X_u) du} ds \right].$$

Notiamo ora che, per omogeneità temporale, la soluzione $\{X_s\}_{s \in [T-t, T]}$ dell'equazione (7.15) con dato iniziale $X_{T-t} = x$ non è altro che la traslazione temporale della soluzione $\{X_s\}_{s \in [0, t]}$ con dato iniziale $X_0 = x$. In altre parole, il processo $\{X_s\}_{s \in [T-t, T]}$ rispetto a $P_{T-t, x}$ ha la stessa legge del processo $\{X_s\}_{s \in [0, t]}$ rispetto a $P_{0, x}$. Si ottiene dunque la seguente versione della formula di Feynman-Kac:

$$\Psi(t, x) = E_{0, x} \left[f(X_t) e^{-\int_0^t k(X_u) du} + \int_0^t g(X_s) e^{-\int_0^s k(X_u) du} ds \right]. \quad (7.21)$$

Si noti che è sparita la dipendenza da T : la funzione $\Psi(t, x)$ definita dalla relazione (7.21) è dunque soluzione dell'equazione differenziale (7.20) per ogni $t \in [0, \infty)$.

Il caso più semplice è quello in cui $g(x) \equiv 0$, $b_i(x) \equiv 0$ e $\sigma_{ij}(x) \equiv \delta_{ij}$, per cui anche $(\sigma\sigma^*)_{ij}(x) \equiv \delta_{ij}$. In tal caso l'equazione (7.20) si riduce all'*equazione del calore* con un potenziale:

$$\begin{cases} \frac{\partial}{\partial t} \Psi(t, x) = \frac{1}{2} \Delta \Psi(t, x) - k(x) \Psi(t, x) & \forall t \geq 0, x \in \mathbb{R}^n \\ \Psi(0, x) = f(x) & \forall x \in \mathbb{R}^n \end{cases}. \quad (7.22)$$

Dato che l'equazione differenziale stocastica (7.15) per $s = 0$ si riduce banalmente a

$$\begin{cases} dX_t = dB_t \\ X_0 = x \end{cases},$$

la soluzione è data da $X_t = x + B_t$. La formula di rappresentazione (7.21) diventa dunque

$$F(t, x) = E \left[f(x + B_t) e^{-\int_0^t k(x + B_u) du} \right].$$

Nel caso speciale in cui anche $k(x) \equiv 0$, usando la densità della legge normale possiamo scrivere

$$F(t, x) = \int_{\mathbb{R}^n} f(x + z) \frac{e^{-\frac{1}{2t}|z|^2}}{(2\pi t)^{n/2}} dz = \int_{\mathbb{R}^n} f(z) \frac{e^{-\frac{1}{2t}|z-x|^2}}{(2\pi t)^{n/2}} dz,$$

ritrovando una formula di rappresentazione classica per l'equazione del calore standard:

$$\begin{cases} \frac{\partial}{\partial t} \Psi(t, x) = \frac{1}{2} \Delta \Psi(t, x) & \forall t \geq 0, x \in \mathbb{R}^n \\ \Psi(0, x) = f(x) & \forall x \in \mathbb{R}^n \end{cases}.$$

8. RIMORSI

Concludiamo con qualche breve cenno ad alcuni argomenti che, per esigenze di tempo, non sono stati trattati nel corso.

8.1. LE DIFFUSIONI COME PROCESSI DI MARKOV

8.1.1. PROCESSI DI MARKOV. Un processo stocastico $X = \{X_t\}_{t \geq 0}$ a valori in \mathbb{R}^n è detto *processo di Markov* se per ogni $0 \leq s \leq t < \infty$ e $A \in \mathcal{B}(\mathbb{R}^n)$ vale la relazione[†]

$$P(X_t \in A | \mathcal{F}_s) = P(X_t \in A | X_s).$$

Questo significa che la legge condizionale di X_t sapendo tutta la storia $\sigma(\{X_u\}_{u \in [0, s]}) \subseteq \mathcal{F}_s$ del processo fino all'istante s è funzione soltanto di X_s : il futuro dipende dal passato solo attraverso lo stato presente.

