Un'introduzione al processo di Wiener e alle equazioni stocastiche

16 Maggio 2005 (*)

Enrico Priola
Dipartimento di Matematica, Università di Torino,
via Carlo Alberto 10, 10123, Torino, Italy.
e-mail priola@dm.unito.it

Indice

1 Richiami di Probabilità **2**

2 Prime nozioni sui processi stocastici **10**

3 Il processo di Wiener **11**
3.1 Definizione e costruzione del processo di Wiener 11
3.2 Proprietà delle traiettorie del processo di Wiener 17
3.3 Osservazioni finali 22

4 Equazioni stocastiche con rumore additivo **24**
4.1 Un teorema di esistenza e unicità 25

5 L'integrale di Ito **30**

6 La formula di Ito **32**
6.1 La formula di Ito multidimensionale 32
6.2 Dimostrazione della formula di Ito unidimensionale 34

7 Cenni sui semigruppi di Markov e le equazioni paraboliche di Kolmogorov **38**

*Ciclo di Lezioni tenute presso il Dipartimento di Matematica "Ulisse Dini" di Firenze (dal 2/05 al 13/05/2005).

1 Richiami di Probabilità

Per maggiori dettagli ed esempi relativi a questa sezione, si possono consultare dei testi di Probabilità come, per esempio, Baldi [1], Billingsley [2], Dudley [3], Durret [4], Flandoli [9]. Assumiamo che il lettore sia familiare con la teoria della misura e l'integrale di Lebesgue.

Uno **spazio di probabilità** è una terna $(\Omega, \mathcal{F}, \mathbb{P})$, dove Ω è un insieme, $\mathcal{F}$ una σ-algebra su Ω e $\mathbb{P} : \mathcal{F} \to [0,1]$, una misura di probabilità su Ω (cioè $\mathbb{P}$ è una misura (non negativa) σ-additiva e tale che $\mathbb{P}(\Omega) = 1$).

Non è restrittivo supporre che $\mathbb{P}$ sia **completa.** Ricordiamo che $\mathbb{P}$ è completa se preso $A \in \mathcal{F}$ con $\mathbb{P}(A) = 0$, allora per ogni $B \subset A$ risulta $B \in \mathcal{F}$ (e quindi anche $\mathbb{P}(B) = 0$).

Se $\mathbb{P}$ non è completa, si può completarla nel modo seguente. Si considera $\mathcal{N} = \{B \subset \Omega$ tali che esiste $A \in \mathcal{F}$, con $\mathbb{P}(A) = 0$ e $B \subset A\}$. Si può sostituire $\mathcal{F}$ con la $\sigma-$algebra $\dot{\mathcal{F}}$, la più piccola σ-algebra che contiene $\mathcal{F}$ ed $\mathcal{N}$. Si dimostra che $\dot{\mathcal{F}} = \{C \subset \Omega \, : \, C = A \cup B$, con $A \in \mathcal{F}$, $B \in \mathcal{N}\}$.

Si dice che una certa proprietà che coinvolge gli $\omega \in \Omega$ vale **quasi certamente (o q.c.)** su Ω se essa è vera tranne che su un $\Omega' \in \mathcal{F}$ con $\mathbb{P}(\Omega') = 0$ (Ω' si chiama anche evento trascurabile).

Prima di presentare alcuni esempi, fissiamo altre notazioni.

Sia $\mathcal{A}$ una famiglia di sottinsiemi di Ω; la più piccola σ-algebra che contiene $\mathcal{A}$ si chiama la **σ-algebra generata** da $\mathcal{A}$ e si indica con $\sigma(\mathcal{A})$.

Dato $A \subset \Omega$, 1_A indica la **funzione indicatrice** di A ($1_A(\omega) = 0$, se $\omega \notin A$, $1_A(\omega) = 1$, se $\omega \in A$).

Dato uno spazio metrico E, indichiamo con $B(E)$ la **σ-algebra di Borel** di E (è la più piccola σ-algebra che contiene tutti gli aperti in E).

Esempio 1.1. *Esempi di Spazi di probabilità*

(i) *Lancio di due dadi non truccati.* Si può scegliere $\Omega = \{1, 2, \ldots, 6\} \times \{1, 2, \ldots, 6\}$; $\mathcal{F}$ è la famiglia di tutti i sottinsiemi di Ω. Possiamo porre $\mathbb{P}(A) = \frac{|A|}{36}$, $A \in \mathcal{F}$ (dove $|A|$ è la cardinalità di A).

(ii) $\Omega = [0,1]$; $\mathcal{F} = B([0,1])$ è la $\sigma-$algebra di Borel di $[0,1]$; $\mathbb{P} = \lambda$ (misura di Lebesgue su [0,1]).

(iii) $\Omega = \mathbb{R}$; $\mathcal{F} = B(\mathbb{R})$;

$$\mathbb{P}(A) = \int_A \frac{e^{\frac{-(x-a)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} dx, \quad A \in \mathcal{F}, \, a \in \mathbb{R}, \; \sigma > 0. \tag{1.1}$$

La misura $\mathbb{P}$ si dice misura di **probabilità normale (o Gaussiana)** $N(a, \sigma^2)$ (o $N(a,q)$, con $q > 0$). La funzione $x \mapsto \frac{e^{\frac{-(x-a)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma}$ è la densità di $N(a, \sigma^2)$. Si definisce anche $N(a, 0)$ come la misura di Dirac concentrata in a.

(iv) $\Omega = C([0,\infty), \mathbb{R}^n) = \{\omega : [0,\infty) \to \mathbb{R}^n$ continue $\}$. Ω diventa uno spazio metrico, completo e separabile, con la metrica:

$$d(f, g) = \sum_{n \geq 1} \frac{1}{2^n} \frac{\|f - g\|_\infty^n}{1 + \|f - g\|_\infty^n}, \quad f, g \in \Omega,$$
$$\|f - g\|_\infty^n = \sup_{t \in [0,n]} |f(t) - g(t)|, \quad n \geq 1. \tag{1.2}$$

Nel seguito incontreremo una misura di probabilità $\mathbb{P}$ su $(\Omega, B(\Omega))$, la *misura di Wiener.*
∎

Sia assegnato uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$.

Una funzione misurabile $X : \Omega \to \mathbb{R}^n$ si chiama **variabile aleatoria** (brevemente v.a.). Quindi si richiede che per ogni $A \in B(\mathbb{R}^n)$, $X^{-1}(A) \in \mathcal{F}$. Per indicare esplicitamente la σ-algebra a volte scriveremo che la v.a. X è $\mathcal{F}$-misurabile. Se $\mathbb{R}^n = \mathbb{R}$, si parla di v.a. reale.

Sia $X : \Omega \to \mathbb{R}^n$ una funzione. La più piccola σ-algebra che rende misurabile X si chiama **σ-algebra generata da** X e si indica con $\sigma(X)$**.** E' chiaro che $\sigma(X) = \{X^{-1}(B)\}_{B \in B(\mathbb{R}^n)}$. Se X è una v.a. risulta $\sigma(X) \subset \mathcal{F}$.

Più in generale se $(X_i)_{i \in J}$ è una famiglia di funzioni da Ω in $\mathbb{R}^n$, La più piccola σ-algebra che rende misurabili tutte le X_i si chiama **σ-algebra generata da** $(X_i)_{i \in J}$ e si indica con $\sigma(X_i,\ i \in J)$**.**

Sia $X : \Omega \to \mathbb{R}^n$ una v.a.; se $A \in B(\mathbb{R}^n)$, useremo spesso la notazione $(X \in A)$ o $\{X \in A\}$, per indicare l'insieme (o evento) $X^{-1}(A) = \{\omega \in \Omega : X(\omega) \in A\}$. Inoltre poniamo

$$\mathbb{P}(X \in A) := \mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\}). \tag{1.3}$$

La legge (o distribuzione) di X è la misura di probabilità μ_X su $B(\mathbb{R}^n)$ data da

$$\mu_X(B) = \mathbb{P}(X \in B), \quad B \in B(\mathbb{R}^n).$$

(per esempio se $X = c \in \mathbb{R}^n$, q.c., allora $\mu_X = \delta_c$, la misura di Dirac concentrata in c).

Una v.a. reale si dice v.a. **normale o Gaussiana,** se la sua legge è $N(a, \sigma^2)$, per qualche $a \in \mathbb{R}$, $\sigma \geq 0$.

Sia $X : \Omega \to \mathbb{R}$ una v.a.; Se X è integrabile secondo Lebesgue su $(\Omega, \mathcal{F})$, cioè se esiste finito l'integrale di Lebesgue $\int_\Omega |X(\omega)| \mathbb{P}(d\omega)$, diciamo che X è una **v.a. integrabile.** In tale caso si pone

$$\mathbb{E}X := \int_\Omega X(\omega) \mathbb{P}(d\omega).$$

$\mathbb{E}X$ is dice **attesa (o media o speranza) di** X (osserviamo che se $X \geq 0$ allora ha sempre senso $\mathbb{E}X$; essa può essere finita o infinita).

In modo simile sia $X = (X_i)_{i=1,\ldots n} : \Omega \to \mathbb{R}^n$ una v.a. (n-dimensionale). Se $\mathbb{E}|X| < \infty$ (dove $|\cdot|$ indica la norma euclidea di $\mathbb{R}^n$, $n \geq 1$) allora si pone $\mathbb{E}X = (\mathbb{E}X_i)_{i=1,\ldots n}$.

Il calcolo delle attese si può ricondurre al calcolo di integrali su $\mathbb{R}^n$, grazie al seguente risultato:

Teorema 1.2. *(Formula di cambiamento di variabile) Sia $X : \Omega \to \mathbb{R}^n$ una v.a. con legge μ_X. Sia $f : \mathbb{R}^n \to \mathbb{R}$ una funzione Borelliana, tale che o $f(y) \geq 0$, $y \in \mathbb{R}^n$, o $\int_{\mathbb{R}^n} |f(y)|\mu_X(dy) < \infty$. Allora si ha:*

$$\mathbb{E}f(X) = \int_{\mathbb{R}^n} f(y)\mu_X(dy). \tag{1.4}$$

Per esempio, se X è una variabile aleatoria con legge $N(a, \sigma^2)$, risulta:

$$\mathbb{E}X = \int_{\mathbb{R}} x \frac{e^{\frac{-(x-a)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} dx = a.$$

Sia X una v.a. reale tale che $\mathbb{E}X^2 < \infty$. La **varianza di** X è $\text{var}(X) = \mathbb{E}(X - m)^2 = EX^2 - m^2$, dove $m = EX$.

Per esempio se X è normale con legge $N(a, \sigma^2)$, si trova che $\text{var}(X) = \sigma^2$.

Lemma 1.3. *(Disuguaglianze di Chebyshev)*

1. *Sia X una v.a. reale e non negativa. Allora per ogni $\epsilon > 0$, $p > 0$, risulta:*

$$\mathbb{P}(X \geq \epsilon) \leq \frac{\mathbb{E}X^p}{\epsilon^p}.$$

2. *Sia X una v.a. reale con $\mathbb{E}X^2 < \infty$. Poniamo $m = \mathbb{E}X$. Allora risulta:*

$$\mathbb{P}(|X - m| \geq \epsilon) \leq \frac{var(X)}{\epsilon^2}, \quad \epsilon > 0.$$

Parliamo ora del concetto di **indipendenza.**

Partiamo dal concetto elementare di **probabilità condizionale.** Siano $A, B \in \mathcal{F}$, con $\mathbb{P}(B) > 0$. Quale può essere una ragionevole definizione di $\mathbb{P}(A/B)$, la probabilità che capiti l'evento A sapendo che B si è verificato? Poichè sappiamo che $\omega \in B$, possiamo guardare a B come ad un nuovo spazio di probabilità $\dot{\Omega}$ con $\dot{\mathcal{F}} = \{F \cap B,\ F \in \mathcal{F}\}$ e $\dot{\mathbb{P}} = \frac{\mathbb{P}}{\mathbb{P}(B)}$. In questo nuovo spazio $(\dot{\Omega}, \dot{\mathcal{F}}, \dot{\mathbb{P}})$ la probabilità che ω sia in A è $\dot{\mathbb{P}}(A \cap B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$. Pertanto scriviamo

$$\mathbb{P}(A/B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \quad \mathbb{P}(B) > 0.$$

Dire che A è indipendente da B significa dire che $\mathbb{P}(A/B) = \mathbb{P}(A)$, cioè che $\mathbb{P}(A \cap B) = \mathbb{P}(B)\mathbb{P}(A)$.

Introduciamo le seguenti definizioni.

(1) Diciamo che due eventi $A, B \in \mathcal{F}$ sono **indipendenti** se $\mathbb{P}(A \cap B) = \mathbb{P}(B)\mathbb{P}(A)$. La definizione si estende ad un numero finito di eventi. Siano A_1, ..., $A_n \in \mathcal{F}$; essi si dicono indipendenti se per ogni scelta di $i_1, \ldots, i_m$ tra 1 ed n, risulta

$$\mathbb{P}(A_{i_1} \cap \ldots \cap A_{i_m}) = \mathbb{P}(A_{i_1}) \cdots \mathbb{P}(A_{i_m}).$$

(2) Date due v.a. reali $X, Y : \Omega \to \mathbb{R}^n$. Esse si dicono **v.a. indipendenti** se per ogni $F, G \in B(\mathbb{R}^n)$, si ha:

$$\mathbb{P}(X \in F, Y \in G) = \mathbb{P}(X \in F)\, \mathbb{P}(Y \in G),$$

dove $\mathbb{P}(X \in F, Y \in G) = \mathbb{P}(\{\omega \in \Omega \,:\, X(\omega) \in F\} \cap \{\omega \in \Omega \,:\, Y(\omega) \in G\})$.
(3) Due $\sigma-$algebre $\mathcal{G}$ e $\mathcal{H}$ contenute in $\mathcal{F}$ si dicono **indipendenti** se per ogni scelta di $G \in \mathcal{G}$ e $H \in \mathcal{H}$, gli eventi G e H sono indipendenti.

La definizione (1) è un caso particolare della definizione (2) (infatti siano $A, B \in \mathcal{F}$; si verifica che A e B sono indipendenti se e solo se le v.a. 1_A e 1_B sono indipendenti).

La definizione (2) è un caso particolare della definizione (3) (infatti siano X, Y v.a.; si verifica che X e Y sono indipendenti se e solo se le $\sigma-$algebra $\sigma(X)$ e $\sigma(Y)$ sono indipendenti).

Le precedenti definizioni di indipendenza si estendono a **famiglie di oggetti.** Le definizioni (2) e (3) diventano:

(2') Siano $(X_j)_{j\in J}$ v.a. a valori in $\mathbb{R}^n$. Esse si dicono indipendenti se per ogni scelta (finita) di j_1, .. $j_m \in J$, risulta che X_{j_1}, ... X_{j_m} sono indipendenti; ovvero per ogni scelta di $A_{j_i} \in B(\mathbb{R}^n)$, risulta che

$$\mathbb{P}(X_{j_1} \in A_{j_1}, \dots, X_{j_m} \in A_{j_m}) = \mathbb{P}(X_{j_1} \in A_{j_1}) \cdots \mathbb{P}(X_{j_m} \in A_{j_m}).$$

(3') Siano $(\mathcal{F}_j)_{j\in J}$ σ-algebre contenute in $\mathcal{F}$. Esse si dicono indipendenti se per ogni scelta (finita) di j_1, .. $j_m \in J$, risulta che $\mathcal{F}_{j_1}$, ... $\mathcal{F}_{j_m}$ sono indipendenti (ovvero per ogni $F_{i_1} \in \mathcal{F}_{j_1}$, ..., $F_{j_m} \in \mathcal{F}_{j_m}$, gli eventi F_{i_1}, ..., F_{j_m} sono indipendenti).

È utile il π-λ **lemma di Dynkin** che ora introduciamo (vedere [2] o [22]).

Una famiglia $\mathcal{L}$ di sottinsiemi di Ω si dice π**-system** se presi $A, B \in \mathcal{L}$, risulta che $A \cap B \in \mathcal{L}$. Una famiglia $\mathcal{M}$ di sottinsiemi di Ω si dice λ**-system** se verifica

1. $\Omega \in \mathcal{M}$;
2. se $A \in \mathcal{M}$ allora $A^c \in \mathcal{M}$ (dove A^c è il complementare di A);
3. se $A_j \in \mathcal{M}$, $j \in \mathbb{N}$, e $A_j \cap A_i = \varnothing$, per $i \neq j$, allora $\cup_{j\in\mathbb{N}} A_j \in \mathcal{M}$ ($\mathcal{M}$ è stabile per unioni disgiunte).

Lemma 1.4. *($\pi - \lambda$ Lemma) Se $\mathcal{A}$ è un π-system e $\mathcal{M}$ un $\lambda-$system allora $\mathcal{A} \subset \mathcal{M}$ implica che $\sigma(\mathcal{A}) \subset \mathcal{M}$ (dove $\sigma(\mathcal{A})$ è la più piccola σ-algebra contenente $\mathcal{A}$).*

Con il precedente risultato si può provare la seguente proprietà.

Proposizione 1.5. *Siano X_1, .. X_m v.a. su Ω a valori in $\mathbb{R}^n$. Sia $X = (X_1, \dots, X_m) : \Omega \to \mathbb{R}^{mn}$. Allora le seguenti affermazioni sono equivalenti:*

(i) X_1, ..., X_m sono indipendenti;

(ii) $\mu_X = \mu_{X_1}\times$ ... $\times\mu_{X_m}$ su $B(\mathbb{R}^{nm})$ (dove μ_X indica le legge di X e $\mu_{X_1}\times$ ... $\times\mu_{X_m}$ indica la misura prodotto);

Dimostrazione. Diamo un'idea della dimostrazione che (i) implica (ii). Si ha che

$$\mu_X(A_1 \times \ldots \times A_m) = \mu_{X_1}(A_1) \cdots \mu_{X_m}(A_m), \;\; A_i \in B(\mathbb{R}^n), \; i = 1, \dots, m.$$

Dunque le due misure di probabilità $\mu_1 = \mu_X$ e $\mu_2 = \mu_{X_1}\times$ ... $\times\mu_{X_m}$ coincidono sui rettangoli $A_1 \times \ldots \times A_m$. I rettangoli formano un π-system e generano $B(\mathbb{R}^{nm})$ (cioè la più piccola σ-algebra che contiene i rettangoli è $B(\mathbb{R}^{nm})$). La dimostrazione si conclude osservando che la famiglia

$$\mathcal{M} = \{C \in B(\mathbb{R}^{nm}) \;:\; \mu_1(C) = \mu_2(C)\}$$

è un $\lambda-$system. ∎

Corollario 1.6. *Siano X_1, .. X_m v.a. reali, indipendenti e integrabili su Ω. Allora risulta:*

$$(*)\ \ \mathbb{E}(\prod_{k=1}^{m} X_k) = \prod_{k=1}^{m} \mathbb{E}(X_k).$$

Osserviamo che (*) non implica che le v.a. X_1, ... X_m sono indipendenti.

Proposizione 1.7. *Siano X_1, .. X_m v.a. reali, indipendenti e a quadrato integrabile su Ω. Allora risulta:*

$$var\Big(\sum_{k=1}^{m} X_k\Big) = \sum_{k=1}^{m} var(X_k).$$

Dimostrazione. Sia $\mathbb{E}X_i = m_i$, $i = 1, \dots, m$. Si ha, usando il corollario precedente,

$$var\Big(\sum_{k=1}^{m} X_k\Big) = \mathbb{E}\Big(\sum_{k=1}^{m}(X_k - m_k)\Big)^2$$
$$= \sum_{k=1}^{m} \mathbb{E}(X_k - m_k)^2 + 2\sum_{k<j} \mathbb{E}((X_k - m_k)(X_j - m_j)) = \sum_{k=1}^{m} var(X_k).$$

■

Nel seguito utilizzeremo il seguente risultato.

