Markov chains, presented in the previous chapter, belong to a wide class of mathematical objects, the stochastic processes, which will introduced in the present chapter. Stochastic processes form the basis for the mathematical modeling of time-dependent phenomena. We will begin focussing our attention on the randomly-driven motion of a pollen grain, the well-known *brownian motion*, historically the first stochastic process to be stated in rigorous mathematical terms. The brownian motion will represent our starting point for introducing the fundamental notions of stochastic processes. It will be introduced retracing the pioneering work of A. Einstein.

## I. BROWNIAN MOTION: A HEURISTIC INTRODUCTION

For the sake of simplicity, the one-dimensional case will be considered first, the generalization to higher dimensionality being straightforward. We will assume that in a given time interval $\tau$ a pollen grain, due to random collisions with water molecules, undergoes a random variation $\zeta$ in its position. We will denote through:

$$\Phi(\tau, \zeta) \tag{1}$$

the probability density for this transition phenomenon, in such a way as:

$$\int_a^b \Phi(\tau, \zeta)\, d\zeta \tag{2}$$

can be interpreted as the probability that the particle, being in any given position $x$ at time $t$, is found in the interval $y \in [x + a, x + b]$ at time $s = t + \tau$; we are assuming in particular that the *transition probability density* $\Phi(\tau, \zeta)$ depends only on the quantities $\tau = t - s$ and $\zeta = y - x$.

This object remarkably resembles the transition probability density of Markov chains, except for the fact that, in the present context, time and sample space are continuous. We will assume that the following property:

$$\Phi(\tau, \zeta) = \Phi(\tau, -\zeta) \tag{3}$$

holds, which expresses the fact that moving rightwards of leftwards is equally probable, and implies that the "mean jump" vanishes:

$$\int_{\mathbb{R}} d\zeta\, \zeta\, \Phi(\tau, \zeta) = 0 \tag{4}$$

Let now $\int_{\mathcal{A}} p(x, t)\, dx$ denote the probability that, at time $t$, the particle lies in the interval $\mathcal{A}$. We will assume that the following very natural *continuity equation* holds:

$$p(x, t + \tau) = \int_{\mathbb{R}} d\zeta\, \Phi(\tau, \zeta)\, p(x - \zeta, t) \tag{5}$$

expressing the fact that the position at time $s = t + \tau$ is the sum of the position at time $t$ and of the transition occurred between times $t$ and $s$. If

$\tau$ is *small*, provided that $\Phi(\tau, \zeta)$ tends to 0 rapidly enough as $|\zeta| \to +\infty$, we can expand in Taylor series:

$$p(x, t + \tau) = p(x, t) + \tau \frac{\partial p}{\partial t}(x, t) + \dots \quad , \tag{6}$$

$$p(x - \zeta, t) = p(x, t) - \zeta \frac{\partial \Phi}{\partial x}(x, t) + \frac{1}{2}\zeta^2 \frac{\partial^2 \Phi}{\partial^2 x}(x, t) + \dots \tag{7}$$

Substituting the above expansions in the continuity equation (5) and retaining non vanishing low-order terms only:

$$p(x, t+\tau) = p(x, t)+\tau \frac{\partial p}{\partial t}(x, t)+\dots = p(x, t)+\frac{1}{2}\frac{\partial^2 p}{\partial^2 x}(x, t) \int_{\mathbb{R}} d\zeta \, \zeta^2 \Phi(\tau, \zeta)+\dots \tag{8}$$

where we have taken into account the fact (4) that the mean transition is vanishing, enabling us to retain terms of at most first order in $\tau$ and second in $\zeta$, and that the transition probability density is normalized to 1. Inspired by the microscopic theory of diffusion, we are induced to interpret the quantity:

$$D = \frac{1}{2\tau} \int_{\mathbb{R}} d\zeta \, \zeta^2 \Phi(\tau, \zeta) = \frac{\sigma^2(\tau)}{2\tau} \tag{9}$$

as macroscopic diffusion coefficient; for this quantity to be *constant*, it is necessary to assume that $\sigma^2(\tau)$ is proportional to $\tau$. Under this additionally hypothesis, one arrives to the **diffusion equation**:

$$\frac{\partial p}{\partial t}(x, t) = D \frac{\partial^2 p}{\partial^2 x}(x, t) \tag{10}$$

To complete the description, it is necessary to equip the diffusion equation (10) with an initial condition. A possible choice is:

$$p(x, 0) = \delta(x) \tag{11}$$

representing a particle at the origin of a suitable reference frame. Were that the case, the solution of (10) would be a gaussian probability distribution:

$$p(x, t) = \frac{1}{\sqrt{4\pi D t}} \exp\left(-\frac{x^2}{4Dt}\right) \tag{12}$$

which coincides, by virtue of the continuity equation (5), with the transition probability density:

$$\Phi(\tau, \zeta) = \frac{1}{\sqrt{4\pi D \tau}} \exp\left(-\frac{\zeta^2}{4Dt}\right) \tag{13}$$

This expression will represent a natural starting point for the rigorous mathematical treatment of the brownian motion.