Per $0 \leq s \leq t < \infty$, si definisce *nucleo di transizione* $Q_{s,t}(x, dy)$ una (qualunque) versione regolare della legge condizionale di X_t rispetto a X_s (equivalentemente, rispetto a \mathcal{F}_s). Più precisamente, $Q_{s,t}$ è un nucleo di probabilità tale che

$$P(X_t \in dy | \mathcal{F}_t) = Q_{s,t}(X_s, dy).$$

Il processo di Markov X è detto *omogeneo* se il corrispondente nucleo di transizione dipende da s, t solo attraverso la differenza $t - s$, cioè $Q_{s,t}(x, dy) = Q_{t-s}(x, dy)$. In altre parole, la legge condizionale di X_{t+h} rispetto a X_t non dipende da t , ma solo da h . Per $t, h \geq 0$ si ha dunque

$$P(X_{t+h} \in dy | \mathcal{F}_t) = Q_h(X_t, dy).$$

In questo caso per ogni $s, t \geq 0$ vale la relazione

$$\int_{y \in \mathbb{R}^n} Q_s(x, dy) Q_t(y, dz) = Q_{t+s}(x, dz), \quad (8.1)$$

nota come *equazione di Chapman-Kolmogorov*.

Sia ora $X = \{X_t\}_{t \geq 0}$ una *diffusione*, cioè un processo stocastico definito su uno spazio filtrato standard $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ a valori in \mathbb{R}^n che risolve l'equazione

$$\begin{cases} dX_t = \sigma(X_t) dB_t + b(X_t) dt \\ X_0 = x \end{cases}, \quad (8.2)$$

dove $x \in \mathbb{R}^n$, $B = \{B_t\}_{t \geq 0}$ è un $\{\mathcal{F}_t\}_{t \geq 0}$ -moto browniano d -dimensionale e $\sigma : \mathbb{R}^n \rightarrow \mathcal{M}(n \times d, \mathbb{R})$, $b : \mathbb{R}^n \rightarrow \mathbb{R}^n$ sono funzioni che supporremo per semplicità globalmente Lipschitziane: $|b(x) - b(y)| + |\sigma(x) - \sigma(y)| \leq L|x - y|$ per ogni $x, y \in \mathbb{R}^n$, con $L < \infty$. Indicheremo la dipendenza da x nella probabilità $P = P_x$.

Una proprietà fondamentale delle diffusionsi è che sono processi di Markov omogenei:

$$P_x(X_{t+h} \in A | \mathcal{F}_t) = P_z(X_h \in A) |_{z=X_t} = Q_h(X_t, A),$$

dove $Q_h(x, dy) = P_x(X_h \in dy)$ è il nucleo di transizione. Questa proprietà vale in particolare per il moto browniano, come si può dimostrare usando il Teorema 3.27.

[†]Si pone $P(B | X_s) := E(1_B | \sigma(X_s))$ per $B \in \mathcal{F}$.

8.1.2. GENERATORE E SEMIGRUPPO. Un processo di Markov omogeneo $X = \{X_t\}_{t \geq 0}$ è caratterizzato dal suo “comportamento infinitesimale”. Più precisamente, introducendo lo spazio $C_0(\mathbb{R}^n, \mathbb{R})$ delle funzioni $f(x)$ continue da \mathbb{R}^n in \mathbb{R} che tendono a zero per $|x| \rightarrow \infty$, si definisce il *generatore* L di X ponendo

$$Lf(x) := \lim_{t \downarrow 0} \frac{E_x(f(X_t)) - f(x)}{t}, \quad (8.3)$$

per ogni funzione $f \in C_0(\mathbb{R}^n, \mathbb{R})$ per cui tale limite esiste. L'insieme di tali funzioni è detto *dominio* del generatore. È possibile mostrare che il generatore (insieme con il suo dominio) determina la legge del processo di Markov X .