Teorema 1.8. *Sia (μ_k) una successione di misure di probabilità (di Borel) su $\mathbb{R}$. Allora esiste uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ su cui esistono variabili aleatorie indipendenti (X_k) tali che la legge di X_k è μ_k, per $k \in \mathbb{N}$.*

Una sua dimostrazione utilizza un teorema sulle misure prodotto che ora spieghiamo (vedere anche [3]).

Siano $(\Omega_k, \mathcal{F}_k)$, spazi misurabili, $k \in \mathbb{N}$. Consideriamo lo spazio prodotto $\Omega = \prod_{k\geq 0} \Omega_k$, munito della σ-algebra $\mathcal{F} = \prod_{k\geq 0} \mathcal{F}_k$. $\mathcal{F}$ è per definizione la σ-algebra generata da tutti gli *insiemi rettangoli* , cioè del tipo

$$A = \prod_{k\geq 0} A_k, \ \text{ con } \ A_k \in \mathcal{F}_k$$

e $A_k = \Omega_k$, tranne al più un numero finito di indici k.

Teorema 1.9. *Siano $(\Omega_k, \mathcal{F}_k, \mathbb{P}_k)$, spazi di probabilità, $k \in \mathbb{N}$. Consideriamo lo spazio misurabile prodotto $(\Omega, \mathcal{F})$ appena introdotto. Esiste un'unica misura di probabilità $\mathbb{P}$ su $(\Omega, \mathcal{F})$ tale che per ogni rettangolo $A = \prod_{k=0}^{\infty} A_k$, vale*

$$\mathbb{P}(A) = \prod_{k=0}^{\infty} \mathbb{P}_k(A_k).$$

Dimostrazione del Teorema 1.8. In accordo al Teorema 1.9, si considerano gli spazi di probabilità $(\Omega_k, \mathcal{F}_k, \mathbb{P}_k) = (\mathbb{R}, B(\mathbb{R}), \mu_k)$, $k \in \mathbb{N}$, e lo spazio di probabilità prodotto $(\Omega, \mathcal{F}, \mathbb{P})$. Si definisce $X_k(\omega) = \omega_k$, $\ \omega = (\omega_k) \in \Omega$, $\ k \in \mathbb{N}$.

Verifichiamo che $X_{k_1}, \dots, X_{k_n}$ sono indipendenti. Si ha

$$\mathbb{P}(X_{k_1} \in A_{k_1}, \dots, X_{k_n} \in A_{k_n}) =$$
$$= \prod_{i=1}^{n} \mu_{k_i}(A_{k_i}) = \prod_{i=1}^{n} \mathbb{P}(X_{k_i} \in A_{k_i}),$$

con $A_{k_i} \in B(\mathbb{R})$, $i = 1, \dots, n$. ■

Si può provare anche il Teorema 1.8, prendendo come spazio Ω l'intervallo $[0, 1]$ con la misura di Lebesgue e usando un metodo che risale a Steinhaus, vedere [22].

Sia data una misura di probabilità (o legge o distribuzione) μ su $B(\mathbb{R}^n)$. Definiamo la sua **funzione caratteristica o trasformata di Fourier** $\hat{\mu} : \mathbb{R}^n \to \mathbb{C}$, nel modo seguente:

$$\hat{\mu}(h) = \int_{\mathbb{R}^n} e^{i\langle x,h\rangle} \mu(dx), \quad h \in \mathbb{R}^n, \tag{1.5}$$

dove $\langle \cdot, \cdot \rangle$ indica il prodotto interno di $\mathbb{R}^n$. Si può dimostrare che $\hat{\mu}$ è uniformente continua e limitata su $\mathbb{R}^n$. Il nome di funzione caratteristica è giustificato dal seguente risultato.

Teorema 1.10. *Siano μ_1 e μ_2 due misura di probabilità su $B(\mathbb{R}^n)$. Se $\hat{\mu}_1 = \hat{\mu}_2$ su $\mathbb{R}^n$ allora le due misure μ_1 e μ_2 coincidono (su $B(\mathbb{R}^n)$).*

Data una v.a. $X : \Omega \to \mathbb{R}^n$, si definisce **funzione caratteristica ϕ_X di X** la funzione caratteristica della legge μ_X di X; dunque

$$\phi_X(h) = \hat{\mu}_X(h) = \int_{\mathbb{R}^n} e^{i\langle x,h\rangle} \mu_X(dx) = \mathbb{E} e^{i\,\langle X,h\rangle}, \quad h \in \mathbb{R}^n.$$

Usando il Teorema 1.10 e la Proposizione 1.5 si ottiene il seguente utile criterio.

Proposizione 1.11. *Siano X_1, .. X_m v.a. su Ω a valori in $\mathbb{R}^n$. Sia $X = (X_1, \dots, X_m) : \Omega \to \mathbb{R}^{mn}$. Allora le seguenti affermazioni sono equivalenti:*

(i) X_1, ..., X_m *sono indipendenti;*

(ii) $\phi_X(h_1, \dots, h_m) = \phi_{X_1}(h_1) \cdots \phi_{X_m}(h_m), \quad h = (h_1, \dots, h_m) \in \mathbb{R}^{nm}$
(la funzione caratteristica di X si fattorizza nel prodotto delle ϕ_{X_i}).

Parliamo adesso di **leggi normali (o Gaussiane) multidimensionali.**

Sia $a \in \mathbb{R}^n$ e Q una matrice $n \times n$ simmetrica e semidefinita positiva. Una misura di probabilità μ su $B(\mathbb{R}^n)$ si dice **normale (o Gaussiana)** $N(a, Q)$**,** se la sua funzione caratteristica è

$$\hat{\mu}(h) = e^{i\langle h,a\rangle}\, e^{-\frac{1}{2}\langle Qh,h\rangle}, \quad h \in \mathbb{R}^n.$$

Un teorema di Bochner assicura che la definizione precedente è ben data (cioè $\hat{\mu}$ è effettivamente la funzione caratteristica di una misura di probabilità su $\mathbb{R}^n$). $N(a, Q)$ si dice anche misura Gaussiana di media a e matrice di covarianza Q. Si osservi che

$$\widehat{N(a, \sigma^2)}(h) = \int_{\mathbb{R}} e^{ihx}\, \frac{e^{\frac{-(x-a)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} dx = e^{iha}\, e^{-\,h^2\sigma^2/2}, \quad h \in \mathbb{R}.$$

Si può verificare che se Q è definita positiva allora $N(a, Q)$ ammette densità

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \det Q}} e^{-\frac{1}{2}\langle Q^{-1}(x-a),(x-a)\rangle}, \quad \text{cioè} \;\; N(a, Q)(B) = \int_B f(x)dx, \quad B \in B(\mathbb{R}^n).$$

Una v.a. X a valori in $\mathbb{R}^n$ si dice **v.a. normale** se la sua legge è $N(a, Q)$, per qualche scelta di a e Q. Si può allora verificare che $\mathbb{E}X = a$.

Usando la funzione caratteristica si possono provare i seguenti risultati.

Proposizione 1.12. *Siano X_1, ..., X_n v.a. reali indipendenti e normali con leggi $N(a_i, q_i)$, $i = 1, \dots, n$, rispettivamente. Allora la v.a. congiunta $X = (X_1, .. X_n)$ (a valori in $\mathbb{R}^n$) è normale. Inoltre la v.a. somma $Y = \sum_{i=1}^n X_i$ è normale con legge $N(\sum_{i=1}^n a_i, \sum_{i=1}^n q_i)$.*

Proposizione 1.13. *Sia X una v.a. a valori in $\mathbb{R}^n$, normale con legge $N(a, Q)$. Sia L una matrice reale $k \times n$ e sia $b \in \mathbb{R}^k$. Allora la v.a. $LX + b$ (a valori in $\mathbb{R}^k$) è normale con legge $N(La + b, LQL^T)$, dove L^T è la matrice trasposta di L.*

Siano X_1, ..., X_n v.a. reali e a quadrato integrabile su Ω. La **matrice di covarianza di** $X = (X_1$**, ...,** $X_n)$ è la matrice C, $n \times n$, i cui elementi sono

$$C_{ij} = \mathbb{E}[(X_i - \mathbb{E}X_i)(X_j - \mathbb{E}X_j)].$$

Si verifica che C è simmetrica e semidefinita positiva. Le v.a. X_1, .. X_n si dicono v.a. **scorrelate** se C è una matrice diagonale. In generale X e Y v.a. reali possono essere scorrelate senza essere indipendenti. Tuttavia vale la seguente

Proposizione 1.14. *Sia $X = (X_1, \ ..., \ X_n) : \Omega \to \mathbb{R}^n$ una v.a. normale con legge $N(a, Q)$. Valgono le seguenti proprietà:*

(i) la matrice di covarianza di X è Q;

(ii) X_1, ..., X_n sono v.a. scorrelate se e solo se sono indipendenti.

Passiamo ora in rassegna la **convergenza di variabili aleatorie.**

Sia (X_n) una successione di variabili aleatorie a valori in $\mathbb{R}^k$, sia $X : \Omega \to \mathbb{R}^k$ una v.a..

(1) Diciamo che (X_n) **converge q.c. (o quasi certamente o quasi ovunque)** a X, se $\lim_{n\to\infty} X_n(\omega) = X(\omega)$, per ogni $\omega \notin \Omega'$, con $\Omega' \in \mathcal{F}$ e $\mathbb{P}(\Omega') = 0$.

(2) Sia $\mathbb{E}(|X_n|^p + |X|^p) < \infty$, $n \in \mathbb{N}$, $p \geq 1$; diciamo che (X_n) **converge in L^p a X,** se

$$\lim_{n\to\infty} \mathbb{E}|X_n - X|^p = 0 \quad \text{(dove } |\cdot| \text{ indica la norma euclidea in } \mathbb{R}^k).$$

(osserviamo che poichè $\mathbb{P}(\Omega) = 1$, se (X_n) converge a X in L^p allora (X_n) converge a X in L^r, per ogni $1 \leq r \leq p$).

(3) Diciamo che (X_n) **converge in probabilità a** X se per ogni $\epsilon > 0$, si ha:

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \epsilon) = 0.$$

(4) Diciamo che (X_n) **converge in legge a** X se per ogni funzione $f : \mathbb{R}^k \to \mathbb{R}$, continua e limitata, si ha:

$$\lim_{n\to\infty} |\mathbb{E}f(X_n) - \mathbb{E}f(X)| = 0.$$

(si dice anche che le leggi μ_{X_n} convergono debolmente a μ_X). Un teorema afferma che (X_n) converge in legge a X se e solo se

$$\lim_{n\to\infty} \mathbb{E}e^{i\langle X_n, h\rangle} = \mathbb{E}e^{i\langle X, h\rangle}, \quad h \in \mathbb{R}^n.$$

Si può provare che *la convergenza in L^p e la convergenza q.c. implicano entrambe la convergenza in probabilità che a sua volta implica la convergenza in legge.*

Esempi mostrano che non ci sono altre implicazioni in generale.

Enunciamo ora il **lemma di Borel-Cantelli.**

Sia $(A_n)_{n\in\mathbb{N}} \subset \Omega$. Definiamo

$$\limsup A_n := \bigcap_{n\geq 0}\bigcup_{k\geq n} A_k = \{\omega\in\Omega \,:\, \text{per ogni } \; n\geq 0, \;\; \textit{esiste}\, k_0 = k_0(\omega) \geq n \;:\, \omega \in A_{k_0}\}.$$

e analogamente $\liminf A_n := \bigcup_{n\geq 0}\bigcap_{k\geq n} A_k = \{\omega\in\Omega \,:\, \textit{esiste} \;\, n_0 = n_0(\omega) \;\; : \omega \in A_k, \; \textit{per}\; k\geq n_0\}$. E' chiaro che $(\limsup A_n)^c = \liminf(A_n^c)$.

Lemma 1.15. *(Lemma di Borel-Cantelli) In uno spazio di probabilità* $(\Omega,\mathcal{F},\mathbb{P})$*, sia* $(A_n)_{n\in\mathbb{N}}\subset\mathcal{F}$*. Consideriamo* $A = \limsup A_n$*.*

(i) se $\sum_{k\geq 0}\mathbb{P}(A_k) < \infty$ *allora* $\mathbb{P}(A) = 0$*;*

(ii) se $\sum_{k\geq 0}\mathbb{P}(A_k) = \infty$ *e se gli* A_n *sono indipendenti allora* $\mathbb{P}(A) = 1$*.*

Dimostrazione. Proviamo solo la (i). Si ha

$$\mathbb{P}(A) = \mathbb{P}(\bigcup_{n\geq 0}\bigcap_{k\geq n} A_k) \leq \mathbb{P}(\bigcap_{k\geq n} A_k) \leq \sum_{k\geq n}\mathbb{P}(A_k),$$

per ogni $n\in\mathbb{N}$. Per ipotesi $\lim_{n\to\infty}\sum_{k\geq n}\mathbb{P}(A_k) = 0$, da cui la tesi. ■

Corollario 1.16. *Sia* (X_n) *una successione di variabili aleatorie a valori in* $\mathbb{R}^k$*, e sia* $X:\Omega\to\mathbb{R}^k$ *una v.a..; se* (X_n) *converge in probabilità ad* X*, allora esiste una sottosuccessione* (X_{n_k}) *che converge q.c. ad* X*.*

Dimostrazione. Fissato $n\in\mathbb{N}$, scegliamo k_n per cui

$$\mathbb{P}(|X_{k_n} - X| > 1/n) \leq 1/n^2.$$

Possiamo ordinare k_n in modo crescente. Sia $A_n = \{|X - X_{k_n}| > 1/n\}$. Si ha $\sum_{n\geq 1}\mathbb{P}(A_n) < \infty$. Dunque per il lemma precedente $\mathbb{P}(\limsup A_n) = 0$, cioè $\mathbb{P}(\liminf(A_n^c)) = 1$.

Perciò per ogni ω q.c., esiste $n_0 = n_0(\omega)$, tale che se $n\geq n_0$, $\omega\notin A_n$, cioè

$$|X_{k_n}(\omega) - X(\omega)| \leq 1/n, \quad n\geq n_0,$$

da cui la tesi. ■

Ricordiamo la definizione di **attesa condizionale.**

Proposizione 1.17. *Sia* X *una v.a. reale e integrabile. Sia* $\mathcal{G}\subset\mathcal{F}$ *una sotto* σ*-algebra. Allora esiste una v.a. reale* Z*,* $\mathcal{G}$*-misurabile, tale che*

$$\mathbb{E}(Z\,1_G) = \mathbb{E}(X\,1_G), \quad \textit{per}\; G\in\mathcal{G}. \tag{1.6}$$

Inoltre se Z_1 *e* Z_2 *sono due v.a.* $\mathcal{G}$*-misurabili che verificano* (1.6)*, allora si ha* $Z_1 = Z_2$*, q.c..*

Dimostrazione. Si usa il teorema di Radon-Nikodym. Infatti se si pone $\nu(G) = \mathbb{E}(X1_G)$, per $G\in\mathcal{G}$, si trova che ν è una misura finita con segno su $(\Omega,\mathcal{G})$, assolutamente continua rispetto a $\mathbb{P}$ (pensando $\mathbb{P}$ ristretta a $(\Omega,\mathcal{G})$). Dunque per il teorema di Radon-Nikodym, esiste una v.a. $Z\in L^1(\Omega,\mathcal{G},\mathbb{P})$, tale che

$$\nu(G) = \int_G Z(\omega)\mathbb{P}(d\omega),$$

ovvero $\nu(G) = \mathbb{E}(Z1_G)$, per ogni $G\in\mathcal{G}$. L'unicità segue dal fatto che $\mathbb{E}[(Z_1 - Z_2)1_G] = 0$, per ogni $G\in\mathcal{G}$, implica che $Z_1 = Z_2$, essendo Z_1 e Z_2 entrambe $\mathcal{G}$-misurabili. ■

Sia X una v.a. reale e integrabile. Sia $\mathcal{G}\subset\mathcal{F}$ una sotto σ-algebra. Si dice **attesa condizionale di** X **rispetto a** $\mathcal{G}$ e si indica con $\mathbb{E}(X/\mathcal{G})$, la classe di equivalenza di v.a. Z $\mathcal{G}$-misurabili che verificano (1.6). La definizione si estende facilmente a v.a. a valori in $\mathbb{R}^n$.

2 Prime nozioni sui processi stocastici

Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità. **Un processo stocastico** a tempo continuo (o semplicemente un processo) su $(\Omega, \mathcal{F}, \mathbb{P})$ a valori in $\mathbb{R}^n$ è una famiglia di variabili aleatorie $(X_t)_{t\in I}$, dove I è un **intervallo chiuso** di $[0, +\infty)$. Dunque

$$X_t : \Omega \to \mathbb{R}^n, \;\; t \in I.$$

Quando non ci sarà pericolo di confusione, scriveremo (X_t) al posto di $(X_t)_{t\in I}$. In altre parole, un processo stocastico è un'applicazione $X_{(\cdot)}(\cdot) : I\times\Omega \to \mathbb{R}^n$, che è $\mathcal{F}-$misurabile nella seconda variabile (vedere per esempio [15], [2] o [3] per maggiori dettagli).

Dato un processo (X_t), la mappa: $t \mapsto X_t(\omega)$, definita su I, a valori in $\mathbb{R}^n$, si dice **traiettoria** relativa a $\omega \in \Omega$. Le traiettorie di un processo (X_t) sono tutte le funzioni $t \mapsto X_t(\omega)$, al variare di $\omega \in \Omega$.

Un processo (X_t) si dice **continuo** se le sue traiettorie sono continue q.c. (cioè esiste un $\Omega' \in \mathcal{F}$ con $\mathbb{P}(\Omega') = 0$, tale che se $\omega \notin \Omega'$, la funzione: $t \mapsto X_t(\omega)$ è continua in I).

Definizione 2.1. *Dati due processi* (X_t) *e* (Y_t) *su* $(\Omega, \mathcal{F}, \mathbb{P})$*.*

(1) Diciamo che (X_t) *è una* **versione (o modificazione)** *di* (Y_t)*, se*

$$\mathbb{P}(X_t = Y_t) = \mathbb{P}(\{\omega \,:\, X_t(\omega) = Y_t(\omega)\}) = 1, \quad t \in I.$$

(2) Diciamo che (X_t) *e* (Y_t) *sono* **indistinguibili** *se* $\{X_t = Y_t, \;\; t \in I\} = \cap_{t\in I}\{X_t = Y_t\} \in \mathcal{F}$ *e risulta*

$$\mathbb{P}(X_t = Y_t, \;\; t \in I) = 1.$$

In generale (2) è strettamente più forte di (1). *Però nel caso in cui* (X_t) *e* (Y_t) *sono entrambi continui, se* (X_t) *è una versione di* (Y_t) *risulta che* (X_t) *e* (Y_t) *sono indistinguibili* (si noti che in questo caso $\{X_t = Y_t, \;\; t \in I\} = \{X_t = Y_t, \;\; t \in I \cap \mathbb{Q}\} \in \mathcal{F}$).

Dato un processo $(X_t)_{t\in I}$ possiamo considerare le sue **distribuzioni finito dimensionali** $\mu^X_{t_1,\ldots,t_m}$, al variare di $t_1 < \ldots < t_m$, $t_i \in I$, $i = 1, .., m$, e di $m \in \mathbb{N}$. Ciascuna $\mu^X_{t_1,\ldots,t_m}$ è per definizione la legge della v.a.

$$(X_{t_1}, ..., X_{t_m}) : \Omega \to \mathbb{R}^{nm}$$

(quindi $\mu^X_{t_1,\ldots,t_m}$ è una misura di probabilità su $B(\mathbb{R}^{nm})$). Due processi (anche definiti su spazi di probabilità diversi) a valori in $\mathbb{R}^n$ si dicono *equivalenti* se hanno le stesse distribuzioni finito dimensionali.

Parliamo ora di **filtrazioni**.

Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità. Una **filtrazione** è una famiglia crescente di sotto σ-algebre $(\mathcal{F}_t)_{t\in I}$ di $\mathcal{F}$, cioè $\mathcal{F}_t \subset \mathcal{F}_s \subset \mathcal{F}$, se $t \leq s$, $t, s \in I$.