## II. THEORY OF STOCHASTIC PROCESSES

The heuristic discussion of the preceding paragraph has put in evidence several remarkable aspects: first, the motion of the pollen grain is treated with probabilistic arguments. This implies that a probability space $(\Omega, \mathcal{F}, P)$ has to be introduced, on which it must be possible to define the random variables:

$$X_t : (\Omega, \mathcal{F}, P) \to \left(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right)\right) \tag{14}$$

with the interpretation of position of the pollen grain at time $t$. The *time parameter* $t$ takes values on a suitable *time interval* $T \subseteq [0, +\infty)$. The movement of the pollen grain will be therefore described by the family:

$$\{X_t\}_{t \in T} \tag{15}$$

of random variables. For all $\omega \in \Omega$, one obtains the *trajectory*:

$$t \rightsquigarrow X_t(\omega) \tag{16}$$

corresponding to a possible motion of the system. Moreover, in the above outlined formalism it makes sense to compute expressions of the form:

$$P\left(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\right), \quad A_1, \ldots, A_n \in \mathcal{B}\left(\mathbb{R}^d\right) \tag{17}$$

As it is well known, the $\sigma$-algebra $\mathcal{F}$ contains all possible events, corresponding to all the possible statements one can formulate once an outcome of the experiment is registered; in the present context, however, the information obtained through an observation increases with time: at time $t$ one has gained knowledge of the position of the particle for past times $s \leq t$, but not for future times $s > t$. This means that a key ingredient for the probabilistic description of the brownian motion is represented by the following definition:

**Definizione 1** *We call* **filtration** $\{\mathcal{F}_t\}_{t \in T}$*,* $T \subset \mathbb{R}^+$*, a family of sub-*$\sigma$*-algebras of* $\mathcal{F}$ *increasing with* $t$*, i.e. such that* $\mathcal{F}_s \subset \mathcal{F}_t$ *if* $s \leq t$*. We call* **stochastic basis** *a probability space endowed with a filtration, i.e. an object of the form:*

$$\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in T}, P\right) \tag{18}$$

Intuitively, $\mathcal{F}_t$ contains all events that it is possible to *discriminate*, that is to conclude whether have occurred or not, once the system has been observed up to the instant $t$. In other words, it represents all the information available up to time $t$ and including time $t$.

We can now give the following general definition:

**Definizione 2** *A* **stochastic process** *is an object of the form:*

$$X = \left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in T}, \{X_t\}_{t \in T}, P\right) \tag{19}$$

*where:*

1. $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in T}, P\right)$ *is a stochastic basis*

2. $\{X_t\}_{t \in T}$ *is a family of random variables taking values in a measurable space* $(E, \mathcal{E})$*:*

$$X_t : \Omega \to E \tag{20}$$

*and such that, for all* $t$*,* $X_t$ *is* $\mathcal{F}_t$*-measurable. To express this circumstance, we say that* $\{X_t\}_{t \in T}$ *is* **adapted** *to the filtration.*

The set $E$ is called **state space** of the process: for the purpose of our applications, it will coincide with $\mathbb{R}^d$, endowed with the $\sigma$-algebra of Borel sets. Given $\omega \in \Omega$, the map:

$$T \ni t \rightsquigarrow X_t(\omega) \in E \tag{21}$$

is called **trajectory** of the process, which is called **(almost certainly) continuous** if the set of points $\omega$ such that the corresponding trajectory is continuous has probability 1.

## 1. *Finite-dimensional Laws*

Let $X = \left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in T}, \{X_t\}_{t \in T}, P\right)$ be a stochastic process and $\pi = (t_1, \dots, t_n)$ a finite set of instants in $T$, such that $t_1 < \dots < t_n$. The map:

$$\Omega \ni \omega \rightsquigarrow (X_{t_1}(\omega), \dots, X_{t_n}(\omega)) \in E^n \tag{22}$$

is clearly a random variable; let $\mu_\pi$ be the image of $P$ through such map:

$$\mu_\pi\left(A_1 \times \dots \times A_n\right) = P\left(X_{t_1} \in A_1, \dots, X_{t_n} \in A_n\right), \quad A_1, \dots, A_n \in \mathcal{E}. \tag{23}$$

for all collections $\pi$; $\mu_\pi$ is called a **finite-dimensional law** of $X$. Two processes sharing the same finite-dimensional laws are said **equivalent**.

Given $\pi$, $\mu_\pi$ is a probability measure on $(E^n, \otimes^n \mathcal{E})$, $\otimes^n \mathcal{E}$ being the $\sigma$-algebra generated by set of the form $A_1 \times \dots \times A_n$, $A_j \in \mathcal{E}$. Finite-dimensional laws, by construction, satisfy a simple **consistency** property:

$$\mu_\pi\left(A_1 \times ..A_{i-1} \times E \times A_{i+1} .. \times A_n\right) = \mu_{\pi'}\left(A_1 \times ..A_{i-1} \times A_{i+1} .. \times A_n\right) \tag{24}$$

for all $t$ and $\pi = (t_1, \dots, t_n)$, provided that $\pi' = (t_1, \dots t_{i-1}, t_{i+1}, \dots, t_n)$. On the contrary, we will say that a family of probability measures $\{\mu_\pi\}_\pi$ is **consistent** if it satisfies the property (24). We state without proof, that the interested reader can find in [**?** ], the following fundamental result:

**Teorema 3** *(***Kolmogorov***) Let* $E$ *be a complete and separable metric space,* $\mathcal{E}$ *the* $\sigma$*-algebra of Borel sets and* $\{\mu_\pi\}_\pi$ *a family of consistent probability measures. Let:*

$$\Omega \stackrel{def}{=} E^T = \{\omega : T \to E\} \tag{25}$$

*be the set of all functions* $\omega$ *from* $T$ *to* $E$*, and:*

$$\mathcal{F} \stackrel{def}{=} \mathcal{B}(E)^T \tag{26}$$

*the $\sigma$-algebra generated by the cylindrical sets:*

$$\{\omega \in \Omega \mid \omega(t_1) \in A_1, \ldots, \omega(t_n) \in A_n\} \tag{27}$$

*let moreover:*

$$X_t(\omega) \stackrel{def}{=} \omega(t) \tag{28}$$

*and $\mathcal{F}_t = \sigma(X_s,\, s \leq t)$, be the* **natural filtration**.