Nel caso di una diffusione X (soluzione dell'equazione (8.2)), il generatore L è proprio l'operatore differenziale del secondo ordine L introdotto nel capitolo precedente, che riscriviamo per comodità:

$$Lf(x) = \frac{1}{2} \sum_{i,j=1}^n (\sigma\sigma^*)_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} f(x) + \sum_{i=1}^n b_i(x) \frac{\partial}{\partial x_i} f(x).$$

In questo caso si può mostrare che la relazione (8.3) vale per tutte le funzioni f di classe C^2 a supporto compatto, che dunque appartengono al dominio del generatore.

A un processo di Markov X è associata una famiglia di operatori $\{P_t\}_{t \geq 0}$, detta *semigrupp* del processo, che agiscono sulle funzioni $f \in C_0(\mathbb{R}^n, \mathbb{R})$ nel modo seguente:

$$P_t f(x) := E_x(f(X_t)) = \int_{\mathbb{R}^n} f(y) Q_t(x, dy). \quad (8.4)$$

È possibile mostrare che $P_t f \in C_0(\mathbb{R}^n, \mathbb{R})$ per ogni $f \in C_0(\mathbb{R}^n, \mathbb{R})$. Il nome semigrupp deriva dal fatto che $P_t \circ P_s = P_{t+s}$, cioè $P_t(P_s f) = P_{t+s} f$ per ogni $s, t \geq 0$ e per ogni $f \in C_0(\mathbb{R}^n, \mathbb{R})$, come segue dall'equazione (8.1) di Chapman-Kolmogorov.

Si noti che $Lf = \lim_{t \downarrow 0} \frac{1}{t} (P_t f - f) = \frac{d}{dt} P_t(f)|_{t=0}$, per ogni f nel dominio di L . Più in generale, per ogni f nel dominio di L , $P_t f$ è ancora nel dominio di L e vale la relazione

$$\frac{d}{dt} P_t f = L(P_t f) = P_t(Lf), \quad \forall t \geq 0.^\dagger \quad (8.5)$$

Il generatore L determina il semigrupp. Ispirati dalla relazione (8.5), si scrive spesso $P_t = e^{tL}$.

8.1.3. EQUAZIONI DI KOLMOGOROV. Sia ora X una diffusione per cui il nucleo di transizione $Q_t(x, dy) = P_x(X_t \in dy)$ è assolutamente continuo, per ogni $t > 0$. Supponiamo cioè che esista una funzione misurabile $q_t : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ tale che

$$Q_t(x, dy) = q_t(x, y) dy, \quad \forall t > 0, \forall x, y \in \mathbb{R}^n.$$

Imponendo opportune ipotesi di regolarità, è possibile mostrare che la densità $q_t(x, y)$ soddisfa l'*equazione di Kolmogorov backward*: per ogni $y \in \mathbb{R}^n$ fissato si ha

$$\frac{\partial}{\partial t} q_t(x, y) = L_x q_t(x, y), \quad \forall t > 0, \forall x \in \mathbb{R}^n, \quad (8.6)$$

dove con L_x si indica che l'operatore L agisce sulla variabile x di $q_t(x, y)$. Vale inoltre l'*equazione di Kolmogorov forward*, nota anche come *equazione di Fokker-Plank*: per ogni $x \in \mathbb{R}^n$ fissato si ha

$$\frac{\partial}{\partial t} q_t(x, y) = L_y^* q_t(x, y), \quad \forall t > 0, \forall y \in \mathbb{R}^n, \quad (8.7)$$

dove l'operatore L^* è l'aggiunto di L , definito da

$$L^* f(y) = \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2}{\partial y_i \partial y_j} [(\sigma\sigma^*)_{ij}(y) f(y)] - \sum_{i=1}^n \frac{\partial}{\partial y_i} [b_i(y) f(y)].$$

Una derivazione informale delle equazioni (8.6) e (8.7) si ottiene facilmente a partire dalle relazioni (8.4) e (8.5), scambiando l'ordine di derivate e integrali.