Intuitivamente $\mathcal{F}_t$ raccoglie tutte le informazioni (o eventi) conosciute (fino) al tempo t.

Considereremo sempre **filtrazioni complete.** Cioè se

$$\mathcal{N}' = \{F \in \mathcal{F}, \; \mathbb{P}(F) = 0\},$$

allora assumiamo che $\mathcal{N}$' $\subset \mathcal{F}_t$, per ogni $t \in I$. Possiamo sempre estendere un filtrazione assegnata $(\mathcal{F}_t)$ in modo da ottenere una filtrazione completa: basta sostituire $\mathcal{F}_t$ con $\dot{\mathcal{F}}_t = \sigma(\mathcal{F}_t, \mathcal{N}')$ (la σ-algebra generata da $\mathcal{N}$' e $\mathcal{F}_t$), $t \in I$.

Uno spazio di probabilità su cui è definita anche una filtrazione completa, cioè un oggetto del tipo $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in I}, \mathbb{P})$, si chiama anche **base stocastica.** ([1])

Sia assegnato un processo $(X_t)_{t\in I}$ su una base stocastica. Diciamo che (X_t) è **adattato alla filtrazione o che è $\mathcal{F}_t$-adattato,** se per ogni $t \in I$, $X_t : (\Omega, \mathcal{F}_t) \to \mathbb{R}^n$ è misurabile.

Un processo è sempre adattato rispetto alla sua **filtrazione naturale** $(\mathcal{F}_t^X)_{t\in I}$**,** dove $\mathcal{F}_t^X = \sigma(X_s,\ s \leq t)$ (la più piccola σ-algebra che rende misurabili tutte le v.a. X_s, con $s \leq t$). Si noti anche che (X_t) è $\mathcal{F}_t$-adattato se e solo se $\mathcal{F}_t^X \subset \mathcal{F}_t$, $t \in I$.

3 Il processo di Wiener

3.1 Definizione e costruzione del processo di Wiener

Definizione 3.1. *(Processo di Wiener reale) Un processo stocastico $(W_t)_{t\geq 0}$ definito su uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ e a valori in $\mathbb{R}$, si dice processo di Wiener o moto Browiano, se verifica le seguenti proprietà:*

1. $W_0 = 0$, *q. c.;*
2. *per ogni scelta di* $0 = t_0 < t_1 < t_2 < \ldots < t_n$, $n \in \mathbb{N}$, *gli incrementi* $W_{t_1} - W_{t_0}, W_{t_2} - W_{t_1}, \ldots, W_{t_n} - W_{t_{n-1}}$ *sono indipendenti (si dice anche che il processo* (W_t) *è a incrementi indipendenti);*
3. *la variabile aleatoria* $W_t - W_s$, *per ogni* $0 \leq s < t$, *ha legge normale* $N(0, t-s)$ *;*
4. (W_t) *ha traiettorie continue q.c..*

Teorema 3.1. *Esiste un processo di Wiener (ovvero esiste uno spazio di probabilità dove è definito un processo di Wiener).*

Ci sono vari modi per dimostrare il precedente teorema. Qui presentiamo la **costruzione di Lévy-Ciesielski** del processo di Wiener, vedere per esempio [19], [8] o [22].

In questa costruzione si utilizza anche il Teorema 1.8. Si definisce prima (W_t) per $t \in [0,1]$, come limite di una particolare serie di funzioni. A questo fine si devono usare le funzioni di Haar e di Schauder.

Le **funzioni di Haar** (h_k), $k \in \mathbb{N}$, $h_k : [0,1] \to \mathbb{R}$, sono definite nel modo seguente:

$$h_0(t) = 1, \quad t \in [0,1],$$

$$h_1(t) = \begin{cases} 1 \text{ per } t \in [0, 1/2] \\ -1 \text{ per } t \in (1/2, 1], \end{cases}$$

[1]Nei testi di calcolo stocastico si richiede usualmente che una base stocastica abbia anche la filtrazione $(\mathcal{F}_t)$ continua a destra, cioè che valga $\mathcal{F}_t = \cap_{s>t}\mathcal{F}_s$, per ogni $t \in I$. Nei risultati che incontreremo non avremmo bisogno di tale ipotesi.

e se $2^n \leq k < 2^{n+1}$, $n \in \mathbb{N}$,

$$h_k(t) = \begin{cases} 2^{n/2} \text{ per } t \in \left[\dfrac{k-2^n}{2^n}, \dfrac{k-2^n+1/2}{2^n}\right], \\ -2^{n/2} \text{ per } t \in \left(\dfrac{k-2^n+1/2}{2^n}, \dfrac{k-2^n+1}{2^n}\right], \\ 0 \text{ altrimenti.} \end{cases}$$

Lemma 3.2. *Le funzioni (h_k) formano un sistema ortonormale e completo in $L^2(0,1)$.*

Dimostrazione. (i) Si ha: $\int_0^1 h_k^2(t)dt = 2^n(1/2^{n+1} + 1/2^{n+1}) = 1$, con $2^n \leq k < 2^{n+1}$, $n \in \mathbb{N}$.

Sia ora $l > k$. Allora o $h_k h_l = 0$, per ogni t (cioè h_k e h_l hanno supporti disgiunti), oppure h_k è costante sul supporto di h_l (si osservi che le h_k hanno supporto disgiunto se $2^n \leq k < 2^{n+1}$). In questo secondo caso si ha:

$$\int_0^1 h_k h_l dt = \pm 2^{n/2} \int_0^1 h_l dt = 0 \ \ (l > k).$$

(ii) Sia $f \in L^2(0,1)$ tale che $\int_0^1 f(t)h_k(t)dt = 0$, per $k \in \mathbb{N}$. Dobbiamo verificare che $f = 0$.

Proviamo prima di tutto che

$$\int_r^s f(t)dt = 0, \tag{3.1}$$

per tutti i numeri diadici r, s, con $r < s$.

Si ha $\int_0^1 f(t)h_0(t)dt = 0 = \int_0^1 f(t)dt$ (con $k = 0$). Sia ora $k = 1$, si ha

$$\int_0^{1/2} f dt = \int_{1/2}^1 f dt.$$

Perciò $2\int_0^{1/2} f dt = \int_0^1 f dt = 0$ e così $\int_0^{1/2} f dt = \int_{1/2}^1 f dt = 0$. Procedendo in modo simile si trova che

$$\int_{k/2^{n+1}}^{\frac{k+1}{2^{n+1}}} f dt = 0, \ \text{ per ogni } \ 0 \leq k < 2^{n+1}.$$

Così (3.1) è provata. Usando (3.1) ed il teorema di Lebesgue sulla derivazione della funzione integrale, si trova che

$$f(r) = \frac{d}{dr}\int_0^r f(t)dt = 0, \ \text{ q.ov. in } \ r \in [0,1],$$

da cui la tesi. ■

Le **funzioni di Schauder** (s_k), $k \in \mathbb{N}$, $s_k : [0,1] \to \mathbb{R}$, sono primitive delle funzioni di Haar:

$$s_k(t) = \int_0^t h_k(s)ds, \ \ k \in \mathbb{N}.$$

Le s_k sono tutte non negative; i grafici delle s_k sono "dei denti " che si trovano sugli intervalli $[\frac{k-2^n}{2^n}, \frac{k-2^n+1}{2^n}]$, se $2^n \leq k < 2^{n+1}$. Inoltre risulta:

$$\|s_k\|_\infty = \sup_{t\in[0,1]} |s_k(t)| = 2^{n/2}\frac{1}{2^{n+1}} = 2^{-n/2-1}, \;\; 2^n \leq k < 2^{n+1}. \tag{3.2}$$

Per continuare serve il seguente lemma.

Lemma 3.3. *Siano (X_k) una successione di variabili aleatorie su $(\Omega, \mathcal{F}, \mathbb{P})$ normali con legge $N(0,1)$. Risulta che $|X_k(\omega)|$ è $O(\sqrt{\log k})$, per $k \to \infty$, q.c..*

Dimostrazione. Per $x > 0$, si ha:

$$\mathbb{P}(|X_k| > x) = \mathbb{P}(\omega \in \Omega \,:\, |X_k| > x) = \frac{2}{\sqrt{2\pi}} \int_x^\infty e^{-s^2/4} e^{-s^2/4} ds$$
$$\leq \frac{2}{\sqrt{2\pi}} e^{-x^2/4} \int_x^\infty e^{-s^2/4} ds \leq C e^{-x^2/4}.$$

Sia ora $x = 4\sqrt{\log k}$. Si trova $\mathbb{P}(|X_k| > 4\sqrt{\log k}) \leq c e^{-4\log k} = \frac{c}{k^4}$, $k \in \mathbb{N}$. Per il lemma di Borel-Cantelli, per ogni ω (q.c.) esiste $k_0 = k_0(\omega)$ tale che per ogni $k \geq k_0$, vale

$$|X_k(\omega)| \leq 4\sqrt{\log k},$$

da cui la tesi. ■

Dimostrazione del Teorema 3.1. Consideriamo uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$, dove sia definita una successione di variabili aleatorie (X_k) indipendenti e normali con legge $N(0,1)$. Proviamo che

$$W_t(\omega) = \sum_{k\geq 0} X_k(\omega) s_k(t), \;\; \omega \in \Omega, \tag{3.3}$$

definisce un processo di Wiener su $[0,1]$ (ovviamente $W_0(\omega) = 0$ per ogni ω).

(i) Per prima cosa proviamo che, q.c., la serie converge uniformemente su $[0,1]$ (questo proverà anche che $t \mapsto W_t(\omega)$ è continua su $[0,1]$, q.c.). A questo fine, mostriamo che per ogni $\eta > 0$, esiste $n_0 \in \mathbb{N}$, tale che se $n \geq n_0$ e $p \geq 1$,

$$\Big\| \sum_{k=0}^{n} X_k(\omega)\, s_k - \sum_{k=0}^{n+p} X_k(\omega)\, s_k \Big\|_\infty = \sup_{t\in[0,1]} \Big| \sum_{k=n+1}^{n+p} X_k(\omega)\, s_k(t) \Big| < \eta. \tag{3.4}$$

Osserviamo anzitutto che il lemma precedente implica, per ogni $\epsilon > 0$,

$$|X_k(\omega)| \text{ è } O(k^\epsilon), \;\; q.c.$$

Usando (3.2) e il fatto che le s_k hanno supporto disgiunto se $2^n \leq k < 2^{n+1}$, si trova

$$\sup_{t\in[0,1]} \sum_{2^n \leq k < 2^{n+1}} s_k(t) \leq 2^{-n/2-1}.$$

Perciò, fissato ω (q.c.), per $t \in [0,1]$, $m \in \mathbb{N}$, $\epsilon < 1/2$, si ha:

$$\begin{aligned}
&\sum_{k \geq 2^m} |X_k(\omega)|\, s_k(t) = \sum_{n \geq m} \sum_{2^n \leq k < 2^{n+1}} |X_k(\omega)|\, s_k(t) \\
&\leq C \sum_{n \geq m} 2^{n\epsilon} \sup_{t \in [0,1]} \sum_{2^n \leq k < 2^{n+1}} s_k(t) \\
&\leq C' \sum_{n \geq m} 2^{n(\epsilon - 1/2)}.
\end{aligned}$$

Essendo $\sum_{n \geq 0} 2^{n(\epsilon-1/2)} < \infty$, possiamo prendere m_0 tale che $C' \sum_{n \geq m_0} 2^{n(\epsilon-1/2)} < \eta$. Scegliendo $n_0 = 2^{m_0}$ si trova (3.4).

Questo prova la convergenza uniforme della serie di funzioni.

(ii) Proviamo che $W_t - W_s$ ha legge $N(0, t-s)$, per $t > s \geq 0$.

Basta verificare che

$$\mathbb{E}(e^{i(W_t - W_s)h}) = e^{-(t-s)h^2/2}, \quad h \in \mathbb{R}.$$

Usando l'indipendenza degli X_k e il teorema di convergenza dominata, si trova

$$\begin{aligned}
&\mathbb{E}(e^{i(W_t - W_s)h}) = \mathbb{E}(e^{ih \sum_{k \geq 0} X_k (s_k(t) - s_k(s))}) \\
&= \prod_{k \geq 0} \mathbb{E}(e^{ihX_k(s_k(t) - s_k(s))}) \\
&= \prod_{k \geq 0} (e^{-\frac{h^2}{2}(s_k(t) - s_k(s))^2}) \qquad \text{(poichè ogni } X_k \text{ è normale)} \\
&= e^{-\frac{h^2}{2} \sum_{k \geq 0} (s_k(t) - s_k(s))^2} = e^{-\frac{h^2}{2} \sum_{k \geq 0} (s_k^2(t) + s_k^2(s) - 2 s_k(t) s_k(s))}.
\end{aligned}$$

Per continuare, osserviamo che, per $t \geq s$,

$$\begin{aligned}
\sum_{k \geq 0} s_k(t) s_k(s) &= \sum_{k \geq 0} \Big(\int_0^1 h_k 1_{[0,t]} dr \Big) \Big(\int_0^1 h_k 1_{[0,s]} dr \Big) \\
&= \Big(\int_0^1 1_{[0,t]}(r) dr \Big) \Big(\int_0^1 1_{[0,s]}(r) dr \Big) = s,
\end{aligned} \tag{3.5}$$

essendo (h_k) un sistema ortonormale e completo in $L^2(0,1)$. Dunque si ha:

$$\mathbb{E}(e^{i(W_t - W_s)h}) = e^{-\frac{h^2}{2} \sum_{k \geq 0} (s_k^2(t) + s_k^2(s) - 2 s_k(t) s_k(s))} = e^{-\frac{h^2}{2}(t - 2s + s)} = e^{-\frac{h^2}{2}(t-s)}.$$

(iii) Proviamo che (W_t) ha gli incrementi indipendenti.

Si tratta di provare che, per $h_1, \ldots h_n \in \mathbb{R}$, $0 = t_0 < t_1 < t_2 < \ldots < t_n \leq 1$,

$$\mathbb{E}\Big(e^{i \sum_{j=1}^n h_j (W_{t_j} - W_{t_{j-1}})}\Big) = \prod_{j=1}^n e^{-h_j^2 (t_j - t_{j-1})/2} = \prod_{j=1}^n \mathbb{E}\Big(e^{i\, h_j (W_{t_j} - W_{t_{j-1}})}\Big), \tag{3.6}$$

vedere la Proposizione 1.11.

Proviamo la tesi solo per $n = 2$, cioè mostriamo che $W_t - W_s$ è indipendente da W_s, con $t > s > 0$. Il caso generale si può trattare in modo simile, partendo dall'identità

$$\mathbb{E}\Big(e^{i\sum_{j=1}^{n}(h_j - h_{j+1})W_{t_j}}\Big) = \mathbb{E}\Big(e^{i\sum_{j=1}^{n}h_j(W_{t_j} - W_{t_{j-1}})}\Big), \quad \text{ponendo } h_{n+1} = 0.$$

Sia $t_2 = t$, $t_1 = s$, ragionando come in precedenza,

$$\begin{aligned} &\mathbb{E}(e^{i((W_t - W_s)h_2 + W_s h_1)}) = \mathbb{E}(e^{i(h_2 W_t + W_s(h_1 - h_2))}) \\ &= \prod_{k\geq 0}\mathbb{E}(e^{iX_k[h_2 s_k(t) + s_k(s)(h_1-h_2)]}) = \prod_{k\geq 0} e^{-\frac{1}{2}(h_2 s_k(t) + s_k(s)(h_1-h_2))^2} \\ &= e^{-\frac{1}{2}(h_2^2 t + 2(h_1-h_2)h_2 s + (h_1-h_2)^2 s)} = e^{-\frac{1}{2}[h_1^2 s + h_2^2(t-s)]}. \end{aligned}$$

Per concludere la dimostrazione occorre estendere la definizione di (W_t) in (3.3) a $t \in [0, \infty)$.

Consideriamo una successione di spazi di probabilità $(\Omega_k, \mathcal{F}_k, \mathbb{P}_k) = (\Omega, \mathcal{F}, \mathbb{P})$, $k \geq 0$, dove sullo spazio $(\Omega, \mathcal{F}, \mathbb{P})$ è definito un processo di Wiener W_t, $t \in [0, 1]$. Introduciamo la spazio di probabilità prodotto

$$(\dot\Omega, \dot{\mathcal{F}}, \dot{\mathbb{P}}), \quad \dot\Omega = \prod_{k\geq 0}\Omega_k, \quad \dot{\mathcal{F}} = \prod_{k\geq 0}\mathcal{F}_k, \quad \dot{\mathbb{P}} = \prod_{k\geq 0}\mathbb{P}_k,$$

vedere Teorema 1.9. Indichiamo con W_t^k il processo di Wiener su $(\Omega_k, \mathcal{F}_k, \mathbb{P}_k)$. Definiamo in modo ricorsivo: $\dot W_t = W_t^1$, *se* $t \in [0, 1]$,

$$\dot W_t = \dot W_n + W_{t-n}^{n+1}, \quad \textit{se} \;\; t \in [n, n+1],\; n \in \mathbb{N},\; n \geq 1$$

(dove identifichiamo W_t^k con $\dot W_t^k$, $\dot W_t^k : \dot\Omega \to \mathbb{R}$, $\dot W_t^k(\omega) := W_t^k(\omega_k)$, $\omega = (\omega_k) \in \dot\Omega$, $k \in \mathbb{N}$, $t \in [0, 1]$).

Ovvero, se $\omega = (\omega_k) \in \dot\Omega$,

$$\begin{aligned} &\dot W_t(\omega) = W_t^1(\omega_1), \quad t \in [0,1], \\ &\dot W_t(\omega) = W_1^1(\omega_1) + W_{t-1}^2(\omega_2), \quad t \in [1,2], \\ &\dots \\ &\dot W_t(\omega) = \sum_{k=1}^{n} W_1^k(\omega_k) + W_{t-n}^{n+1}(\omega_{n+1}), \;\; t \in [n, n+1]. \end{aligned}$$

Usando le proprietà prima stabilite per $t \in [0, 1]$ e la definizione di misura prodotto, si verifica che $(\dot W_t)$ $(t \in [0, \infty))$ è un processo di Wiener su $(\dot\Omega, \dot{\mathcal{F}}, \dot{\mathbb{P}})$. Per stabilire la continuità delle traiettorie di $(\dot W_t)$ si usa anche che $W_0^k = 0$, q.c., $k \in \mathbb{N}$.

Proviamo solo che $\dot W_t - \dot W_s$, $t > s$, ha legge $N(0, t - s)$. Nel seguito non indichiamo la dipendenza da ω. Siano $m \leq s < m + 1$, $n \leq t < n + 1$. Consideriamo il caso non ovvio in cui $m < n$. Si ha:

$$\dot W_t - \dot W_s = \sum_{k=1}^{n} W_1^k + W_{t-n}^{n+1} - (\sum_{k=1}^{m} W_1^k + W_{s-m}^{m+1})$$

$$= \sum_{k=m+1}^{n} W_1^k + (W_{t-n}^{n+1} - W_{s-m}^{m+1}) = Y_1 + Y_2 + Y_3, \;\; \text{con}$$

$$Y_1 = \sum_{k=m+2}^{n} W_1^k \;\; , Y_2 = (W_1^{m+1} - W_{s-m}^{m+1}), \quad Y_3 = W_{t-n}^{n+1},$$

dove Y_1, Y_2, Y_3 sono v.a. normali e indipendenti, grazie alle proprietà della misura prodotto $\dot{\mathbb{P}}$. Usando la Proposizione 1.12, si trova che $Y_1 + Y_2 + Y_3$ ha legge $N(0, n - m - 1 + (1 - s + m)) + t - n) = N(0, t - s)$.