*Then there exists a* **unique** *probability measure $P$ on $(\Omega, \mathcal{F})$ such that the measures $\{\mu_\pi\}_\pi$ are the finite-dimensional laws of the stochastic process $X = \left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in T}, \{X_t\}_{t\in T}, P\right)$.*

Despite its complexity, the meaning of Kolmogorov's Theorem is clear: for any generic family of consistent probability distributions there exists a unique stochastic process having precisely that probability distributions as finite-dimensional laws. In many contexts, it is interesting to draw some conclusions about the regularity of the trajectories of a stochastic process.

To this purpose, we preliminary mention that two processes $X = \left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in T}, \{X_t\}_{t\in T}, P\right)$ and $X' = \left(\Omega', \mathcal{F}', \{\mathcal{F}'_t\}_{t\in T}, \{X'_t\}_{t\in T}, P'\right)$ are called **modifications** of each other if $(\Omega, \mathcal{F}, P) = (\Omega', \mathcal{F}', P')$ and if, for all $t \in T$, $X_t = X'_t$ almost surely.

We state without proof, that the interested reader can find in [**?** ], the following:

**Teorema 4** *(***Kolmogorov***) Let $X$ be a process taking values in $\mathbb{R}^d$, and such that for suitable $\alpha > 0$, $\beta > 0$, $c > 0$ and for all instants $s, t$:*

$$E\left[|X_t - X_s|^\beta\right] \leq c|t-s|^{1+\alpha} \tag{29}$$

*Then there exists a modification $X'$ of $X$ which is continuous.*

### A. Construction of the Brownian Motion

We now have all instruments for giving a precise definition of the brownian motion. Motivated by Einstein's discussion, we consider $\left(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right)\right)$ as state space of the process. We recall that in the historical tractation of the brownian motion a key role was played by the transition probability density, which follows a normal law. Moreover, in Einstein's discussion, the assumption that the transition of the pollen grain be independent on its position at current and past time was implicit. Summing up these observations and completing the description of the brownian motion with a suitable initial condition we formulate the following:

**Definizione 5** *A process $B = \left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \{B_t\}_{t\geq 0}, P\right)$ is called* **brownian motion** *if:*

1. *$B_0 = \underline{0}$ almost certainly*
2. *for all $0 \leq s \leq t$ the random variable $B_t - B_s$ is independent on $B_u$ for all $u \leq s$, and in particular it is independent on $\mathcal{F}_s$;*

3. *for all* $0 \leq s \leq t$ *the random variable* $B_t - B_s$ *has law* $N(\underline{0}, (t-s)\mathbb{I})$, *where* $\mathbb{I}$ *is the* $d \times d$ *identity matrix.*

The first property corresponds to the requirement that the particle starts its motion at the origin of a suitable reference frame. The second property is commonly resumed saying that **the increments of the brownian motion are independent of the past**, and is strongly related to the memoryless condition for Markov chains, as it will be soon explained in detail. The third property makes the introduction of a gaussian transition probability density rigorous.

As the reader might have observed, so far the stochastic basis of the brownian motion has not been defined. In fact, no arguments have been presented to guarantee the existence of a brownian motion: in the previous definition (5) the requirements formulated in the heuristic introduction to the brownian motion have just been formulated in the language of stochastic processes.

To the purpose of proving the existence of the brownian motion, we first observe that the definition (5) of brownian motion corresponds to assigning the finite-dimensional laws of the process. We will limit our treatment to the one-dimensional case $d = 1$, as the multidimensional case results from a straightforward generalization left to the reader.

First, we prove that the vector random variable $(B_{t_1}, \dots, B_{t_n})$ is normal. By virtue of the second and third property of the brownian motion the vector random variable $(Y_{t_1}, \dots, Y_{t_n})$ given by:

$$Y_{t_1} = B_{t_1} \quad Y_{t_k} = B_{t_k} - B_{t_{k-1}} \quad k = 2 \dots n \tag{30}$$

has independent and normally distributed components with zero mean and covariance matrix:

$$\Delta = \begin{pmatrix} t_1 & 0 & \dots & 0 \\ 0 & t_2 - t_1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & t_n - t_{n-1} \end{pmatrix} \tag{31}$$

Since it is related to the vector random variable $(B_{t_1}, \dots, B_{t_n})$ by the linear transformation:

$$\begin{pmatrix} B_{t_1} \\ B_{t_2} \\ \dots \\ B_{t_n} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 1 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} Y_{t_1} \\ Y_{t_2} \\ \dots \\ Y_{t_n} \end{pmatrix} = A \begin{pmatrix} Y_{t_1} \\ Y_{t_2} \\ \dots \\ Y_{t_n} \end{pmatrix} \tag{32}$$

the former is also normally distributed with covariance matrix:

$$\Gamma = A \, \Delta \, A^T \tag{33}$$

Since the covariance matrix $\Gamma$ is invertible with inverse $\Gamma^{-1} = A^{-1} \, \Delta^{-1} \, (A^{-1})^T$ where, as a simple calculation shows:

$$A^{-1} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ -1 & 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -1 & 1 \end{pmatrix} \tag{34}$$

we can write the finite-dimensional laws of the brownian motion:

$$\mu_\pi\left(A_1\times\cdots\times A_n\right)=\int_{A_1}dx_1\ldots\int_{A_n}dx_n\,\frac{\exp\left(-\frac{1}{2}\left(\sum_{i,j=1}^n x_i\Gamma^{-1}_{ij}x_j\right)\right)}{(2\pi)^{n/2}\,\det(\Gamma)^{1/2}} \tag{35}$$

where, due to equations (33) and (34), the following simplifications occur:

$$\sum_{i,j=1}^n x_i\Gamma^{-1}_{ij}x_j=\frac{x_1^2}{t_1}+\sum_{k=2}^n\frac{(x_k-x_{k-1})^2}{(t_k-t_{k-1})}\qquad \det(\Gamma)=t_1(t_2-t_1)\ldots(t_n-t_{n-1}) \tag{36}$$

leading to the identity:

$$\begin{aligned}&\mu_\pi\left(A_1\times\cdots\times A_n\right)= \qquad (37)\\ &=\textstyle\int_{A_1}dx_1\,p(x_1,t_1|0,0)\int_{A_2}dx_2\,p(x_2,t_2|x_1,t_1)\ldots\int_{A_n}dx_n p(x_n,t_n|x_{n-1},t_{n-1})\end{aligned}$$

where the transition probability density is given by:

$$p(y,t|x,s)=\frac{1}{\sqrt{2\pi(t-s)}}\exp\left(-\frac{(y-x)^2}{2(t-s)}\right),\quad s<t \tag{38}$$

In the multidimensional case one has:

$$p(\underline{y},t|\underline{x},s)=\frac{1}{[2\pi(t-s)]^{d/2}}\exp\left(-\frac{|\underline{y}-\underline{x}|^2}{2(t-s)}\right),\quad s<t \tag{39}$$

It is also immediate to realize that the expression (38) implies the three properties in the definition (5) of brownian motion. The first property is obvious. The second and third are retrieved choosing $n=2$ and $A_1=\mathbb{R}$:

$$\mu_\pi\left(A_2\right)=\int_{\mathbb{R}}dx_1\,p(x_1,t_1|0,0)\int_{A_2}dx_2\,p(x_2,t_2|x_1,t_1) \tag{40}$$

which shows that $B_{t_2}=B_{t_1}+B_{t_2}-B_{t_1}$, with the increment $B_{t_2}-B_{t_1}$ independent on $B_{t_1}$ and normally distributed with mean 0 and variance $t_2-t_1$.

## B. Transition Probability and Existence of the Brownian Motion

The definition (5) of brownian motion given above is equivalent to the assignment of the finite-dimensional laws. The latter present an interesting structure, and involve an object closely recalling the transition matrices introduced in the context of Markov chains. Consider in fact the following map, called **markovian transition function of the brownian motion**:

$$p(A,t\,|\underline{x},s)\stackrel{def}{=}\frac{1}{[2\pi(t-s)]^{d/2}}\int_A d\underline{y}\,\exp\left(-\frac{|\underline{y}-\underline{x}|^2}{2(t-s)}\right) \tag{41}$$

where $s,t\in T$, $s<t$, $\underline{x}\in\mathbb{R}^d$ and $A\in\mathcal{B}\left(\mathbb{R}^d\right)$. If $s=t$, we put instead:

$$p(A,s\,|\underline{x},s)\stackrel{def}{=}\delta_{\underline{x}}(A) \tag{42}$$

Intuitively, we interpret $p(A,t\,|\underline{x},s)$ as the probability that the pollen grain is in the Borel set $A$ at time $t$, given the fact that it was at the point $\underline{x}$ at time $s$. Nevertheless we observe that the probability of finding a particle at a precise point vanishes at any instant $s>0$, so that the map (42) cannot be properly interpreted as conditional probability. On the other hand, we realize that, by definition, $B_t - B_0 \sim N(0,t)$, and therefore:

$$P(B_t \in A) = p(A,t\,|\underline{0},0) = \frac{1}{(2\pi t)^{d/2}} \int_A d\underline{y}\, \exp\left(-\frac{|\underline{y}|^2}{2t}\right) \tag{43}$$

recalling that, by definition, $B_0 = \underline{0}$ almost surely. In general:

$$P(\underline{x} + B_t - B_s \in A) = p(A,t\,|\underline{x},s) \tag{44}$$

We now introduce some observations, that will turn out to be useful later in the discussion:

1. for fixed $s,t,A$ the function $\underline{x} \mapsto p\left(A,t\,|\,\underline{x},s\right)$ is Borel-measurable

2. for fixed $s,t,\underline{x}$, the function $A \mapsto p\left(A,t\,|\,\underline{x},s\right)$ is a probability measure on $\left(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right)\right)$;

3. $p$ satisfies the following **Chapman-Kolmogorov equation**:

$$p\left(A,t\,|\,\underline{x},s\right) = \int_{\mathbb{R}^d} p\left(A,t\,|\,\underline{y},u\right)\, p\left(dy,u\,|\,\underline{x},s\right) \tag{45}$$

for all $s<u<t$, $\underline{x}\in\mathbb{R}^d$and $A\in\mathcal{B}\left(\mathbb{R}^d\right)$.

The reader is invited to derive the last equation (45) through a direct calculation, which is a simple exercise on gaussian integrals.