[†]La prima uguaglianza in (8.5) segue dalle equazioni (7.20) e (7.21) (Feynman-Kac) con $k, g \equiv 0$.

8.2. L'INTEGRALE DI STRATONOVICH

Sia $X = \{X_t\}_{t \in [0, T]}$ un processo di Itô reale con decomposizione $dX_t = \varphi_t dB_t + \psi_t dt$, dove B è un moto browniano reale e $\varphi \in M_{loc}^2[0, T]$, $\psi \in M_{loc}^1[0, T]$. Definiamo l'integrale di Stratonovich di X rispetto al moto browniano B ponendo, per $t \in [0, T]$,

$$\int_0^t X_s \circ dB_s := \int_0^t X_s dB_s + \frac{1}{2} \langle X, B \rangle_t = \int_0^t X_s dB_s + \frac{1}{2} \int_0^t \varphi_s ds. \quad (8.8)$$

Nel caso in cui il processo X sia continuo, si può ottenere tale integrale come limite in probabilità:

$$\int_0^t X_s \circ dB_s = \lim_{n \rightarrow \infty} \sum_{i=0}^{k_n-1} \frac{X_{t_i^{(n)}} + X_{t_{i+1}^{(n)}}}{2} (B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}}),$$

dove $\{0 = t_0^{(n)} < t_1^{(n)} < \dots < t_{k_n}^{(n)} = t\}$ è una successione di partizioni di $[0, t]$ con passo che tende verso zero (per esempio $t_i^{(n)} = \frac{i}{n}t$ per $0 \leq i \leq k_n = n$).

L'utilità di questa definizione è che per questo integrale vale un perfetto analogo della chain rule del calcolo ordinario. Più precisamente, per ogni funzione $\Phi : \mathbb{R} \rightarrow \mathbb{R}$ di classe C^3 vale la relazione

$$\Phi(B_t) - \Phi(B_0) = \int_0^t \Phi'(B_s) \circ dB_s. \quad (8.9)$$

Questa proprietà rende l'integrale di Stratonovich più conveniente dell'integrale di Itô in alcuni ambiti, come ad esempio la teoria delle diffusioni su varietà differenziabili.

La dimostrazione della formula (8.9) è semplice: per ipotesi la funzione $\Phi' : \mathbb{R} \rightarrow \mathbb{R}$ è di classe C^2 , quindi $\Phi'(B)$ è un processo di Itô con decomposizione $d\Phi'(B_t) = \Phi''(B_t) dB_t + \frac{1}{2} \Phi'''(B_t) dt$, grazie alla formula di Itô. Ricordando la definizione (8.8), si vede immediatamente che la relazione (8.9) non è altro che l'ordinaria formula di Itô applicata a $\Phi(B_t)$.

Sono possibili diverse estensioni: in particolare, si può definire l'integrale di Stratonovich rispetto a un arbitrario processo di Itô, al posto del moto browniano. Tuttavia l'integrando deve essere sempre un processo di Itô.[†] L'integrale di Stratonovich è dunque definito per una classe più ristretta di integrandi, rispetto all'integrale di Itô. Sottolineiamo anche che, per poter applicare la formula (8.9), la funzione Φ deve essere di classe C^3 e non solo C^2 , come per l'ordinaria formula di Itô.

8.3. TEOREMI DI RAPPRESENTAZIONE PER MARTINGALE

Concludiamo con due interessanti risultati di rappresentazione per martingale. Il primo risultato afferma che ogni martingala locale continua $M = \{M_t\}_{t \geq 0}$ è un moto browniano cambiato di tempo.