Questo conclude la dimostrazione. ∎

Ora introduciamo il **processo di Wiener n-dimensionale.**

Definizione 3.2. *(Processo di Wiener n-dimensionale) Un processo stocastico* $(W_t)_{t\geq 0}$ *definito su uno spazio di probabilità* $(\Omega, \mathcal{F}, \mathbb{P})$ *e a valori in* $\mathbb{R}^n$*, si dice processo di Wiener n-dimensionale, se valgono le seguenti proprietà:*

1. $W_0 = 0$, *q. c.;*
2. *per ogni scelta di* $0 = t_0 < t_1 < t_2 < \ldots < t_n$, $n \in \mathbb{N}$, *gli incrementi* $W_{t_1} - W_{t_0}, W_{t_2} - W_{t_1}, \ldots, W_{t_n} - W_{t_{n-1}}$ *sono indipendenti;*
3. *la variabile aleatoria (n-dimensionale)* $W_t - W_s$, *per ogni* $0 \leq s < t$, *ha legge normale* $N(0, (t - s)I)$ *;*
4. (W_t) *ha traiettorie continue q.c..*

Un processo di Wiener $n-$dimensionale, si può costruire partendo da un processo di Wiener reale (W_t) definito su qualche spazio $(\Omega, \mathcal{F}, \mathbb{P})$. Si introduce prima lo spazio di probabilità prodotto

$$(\dot{\Omega}, \dot{\mathcal{F}}, \dot{\mathbb{P}}), \quad \dot{\Omega} = \prod_{k=1}^{n} \Omega_k, \quad \dot{\mathcal{F}} = \prod_{k=1}^{n} \mathcal{F}_k, \quad \dot{\mathbb{P}} = \prod_{k=1}^{n} \mathbb{P}_k,$$

dove $(\Omega, \mathcal{F}, \mathbb{P}) = (\Omega_k, \mathcal{F}_k, \mathbb{P}_k)$, per $k = 1, ..., n$. In $(\dot{\Omega}, \dot{\mathcal{F}}, \dot{\mathbb{P}})$, si pone:

$$\dot{W}_t(\omega) := (W_t(\omega_1), ..., W_t(\omega_n)), \quad \omega = (\omega_1, ..., \omega_n) \in \dot{\Omega},\ t \geq 0.$$

Usando le proprietà della misura prodotto si può verificare che $(\dot{W}_t)$ è un processo di Wiener n-dimensionale.

Dato un processo di Wiener n-dimensionale (W_t) con $W_t = (W_t^k)$, $k = 1, ..., n$, $t \geq 0$, si verifica facilmente che ogni (W_t^k) è un processo di Wiener reale.

Si può dare una definizione più generale di processo di Wiener (reale o n-dimensionale) che fa intervenire le *filtrazioni.*

Definizione 3.3. *(Processo di Wiener n-dimensionale rispetto ad una filtrazione) Un processo stocastico* $(W_t)_{t\geq 0}$, *a valori in* $\mathbb{R}^n$, *definito su una base stocastica* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ *e* $\mathcal{F}_t$*-adattato, si dice processo di Wiener* `rispetto a` $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, *se verifica le seguenti proprietà:*

1. $W_0 = 0$, *q. c.;*
2. $W_t - W_s$ *è indipendente da* $\mathcal{F}_s$ *per ogni scelta di* $0 \leq s < t$;
3. *la variabile aleatoria* $W_t - W_s$, *per ogni* $0 \leq s < t$, *ha legge normale* $N(0, (t - s)I)$;
4. (W_t) *ha traiettorie continue q.c..*

Questa definizione è più generale di quella precedente per il fatto che si ammette la presenza di un filtrazione diversa da quella naturale.

Più precisamente, sia (W_t) un processo di Wiener, secondo la Definizione 3.1; si definisce la **filtrazione naturale** $(\mathcal{F}_t^W)_{t\geq 0}$ **(**$\mathcal{F}_t^W = \sigma(W_s,\, s \leq t)$**,** con $t \geq 0$). Si prova che (W_t) è di Wiener rispetto alla base stocastica $(\Omega, \mathcal{F}, (\mathcal{F}_t^W)_{t\geq 0}, \mathbb{P})$, secondo la Definizione 3.3. Per la dimostrazione si può usare il seguente risultato.

Lemma 3.4. *Sia Y una v.a. e $(X_t)_{t\in I}$ un processo su $(\Omega, \mathcal{F}, \mathbb{P})$. Se per ogni scelta di $t_1 < t_2 ... < t_m$ in I, risulta che Y è indipendente dalla v.a. $(X_{t_1}, ..., X_{t_m})$ allora $\sigma(Y)$ è indipendente da $\mathcal{G} = \sigma(X_t,\ t \in I)$.*

Dimostrazione. Osserviamo che

$$\mathcal{H} = \{H \in \mathcal{F} \,:\, \mathbb{P}(Y \in A, H) = \mathbb{P}(Y \in A)\mathbb{P}(H), \text{ per ogni } A \in B(\mathbb{R}^n)\}$$

è un λ-system. Inoltre per ipotesi $\mathcal{H}$ contiene gli eventi $\{(X_{t_1}, ..., X_{t_m}) \in A_{t_1} \times ... \times A_{t_m}\}$ (con $A_{t_i} \in B(\mathbb{R}^n)$). Questi eventi formano un π-system. Per il Lemma 1.4 si conclude che $\mathcal{G} \subset \mathcal{H}$, da cui la tesi. ∎

3.2 Proprietà delle traiettorie del processo di Wiener

In questa sezione consideriamo processi di Wiener reali. Questo non è restrittivo ai fini dello studio delle proprietà delle traiettorie.

Una funzione $f : [a, b] \to \mathbb{R}^n$, si dice **$\gamma$-Hölderiana** (su $[a, b]$), con $\gamma \in (0, 1]$, se

$$[f]_\gamma := \sup_{t,s\in[a,b],\, t\neq s} \frac{|f(t) - f(s)|}{|t - s|^\gamma} < \infty$$

($|\cdot|$ indica la norma euclidea di ogni $\mathbb{R}^n$). Lo spazio di tutte le funzioni $\gamma-$Hölderiane da $[a, b]$ in $\mathbb{R}^n$ è uno spazio di Banach con la norma $\|\cdot\|_\gamma = \|\cdot\|_\infty + [\cdot]_\gamma$; si indica con $C^\gamma[a, b]$. Osserviamo che se $f \in C^\gamma[a, b]$, allora $f \in C^{\gamma'}[a, b]$, per ogni $\gamma' \leq \gamma$.

Proposizione 3.5. *Le traiettorie del processo di Wiener definito nel Teorema 3.1, q.c., su ogni $[0, T]$, $T > 0$, sono $\gamma-$Hölder per qualunque $\gamma \in (0, 1/2)$.*

Dimostrazione. I Passo. Per semplicità prendiamo $T = 1$. Sia $\|\cdot\|_\gamma = \|\cdot\|_\infty + [\cdot]_\gamma$ la norma di $C^\gamma[0, 1]$, Ragioneremo in modo simile alla (i) della dimostrazione del Teorema 3.1.

Fissato ω (q.c.), si tratta di considerare, per $m \in \mathbb{N}$,

$$\Big\| \sum_{k\geq 2^m} X_k(\omega)\, s_k \Big\|_\gamma = \Big\| \sum_{n\geq m} \sum_{2^n\leq k<2^{n+1}} X_k(\omega)\, s_k \Big\|_\gamma \leq \sum_{n\geq m} \Big\| \sum_{2^n\leq k<2^{n+1}} X_k(\omega)\, s_k \Big\|_\gamma . \tag{3.7}$$

Adesso, per come sono costruite le funzioni $X_k(\omega)s_k$ (notare che hanno anche supporto disgiunto), si trova, per $2^n \leq k < 2^{n+1}$,

$$\begin{aligned} &\Big\| \sum_{2^n\leq k<2^{n+1}} X_k(\omega)s_k \Big\|_\gamma \leq \max_{2^n\leq k<2^{n+1}} \|X_k(\omega)s_k\|_\gamma \\ &\leq \max_{2^n\leq k<2^{n+1}} |X_k(\omega)|\, \|s_k\|_\gamma \leq 2^{(n+1)\epsilon} \max_{2^n\leq k<2^{n+1}} \|s_k\|_\gamma \\ &= 2^{(n+1)\epsilon} \|s_{2^n}\|_\gamma. \end{aligned}$$

Ora $\|s_k\|_\gamma = \|s_k\|_\infty + [s_k]_\gamma$. Per calcolare $[s_{2^n}]_\gamma$ si può procedere così. Data una funzione lineare $f(t) = \frac{b}{a}\, t$, $t \in [0, a]$, a, $b > 0$, si ha

$$[f]_\gamma = \sup_{t \in [0,a]} \frac{b}{a} \frac{t}{t^\gamma} = \frac{b}{a^\gamma}.$$

Dunque, usando anche le proprietà di simmetria delle s_k, si ha:

$$[s_k]_\gamma = 2^{-n/2-1} 2^{(n+1)\gamma} = c_\gamma 2^{n(\gamma-1/2)}, \quad 2^n \leq k < 2^{n+1}.$$

Ritornando alla stima (3.7), si trova che

$$\sum_{n \geq m} \Big\| \sum_{2^n \leq k < 2^{n+1}} X_k(\omega)\, s_k \Big\|_\gamma \leq c_\gamma \sum_{n \geq m} 2^{n\epsilon} (2^{n(\gamma-1/2)} + 2^{-n/2}).$$

Fissato $\gamma \in (0, 1/2)$, basta scegliere $\epsilon > 0$ abbastanza piccolo per ottenere che

$$\sum_{n \geq 0} 2^{n\epsilon} (2^{n(\gamma-1/2)} + 2^{-n/2}) < \infty.$$

Perciò $\sum_{k \geq 0} X_k(\omega)\, s_k(t)$ converge nella norma di $C^\gamma[0, 1]$. Questo prova la tesi.

II Passo. Considero tutti gli intervalli $[0, n]$, $n \in \mathbb{N}$. Per ogni $[0, n]$, esiste un $\Omega_n \in \mathcal{F}$ con $\mathbb{P}(\Omega_n) = 0$, tale che se $\omega \notin \Omega_n$, la funzione: $t \mapsto W_t(\omega)$ è γ-Hölder su $[0, n]$, per ogni $\gamma \in (0, 1/2)$. Consideriamo

$$\Omega_0 = \cup_{n \geq 0} \Omega_n. \tag{3.8}$$

Certamente $\mathbb{P}(\Omega_0) = 0$ e se $\omega \notin \Omega_0$, la funzione $t \mapsto W_t(\omega)$, definita su $[0, +\infty)$, è localmente γ-Hölder, per ogni $\gamma \in (0, 1/2)$. ■

Si può anche dimostrare, usando il seguente teorema, che **ogni** processo di Wiener (W_t) (definito su un qualche spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$) ha traiettorie che sono, q.c., localmente $\gamma-$Hölderiane, per ogni $\gamma \in (0, 1/2)$.

Teorema 3.6. *(Teorema di regolarità di Kolmogorov) Sia $(X_t)_{t \in I}$, $I = [h, k] \subset \mathbb{R}_+$, un processo definito su $(\Omega, \mathcal{F}, \mathbb{P})$ a valori in $\mathbb{R}^n$. Supponiamo che esistano costanti positive a, b e c tali che*

$$\mathbb{E}|X_t - X_s|^b \leq c|t - s|^{1+a}, \quad t,\, s \in I.$$

Allora esiste una versione (Y_t) di (X_t), tale che le traiettorie di (Y_t) sono q.c. $\gamma-$Hölderiane su I, per ogni esponente $\gamma < \frac{a}{b}$.

Per la dimostrazione si possono consultare [15], [3], [16], [9], [13]. Osserviamo che se richiediamo in più che (X_t) sia continuo allora otteniamo che (X_t) e (Y_t) sono indistinguibili. Quindi se (X_t) è continuo e valgono le ipotesi del teorema precedente, le traiettorie di (X_t) sono q.c. $\gamma-$Hölderiane su I, per $\gamma < \frac{a}{b}$.

Corollario 3.7. *Le traiettorie di un (qualunque) processo di Wiener (W_t), q.c., su ogni $[0, T]$, $T > 0$, sono $\gamma-$Hölder per qualunque $\gamma \in (0, 1/2)$.*

Dimostrazione. Si tratta di applicare il precedente teorema di Kolmogorov. Si ha:

$$\mathbb{E}|W_t - W_s|^{2m} = \mathbb{E}|W_{t-s}|^{2m} = \int_{\mathbb{R}} x^{2m} \frac{e^{-\frac{x^2}{2(t-s)}}}{\sqrt{2\pi(t-s)}} dx$$

$$= (t-s)^m \int_{\mathbb{R}} y^{2m} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy = C_m |t-s|^m, \quad m \in \mathbb{N}, \quad 0 \le s < t \le T$$

(si potrebbe mostrare che $C_m = (2m-1)(2m-3)...$). Perciò $a = m-1$, $b = 2m$ e $\gamma < a/b < \frac{1}{2} - \frac{1}{2m}$, per ogni $m \in \mathbb{N}$. La tesi segue facilmente. ■

Il valore $\gamma = 1/2$ è un valore critico per le traiettorie di un processo di Wiener e non può essere raggiunto. Questo fatto segue da un notevole risultato, molto più preciso della precedente corollario. Questo risultato è stato scoperto da Paul Lévy (per la dimostrazione vedere per esempio [15]).

Teorema 3.8. *(di P. Lévy) Sia (W_t) una processo di Wiener (definito su un qualche spazio di probabilità). Per ogni ω, q.c., risulta:*

$$\limsup_{\delta \to 0^+} \sup_{0 \le s < t \le 1,\ |t-s| \le \delta} \frac{|W_t(\omega) - W_s(\omega)|}{(2\delta \log(1/\delta))^{1/2}} = 1.$$

Ora proveremo che le traiettorie di un processo di Wiener (W_t) non sono a variazione limitata (o BV) su ogni intervallo limitato di $[0, \infty)$.

Ricordiamo la definizione di **funzione BV o a variazione limitata**. Sia $[a,b] \subset \mathbb{R}$ e $f : [a,b] \to \mathbb{R}^n$ una funzione. La **variazione** di f su $[a,b]$ è

$$V_a^b(f) = \sup_{\pi} \sum_{t_i < t_{i+1},\ t_i, t_{i+1} \in \pi} |f(t_i) - f(t_{i+1})|,$$

dove π varia tra tutte le possibile partizioni (finite) di $[a,b]$, cioè $\pi = \{t_0, ..., t_m\}$, con $a = t_0 < t_1 ... < t_m = b$ (al variare anche di $m \in \mathbb{N}$).

f si dice BV su $[a,b]$ se $V_a^b(f) < \infty$.

Data una partizione π di $[a,b]$, introduciamo

$$|\pi| = \max_{0 \le k \le m-1} |t_{k+1} - t_k|.$$

Teorema 3.9. *Le traiettorie di un processo di Wiener (W_t), q.c., non sono BV su ogni $[a,b] \subset \mathbb{R}_+$.*

Dimostrazione. La dimostrazione è in 3 passi.

I Passo. Sia $\pi = \{t_0, ..., t_m\}$, una partizione di $[a,b]$. Introduciamo

$$S_\pi(\omega) = \sum_{k=0}^{m-1} |W_{t_{k+1}}(\omega) - W_{t_k}(\omega)|^2$$

Proviamo che, per ogni $\epsilon > 0$, esiste un $\delta > 0$, tale che se $|\pi| < \delta$ allora

$$\mathbb{E}|S_\pi - (b-a)|^2 < \epsilon \quad (\text{cioè } \lim_{|\pi| \to 0} S_\pi = b - a \text{ in } L^2) \tag{3.9}$$

(si dice anche che la *variazione quadratica di* (W_t) *è* $b-a$).

Fissata di una partizione $\pi = \{t_0, ..., t_m\}$, consideriamo

$$\mathbb{E}|S_\pi - (b-a)|^2 = \mathbb{E}\Big|\sum_{k=0}^{m-1} |W_{t_{k+1}} - W_{t_k}|^2 - \sum_{k=0}^{m-1}(t_{k+1}-t_k)\Big|^2$$
$$= \mathbb{E}\Big|\sum_{k=0}^{m-1}[|W_{t_{k+1}} - W_{t_k}|^2 - (t_{k+1}-t_k)]\Big|^2.$$

Poichè le v.a. $X_k = |W_{t_{k+1}} - W_{t_k}|^2 - (t_{k+1}-t_k)$ sono indipendenti, con media 0, si ha (vedere Proposizione 1.7)

$$\mathbb{E}|S_\pi - (b-a)|^2 = var\Big(\sum_{k=0}^{m-1} X_k\Big) = \sum_{k=0}^{m-1} var(X_k)$$
$$= \sum_{k=0}^{m-1} \mathbb{E}[|W_{t_{k+1}} - W_{t_k}|^2 - (t_{k+1}-t_k)]^2$$
$$= \sum_{k=0}^{m-1} \mathbb{E}\Big[(t_{k+1}-t_k)\Big(\frac{|W_{t_{k+1}} - W_{t_k}|^2}{t_{k+1}-t_k} - 1\Big)\Big]^2 = \sum_{k=0}^{m-1}(t_{k+1}-t_k)^2\mathbb{E}\Big[\frac{|W_{t_{k+1}} - W_{t_k}|^2}{t_{k+1}-t_k} - 1\Big]^2$$

Ora per ogni $k = 1, ..., m$, $\frac{W_{t_{k+1}} - W_{t_k}}{\sqrt{t_{k+1}-t_k}}$ è una v.a. normale con legge $N(0,1)$.

Pertanto $\mathbb{E}\Big[\frac{|W_{t_{k+1}} - W_{t_k}|^2}{t_{k+1}-t_k} - 1\Big]^2 = c$, con c indipendente da k (si verifica che $c = 2$). Si ha:

$$\mathbb{E}|S_\pi - (b-a)|^2 = c\sum_{k=0}^{m-1}(t_{k+1}-t_k)^2 \le c|\pi|\sum_{k=0}^{m-1}(t_{k+1}-t_k) = c|\pi|(b-a) \;\to\; 0, \quad (3.10)$$

quando $|\pi| \to 0$. Dunque (3.9) è provata.

II Passo. Verifichiamo che $W_t(\omega)$ non è a variazione limitata su $[a,b]$, per $\omega \notin \Omega_{a,b}$, con $\Omega_{a,b} \in \mathcal{F}$ e $\mathbb{P}(\Omega_{a,b}) = 0$. Si ha

$$(*)\;\; S_\pi(\omega) \le \max_{0\le i\le m-1} |W_{t_{i+1}}(\omega) - W_{t_i}(\omega)| \sum_{i=0}^{m-1} |W_{t_{i+1}}(\omega) - W_{t_i}(\omega)|.$$

Ora per la continuità delle traiettorie si ha, q.c.,

$$\lim_{|\pi|\to 0} \max_{0\le i\le m-1} |W_{t_{i+1}}(\omega) - W_{t_i}(\omega)| = 0.$$

Se per assurdo, fosse $\sum_{i=0}^{m-1} |W_{t_i+1}(\omega) - W_{t_i}(\omega)| \le C(\omega) < \infty$, per ogni π, con $\omega \in \Omega'$, $\Omega' \in \mathcal{F}$ e $\mathbb{P}(\Omega') > 0$, allora da (*) si avrebbe che

$$\lim_{|\pi|\to 0} S_\pi(\omega) = 0, \;\; \omega \in \Omega'.$$

Questo contraddice (3.9).