As proved earlier, the finite-dimensional laws of the brownian motion are given by:

$$\mu_\pi\left(A_1\times\cdots\times A_n\right) = \int_{A_1} p(d\underline{x}_1,t_1\,|\underline{0},0)\int_{A_2}\cdots\int_{A_n} p(d\underline{x}_n,t_n\,|\underline{x}_{n-1},t_{n-1}) \tag{46}$$

where $\pi = (t_1,\dots,t_n)$, $0\le t_1<\cdots<t_n$. The consistence of these finite-dimensional laws is an immediate consequence of the Chapman-Kolmogorov equation.

Kolmogorov's Theorem (3) ensures the **existence** of a stochastic process:

$$B = \left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\ge 0}, \{B_t\}_{t\ge 0}, P\right) \tag{47}$$

having (46) as finite-dimensional laws, and makes it possible to construct the stochastic basis for such process. $\Omega$ is the set of trajectories:

$$\Omega = \{\omega : [0,+\infty) \to \mathbb{R}^d\} \tag{48}$$

$\mathcal{F}$ the $\sigma$-algebra $\mathcal{B}\left(\mathbb{R}^d\right)^{[0,+\infty)}$, appearing in Kolmogorov's Theorem, and $\mathcal{F}_t = \sigma\left(X_s,\, s\le t\right)$ the natural filtration. The process is defined by:

$$\Omega \ni \omega \rightsquigarrow B_t(\omega) \stackrel{def}{=} \omega(t) \tag{49}$$

The second Kolmogorov's Theorem guarantees that a large class of stochastic processes can be modified in such a way as to be turned into a continuous process. To prove that the second Kolmogorov's Theorem applies to the brownian motion, we consider $\beta = 2n$ for some integer $n \geq 2$ and $t > s$:

$$E\left[|B_t - B_s|^\beta\right] = \frac{1}{\sqrt{2\pi(t-s)}} \int_{\mathbb{R}} dx\, x^{2n} \exp\left(-\frac{x^2}{2(t-s)}\right) = \tag{50}$$

$$= \frac{2^n}{\sqrt{4\pi}} (t-s)^n \int_0^\infty du\, u^{n+\frac{1}{2}} \exp(-u) = \frac{2^n\,\Gamma\left(n+\frac{1}{2}\right)}{\sqrt{4\pi}} (t-s)^n = c\,(t-s)^{1+\alpha}$$

From now on, we will always assume to work with a continuous brownian motion.

Finally, it is often useful to chose a larger filtration than the natural one, so that the stochastic basis:

$$\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P\right) \tag{51}$$

satisfies the so-called **usual hypotheses**, that is:

1. the filtration is right continuous, i.e. $\mathcal{F}_t = \mathcal{F}_{t_+} \stackrel{def}{=} \bigcap_{s>t} \mathcal{F}_s$

2. each sub-$\sigma$-algebra $\mathcal{F}_t$ contains all the events of $\mathcal{F}$ with vanishing probability

A simple way to satisfy the usual hypotheses is add to all the $\sigma$-algebras $\mathcal{F}_t = \sigma(X_s,\, s \leq t)$ all the events of $\mathcal{F}$ with vanishing probability. The so-obtained filtration is referred to as **completed natural filtration**.

### C. The Martingale Property and the Markov Property

Due to the fact that the increments of the brownian motion are independent on the past, the stochastic process exhibits the following remarkable property:

$$\begin{aligned} E[B_t|\mathcal{F}_s] &= E[B_t - B_s + B_s|\mathcal{F}_s] = \\ &= E[B_t - B_s|\mathcal{F}_s] + E[B_s|\mathcal{F}_s] = \\ &= E[B_t - B_s] + B_s = B_s \end{aligned} \tag{52}$$

where the independence property and the fact that $B_s$ is $\mathcal{F}_s$-measurable have been recalled. The above equation means that $B_s$ is the best prediction for $B_t$ given the information obtained observing the system up to time $s$. The brownian motion is therefore a **martingale**.

The problem of determining $E[f(B_t)|\mathcal{F}_s]$, if $f : \mathbb{R}^d \to \mathbb{R}$ is a limited Borel function $f : \mathbb{R}^d \to \mathbb{R}$, has great interest. To this purpose, we will refer to theorem (**??**), which we recall here for the sake of clarity:

**Teorema 6** *Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\mathcal{G}$ and $\mathcal{H}$ mutually independent sub-$\sigma$-algebras of $\mathcal{F}$. Let $X : \Omega \to E$ be a $\mathcal{G}$-measurable random variable taking values in a measurable space $(E, \mathcal{E})$ and $\psi$ a function*

*$\psi : E \times \Omega \to \mathbb{R}$ $\mathcal{E} \otimes \mathcal{H}$-measurable and such that $\omega \mapsto \psi(X(\omega), \omega)$ is integrable. Then:*

$$E\left[\psi(X,\cdot)|\mathcal{G}\right] = \Phi(X), \quad \Phi(x) \stackrel{def}{=} E\left[\psi(x,\cdot)\right] \tag{53}$$

We know that $\mathcal{G} \equiv \mathcal{F}_s$ and $\mathcal{H} \equiv \sigma(B_t - B_s)$ are mutually independent, and that the random variable $X \equiv B_s : \Omega \to E \equiv \mathbb{R}^d$ is $\mathcal{G}$-measurable. Moreover the function $\psi : \mathbb{R}^d \times \Omega \to \mathbb{R}$ given by $\psi(\underline{x}, \omega) \equiv f(\underline{x} + B_t(\omega) - B_s(\omega))$ is $\mathcal{H}$-measurable. The function $\Phi(\underline{x})$ appearing in theorem (**??**) is therefore:

$$\Phi(\underline{x}) = E\left[f(\underline{x} + B_t - B_s))\right] \tag{54}$$

and since the random variable $\omega \rightsquigarrow \underline{x} + B_t(\omega) - B_s(\omega)$ has law $N(\underline{x}, (t-s)\,\mathbb{I})$:

$$\Phi(\underline{x}) = \frac{1}{\left[2\pi(t-s)\right]^{d/2}} \int_{\mathbb{R}^d} d\underline{y}\, f(\underline{y}) \exp\left(-\frac{|\underline{y}-\underline{x}|^2}{2(t-s)}\right) \tag{55}$$