TEOREMA 8.1 (DAMBIS, DUBINS&SCHWARZ). Sia $M = \{M_t\}_{t \geq 0}$ una martingala continua, definita su uno spazio filtrato standard $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$, tale che $\lim_{t \rightarrow \infty} \langle M \rangle_t = \infty$ q.c.. Allora sullo stesso spazio è possibile definire un moto browniano reale $B = \{B_t\}_{t \geq 0}$ per cui vale la relazione

$$M_t = B_{\langle M \rangle_t}, \quad \forall t \geq 0.$$

Per la dimostrazione, si veda il Teorema 4.6 del capitolo 3 in [Karatzas e Shreve, 1998]. Sottolineiamo che la condizione $\lim_{t \rightarrow \infty} \langle M \rangle_t = \infty$ q.c. serve essenzialmente a garantire che lo spazio di probabilità (Ω, \mathcal{F}, P) sia abbastanza ricco per poterci costruire il moto browniano B . Nel caso in cui tale condizione non sia verificata, il teorema resta comunque valido, a patto di ampliare lo spazio di probabilità. Osserviamo che esiste anche una generalizzazione multidimensionale non banale di questo teorema, nota come Teorema di Knight (si veda il Teorema 4.13 nel capitolo 3 in [Karatzas e Shreve, 1998]).

[†]O, più in generale, una *semimartingala*, cioè la somma di una martingala locale e di un processo a variazione finita.

Il secondo risultato che presentiamo afferma che ogni martingala di quadrato integrabile adattata alla filtrazione completata del moto browniano (che sia cioè funzione del moto browniano) è l'integrale stocastico di un processo in \mathcal{M}^2 . Per la dimostrazione, si veda il Teorema 4.15 del capitolo 3 in [Karatzas e Shreve, 1998].

TEOREMA 8.2. Sia $B = \{B_t\}_{t \geq 0}$ un moto browniano reale, definito su uno spazio di probabilità (Ω, \mathcal{F}, P) , e indichiamo con $\{\bar{\mathcal{G}}_{t+}\}_{t \geq 0}$ l'ampliamento standard della filtrazione naturale di B . Per ogni $\{\bar{\mathcal{G}}_{t+}\}_{t \geq 0}$ -martingala $M = \{M_t\}_{t \geq 0}$ definita su Ω , tale che $E(M_t^2) < \infty$ per ogni $t \geq 0$, esiste un (unico) processo $X \in \mathcal{M}^2$ tale che

$$M_t = M_0 + \int_0^t X_s dB_s. \quad (8.10)$$

In particolare, per ogni variabile Z aleatoria \mathcal{G}_T -misurabile e di quadrato integrabile, esiste $X \in \mathcal{M}^2[0, T]$ tale che $Z = E(Z) + \int_0^T X_s dB_s$. Basta infatti applicare il teorema alla martingala $M_t = E(Z|\mathcal{G}_t)$ e porre $t = T$ in (8.10).

BIBLIOGRAFIA

- Asmussen S. (2003). *Applied Probability and Queues*. Second Edition. Springer.
- Baldi P. (2000). *Equazioni differenziali stocastiche e applicazioni*. Pitagora Editrice.
- Billingsley P. (1995). *Probability and Measure*. Third Edition. John Wiley & Sons.
- Comets F., Meyre T. (2006). *Calcul stochastique et modèles de diffusions*. Dunod.
- Durrett R. (1996). *Stochastic Calculus, A Practical Introduction*. CRC Press.
- Itô K. (1944). Stochastic integral. *Proc. Imp. Acad. Tokyo*, **20**(8), 519–524.
- Karatzas I., Shreve S. E. (1998). *Brownian Motion and Stochastic Calculus*. Second Edition. Springer.
- Le Gall J.-F. (2008). *Mouvement brownien et calcul stochastique*. Notes de cours de DEA. Disponibile sulla pagina web dell'autore: <http://www.dma.ens.fr/~legall/>.
- Mörters P., Peres Y. (2009). *Brownian Motion*. Cambridge University Press. Una versione preliminare è disponibile sulle pagine web degli autori: <http://people.bath.ac.uk/maspm>, <http://www.stat.berkeley.edu/~peres>.
- Protter P. E. (2004). *Stochastic Integration and Differential Equations*. Second Edition. Springer.
- Williams D. (1991). *Probability with Martingales*. Cambridge University Press.