III Passo. Proviamo la tesi. Consideriamo tutti gli intervalli $[a,b]$, con $a,b \in \mathbb{Q}$. Introduciamo

$$\Omega' = \bigcup_{a,b\in\mathbb{Q}, a\le b} \Omega_{a,b}.$$

Risulta che $\Omega' \in \mathcal{F}$ e $\mathbb{P}(\Omega') = 0$. Dunque se $\omega \notin \Omega'$, risulta che $t \mapsto W_t(\omega)$ non è BV su ogni $[c,d] \subset \mathbb{R}_+$. ■

Corollario 3.10. *Le traiettorie di un processo di Wiener (W_t), q.c., non sono γ-Hölder, con $\gamma > 1/2$, su ogni $[a,b] \subset \mathbb{R}_+$.*

Dimostrazione. Fissiamo $[a,b] \subset \mathbb{R}_+$ e consideriamo $\Omega' \in \mathcal{F}$ con $\mathbb{P}(\Omega') = 1$ per cui esiste una successione di partizioni (π_n) di $[a,b]$ tali che $\lim_{n\to\infty} S_{\pi_n}(\omega') = b - a, \quad \omega \in \Omega'$ (tale Ω' esiste grazie a (3.9)). Proviamo che per ogni $\omega' \in \Omega'$, la traiettoria $t \mapsto W_t(\omega')$ non è γ-Hölder se $\gamma > 1/2$. Se per assurdo fissato $\omega' \in \Omega'$ esistesse $\gamma = \gamma(\omega') \in (1/2, 1]$ per cui $t \mapsto W_t(\omega')$ è γ-Hölder su $[a,b]$ avremmo

$$S_{\pi_n}(\omega') \le C \sum_{k=0}^{m_n-1} |t_{k+1}^n - t_k^n|^{2\gamma} \le |\pi_n|^{2\gamma-1}(b-a) \ \to\ 0, \ \ \textit{per } n \to \infty,$$

che dà una contraddizione. ■

Dimostriamo ora un teorema che è più preciso del precedente corollario. La dimostrazione segue [10] (si utilizza un metodo dovuto a Dvoretzky, Erdös e Kakutani).

Una funzione $f : [0,\infty) \to \mathbb{R}^n$, si dice **$\gamma$-Hölderiana nel punto** $s \ge 0$, con $\gamma \in (0,1]$, se esiste $\eta > 0$, $b > 0$, per cui $|t-s| < \eta$, $t \ge 0$, implica

$$\frac{|f(t) - f(s)|}{|t-s|^\gamma} \le b.$$

Teorema 3.11. *Sia (W_t) un processo di Wiener reale. Per ogni $\gamma > 1/2$, q.c., le traiettorie di (W_t) non sono γ-Hölderiane in nessun punto $s \ge 0$.*

Dimostrazione. Fissiamo $T \in \mathbb{N}$, $N \in \mathbb{N}$, $N \ge 1$ (sufficientemente grande), per cui $N(\gamma - 1/2) > 1$. Sia $b > 0$, consideriamo

$$A_n^b = \{\omega \ :\ \textit{esiste } s \in [0,T] \textit{ per cui } |W_t(\omega) - W_s(\omega)| < b|t-s|^\gamma, \textit{ se } |t-s| \le N/n\}.$$

Se per assurdo $t \mapsto W_t(\omega)$ è $\gamma-$Hölder in un punto $s \in [0,T]$, avremmo

$$|W_t(\omega) - W_s(\omega)| < b_0|t-s|^\gamma, \ \ \textit{per } |t-s| \le N/n_0,$$

per qualche $b_0 > 0$ e $n_0 \in \mathbb{N}$. Dunque $\omega \in A_{n_0}^{b_0}$.

Ora $A_n^b \subset A_{n+1}^b$, $n \in \mathbb{N}$. Sia $A^b = \cup_{n\in\mathbb{N}} A_n^b$. Se proviamo che $\mathbb{P}(A^b) = 0$, per ogni $b > 0$, abbiamo trovato la tesi (data anche l'arbitrarietà di T).

D'ora in poi, supponiamo b fissato e non indichiamo più la dipendenza di A_n da b. Introduciamo

$$Z_k = \max_{1\le i\le N} |W_{(\frac{k+i}{n})} - W_{(\frac{k+i-1}{n})}|,$$

$k = 0, ..., nT$ e

$$B_n = \{\omega \ :\ Z_k(\omega) \le\ 2bN^\gamma/n^\gamma, \textit{ per qualche } k = 0, ..., nT\}.$$

Proviamo che

$$A_n \subset B_n, \quad n \in \mathbb{N}. \tag{3.11}$$

Sia $\omega \in A_n$, allora esiste $s = s(\omega) \in [0,T]$, per cui $|W_t(\omega) - W_s(\omega)| < b|t-s|^\gamma$, $|t-s| \le N/n$.

Sia k_0 il più grande intero per cui $k_0/n \le s$. Risulta che

$$|W_{(\frac{k_0+1}{n})}(\omega) - W_{(\frac{k_0}{n})}(\omega)| \le 2b(1/n)^\gamma$$

e $|W_{(\frac{k_0+N}{n})}(\omega) - W_{(\frac{k_0+N-1}{n})}(\omega)| \le 2bN^\gamma/n^\gamma$ (infatti per construzione $k_0/n \ge s - N/n$). Perciò $Z_{k_0}(\omega) \le 2bN^\gamma/n^\gamma$ e cosi' $\omega \in B_n$. Questo prova (3.11).

Usando (3.11), si trova

$$\mathbb{P}(A) = \lim_{n\to\infty} \mathbb{P}(A_n) \le \liminf_{n\to\infty} \mathbb{P}(B_n).$$

Ora

$$\mathbb{P}(B_n) = \mathbb{P}\Big(\bigcup_{k=0}^{nT}(Z_k \le 2bN^\gamma/n^\gamma)\Big) \le \sum_{k=0}^{nT} \mathbb{P}(Z_k \le 2bN^\gamma/n^\gamma)$$
$$\le (nT+1)\Big(\mathbb{P}(|W_{1/n}| \le 2bN^\gamma/n^\gamma)\Big)^N.$$

Infatti si ha (ponendo $W_{(\frac{k+i}{n})} - W_{(\frac{k+i-1}{n})} = X_{ki}$):

$$\mathbb{P}(Z_k \le 2bN^\gamma/n^\gamma) \le \mathbb{P}(\max_{1\le i\le N} |X_{ki}| \le 2bN^\gamma/n^\gamma)$$
$$= \mathbb{P}(|X_{k1}| \le 2bN^\gamma/n^\gamma, ..., |X_{kN}| \le 2bN^\gamma/n^\gamma)$$
$$= \prod_{i=1}^N \mathbb{P}(|X_{ki}| \le 2bN^\gamma/n^\gamma) = \Big(\mathbb{P}(|W_{1/n}| \le 2bN^\gamma/n^\gamma)\Big)^N, \quad k \in 0, ..., nT.$$

Dunque, ponendo $c = 2bN^\gamma$,

$$\mathbb{P}(B_n) \le (nT+1)\Big(\mathbb{P}(|W_{1/n}| \le c/n^\gamma)\Big)^N$$
$$\le (nT+1)\Big(\int_{-c/n^\gamma}^{c/n^\gamma} e^{-nx^2/2}\sqrt{n/2\pi}\, dx\Big)^N \le \frac{1}{(2\pi)^{N/2}}(nT+1)\,(2c\sqrt{n}/n^\gamma)^N.$$

Essendo $(\gamma - 1/2)N > 1$, si trova $\lim_{n\to\infty} \mathbb{P}(B_n) = 0$, da cui la tesi. ∎

Corollario 3.12. *Le traiettorie di un processo di Wiener (W_t), q.c., non sono derivabili in nessun $t \ge 0$.*

La traiettorie di un processo di Wiener godono di numerose altre proprietà, vedere per esempio [15], [3], [16], [11], [19].

3.3 Osservazioni finali

(1) Un processo (X_t), definito e adattato rispetto ad una base stocastica $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in I}, \mathbb{P})$, si chiama **$\mathcal{F}_t$-martingala** se ogni v.a. $X_t \in L^1(\Omega)$, $t \in I$, e inoltre

$$\mathbb{E}(X_t - X_s/\mathcal{F}_s) = 0, \quad q.c., \ t > s, \quad t, s \in I.$$

Si verifica facilmente che un *processo di Wiener* (W_t) *definito e adattato rispetto ad una base stocastica* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ *è una* $\mathcal{F}_t$*-martingala.* Infatti si ha:

$$\mathbb{E}(W_t - W_s/\mathcal{F}_s) = 0, \quad t > s \geq 0,$$

usando il fatto che (W_t) ha incrementi indipendenti.

(2) Dato un processo di Wiener reale (W_t) possiamo calcolare le sue *distribuzioni finito dimensionali* $\mu^W_{t_1,...,t_n}$, al variare di $0 \leq t_1 < ... < t_n$ e di $n \in \mathbb{N}$. Si tratta di applicare le Proposizioni 1.13 e 1.14.

Possiamo limitarci a considerare $t_1 > 0$ (infatti se $t_1 = 0$, $\mu^W_{0,...,t_n} = \delta_0 \times \mu^W_{t_2,...,t_n}$). Consideriamo le v.a.

$$X_n = (W_{t_1}, ..., W_{t_n}), \quad Y_n = (W_{t_1}, W_{t_2} - W_{t_1}, ..., W_{t_n} - W_{t_{n-1}}).$$

Dalle proprietà di (W_t), sappiamo che Y_n ha legge normale $N(0, D_n)$, dove D_n è una matrice $n \times n$ diagonale, avente autovalori $t_1, t_2 - t_1, ..., t_n - t_{n-1}$. È facile costruire una matrice invertibile L tale che $Y_n = LX_n$; per esempio se n = 2, si ha

$$L = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

Dunque la legge di X_n, cioè $\mu^W_{t_1,...,t_n}$, è $N(0, L^{-1}D_n(L^{-1})^T)$. Si può determinare facilmente $Q^n = L^{-1}D_n(L^{-1})^T$, osservando che

$$Q^n_{ij} = \mathbb{E}(W_{t_i}W_{t_j}) = t_i \wedge t_j, \quad i, j = 1, ..., n, \quad \text{dove } \ t_i \wedge t_j = \min(t_i, t_j)$$

(infatti se $t_i < t_j$, $\mathbb{E}(W_{t_i}W_{t_j}) = \mathbb{E}[(W_{t_j} - W_{t_i})W_{t_i}] + \mathbb{E}W^2_{t_i} = t_i$).

Viceversa ogni processo stocastico (X_t) che ha come distribuzioni finito dimensionali le precedenti $\mu^W_{t_1,...,t_n} = N(0, Q^n)$, è *quasi* un processo di Wiener. Infatti (X_t) verifica le proprietà 1, 2 e 3 della Definizione 3.1, ma non necessariamente la 4 (non è detto che questo processo abbia le traiettorie continue). Tuttavia, usando il teorema di regolarità di Kolmogorov, si può mostrare che esiste una sua versione con traiettorie continue. Dunque esiste una versione di (X_t) che è un processo di Wiener.

Ogni processo stocastico le cui distribuzioni finito dimensionali sono leggi normali, si dice *processo Gaussiano.* Il processo di Wiener è un particolare processo Gaussiano (vedere anche [15]).

Il precedente discorso vale in modo analogo per processi di Wiener n-dimensionali.

(3) Accenniamo alla **misura di Wiener.** Sia (W_t) un processo di Wiener a valori in $\mathbb{R}^d$, definito su $(\Omega, \mathcal{F}, \mathbb{P})$.

Consideriamo $E = C([0,\infty), \mathbb{R}^d) = \{\omega : [0,\infty) \to \mathbb{R}^d$ continue $\}$; E è uno spazio metrico, separabile e completo, vedere anche (1.2). Indichiamo con $B(E)$ la σ-algebra di Borel su E.

Si può provare che $B(E)$ coincide con la σ-algebra generata da tutti i rettangoli

$$B_{t_1,...,t_n}(A_1 \times ... \times A_n) = \{\omega \in E : \omega(t_1) \in A_1, ..., \omega(t_n) \in A_n\},$$

al variare di $t_1 < ... < t_n$ in $[0,\infty)$, $n \in \mathbb{N}$ e $A_k \in B(\mathbb{R}^d)$, vedere per esempio [13]. Usando questo fatto si verifica subito che la funzione T, $T : \Omega \to E$,

$$T(\omega) = (W_t(\omega))_{t\geq 0}, \quad \omega \in \Omega,$$

è misurabile (da $(\Omega, \mathcal{F})$ in $(E, B(E))$. Notiamo che si ha:

$$T^{-1}(B_{t_1,...,t_n}(A_1 \times ... \times A_n)) = \{W_{t_1} \in A_1, ..., W_{t_n} \in A_n\} \in \mathcal{F}.$$

La misura di probabilità $\mathbb{P}_0$, immagine di $\mathbb{P}$ tramite T, si chiama **misura di Wiener** (su $(E, B(E))$). Per ogni $B \in B(E)$, si ha: $\mathbb{P}_0(B) := \mathbb{P}(T^{-1}(B))$. Si può dimostrare che $\mathbb{P}_0$ è unica, nel senso che se un'altra misura di probabilita $\mathbb{P}_1$ su $B(E)$, verifica

$$\mathbb{P}_0(B_{t_1,...,t_n}(A_1 \times ... \times A_n)) = \mathbb{P}_1(B_{t_1,...,t_n}(A_1 \times ... \times A_n)),$$

per ogni rettangolo $B_{t_1,...,t_n}(A_1 \times ... \times A_n)$, allora $\mathbb{P}_1 = \mathbb{P}_0$. È chiaro che $\mathbb{P}_0$ è concentrata sulle funzioni $\omega \in E$, che sono nulle in $t = 0$.

Sullo spazio di probabilità $(E, B(E), \mathbb{P}_0)$, si definisce il **processo di Wiener canonico** (X_t^0), ponendo

$$X_t^0(\omega) = \omega(t), \quad \omega \in E, \ \ t \geq 0.$$

Si prova facilmente che (X_t^0) verifica tutte le proprietà della Definizione 3.1.

Per maggiori dettagli sulla misura di Wiener, vedere per esempio [15], [9], [3], [16].

4 Equazioni stocastiche con rumore additivo

Cominciamo con un *modello fisico semplificato* in cui interviene un'equazione differenziale stocastica, vedere anche [9] e [12].

Consideriamo una particella microscopica sospesa in un fluido bidimensionale. Supponiamo, per semplicità, di considerare solo la componente nella direzione dell'asse x del vettore spostamento della particella.

Se $x = X(t)$ indica la posizione della particella al tempo t, il suo spostamento tra t e $t + \triangle t$, si potrà indicare con $\triangle X = X(t + \triangle t) - X(t)$. Possiamo esprimere $\triangle X$, nel modo seguente,

$$X(t + \triangle t) - x = b(t, x)\triangle t + \sigma(t, x)(W(t + \triangle t) - W(t)),$$

dove b e σ sono funzioni (regolari) e $W(t)$ è un processo di Wiener reale (o più precisamente una traiettoria di un processo di Wiener).

Considerando l'interpretazione fisica del moto Browniano (come elaborata all'inizio del 1900 da Einstein e Smoluchowski) $\sigma(t, x)(W(t + \triangle t) - W(t))$ indica lo spostamento della particella tra t e $t + \triangle t$ dovuto esclusivamente agli urti con le molecole del fluido; $\sigma(t, x)$ misura l'effetto della temperatura del fluido nel punto x all'istante t.

Invece $b(t, x)\triangle t$ indica lo spostamento della particella dovuto al moto del fluido ($b(t, x)$ misura la velocità del fluido nel punto x all'istante t).

Si dice anche che b è il *coefficiente di drift e σ quello di diffusione.* Passando al limite per $\triangle t \to 0^+$, si trova *formalmente*

$$dX(t) = b(t, X(t))dt + \sigma(t, X(t))dW_t, \tag{4.1}$$

$t > 0$, usualmente con una condizione iniziale, $X_0 = z \in \mathbb{R}$. Resta il problema di dare un senso a (4.1). Non è possibile dividere semplicemente per dt e ragionare ω per ω; infatti il "rumore bianco" $\frac{dW}{dt}$ non è definito ω per ω, vedere Corollario 3.12.

Per dare un senso alla precedente espressione, **Ito** propose di considerare (4.1) in senso integrale, cioè di trattare

$$X_t = z + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s, \quad t \geq 0, \tag{4.2}$$

dando un senso preciso all'*integrale stocastico* $\int_0^t \sigma(s, X_s)dW_s$. Tale integrale non si può definire ω per ω, nel senso di Stieltjes, in quanto le traiettorie di (W_t) non sono localmente BV. Ito definì l'integrale precedente con un procedimento di tipo globale rispetto a ω, ovvero con un passaggio al limite in $L^2(\Omega)$.

Lo studio dell'integrale stocastico richiede nozioni più avanzate di Probabilità, tra cui alcuni teoremi sulle martingale.

Nel seguito considereremo solo *equazioni stocastiche con rumore additivo,* quando cioè σ non dipende ne da x ne da t, cioè equazioni del tipo

$$dX_t = b(X_t)dt + \sigma dW_t$$

(il fatto che non chiediamo che b dipenda da t è solo per semplificare un poco le notazioni e gli enunciati; non è una semplificazione sostanziale). Per definire tali equazioni non è richiesto l'introduzione dell'integrale stocastico di Ito.

Le equazioni stocastiche con rumore additivo costituiscono un'utile introduzione alle equazioni (4.2) che sono anche dette *equazioni stocastiche con rumore moltiplicativo.* Le soluzioni delle equazioni stocastiche (4.2) si chiamano anche *processi di diffusione.*

La trattazione delle equazioni stocastiche generali si può trovare in [7], [16], [17], [9], [10], [19], [13], [14]. Un testo che presenta applicazioni delle equazioni stocastiche alla finanza matematica è [12].

4.1 Un teorema di esistenza e unicità

I riferimenti principali per questa sezione sono [14], [9], [7], [13].

Consideriamo fissato un *processo di Wiener n-dimensionale* (W_t) *definito e adattato rispetto ad una base stocastica* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, vedere la Definizione 3.3 (per semplicità si può anche pensare di avere un processo di Wiener (W_t) su $(\Omega, \mathcal{F}, \mathbb{P})$ e poi considerare come filtrazione $(\mathcal{F}_t)$, la filtrazione naturale completata di (W_t), vedere la Sezione 2).

Considereremo una equazione stocastica (brevemente una EDS) autonoma e con rumore additivo, cioè del tipo

$$dX_t = b(X_t)dt + \sigma dW_t, \;\; X_0 = x \in \mathbb{R}^n, \;\; t \geq 0, \tag{4.3}$$

dove sono assegnate una matrice reale σ, $n \times n$, una funzione continua (o campo vettoriale continuo) $b : \mathbb{R}^n \to \mathbb{R}^n$ e $x \in \mathbb{R}^n$. Abbiamo la seguente nozione di soluzione.

Definizione 4.1. *Un processo continuo* $(X_t)_{t\geq 0}$ *definito su* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$*, si dice* **soluzione (globale)** *di* (4.3) *se valgono le seguenti proprietà:*

(1) (X_t) *è* $\mathcal{F}_t$*-adattato;*

(2) X_t *risolve "*ω *per* ω*", q.c., l'equazione integrale*

$$X_t(\omega) - x = \int_0^t b(X_s(\omega))ds + \sigma W_t(\omega), \;\; t \geq 0. \tag{4.4}$$

Se sarà necessario indicare esplicitamente la dipendenza della soluzione da x, scriveremo (X_t^x).

Osservazioni 4.1. (i) Essendo il processo soluzione (X_t) continuo, possiamo dire che esiste $\Omega' \in \mathcal{F}$, con $\mathbb{P}(\Omega') = 0$, tale che se $\omega \notin \Omega'$, la funzione $t \mapsto X_t(\omega)$ risolve (4.4) per ogni $t \geq 0$.

(ii) Potremmo più in generale considerare EDS nonautonome (con rumore additivo) del tipo $dX_t = b(t, X_t)dt + \sigma dW_t, \quad X_s = Y,\ t \in [s,T]$, dove $T > s \geq 0$, $b : [s,T] \times \mathbb{R}^n \to \mathbb{R}^n$ è continua e Y è una v.a. assegnata. L'equazione precedente scritta in forma integrale diventa

$$X_t(\omega) - Y(\omega) = \int_s^t b(r, X_r(\omega))dr + \sigma(W_t(\omega) - W_s(\omega)), \quad t \geq s. \tag{4.5}$$

La soluzione è un processo continuo e $\mathcal{F}_t$-adattato $(X_t)_{t\in[s,T]}$ che risolve (4.5) q.c.. Occorre però richiedere che Y sia $\mathcal{F}_s$-misurabile. La trattazione di (4.5) si può fare in maniera simile a quella che faremo per (4.4).