Theorem (**??**) implies that:

$$E\left[f(B_t)|\mathcal{F}_s\right] = E\left[f(B_s + B_t - B_s)|\mathcal{F}_s\right] = E\left[\psi(B_s,\cdot)|\mathcal{F}_s\right] = \Phi(B_s) \tag{56}$$

(57)

and:

$$E\left[f(B_t)|\mathcal{F}_s\right] = \frac{1}{\left[2\pi(t-s)\right]^{d/2}} \int_{\mathbb{R}^d} d\underline{y}\, f(\underline{y}) \exp\left(-\frac{|\underline{y}-B_s|^2}{2(t-s)}\right) \tag{58}$$

This result is extremely important: the observation of the system up to time $s$ corresponds to the knowledge of $B_s$, a circumstance already reported in Markov chains. In particular, if $f = 1_A$ for some $A \in \mathcal{B}(\mathbb{R}^d)$, the above expression becomes:

$$P\left(B_t \in A|\mathcal{F}_s\right) = p(A, t \,|\, B_s, s) \tag{59}$$

(59) is referred to as **Markov property** of the brownian motion. The importance of this property will be discussed in the following.

## III. I PROCESSI DI MARKOV

Moving from the discussion of the previous paragraph, we will now abstract the general definition of Markov process, which extends to continuous time and state space the notion of Markov chain, introduced in the previous chapters. The discussion will begin with the following:

**Definizione 7** *A* **markovian transition function** *on a measurable space $(E, \mathcal{E})$ is a real-valued function $p\,(A, t \,|\, x, s)$, where $s, t \in \mathbb{R}^+$, $s \leq t$, $x \in E$ and $A \in \mathcal{E}$ is such that:*

1. *for fixed $s, t, A$ the function $x \mapsto p\,(A, t \,|\, x, s)$ is $\mathcal{E}$-measurable*

2. *for fixed $s, t, x$ the function $A \mapsto p(A, t \,|\, x, s)$ is a probability measure on $(E, \mathcal{E})$*

3. *$p$ satisfies the following* **Chapman-Kolmogorov equation***:*

$$p(A, t \,|\, x, s) = \int_E p(A, t \,|\, y, u)\, p(dy, u \,|\, x, s) \tag{60}$$

*for all $s < u < t$, $x \in E$ and $A \in \mathcal{E}$;*

4. *if $s = t$, $p(A, s \,|\, x, s) = \delta_x(A)$ for all $x \in E$.*

The reader is invited to observe that $p$ generalizes the $n$-step transition matrices of Markov chains.

**Definizione 8** *Let $(E, \mathcal{E})$ be a measurable space. Given a markovian transition function $p$ on $(E, \mathcal{E})$ and a probability law $\mu$, we call* **Markov process** *associated to $p$, with starting point $u$ and initial law $\mu$, a process $X = \left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in T}, \{X_t\}_{t\in T}, P\right)$, with time domain $T = [u, +\infty)$ and state space $(E, \mathcal{E})$, such that:*

1. *$X_u$ has law $\mu$;*

2. *the following* **Markov property** *holds:*

$$P(X_t \in A | \mathcal{F}_s) = p(A, t \,|\, X_s, s) \tag{61}$$

*almost surely for all $A \in \mathcal{E}$ and $t > s \geq u$.*

A deep and surprising relationship between theory of stochastic processes and theory of partial differential equation exists. In the remainder of the present chapter a first discussion providing an insight into this topic will be raised, and completed once the formalism of stochastic differential equations will have been introduced.

### A. Semigroup Associated to a Markovian Transition Function

Let $p$ be a markovian transition function on $\left(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)\right)$, which for the moment will be assumed **time-homogeneous**, i.e. depending only on $t - s$. Denoting through $M_b(\mathbb{R}^d)$ the set of real-valued and bounded Borel functions on $\mathbb{R}^d$ and defining the family of operators $\{T_t\}_{t\geq 0}$, $T_t : M_b(\mathbb{R}^d) \to M_b(\mathbb{R}^d)$ through the position:

$$(T_t f)(\underline{x}) \stackrel{def}{=} \int_{\mathbb{R}^d} d\underline{y}\, f(\underline{y})\, p(d\underline{y}, t \,|\, \underline{x}, 0) \tag{62}$$

we immediately find that $T_0$ is the identity, and that:

$$T_s \circ T_t = T_{s+t} \tag{63}$$

To prove (63), let us compute:

$$\begin{gathered}((T_s \circ T_t) f)\,(\underline{x}) = \int_{\mathbb{R}^d} d\underline{y}\,\left(T_t f\right)(\underline{y})\,p(d\underline{y}, s \,|\, \underline{x}, 0) = \\ = \int_{\mathbb{R}^d} d\underline{y}\,\int_{\mathbb{R}^d} d\underline{z}\, f(\underline{z}) p(d\underline{z}, t \,|\, \underline{y}, 0) p(d\underline{y}, s \,|\, \underline{x}, 0) = \\ = \int_{\mathbb{R}^d} d\underline{z}\, f(\underline{z})\,\int_{\mathbb{R}^d} d\underline{y} p(d\underline{z}, s+t \,|\, \underline{y}, s)\, p(d\underline{y}, s \,|\, \underline{x}, 0) = \\ = \int_{\mathbb{R}^d} d\underline{z}\, f(\underline{z})\, p(d\underline{z}, s+t \,|\, \underline{x}, 0) = \left(T_{s+t} f\right)(\underline{x})\end{gathered} \tag{64}$$