(iii) Si potrebbe anche considerare nella EDS (4.3) un processo di Wiener d-dimensionale e una matrice σ $n \times d$. Il processo soluzione (X_t) sarebbe ancora n-dimensionale.

(iv) La richiesta che la soluzione (X_t) sia un processo $\mathcal{F}_t$-adattato è giustificata anche dal fatto che vorremo poter applicare a (X_t) la formula di Ito.

(v) La trattazione che faremo si potrebbe anche adattare a EDS con rumore additivo rispetto a processi più generali del processo di Wiener (W_t), per esempio rispetto ai processi di Lévy.

Per studiare l'esistenza è l'unicità di soluzioni per (4.3) faremo le seguenti ipotesi.

Ipotesi 4.1. *(i)* $b \in C^1(\mathbb{R}^n, \mathbb{R}^n)$.

(ii) Esiste una costante $k > 0$ per cui

$$\langle b(x+y) - b(y), x\rangle \leq k(1 + |x|^2), \quad x, y \in \mathbb{R}^n, \tag{4.6}$$

dove $\langle \cdot, \cdot \rangle$ e $|\cdot|$ indicano il prodotto interno e la norma euclidea in $\mathbb{R}^n$.

La condizione (i) si può generalizzare chiedendo che b sia *localmente Lipschitziana.* La condizione (ii) è una *condizione di monotonia,* ben nota nella teoria delle equazioni differenziali ordinarie per assicurare l'esistenza nel futuro delle soluzioni.

Osserviamo che se vale (i) e b è Lipschitziana su $\mathbb{R}^n$ (ovvero $\sup_{x\in\mathbb{R}^n} |Db(x)| < \infty$, dove Db indica il gradiente di b) allora (ii) è soddisfatta con $k = L$, dove L indica la costante di Lipschitz di b. Infatti si ha:

$$\langle b(x+y) - b(y), x\rangle \leq L|x|^2, \quad x, y \in \mathbb{R}^n.$$

Se $\mathbb{R}^n = \mathbb{R}$ e $b \in C^1(\mathbb{R})$, allora (ii) vale se $\sup_{x\in\mathbb{R}} b'(x) = k < \infty$.

Un esempio di EDS in dimensione 1 che verifica le ipotesi precedenti è

$$dX_t = (aX - X^3)dt + dW_t, \quad X_0 = x \in \mathbb{R}, \quad t \geq 0 \ \ (a \in \mathbb{R}\ \textit{parametro}).$$

Teorema 4.2. *(Teorema di esistenza è unicità) L'equazione* (4.3)*, con le Ipotesi 4.1, ammette un'unica soluzione, per ogni $x \in \mathbb{R}^n$.*

Per provare il teorema risolveremo (4.4) in modo "deterministico" (ω per ω). Una difficoltà ulteriore sarà provare che il processo soluzione è $\mathcal{F}_t$-adattato.

L'unicità si basa sul seguente risultato elementare, che si usa spesso.

Lemma 4.3. *(Lemma di Gronwall). Siano $v : [a,b] \to \mathbb{R}_+$ una funzione continua e $c,\, d \geq 0$. Se, per ogni $t \in [a,b]$, risulta*

$$v(t) \leq c + d \int_a^t v(s)ds$$

allora $v(t) \leq ce^{d(t-a)}$, $t \in [a,b]$.

L'esistenza si basa anche sul seguente risultato.

Lemma 4.4. *Sia $f : [0,\infty) \to \mathbb{R}^n$ una funzione continua assegnata. Supponiamo che $b \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ verifichi (4.6) e consideriamo l'equazione integrale (nella funzione incognita $u : [0,\infty) \to \mathbb{R}^n$ continua):*

$$u(t) = f(t) + \int_0^t b(u(s))ds, \;\; t \geq 0. \tag{4.7}$$

(i) Esiste un'unica soluzione u per (4.7) su $[0,\infty[$.

(ii) Se b è in aggiunta Lipschitziana su $\mathbb{R}^n$ allora le seguenti iterate di Picard,

$$u_0(t) = f(t), \;\; u_{n+1}(t) = f(t) + \int_0^t b(u_n(s))ds, \;\; t \geq 0, \; n \in \mathbb{N},$$

convergono alla soluzione u, uniformemente su ogni intervallo $[0,a]$, $a > 0$.

Dimostrazione. (i) Si pone $v(t) := u(t) - f(t)$. Cerchiamo una soluzione v di classe C^1 che verifica

$$v(t) = \int_0^t b(v(s) + f(s))ds, \;\; t \geq 0. \tag{4.8}$$

Questa equivale al problema di Cauchy

$$\dot{v}(t) = b(v(t) + f(t)), \quad v(0) = 0.$$

Poichè $F(t,x) := b(x + f(t))$ è di classe C^1 nella variabile x, sappiamo che esiste un'unica soluzione v per il precedente problema di Cauchy, soluzione definita su un intervallo massimale $[0,T[$. Dunque v verifica (4.8) su $[0,T[$. E' chiaro che

$$u(t) := v(t) + f(t), \;\; t \in [0,T[,$$

risolve (4.7) ed è l'unica soluzione in $[0,T[$. Rimane da provare, usando (4.6), che $T = \infty$. Moltiplicando scalarmente per v si ha:

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}|v(t)|^2 &= \langle \dot{v}(t), v(t)\rangle = \langle b(v(t) + f(t)), v(t)\rangle \\
&= \langle b(v(t) + f(t)) - b(f(t)), v(t)\rangle + \langle b(f(t)), v(t)\rangle \\
&\leq k(1 + |v(t)|^2) + M|v(t)| \leq C(1 + |v(t)|^2), \;\; t \in [0,T[,
\end{aligned}$$

dove $M = \sup_{t\in[0,T]} |b(f(t))|$. Integrando su $[0,t]$, si trova

$$|v(t)|^2 \leq 2Ct + 2C \int_0^t |v(s)|^2 ds.$$

Il Lemma di Gronwall implica che $|v(t)|^2 \leq 2CTe^{2CT}$, $t \in [0,T[$. Dunque $T = +\infty$.

(ii) Sia $a > 0$. Si ha, per $t \in [0,a]$,

$$|u_1(t) - u_0(t)| \leq \int_0^t |b(f(r))|dr \leq Bt,$$

$$|u_2(t) - u_1(t)| \leq \int_0^t |b(u_1(r)) - b(f(r))|dr \leq BL \int_0^t s ds = B\frac{Lt^2}{2},$$

dove $B = \sup_{t\in[0,a]} |b(f(t))|$. Per induzione si ottiene:

$$|u_{n+1}(t) - u_n(t)| \leq B\frac{L^n a^{n+1}}{(n+1)!}, \;\; t \in [0,a], \;\; n \in \mathbb{N}.$$

Dunque $u_{n+1}(t) = f(t) + \sum_{k=0}^n (u_{k+1}(t) - u_k(t))$ converge uniformente su $[0,a]$ ad una funzione continua u. Passando al limite in $u_{n+1}(t) = f(t) + \int_0^t b(u_n(s))ds$ si trova che u è la soluzione dell'equazione su $[0,+\infty[$. ■

Dimostrazione del Teorema 4.2. (*Unicità*) Se (X_t) e (Y_t) sono due soluzioni si ha, per ogni $T > 0$,

$$|X_t(\omega) - Y_t(\omega)| \leq \int_0^t |b(X_s(\omega)) - b(Y_s(\omega))|ds \leq c \int_0^t |X_s(\omega) - Y_s(\omega)|ds, \; t \in [0,T],$$

dove c è la costante di Lipschitz di b sulla palla chiusa $B(0,R)$, di centro 0 e raggio R, dove $R = R(\omega) = \max_{t\in[0,T]} (|X_t(\omega)| + |Y_t(\omega)|)$. Dal Lemma di Gronwall si deduce che $X_t(\omega) = Y_t(\omega)$, $t \in [0,T]$, da cui la tesi per l'arbitrarietà di T e di ω.

Si può provare allo stesso modo che se due processi (X_t) e (Y_t) risolvono (4.4) su un intervallo $[0,T(\omega)]$, allora $X_t(\omega) = Y_t(\omega)$, $t \in [0,T(\omega)]$, per ogni ω.

(*Esistenza*) Sappiamo già che esiste $(X_t(\omega))$ che risolve (4.4) su $[0,\infty[$, q.c.. Basta applicare il Lemma 4.4 con

$$f(t) = x + \sigma W_t(\omega), \;\; \omega \; q.c..$$

Rimane da provare che (X_t) è un processo $\mathcal{F}_t$*-adattato,* cioè che $X_t : (\Omega, \mathcal{F}_t) \to \mathbb{R}^n$ è misurabile, per ogni $t > 0$.

Questo è certamente vero se b è Lipschitziana su $\mathbb{R}^n$. Infatti in tale caso X_t è il limite "ω per ω" (uniformemente in ogni $[0,T]$) dei processi

$$X_0(t) = x + \sigma W_t, \;\; X_{n+1}(t) = x + \sigma W_t + \int_0^t b(X_n(s))ds, \;\; t \geq 0, \; n \in \mathbb{N},$$

vedere (ii) nel Lemma 4.4. Si dimostra facilmente che i processi $(X_n(t))$ sono tutti $\mathcal{F}_t$-adattati, usando l'induzione.

Per trattare il caso di b soltanto localmente Lipschitziana, introduciamo delle opportune approssimazioni Lipschitziane (b_k) di b:

$$b_k(x) = \begin{cases} b(x), & se \ \ |x| \leq k, \\ \\ b(kx/|x|), & se \ \ |x| > k, \end{cases}$$

E' chiaro che $b_k : \mathbb{R}^n \to \mathbb{R}^n$ è Lipschitziana. Dunque le EDS

$$dX_t = b_k(X_t)dt + \sigma dW_t, \ \ X_0 = x \in \mathbb{R}^n, \ \ t \geq 0, \ k \in \mathbb{N},$$

ammettono ciascuna un'unico processo soluzione (X_t^k), che è anche $\mathcal{F}_t$-adattato. Per concludere la dimostrazione, proviamo che

$$(*) \quad \lim_{k\to\infty} X_t^k(\omega) = X_t(\omega),$$

per ogni $t \geq 0$, ω q.c.. ((*) significa che, fissati $s_0 > 0$, ω q.c., per ogni $\epsilon > 0$, esiste $k_0 = k_0(\omega)$ tale che se $k \geq k_0$, $|X_{s_0}^k(\omega) - X_{s_0}(\omega)| < \epsilon$).

Fissiamo $\omega \in \Omega$ e introduciamo l'intero $N = N(\omega) > \sup_{t\in[0,s_0]} |X_t(\omega)|$ (sappiamo che $t \mapsto X_t(\omega)$ è continua, ω q.c.). Per costruzione, so che

$$X_t(\omega) - x = \int_0^t b(X_s(\omega))ds + \sigma W_t(\omega) = \int_0^t b_N(X_s(\omega))ds + \sigma W_t(\omega), \ \ t \in [0, s_0].$$

Per l'unicità della soluzione, si deve avere $X_t(\omega) = X_t^N(\omega)$, $t \in [0, s_0]$ ed anche $X_t(\omega) = X_t^k(\omega)$, $t \in [0, s_0]$, per ogni $k \geq N$. Questo prova (*).

La dimostrazione è finita. ■

Osserviamo che *il teorema precedente vale anche se assumiamo che il coefficiente b sia limitato su $\mathbb{R}^n$ e di classe C^1 (senza richiedere la condizione di monotonia (4.6)).*

Proposizione 4.5. *Sia (X_t^x) il processo soluzione di* (4.4)*, sotto le Ipotesi 4.1. Supponiamo che b sia anche limitata su $\mathbb{R}^n$. Allora, per ogni $T > 0$, si ha:*

$$\sup_{t\in[0,T]} \mathbb{E}|X_t^x|^2 < \infty.$$

In particolare $\int_0^T \mathbb{E}|X_s^x|^2 ds < \infty$.

Dimostrazione. Si ha:

$$\mathbb{E}|X_t^x|^2 \leq 3\mathbb{E}\Big(|x|^2 + C_T + \|\sigma\|^2|W_t|^2\Big) \leq C_{x,T}, \quad t \in [0, T].$$

■

Si potrebbe provare il precedente risultato solo sotto le Ipotesi 4.1. Tuttavia la dimostrazione sarebbe meno banale (richiederebbe anche di introdurre i tempi d'arresto, vedere per esempio [16]).

5 L'integrale di Ito

Per semplicità, in questa sezione trattiamo processi reali.

Consideriamo sempre fissato un *processo di Wiener reale* (W_t) *definito e adattato rispetto ad una base stocastica* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, vedere la Definizione 3.3.

Un **processo elementare** $(X_t) = (X_t)_{t\geq 0}$ è un processo del tipo

$$X_t(\omega) = \sum_{k=0}^{n-1} X_{t_k}(\omega)\, 1_{[t_k, t_{k+1}[}(t), \tag{5.1}$$

dove $0 \leq t_0 < t_1 ... < t_n$, $n \in \mathbb{N}$, le $X_{t_k} \in L^2(\Omega)$ sono variabili aleatorie $\mathcal{F}_{t_k}$-misurabili. Ovviamente risulta che $\int_0^T \mathbb{E}|X_s|^2 ds < \infty$, per ogni $T > 0$. Ogni processo elementare ha traiettorie continue a destra.

Per un processo elementare, definiamo **l'integrale stocastico (alla Ito)** nel modo seguente:

$$\int_0^t X_s dW_s := \sum_{k=0}^{n-1} X_{t_k}\,(W_{t_{k+1}\wedge t} - W_{t_k \wedge t}),\ \ t \geq 0,\ \ \text{dove}\ \ t_k \wedge t = \min(t_k, t).$$

Osserviamo che per un processo elementare (X_t) valgono le seguenti proprietà:

(o) il processo $(\int_0^t X_s dW_s)_{t\geq 0}$ è continuo q.c.;

(i) $\mathbb{E} \int_0^t X_s dW_s = 0$, $t \geq 0$;

(ii) $\mathbb{E}(\int_0^t X_s dW_s)^2 = \mathbb{E} \int_0^t |X_s|^2 ds$, $\ t \geq 0$ (*proprietà di isometria*);

La (o) è chiara. Anche (i) si verifica facilmente. Infatti $\mathbb{E}[X_{t_k}\,(W_{t_{k+1}\wedge t} - W_{t_k\wedge t})] = 0$, poichè X_{t_k} è $\mathcal{F}_{t_k}$-misurabile e integrabile e $(W_{t_{k+1}} - W_{t_k})$ è indipendente da $\mathcal{F}_{t_k}$ ed ha media nulla.

Verifichiamo la (ii), supponendo che $t_n = t$. Questo non è restrittivo. Si ha:

$$\mathbb{E}(\int_0^t X_s dW_s)^2 = \mathbb{E}\Big(\sum_{k=0}^{n-1} X_{t_k}\,(W_{t_{k+1}} - W_{t_k})\Big)^2$$
$$= \sum_{k,j=1}^{n-1} \mathbb{E}(X_{t_k} X_{t_j}\,(W_{t_{k+1}} - W_{t_k})\,(W_{t_{j+1}} - W_{t_j}))$$
$$= \sum_{k=0}^{n-1} \mathbb{E}[X_{t_k}^2\,(W_{t_{k+1}} - W_{t_k})^2] = \sum_{k=0}^{n-1} \mathbb{E}[X_{t_k}^2]\,\mathbb{E}(W_{t_{k+1}} - W_{t_k})^2$$
$$= \sum_{k=0}^{n-1} \mathbb{E}[X_{t_k}^2]\,(t_{k+1} - t_k) = \int_0^t \mathbb{E}|X_s|^2 ds.$$

Abbiamo usato che $\mathbb{E}[X_{t_k} X_{t_j}\,(W_{t_{k+1}} - W_{t_k})\,(W_{t_{j+1}} - W_{t_j})] = 0$, se $t_k < t_j$. Questo segue dalle proprietà di (W_t) e dal fatto che le $X_{t_k} \in L^2(\Omega)$ (si può usare il Corollario 1.6).

Usando (i) e (ii) si può estendere la definizione di integrale stocastico a *processi che si possono approssimare attraverso processi elementari.*

Introduciamo la classe L^2_{ad}. Un processo $(X_t)_{t\geq 0}$ tale che, per ogni $T > 0$, la funzione: $[0,T] \times \Omega \to \mathbb{R}$, $(t,\omega) \mapsto X_t(\omega)$ sia $B([0,T]) \times \mathcal{F}_T$-misurabile, si dice **progressivo**.

Certamente ogni processo progressivo è anche $\mathcal{F}_t$-adattato. Infatti $\{t\} \times \{X_t \in A\}$ $= (\{(s,\omega) \in [0,t] \times \Omega \ : \ X_s(\omega) \in A\} \cap \{t\} \times \Omega) \in B([0,t]) \times \mathcal{F}_t$, per ogni $t > 0$, $A \in B(\mathbb{R})$.

E' importante tenere presente che *ogni processo con traiettorie continue a destra che sia $\mathcal{F}_t$-adattato è anche progressivo,* vedere [15].

Un processo progressivo (X_t) si dice **di classe L^2_{ad}** se verifica

$$\mathbb{E}\int_0^T |X_s|^2 ds < \infty, \ \ \textit{per ogni} \ T > 0. \tag{5.2}$$

Se (X_t) è un *processo di classe L^2_{ad}, si può dimostrare che, per ogni $T > 0$, esiste una successione di processi elementari (X^n_t) tale che*

$$\lim_{n\to\infty} \mathbb{E}\int_0^T |X_s - X^n_s|^2 ds = 0, \tag{5.3}$$

vedere per esempio [15], [17] o [13]. Fissato $T > 0$, usando (ii), si ottiene

$$\mathbb{E}\Big(\int_0^T X^n_s dW_s - \int_0^T X^m_s dW_s\Big)^2 = \mathbb{E}\int_0^T |X^n_s - X^m_s|^2 ds, \ \ n, \ m \in \mathbb{N}.$$

Quindi le v.a. $\int_0^T X^n_s dW_s$ formano una successione di Cauchy in $L^2(\Omega)$ che converge, in $L^2(\Omega)$, ad una v.a. che chiamiamo Y_T. Dato (X_t) di classe L^2_{ad}, si pone

$$Y_T := \int_0^T X_s dW_s \ \ \textbf{(integrale stocastico di } (X_t) \textbf{ su } [0,T]\textbf{)}$$

e si prova (senza difficoltà) che la definizione non dipende dalla particolare scelta dei processi elementari (X^n_t) che approssimano (X_t), nel senso di (5.3).

Al variare di $t \geq 0$ si trova il **processo integrale stocastico** $(Y_t)_{t\geq 0}$,

$$Y_t = \int_0^t X_s dW_s, \ \ t \geq 0.$$

Questo è $\mathcal{F}_t$-adattato ed è unico a meno di versioni.

Si può provare che

(a) (Y_t) *ammette una versione con traiettorie continue;*

(b) $\mathbb{E}(Y_t)^2 = \mathbb{E}(\int_0^t X_s dW_s)^2 = \mathbb{E}\int_0^t |X_s|^2 ds, \ \ t \geq 0.$

(c) (Y_t) *è una $\mathcal{F}_t$-martingala (cioè $\mathbb{E}(Y_t/\mathcal{F}_s) = Y_s, \ \ t > s \geq 0$),*

vedere per esempio [15] o [19]. Inoltre *se (X_t) è un processo continuo di classe L^2_{ad} allora, per ogni $T > 0$,*

$$\int_0^T X_s dW_s = \lim_{|\pi|\to 0^+} \sum_{k=0}^{n-1} X_{t_k}(W_{t_{k+1}} - W_{t_k}). \tag{5.4}$$

dove il limite è in $L^2(\Omega)$ *e* $\pi = \{t_0, ...t_n\}$, *con* $t_0 = 0 < ... < t_n = T$. Infatti se (X_t) è continuo si può approssimarlo in $L^2([0,T] \times \Omega)$ con processi elementari $(\hat{X}_t)$,

$$\hat{X}_t(\omega) = \sum_{k=0}^{n-1} X_{t_k}(\omega)\, 1_{[t_k, t_{k+1}[}(t).$$

Questo non si può fare in generale per ogni processo di classe L^2_{ad}.

Esempio 5.1. Proviamo che

$$Y_t = \int_0^t W_s dW_s = W_t^2/2 - t/2, \quad t \geq 0.$$

Fissiamo $t > 0$. Sia $\pi = \{t_0, ..., t_n\}$, con $0 = t_0 < t_1 ... < t_n = t$ una generica partizione di $[0,t]$. Sia $|\pi| = \max_{0 \leq k \leq n-1} |t_{k+1} - t_k|$.

Si ha:

$$\begin{aligned} 2\sum_{k=0}^{n-1} W_{t_k}\left(W_{t_{k+1}} - W_{t_k}\right) &= \sum_{k=0}^{n-1} [W_{t_k} + (W_{t_{k+1}} - W_{t_k})]^2 - \sum_{k=0}^{n-1} [W_{t_k}]^2 - \sum_{k=0}^{n-1} [W_{t_{k+1}} - W_{t_k}]^2 \\ &= \sum_{k=0}^{n-1} [W_{t_{k+1}}]^2 - \sum_{k=0}^{n-1} [W_{t_k}]^2 - \sum_{k=0}^{n-1} [W_{t_{k+1}} - W_{t_k}]^2 \\ &= W_t^2 - \sum_{k=0}^{n-1} [W_{t_{k+1}} - W_{t_k}]^2. \end{aligned}$$

Passando al limite in $L^2(\Omega)$ con $|\pi| \to 0^+$, si trova la tesi, grazie a (3.9) (la variazione quadratica di (W_t) su $[0,t]$ è proprio t).

L'integrale stocastico rispetto a W_t si può definire per processi più generali, di quelli di classe L^2_{ad}. Lo stesso Ito aveva già definito il suo integrale per processi progressivi (X_t) che verificano la condizione

$$\int_0^t |X_s(\omega)|^2 ds < \infty, \quad \textit{per ogni } t > 0, \quad \omega \textit{ q.c.}. \tag{5.5}$$

In questo caso l'integrale stocastico si ottiene come limite *in probabilità* di integrali stocastici di processi elementari. Tuttavia non gode più in generale delle precedenti proprietà (a), (b) e (c).

6 La formula di Ito

Consideriamo sempre fissato un *processo di Wiener n-dimensionale* (W_t) *definito e adattato rispetto ad una base stocastica* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$, vedere la Definizione 3.3.

6.1 La formula di Ito multidimensionale

Fissiamo $T > 0$ e introduciamo lo spazio di funzioni $C^{1,2}_{b,T} = C^{1,2}_b([0,T] \times \mathbb{R}^n)$.

Diciamo che una funzione $f(t,x)$, $f : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ appartiene a $C^{1,2}_{b,T}$, se f è derivabile con continuità una volta in t e due volte in x su $[0,T] \times \mathbb{R}^n$ (cioè esistono le derivate parziali

$\partial_t f$, $\partial_{x_i} f$ e $\partial_{x_i x_j} f$, con $i, j = 1, ..., n$, continue su $[0, T] \times \mathbb{R}^n$) e inoltre f e tutte le derivate parziali $\partial_t f$, $\partial_{x_i} f$ e $\partial_{x_i x_j} f$, con $i, j = 1, ..., n$, sono limitate su $[0, T] \times \mathbb{R}^n$.

Scriviamo $D_x f = (\partial_{x_1} f, ..., \partial_{x_n} f)$. Lo spazio C_b^2 si definisce in modo simile a $C_{b,T}^{1,2}$; si tratta delle funzioni $f : \mathbb{R}^n \to \mathbb{R}$ continue e limitate, che ammettono tutte le derivate parziali $\partial_{x_i} f$ e $\partial_{x_i x_j} f$, con $i, j = 1, ..., n$, continue e limitate su $\mathbb{R}^n$.

Il seguente risultato è molto importante.

Teorema 6.1. *(Formula di Ito) Sia $f \in C_{b,T}^{1,2}$. Sia $(X_t) = (X_t^x)$, $x \in \mathbb{R}^n$, il processo soluzione di* (4.4)*, sotto le Ipotesi 4.1. Indichiamo con $q = (q_{ij})$ la matrice $\sigma\sigma^T$.*

Considerando il processo continuo $f(t, X_t)$, si ha:

$$\begin{aligned} f(t, X_t(\omega)) - f(0, x) &= \int_0^t \partial_s f(u, X_u(\omega)) du + \int_0^t \langle D_x f(u, X_u(\omega)), b(X_u(\omega)) \rangle du \\ &+ \frac{1}{2} \int_0^t \Big(\sum_{i,j=1}^n q_{ij} \partial_{x_i x_j} f(u, X_u(\omega)) \Big) du + Y_t(\omega), \quad t \in [0, T], \ q.c., \end{aligned}$$

dove Y_t è dato da

$$Y_t = \sum_{i,j=1}^n \int_0^t \partial_{x_i} f(u, X_u) \sigma_{ij} dW_u^j := \int_0^t \langle D_x f(u, X_u), \sigma dW_u \rangle.$$

La formula di Ito vale più in generale per i cosiddetti processi (X_t) di Ito, che non sono necessariamente soluzioni di EDS.

Anche le ipotesi su f si potrebbero indebolire (permettendo per esempio ad f di essere un polinomio). In tale caso occorre generalizzare la nozione che abbiamo dato di integrale stocastico, usando la condizione (5.5).

Rimandiamo ai testi di calcolo stocastico per la presentazione e la dimostrazione della formula di Ito in un contesto più generale.

Motivati dalla formula di Ito, consideriamo il seguente *operatore di Kolmogorov,*

$$Lf(x) = \frac{1}{2} \sum_{i,j=1}^n q_{ij} \partial_{x_i x_j} f(x) + \sum_{i=1}^n \partial_{x_i} f(x) b_i(x) = \frac{1}{2} Tr(q D^2 f(x)) + \langle D_x f(x), b(x) \rangle, \quad x \in \mathbb{R}^n, \tag{6.1}$$

$f \in C_b^2$, dove $Tr(A)$ indica la traccia di una matrice A e $D^2 f(x)$ è la matrice Hessiana di f in x. Notiamo che $q = \sigma\sigma^T$ è una matrice simmetrica e semidefinita positiva.

Corollario 6.2. *Sia $f \in C_b^2$. Supponiamo che (X_t^x) sia il processo soluzione di* (4.4)*. Se assumiamo che $b \in C^1$ sia limitata su $\mathbb{R}^n$ allora si ha:*

$$\mathbb{E}(f(X_t^x)) = f(x) + \mathbb{E} \int_0^t Lf(X_s^x) ds, \quad t \geq 0, \quad e \tag{6.2}$$

$$\lim_{h \to 0} \frac{\mathbb{E}(f(X_{t+h}^x)) - \mathbb{E}(f(X_t^x))}{h} = \mathbb{E} Lf(X_t^x), \quad x \in \mathbb{R}^n, \ t \geq 0. \tag{6.3}$$

Dimostrazione. Per ottenere (6.2), si applica l'attesa nella precedente formula di Ito, senza difficoltà visto che b è limitata e $f \in C_b^2$. Certamente l'integrale stocastico Y_t è ben definito, grazie alle ipotesi su f (e si ha $\mathbb{E} Y_t = 0$, $t \geq 0$).

Anche la (6.3) si dimostra facilmente. ■

La (6.3) per $t=0$, diventa

$$\lim_{h\to 0^+} \frac{\mathbb{E}(f(X_h^x)) - f(x)}{h} = Lf(x), \;\; x \in \mathbb{R}^n, \;\; f \in C_b^2. \tag{6.4}$$

Questa formula esprime il fatto che l'operatore L è il generatore (infinitesimale) del processo di diffusione (X_t^x).

Si potrebbe dimostrare il precedente corollario anche solo assumendo che b sia Lipschitziana su $\mathbb{R}^n$ e $f \in C_b^2$.

Tuttavia il risultato in generale non vale con f soltanto di classe C^2 (non si può ottenere il corollario con la stessa generalità con cui si potrebbe provare la formula di Ito). La formula di Ito è un uguaglianza ω per ω tra processi. Il corollario precedente richiede addizionali condizioni di integrabilità sui processi che si considerano.

Partendo dal Corollario 6.2 si potrebbe provare che la legge $p(t,x,\cdot)$ della v.a. X_t^x è soluzione (in senso generalizzato) di un'equazione parabolica di tipo Fokker-Planck, in cui compare l'aggiunto formale L^* di L,

$$L^* = \frac{1}{2} Tr(qD^2 \cdot) - \sum_{i=1}^n \partial_{x_i}[b_i(x)\cdot].$$

Più precisamente, se indichiamo con $p(t,x)$ la misura $p(t,x,\cdot)$, si ha: $\partial_t p(t,x) = L^* p(t,x), \quad t>0$, $p(0,x) = \delta_x$, $x \in \mathbb{R}^n$; l'equazione precedente è intesa in senso distribuzionale, cioè, per ogni $g \in C_0^2(\mathbb{R}^n)$ (ovvero per ogni g di classe C^2 su $\mathbb{R}^n$ con supporto compatto), deve valere

$$\partial_t \Big(\int_{\mathbb{R}^n} g(y) p(t,x)(dy) \Big) = \int_{\mathbb{R}^n} Lg(y) p(t,x)(dy), \; t>0,$$

vedere per esempio [9], [7] e [16], [20]. Le equazioni di Fokker-Planck si chiamano anche equazioni di Kolmogorov forward.

6.2 Dimostrazione della formula di Ito unidimensionale

Riprendiamo la formula di Ito quando $\mathbb{R}^n = \mathbb{R}$.

Teorema 6.3. *Sia $f \in C_{b,T}^{1,2}$. Sia $(X_t) = (X_t^x)$, $x \in \mathbb{R}$, il processo soluzione di (4.4), sotto le Ipotesi 4.1. Considerando il processo continuo $f(t, X_t)$, si ha:*

$$\begin{aligned} f(t, X_t(\omega)) - f(0,x) = &\int_0^t \partial_s f(u, X_u(\omega)) du + \int_0^t \partial_x f(u, X_u(\omega)) b(X_u(\omega)) du \\ &+ \sigma^2 \frac{1}{2} \int_0^t \partial_{xx} f(u, X_u(\omega)) du + \sigma \int_0^t \partial_x f(u, X_u) dW_u. \end{aligned}$$

Spieghiamo un metodo pratico per ricordarsi la formula di Ito.

Si tratta di scrivere la formula di Taylor per $f(X_t)$, identificando dW_t con $\sqrt{dt}$ (per spiegare meglio l'idea trattiamo f non dipendente da t).

$$\begin{aligned} df(X_t) &= f'(X_t) dX_t + \frac{1}{2} f''(X_t)(dX_t)^2 + o((dX_t)^2) \\ &= f'(X_t)(b(X_t)dt + \sigma dW_t) \; + \; \frac{1}{2} f''(X_t)(b(X_t)dt + \sigma dW_t)^2 + o((dX_t)^2) \\ &= f'(X_t) b(X_t) dt + f'(X_t) \sigma dW_t \; + \; \frac{1}{2} f''(X_t) \sigma^2 dt + o(dt). \end{aligned}$$

Trascurando $o(dt)$ e integrando, si trova proprio

$$f(X_t) - f(x) = \int_0^t (b(X_u)f'(X_u) + \frac{1}{2}\sigma^2 f''(X_u))du + \int_0^t f'(X_u)\sigma dW_u.$$

Esempio 6.4. Sia $(W_t^x) = (W_t + x)$, con (W_t) processo di Wiener reale, $x \in \mathbb{R}$ fissato. Si ha: $dW_t^x = dW_t$, $W_0^x = x$; quindi (W_t^x) risolve (4.4) con $b = 0$ e $\sigma = 1$.

Consideriamo $f(x) = \sin(x)$ e il processo $Y_t = f(W_t^x) = \sin(W_t^x)$. Usando la formula di Ito, si trova

$$dY_t = f'(W_t^x)dW_t + \frac{1}{2}f''(W_t^x)dt = \cos(W_t^x)dW_t - \frac{1}{2}\sin(W_t^x)dt.$$

Quindi

$$\sin(W_t^x) - x = \int_0^t \cos(W_s^x)dW_s - \frac{1}{2}\int_0^t \sin(W_s^x)ds.$$

La formula è in contrasto con la nostra intuizione basata sulle regole del calcolo integrale alla Cauchy-Riemann per la presenza di $-\frac{1}{2}\int_0^t \sin(W_s^x)ds$.

Dimostrazione del Teorema 6.3 Per semplicità di notazione, supponiamo che f non dipenda da t. Inoltre poniamo $\sigma = 1$.

Fissiamo $t > 0$. Consideriamo una generica partizione $\pi = \{s_0, ...s_n\}$, con $s_0 = 0 < ... < s_n = t$; $|\pi| = \max_{0\le k\le n-1} |s_{k+1} - s_k|$.

Dalla formula di Taylor, si ha (ragionando ω per ω):

$$\begin{aligned} f(X_t) - f(x) &= \sum_{k=0}^{n-1}[f(X_{s_{k+1}}) - f(X_{s_k})] \\ &= \sum_{k=0}^{n-1} f'(X_{s_k}) \triangle X_k + \frac{1}{2}\sum_{k=0}^{n-1} f''(\xi_k)[\triangle X_k]^2 = \Gamma_\pi, \end{aligned}$$

dove $\triangle X_k = X_{s_{k+1}} - X_{s_k}$ e $|\xi_k - X_{s_k}| \le |X_{s_{k+1}} - X_{s_k}|$, $k = 0, ..., n-1$.

Proveremo la formula di Ito, mostrando che esiste una successione (π_m) di partizioni di $[0, t]$ (con $|\pi_m| \to 0$, per $m \to \infty$) tale che

$$\lim_{m\to\infty} \Gamma_{\pi_m} = \int_0^t (b(X_u)f'(X_u) + \frac{1}{2}f''(X_u))du + \int_0^t f'(X_u)dW_u, \ \ q.c.. \tag{6.5}$$

Fissata π, risulta che

$$\sum_{k=0}^{n-1} f'(X_{s_k}) \triangle X_k = \sum_{k=0}^{n-1} f'(X_{s_k}) \int_{s_k}^{s_{k+1}} b(X_u)du + \sum_{k=0}^{n-1} f'(X_{s_k})\triangle W_k,$$

dove $\triangle W_k = W_{s_{k+1}} - W_{s_k}$. Ora è chiaro che

$$\begin{aligned} &\Big|\sum_{k=0}^{n-1} f'(X_{s_k}) \int_{s_k}^{s_{k+1}} b(X_u)du - \sum_{k=0}^{n-1} f'(X_{s_k})\, b(X_{s_k})|s_{k+1} - s_k|\Big| \\ &\le \sum_{k=0}^{n-1} |f'(X_{s_k})|\, |b(X_{u_k}) - b(X_{s_k})|\, |s_{k+1} - s_k| \le \epsilon \|f'\|_\infty t, \ \ q.c., \end{aligned}$$

per $|\pi|$ sufficientemente piccola (con $u_k \in [s_k, s_{k+1}]$). Perciò

$$\lim_{|\pi|\to 0^+} \sum_{k=0}^{n-1} f'(X_{s_k}) \int_{s_k}^{s_{k+1}} b(X_u)du = \lim_{|\pi|\to 0^+} \sum_{k=0}^{n-1} f'(X_{s_k})\, b(X_{s_k})|s_{k+1} - s_k|$$
$$= \int_0^t f'(X_u)b(X_u)du, \quad q.c..$$

Trattiamo ora $\sum_{k=0}^{n-1} f'(X_{s_k})\triangle W_k$. Per (5.4) si ha che (in $L^2(\Omega)$)

$$\lim_{|\pi|\to 0^+} \sum_{k=0}^{n-1} f'(X_{s_k})\triangle W_k = \int_0^t f'(X_u)dW_u.$$

Quindi esiste una successione $(\pi_m^1)_{m\in\mathbb{N}}$ di partizioni di $[0,t]$ tale che $\sum_{s_k,\, s_{k+1}\in\pi_m^1} f'(X_{s_k})\triangle W_k$ converge q.c. a $\int_0^t f'(X_u)dW_u$, per $m \to \infty$ (più precisamente, $s_k = s_k^m$, $k = 0, ..., n(m)$).

Consideriamo ora $\sum_{k=0}^{n-1} f''(\xi_k)[\triangle X_k]^2$. Si ha:

$$\sum_{k=0}^{n-1} f''(\xi_k)[\triangle X_k]^2$$
$$= \sum_{k=0}^{n-1} f''(\xi_k)\Big(\int_{s_k}^{s_{k+1}} b(X_u)du\Big)^2 + \sum_{k=0}^{n-1} f''(\xi_k)[\triangle W_k]^2 + 2\sum_{k=0}^{n-1} f''(\xi_k)[\triangle W_k]\int_{s_k}^{s_{k+1}} b(X_u)du.$$

Sia $\sup_{u\in[0,t]} |b(X_u(\omega))| = M(\omega)$. Si ha:

$$2\sum_{k=0}^{n-1} f''(\xi_k)[\triangle W_k]\int_{s_k}^{s_{k+1}} b(X_u)du \le 2\|f''\|_\infty M(\omega)(b-a) \max_{k=0,\dots,n-1} |\triangle W_k| \;\to\; 0,$$

per $|\pi| \to 0$, q.c.. In modo simile si prova che

$$\lim_{\pi\to 0^+} \sum_{k=0}^{n-1} f''(\xi_k)\Big(\int_{s_k}^{s_{k+1}} b(X_u)du\Big)^2 = 0, \quad q.c.$$

Rimane da trattare *il termine critico* $\sum_{k=0}^{n-1} f''(\xi_k)[\triangle W_k]^2$. Proviamo che, passando eventualmente ad una sottosuccessione (π_m) di (π_m^1), risulta

$$\lim_{|\pi_m|\to 0} \sum_{s_k,\, s_{k+1}\in\pi_m} f''(\xi_k)[\triangle W_k]^2 = \int_0^t f''(X_u)du. \tag{6.6}$$

Questo concluderà la dimostrazione (osserviamo che se le traiettorie di (W_t) fossero BV avremmo

$$\sum_{s_k,\, s_{k+1}\in\pi} f''(\xi_k)[\triangle W_k]^2 \le \max_{k=0,\dots,n-1} |\triangle W_k| \|f''\|_\infty \sum_{s_k,\, s_{k+1}\in\pi} |\triangle W_k| \;\to\; 0,$$

q.c., per $|\pi| \to 0$).