where the homogeneity of the markovian transition function and the Chapman-Kolmogorov equation have been recalled. Hence $\{T_t\}_{t\geq 0}$ is a **semigroup** of linear operators on $M_b(\mathbb{R}^d)$. Let us now define $\mathcal{D}(A)$ to be the set of functions $f \in M_b(\mathbb{R}^d)$ such that for all $\underline{x} \in \mathbb{R}^d$, the following limit exists:

$$\lim_{t\to 0^+} \frac{1}{t}\left[(T_t f)(\underline{x}) - f(\underline{x})\right] \tag{65}$$

and define:

$$(\mathcal{L} f)(\underline{x}) \stackrel{def}{=} \lim_{t\to 0^+} \frac{1}{t}\left[(T_t f)(\underline{x}) - f(\underline{x})\right] \tag{66}$$

The so-defined operator $\mathcal{L}$ is called **infinitesimal generator** of the semigroup $\{T_t\}_{t\geq 0}$.

If the markovian transition function is **not time homogeneous**, the same reasoning leads to a family of operators $\{T_{s,t}\}_{s\leq t}$, defined through:

$$\left(T_{s,t} f\right)(\underline{x}) \stackrel{def}{=} \int_{\mathbb{R}^d} d\underline{y}\, f(\underline{y})\, p(d\underline{y}, t \,|\, \underline{x}, s) \tag{67}$$

$T_{s,s}$ is the identity, and:

$$T_{s,u} \circ T_{u,t} = T_{s,t}, \quad s \leq u \leq t \tag{68}$$

Instead of the operator $\mathcal{L}$ a family of infinitesimal generators $\{\mathcal{L}_t\}_t$ appears, defined through the expression:

$$(\mathcal{L}_t f)(\underline{x}) \stackrel{def}{=} \lim_{h\to 0^+} \frac{1}{h}\left[(T_{t,t+h} f)(\underline{x}) - f(\underline{x})\right] \tag{69}$$

whenever it makes sense. In the proceeding we will look for conditions ensuring the existence of such limits, and we will begin writing:

$$\frac{1}{h}\left[(T_{t,t+h} f)(\underline{x}) - f(\underline{x})\right] = \frac{1}{h}\int_{\mathbb{R}^d} d\underline{y}\,(f(\underline{y}) - f(\underline{x}))\, p(d\underline{y}, t+h \,|\, \underline{x}, t) \tag{70}$$

We now assume that for all $\underline{x}$ and $R > 0$, denoting through $B_R(\underline{x})$ the spherical region of center $\underline{x}$ and radius $R$, the following limit holds:

$$\lim_{h\to 0^+} \frac{p(B_R(\underline{x})^c, t+h \,|\underline{x}, t)}{h} = 0 \tag{71}$$

which corresponds to asking that *the probability of going out a sphere of arbitrary radius* starting from its center in a time interval of width $h$ be an infinitesimal of order higher than $h$ for $h \to 0^+$. Under the assumption (71):

$$\lim_{h\to 0^+} \frac{1}{h}\int_{B_R(\underline{x})^c} d\underline{y}\,(f(\underline{y}) - f(\underline{x}))\, p(d\underline{y}, t+h \,|\, \underline{x}, t) = 0 \tag{72}$$

since $f$ is limited. Consider now $f \in M_b(\mathbb{R}^d) \cap C^2(\mathbb{R}^d)$, and perform a Taylor expansion:

$$f(\underline{y})-f(\underline{x}) = \sum_\alpha \frac{\partial f}{\partial x_\alpha}(\underline{x})(y_\alpha - x_\alpha) + \frac{1}{2}\sum_{\alpha,\beta} \frac{\partial^2 f}{\partial x_\alpha \partial x_\beta}(\underline{x})(y_\alpha - x_\alpha)(y_\beta - x_\beta) + o(|\underline{y}-\underline{x}|^2) \tag{73}$$

Assuming that *for all R e t the following limits exist*:

$$\lim_{h\to 0^+} \frac{1}{h}\int_{B_R(\underline{x})} d\underline{y}\,(y_\alpha - x_\alpha)\, p(d\underline{y}, t+h \,|\, \underline{x}, t) \stackrel{def}{=} b_\alpha(\underline{x}, t) \tag{74}$$

$$\lim_{h\to 0^+} \frac{1}{h}\int_{B_R(\underline{x})} d\underline{y}\,(y_\alpha - x_\alpha)(y_\beta - x_\beta)\, p(d\underline{y}, t+h \,|\, \underline{x}, t) \stackrel{def}{=} a_{\alpha\beta}(\underline{x}, t) \tag{75}$$