Per provare (6.6), consideriamo per un momento che $\xi_k = X_{s_k}$, per $k = 0, ...n-1$, e proviamo che esiste una sottosuccessione (π_m^2) di (π_m^1) tale che

$$\lim_{|\pi_m^2|\to 0} \sum_{s_k,\, s_{k+1}\in\pi_m^2} f''(X_{s_k})[\triangle W_k]^2 = \int_0^t f''(X_u)du, \ \ q.c. \tag{6.7}$$

Partiamo dalla formula ben nota

$$\lim_{|\pi|\to 0} \sum_{k=0}^{n-1} f''(X_{s_k})[s_{k+1} - s_k] = \int_0^t f''(X_u)du, \ \ q.c.. \tag{6.8}$$

La formula (6.7) segue se proviamo che

$$\lim_{|\pi|\to 0} \mathbb{E}\Big(\sum_{k=0}^{n-1} f''(X_{s_k})[\triangle W_k]^2 \ - \sum_{k=0}^{n-1} f''(X_{s_k})[s_{k+1} - s_k]\Big)^2 = 0. \tag{6.9}$$

Ponendo $\triangle s_k = [s_{k+1} - s_k]$, si trova:

$$\mathbb{E}\Big(\sum_{k=0}^{n-1} f''(X_{s_k})[\triangle W_k]^2 \ - \sum_{k=0}^{n-1} f''(X_{s_k})\triangle s_k\Big)^2 = \mathbb{E}\Big(\sum_{k=0}^{n-1} f''(X_{s_k})[\triangle W_k^2 \ - \triangle s_k]\Big)^2$$

$$= \mathbb{E} \sum_{k=0}^{n-1} f''^2(X_{s_k})[\triangle W_k^2 \ - \triangle s_k]^2 + 2\mathbb{E} \sum_{k<j}^{n-1} f''^2(X_{s_k})[\triangle W_k^2 \ - \triangle s_k]\, f''^2(X_{s_j})[\triangle W_j^2 \ - \triangle s_j]$$

$$= (\text{usando l'indipendenza e } \ \mathbb{E}[\triangle W_k^2 \ - \triangle s_k] = 0) \ \ \mathbb{E} \sum_{k=0}^{n-1} f''^2(X_{s_k})[\triangle W_k^2 \ - \triangle s_k]^2$$

$$\leq \ (\text{ragionando come in (3.10)}) \ \|f''\|_\infty^2 \sum_{k=0}^{n-1} (\triangle s_k)^2 \mathbb{E}\Big[\frac{\triangle W_k^2}{\triangle s_k} - 1\Big]^2 \leq 2\|f''\|_\infty^2 |\pi| \sum_{k=0}^{n-1} \triangle s_k$$

$$= 2t\|f''\|_\infty^2 |\pi| \ \to \ 0,$$

per $|\pi| \to 0$. Questo prova (6.7). Per ottenere (6.6), occorre verificare che, passando eventualmente ad una sottosuccessione (π_m) di (π_m^2), risulta q.c.

$$\lim_{|\pi_m|\to 0} \sum_{s_k,\, s_{k+1}\in\pi_m} [f''(X_{s_k}) - f''(\xi_k)][\triangle W_k]^2 = 0. \tag{6.10}$$

Fissata una partizione π di $[0, t]$, si ha:

$$\sum_{k=0}^{n-1} |f''(X_{s_k}) - f''(\xi_k)|[\triangle W_k]^2 \leq \max_{k=0,...,n-1} |f''(X_{s_k}) - f''(\xi_k)| \sum_{k=0}^{n-1} [\triangle W_k]^2.$$

Poichè $|X_{s_k} - \xi_k| \leq |X_{s_{k+1}} - X_{s_k}|$, usando anche la continuità di f, si trova

$$\lim_{|\pi|\to 0^+} \max_{k=0,...,n-1} |f''(X_{s_k}) - f''(\xi_k)| = 0, \ \ q.c.. \tag{6.11}$$

Abbiamo già visto che $\lim_{|\pi|\to 0} \sum_{k=0}^{n-1} [\triangle W_k]^2 = t$ in $L^2(\Omega)$. Quindi, passando ad una sottosuccessione di (π_m) di (π_m^2), risulta q.c.

$$\lim_{|\pi_m|\to 0} \sum_{s_k,\, s_{k+1}\in\pi_m} [\triangle W_k]^2 = t.$$

Combinando l'ultima formula con (6.11), si trova (6.10). La dimostrazione è conclusa. ■

Si può rendere più elegante la precedente dimostrazione, senza passare per sottosuccessioni di partizioni, provando direttamente che $\lim_{|\pi|\to 0} \Gamma_\pi = 0 \quad$ in probabilità.

7 Cenni sui semigruppi di Markov e le equazioni paraboliche di Kolmogorov

In questa sezione occorre avere una certa familiarità con le principali proprietà dell'attesa condizionale. Vogliamo introdurre la proprietà di Markov; si tratta di un argomento non facile. Presentiamo solo alcune semplici considerazioni. Per maggiori chiarimenti e dimostrazioni si può consultare per esempio [15], [17] e [19].

Indichiamo con $\mathcal{B}_b(\mathbb{R}^n)$, lo spazio di Banach delle funzione Borelliane $f : \mathbb{R}^n \to \mathbb{R}$ limitate, munito della norma del sup $\|\cdot\|_\infty$.

Sia $(X_t^x)_{t\geq 0}$ una famiglia di processi, dipendenti da $x \in \mathbb{R}^n$, a valori in $\mathbb{R}^n$, definiti e adattati rispetto ad una stessa base stocastica $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. Richiediamo che $\mathbb{P}(X_0^x = x) = 1$, per ogni x ("il processo (X_t^x) parte da $x \in \mathbb{R}^n$").

Introduciamo la *funzione di transizione* p, definita su $[0,\infty) \times \mathbb{R}^n \times B(\mathbb{R}^n)$,

$$p(t,x,A) := \mathbb{P}(X_t^x \in A), \quad x \in \mathbb{R}^n, \ t \geq 0, \ A \in B(\mathbb{R}^n);$$

dunque $p(t,x,\cdot)$ è la legge della v.a. (X_t^x) e $0 \leq p \leq 1$.

Supponiamo che $p(\cdot,\cdot,A)$ sia di Borel su $[0,\infty) \times \mathbb{R}^n$, per ogni $A \in B(\mathbb{R}^n)$.

Diciamo che (X_t^x) è un processo di Markov rispetto a $(\mathcal{F}_t)$, omogeneo nel tempo e con funzione di transizione p (o brevemente che è di Markov), se, per ogni $A \in B(\mathbb{R}^n)$, si ha:

$$(*) \ \ \mathbb{E}(1_A(X_{t+h}^x)/\mathcal{F}_t) = p(h, X_t^x, A), \ \ x \in \mathbb{R}^n, \ A \in B(\mathbb{R}^n), \ t, \ h \ \geq 0.$$

Chiaramente (*) implica che $\mathbb{E}(1_A(X_{t+h}^x)/\mathcal{F}_t) = \mathbb{E}(1_A(X_{t+h}^x)/X_t^x)$ (dove $\mathbb{E}(1_A(X_{t+h})/X_t) := \mathbb{E}(1_A(X_{t+h})/\sigma(X_t))$.

Intuitivamente la proprietà di Markov dice che i valori futuri del processo dipendono solo dal presente e non dai valori passati del processo.

Si dimostra facilmente che (*) equivale a

$$\mathbb{E}(f(X_{t+h}^x))/\mathcal{F}_t) = \int_{\mathbb{R}^n} f(y)p(h, X_t^x, dy), \quad x \in \mathbb{R}^n, \ \ t, \ h \ \geq 0, \ \ f \in \mathcal{B}_b(\mathbb{R}^n).$$

Si usa anche la notazione:

$$\mathbb{P}(X_{t+h}^x \in A/\mathcal{F}_t) := \mathbb{E}(1_A(X_{t+h}^x)/\mathcal{F}_t).$$

Si noti che $p(0,x,A) = \mathbb{P}(X_0^x \in A) = \delta_x(A)$.

Proposizione 7.1. *Il processo di Wiener n-dimensionale (W_t^x), $x \in \mathbb{R}^n$, $W_t^x := x + W_t$, definito e adattato rispetto a $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, è di Markov con funzione di transizione p,*

$$p(h,x,A) := N(x, hI)(A) = \mathbb{P}(W_h^x \in A). \tag{7.1}$$

Dimostrazione. Si ha, per $A \in B(\mathbb{R}^n)$,

$$\mathbb{P}(W^x_{t+h} \in A/\mathcal{F}_t) = \mathbb{E}(1_A([W^x_{t+h} - W^x_t] + W^x_t)/\mathcal{F}_t).$$

Ora si può ragionare cosi': $1_A([W^x_{t+h} - W^x_t] + W^x_t) = f([W^x_{t+h} - W^x_t], W^x_t)$, dove $f(x,y) = 1_A(x+y)$, $x, y \in \mathbb{R}^n$. Quindi

$$\mathbb{E}(f(W^x_{t+h} - W^x_t, W^x_t)/\mathcal{F}_t) = \mathbb{E}(f(W^x_{t+h} - W^x_t, y)\,|_{y=W^x_t}, \tag{7.2}$$

usando il fatto che W^x_t è $\mathcal{F}_t$-misurabile e $W^x_{t+h} - W^x_t$ è indipendente da $\mathcal{F}_t$. La formula (7.2) si verifica facilmente se $f(x,y)$ è del tipo $F(x)G(y)$, con F, $G \in \mathcal{B}(\mathbb{R}^n)$ e poi si estende a tutte le $f \in \mathcal{B}(\mathbb{R}^{2n})$ con un procedimento di approssimazione (vedere per esempio [15, Sezione 0.8]).

Dunque abbiamo che

$$\begin{aligned}\mathbb{P}(W^x_{t+h} \in A/\mathcal{F}_t) = \mathbb{E}(1_A([W^x_{t+h} - W^x_t] + y))|_{y=W^x_t} &= \mathbb{P}(W^x_{t+h} - W^x_t + y \in A)|_{y=W^x_t} \\ &= N(W^x_t, hI)(A) = p(h, W^x_t, A).\end{aligned} \tag{7.3}$$

■

Dato un processo stocastico di Markov (X^x_t), possiamo associargli la famiglia di operatori lineari (P_t), $P_t : \mathcal{B}_b(\mathbb{R}^n) \to \mathcal{B}_b(\mathbb{R}^n)$,

$$P_t f(x) = \mathbb{E} f(X^x_t), \ \ f \in \mathcal{B}_b(\mathbb{R}^n), \ x \in \mathbb{R}^n, \ t \geq 0. \tag{7.4}$$

Ogni operatore P_t è di contrazioni. Infatti si ha:

$$|\mathbb{E} f(X^x_t)| \leq \|f\|_\infty, \ \ f \in \mathcal{B}_b(\mathbb{R}^n), \ x \in \mathbb{R}^n, \ t \geq 0.$$

Inoltre $P_0 = I$. Grazie alla proprietà di Markov, la famiglia (P_t) è un semigruppo di operatori su $\mathcal{B}_b(\mathbb{R}^n)$, cioè risulta $P_t P_h = P_{t+h}, \ \ t, h \geq 0$.

Infatti, si ha, per ogni $A \in B(\mathbb{R}^n)$,

$$\begin{aligned}&P_{t+h} 1_A(x) = \mathbb{E}(1_A(X^x_{t+h})) \\ &= \mathbb{E}[\mathbb{E}(1_A(X^x_{t+h})/\mathcal{F}_t)] = \mathbb{E}[p(h, X^x_t, A)] = P_t(p(h,\cdot,A))(x) = P_t(P_h 1_A)(x),\end{aligned}$$

$x \in \mathbb{R}^n, \ t, \ h \geq 0$. Il semigruppo di operatori (P_t) si chiama **semigruppo di Markov** (associato al processo (X^x_t)). Osserviamo che

$$P_t f(x) = \int_{\mathbb{R}^n} f(y) p(t, x, dy), \ \ \ f \in \mathcal{B}_b(\mathbb{R}^n).$$

Esempio 7.2. Nel caso del processo di Wiener (W^x_t), il semigruppo di Markov associato ha un'espressione esplicita; si tratta del classico *semigruppo del calore:*

$$P^{cal}_t f(x) = \int_{\mathbb{R}^n} f(y) p(t,x,dy) = \int_{\mathbb{R}^n} f(y) N(x, tI) dy = \int_{\mathbb{R}^n} f(y) \frac{e^{\frac{-|x-y|^2}{2t}}}{\sqrt{(2\pi)^n t^n}} dy$$

Ritorniamo alla EDS (4.4). Vale il seguente teorema (per una dimostrazione vedere per esempio [21] o [16]).

Teorema 7.3. *Sia* (X_t^x), $x \in \mathbb{R}^n$, *il processo soluzione di* (4.4), *sotto l'Ipotesi 4.1. Allora* (X_t^x) *è di Markov.*

Dimostrazione. Ricordiamo solo i punti principali della dimostrazione, facendo l'ipotesi ulteriore che b sia Lipschitziana su $\mathbb{R}^n$, vedere anche [7] o [13].

(i) Indichiamo con $(X_t^{s,x})$, la soluzione della EDS (4.4) che in s vale x, cioè

$$X_t^{s,x} = x + \int_s^t b(X_r^{s,x})dr + \sigma(W_t - W_s), \;\; t \geq s \geq 0.$$

Si prova che esiste una versione continua, nei parametri s, t e x di $(X_t^{s,x})$.

(ii) Si prova la legge di flusso

$$X_t^{0,x} = X_t^{s, X_s^{0,x}}$$

(questa è ben nota nel caso di X_t^x deterministico, cioè quando $\sigma = 0$ in (4.4)).

(iii) Si prova che $X_{t+h}^{t,y}$ e $X_h^{0,y}$ hanno la stessa legge (al variare dei parametri t, $h \geq 0$, $y \in \mathbb{R}^n$). Questo segue dal fatto che b non dipende da t.

(iv) Infine, si ha, per $A \in B(\mathbb{R}^n)$,

$$\mathbb{E}(1_A(X_{t+h}^x)/\mathcal{F}_t) = \mathbb{E}(1_A(X_{t+h}^{0,x})/\mathcal{F}_t) = \mathbb{E}(1_A(X_{t+h}^{t,X_t^x})/\mathcal{F}_t) = p(h, X_t^x, A);$$

nell'ultimo passaggio si è usato (iii) e un argomento basato sia sull'indipendenza di $X_{t+h}^{t,y}$ da $\mathcal{F}_t$ e sia sulla misurabilità di X_t^x rispetto a $\mathcal{F}_t$, vedere per esempio [1, Lemma 3.9]. Questo ragionamento raffina quello usato in (7.3). ■

Vogliamo concludere, accennando ad un legame tra le equazioni paraboliche di Kolmogorov e le EDS.

Data una matrice q, $n\times n$, simmetrica e semidefinita positiva, indichiamo con $\sigma = \sqrt{q}$ una matrice reale $n \times n$ tale che $\sigma\sigma^T = q$ (σ non è necessariamente unica, lo sarebbe se richiedessi la proprietà di simmetria).

Proposizione 7.4. *Consideriamo il seguente problema di Cauchy*

$$\begin{cases} \partial_t u(t,x) = \frac{1}{2} Tr(q\, D_x^2 u(t,x)) + \langle D_x u(t,x), b(x)\rangle, \;\; x \in \mathbb{R}^n, \; t \in (0,T], \\ \\ u(0,x) = f(x), \;\; f \in \mathcal{B}_b(\mathbb{R}^n), \end{cases} \tag{7.5}$$

dove b *è limitata e di classe* C^1 *e* q *è simmetrica e semidefinita positiva.*

Supponiamo di sapere che esiste per (7.5) *una soluzione classica e regolare* $u \in C_{b,T}^{1,2}$. *Allora, se indichiamo con* (X_t^x) *la soluzione di* (4.4), *con* $\sigma = \sqrt{q}$, *si ha:*

$$\mathbb{E}(f(X_t^x)) = P_t f(x) = u(t,x), \quad x \in \mathbb{R}^n, \;\; t \geq 0. \tag{7.6}$$

Dimostrazione. Si tratta di applicare la formula di Ito con la funzione $v(s,x) = u(t - s, x)$, $s \in [0,t]$, $x \in \mathbb{R}^n$ e il processo $X_t = X_t^x$. Si ha:

$$
\begin{aligned}
v(t,X_t)-v(0,x) &= u(0,X_t)-u(t,x)\\
=\int_0^t\Big(\partial_s v(r,X_r)+\langle D_x v(r,X_r),b(X_r)\rangle &+\frac{1}{2}Tr(q\,D_x^2 v(r,X_r))\Big)dr\\
&+\int_0^t\langle D_x v(r,X_r),\sigma dW_r\rangle\\
=\int_0^t\Big(-\partial_s u(t-r,X_r)+\langle D_x u(t-r,X_r),b(X_r)\rangle &+\frac{1}{2}Tr(q\,D_x^2 u(t-r,X_r))\Big)dr\\
+\int_0^t\langle D_x u(t-r,X_r),\sigma dW_r\rangle &= \int_0^t\langle D_x u(t-r,X_r),\sigma dW_r\rangle
\end{aligned}
$$

(poichè u è soluzione di (7.5)). Dunque

$$
f(X_t)-u(t,x)=\int_0^t\langle D_x u(t-r,X_r),\sigma dW_r\rangle.
$$

Applicando l'attesa, si trova la tesi, grazie al fatto che $\mathbb{E}(\int_0^t\langle D_x u(t-r,X_r),\sigma dW_r\rangle)=0$. ■

Osserviamo che esiste una soluzione classica $u\in C_{b,T}^{1,2}$ per (7.5), se per esempio b, $f\in C_b^2(\mathbb{R}^n)$. Tuttavia la precedente formula (7.6), vale sotto ipotesi meno restrittive sia su b che su u, vedere per esempio [10].

Si pone anche il problema di capire quando partendo dall'equazione (7.5) e considerata la funzione $v(t,x)=P_t f(x)$, $f\in\mathcal{B}_b(\mathbb{R}^n)$, che si ottiene risolvendo un'associata EDS con $\sigma=\sqrt{q}$, questa funzione v risolve effettivamente (7.5). Questo è una problema non banale. Certamente la risposta dipende dalla regolarità che chiediamo su f (dato iniziale) e sul coefficiente di drift b e anche dal fatto che q può essere degenere, vedere anche [18] e [20].

Bibliografia - Calcolo delle Probabilità

[1] Baldi P., Calcolo delle probabilità e statistica, McGraw-Hill, Milano, 2/ed, 1998.

[2] Billingsley P, Probability and Measure, 3rd Edition, Wiley, 1995.

[3] Dudley, R. M., Real analysis and probability. The Wadsworth Brooks/Cole Mathematics Series. Wadsworth and Brooks/Cole AdvancedBooks, CA, 1989.

[4] Durrett R., Probability: Theory and Examples, 2nd edition, Wadsworth, Belmont,. CA, 1995.

[5] Flandoli F., Introduzione alla Probabilità, appunti, 1993.

Calcolo stocastico - livello introduttivo

[6] Baldi P., Equazioni differenziali stocastiche e applicazioni, 2 ed., Pitagora Editore, 2000.

[7] Da Prato, G., Introduction to differential stochastic equations, Second edition. Appunti dei Corsi Tenuti da Docenti della Scuola. Scuola Normale Superiore, Pisa, 1998.

[8] Evans, L.C., An introduction to Stochastic Differential Equations (vedere http://math.berkeley.edu/ evans/).

[9] Flandoli F., Introduzione alla Probabilità, appunti, 1993.

[10] Friedman A., Stochastic differential equations and applications, I e II, Academic Press, 1975.

[11] Kahane j. P., A century of interplay between Taylor series, Fourier serie and Brownian motion, Bull. London. Math. Soc. 29 (1997), 257-279.

[12] Klebaner, F. C., Introduction to stochastic calculus with applications, Reprint of the 1998, Imperial College Press, London; distributed by World Scientific Publishing Co., 1999.

[13] Priola, E., Introduzione alle equazioni stocastiche, appunti del Corso di Dottorato presso il Dipartimento di Matematica dell'Università di Lecce, 2004.

[14] Varadhan S. R. S., Lectures on diffusion problems and partial differential equations, Tata Institute, Bombay, 1980.

Calcolo stocastico - livello avanzato

[15] Baldi P., Equazioni differenziali stocastiche e applicazioni, 2 ed., Pitagora Editore, 2000.

[16] Durrett R., Stochastic calculus. A practical introduction. Probability and Stochastics Series, 1996.

[17] Ethier, S. N.; Kurtz, T. G., Markov processes. Characterization and convergence, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986.

[18] Ikeda N., Watanabe S., Stochastic differential equations and diffusion processes, II ed., 1989.

[19] Karatzas, I., Shreve, S. E., Brownian motion and stochastic calculus. Graduate Texts in Mathematics, 113. Springer-Verlag, New York, 1988.

[20] Krylov N. V., Introduction to the Theory of Diffusion Processes, A.M.S., 1995.

[21] Stroock D. W., Varadhan S. R. S., Multidimensional diffusion processes, Springer, 1979.

[22] Zabczyk, J., Topics in stochastic processes. Scuola Normale Superiore di Pisa. Quaderni, Pisa, 2004.