Observing that $o(|\underline{y}-\underline{x}|^2) \leq \frac{o(R^2)}{R^2}|\underline{y}-\underline{x}|^2$ in $B_R(\underline{x})$, we find the following inequality:

$$\begin{aligned} &\lim_{h\to 0^+} \tfrac{1}{h}\textstyle\int_{B_R(\underline{x})} d\underline{y}\, o(|\underline{y}-\underline{x}|^2)\, p(d\underline{y}, t+h \,|\, \underline{x}, t) \leq \\ &\leq \tfrac{o(R^2)}{R^2} \lim_{h\to 0^+} \tfrac{1}{h}\textstyle\int_{B_R(\underline{x})} d\underline{y}\, |\underline{y}-\underline{x}|^2\, p(d\underline{y}, t+h \,|\, \underline{x}, t) = \tfrac{o(R^2)}{R^2}\sum_\alpha a_{\alpha\alpha}(\underline{x}, t) \end{aligned} \tag{76}$$

that must hold for all $R$. If we now make $R$ tend to 0, we obtain the following interesting expression:

$$(\mathcal{L}_t f)(\underline{x}) = \frac{1}{2}\sum_{\alpha,\beta=1}^d a_{\alpha\beta}(\underline{x}, t)\frac{\partial^2 f}{\partial x_\alpha \partial x_\beta}(\underline{x}) + \sum_\alpha b_\alpha(\underline{x}, t)\frac{\partial f}{\partial x_\alpha}(\underline{x}) \tag{77}$$

Let us now show that the matrix $\{a_{\alpha\beta}(\underline{x}, t)\}_{\alpha\beta}$ is positive-semidefinite at each point and instant. Given an arbitrary point $\underline{x}_0$, fix $\theta \in \mathbb{R}^d$ and choose a function $f$ such that $f(\underline{x}_0) = 0$ and, in the spherical neighborhood $B_{R_0}(\underline{x}_0)$ of $\underline{x}_0$ it is possible to write:

$$f(\underline{x}) = -\left(\sum_{\gamma=1}^d \theta_\gamma (x_\gamma - x_{0,\gamma})\right)^2 \tag{78}$$

Then it is immediate to evaluate:

$$(\mathcal{L}_t f)(\underline{x}_0) = -\sum_{\alpha,\beta=1}^d \theta_\alpha a_{\alpha\beta}(\underline{x}_0, t)\theta_\beta = \lim_{h\to 0^+}\frac{1}{h}\left[(T_{t,t+h}f)(\underline{x}_0) - f(\underline{x}_0)\right] \tag{79}$$

Since $f(\underline{x}_0) = 0$ and furthermore $f(\underline{x}) \leq 0$ for all $\underline{x} \in B_{R_0}(\underline{x}_0)$:

$$\begin{aligned} &-\textstyle\sum_{\alpha,\beta=1}^d \theta_\alpha a_{\alpha\beta}(\underline{x}_0, t)\theta_\beta = \lim_{h\to 0^+}\tfrac{1}{h}\left[(T_{t,t+h}f)(\underline{x}_0)\right] = \\ &= \lim_{h\to 0^+}\tfrac{1}{h}\textstyle\int_{B_{R_0}(\underline{x}_0)} d\underline{y}\,(f(\underline{y})\, p(d\underline{y}, t+h \,|\, \underline{x}_0, t) \leq 0 \end{aligned} \tag{80}$$

that is, the matrix $\{a_{\alpha\beta}(\underline{x}, t)\}_{\alpha\beta}$ is positive semidefinite at each point and instant.

## IV. THE HEAT EQUATION

As shown in the previous paragraph, under suitable condition, to each Markov process, or equivalently to each markovian transition function, a differential operator can be associated. In the homogeneous case, which we consider only for the sake of simplicity, we know that the function $u(\underline{x}, t) = T_t f(\underline{x})$ satisfies:

$$\frac{\partial u}{\partial t}(\underline{x}, 0) = \lim_{h \to 0} \frac{1}{h}(T_h f(\underline{x}) - f(\underline{x})) = \mathcal{L} f(\underline{x}) \tag{81}$$

and that:

$$\frac{\partial u}{\partial t}(\underline{x}, t) = \lim_{h \to 0} \frac{1}{h}(T_{t+h} f(\underline{x}) - T_t f(\underline{x})) = T_t \mathcal{L} f(\underline{x}) \tag{82}$$

If it happens that $T_t$ and $\mathcal{L}$ commute, the following **heat equation** associated to the markovian transition function is found:

$$\begin{cases} \frac{\partial u}{\partial t}(\underline{x}, t) = (\mathcal{L} u)(\underline{x}, t) \\ u(\underline{x}, 0) = f(\underline{x}) \end{cases} \tag{83}$$

On the other hand, let $q(t, \underline{x}, \underline{y})$ be the *fundamental solution* of the heat equation, i.e. the function satisfying:

$$\begin{cases} \frac{\partial q}{\partial t}(t, \underline{x}, \underline{y}) = (\mathcal{L} q)(t, \underline{x}, \underline{y}) \\ q(0, \underline{x}, \underline{y}) = \delta(\underline{x} - \underline{y}) \end{cases} \tag{84}$$

then, from the theory of partial differential equations, it is known that:

$$u(\underline{x}, t) = \int_{\mathbb{R}^d} d\underline{y}\, q(t, \underline{x}, \underline{y}) f(\underline{y}) \tag{85}$$

But since $u(\underline{x}, t) = T_t f(\underline{x})$, the markovian transition function must be:

$$p(d\underline{y}, s + t \,|\underline{x}, s) = q(t, \underline{x}, \underline{y})\, d\underline{y} \tag{86}$$

This observation sheds light on the possibility of *inverting* this process: given a partial differential equation with fundamental solution $q$, we might ask ourselves whether it is possible to construct a Markov process having a markovian transition function $p$ related to $q$ by (86). As pointed out before, for this issue to be faced adequately, the formalism of stochastic differential equations is required.