

Markov Chains and Dynamical Systems: The Open System Point of View*

Stéphane ATTAL

Université de Lyon
Université de Lyon 1, C.N.R.S.
Institut Camille Jordan
21 av Claude Bernard
69622 Villeurbanne cedex, France

Abstract

This article presents several results establishing connections between Markov chains and dynamical systems, from the point of view of open systems in physics. We show how all Markov chains can be understood as the information on one component that we get from a dynamical system on a product system, when losing information on the other component. We show that passing from the deterministic dynamics to the random one is characterized by the loss of algebra morphism property; it is also characterized by the loss of reversibility. In the continuous time framework, we show that the solutions of stochastic differential equations are actually deterministic dynamical systems on a particular product space. When losing the information on one component, we recover the usual associated Markov semigroup.

1 Introduction

This article aims at exploring the theory of Markov chains and Markov processes from a particular point of view. This point of view is very physical and commonly used in the theory of *open systems*. Open systems are physical systems, in classical or in quantum mechanics, which are not closed, that is, which are interacting with another system. In general the system we are interested in is “small” (for example, it has only a finite number of degrees of freedom), whereas the outside system is very large (often called the “environment”, it may be a heat bath typically).

This is now a very active branch of research to study such systems coupled to an environment. In classical mechanics they are used to study conduction problems (Fourier’s law for example, see [2], [5]) but more generally out of equilibrium dynamics (see [6], [3]). In quantum mechanics, open systems appear fundamentally for the study of decoherence phenomena (see [7]), but also it is the basis of quantum communication (see [8]). Problems

*Work supported by ANR project “HAM-MARK” N° ANR-09-BLAN-0098-01

of dissipation, heat conduction, out of equilibrium dynamics in quantum mechanics (see [9], [10]) lead to very important problems which are mostly non-understood at the time we write this article.

The aim of this article is to make clear several ideas and connections between deterministic dynamics of closed systems, effective dynamics of open systems and Markov processes.

Surprisingly enough these ideas are made rather clear in the literature when dealing with the quantum systems, but not that much with classical ones! Indeed, it is common in quantum mechanics to consider a bipartite system on which one component is not accessible (it might be an environment which is too complicated to be described, or it might be Bob who is sharing the photons of a correlated pair with Alice, in Quantum Information Theory, ...). It is well-known that, tracing out over one component of the system, the unitary Schrödinger dynamics becomes described by completely positive maps, in discrete time, or completely positive semigroups, in continuous time.

In [1], for example, the authors show how every discrete time semigroup of completely positive maps can be described by a realistic physical system, called “Repeated Quantum Interactions”. They show that in the continuous time limit these Hamiltonian dynamics spontaneously converge to a dynamics described by a quantum Langevin equation.

In this article we establish many similar results in the context of classical dynamical systems and Markov chains. The article is structured as follows. In the Section 2, we show that Markov chains appear from any dynamical system on a product space, when averaging out one of the two components. This way, Markov chains are interpreted as what remains on one system when it interacts with some environment but we do not have access to that environment. The randomness appears directly for the determinism, by the only fact that we have lost some information. We show that any Markov chain can be obtained this way. We also show two results which characterize what properties are lost when going from a deterministic dynamical system to a Markov chain: typically the loss of algebra morphism property and the loss of reversibility.

In Section 3 we explore the context of classical Markov process in the continuous time setup. We actually concentrate on stochastic differential equations. Despite of their “random character”, we show that stochastic differential equations are actually deterministic dynamical systems. They correspond to a natural dynamical system which is used to dilate some Markov processes into a deterministic dynamics. The role of the environment is played by the canonical probability space (here the Wiener space), the action of the environment is the noise term in the stochastic differential equation.

2 Markov Chains and Dynamical Systems

2.1 Basic Definitions

Let us recall some basic definitions concerning dynamical systems and Markov chains.

Let (E, \mathcal{E}) be a measurable space. Let \tilde{T} be a measurable application from E to E . We then say that \tilde{T} is a *dynamical system* on E . Such a mapping \tilde{T} induces a natural mapping T on $\mathcal{L}^\infty(E, \mathcal{E})$ defined by

$$Tf(x) = f(\tilde{T}x).$$

Note that this mapping clearly satisfies the following properties (left to the reader).

Proposition 2.1

- i) T is a $*$ -homomorphism of the $*$ -algebra $\mathcal{L}^\infty(E, \mathcal{E})$,
- ii) $T(\mathbb{1}_E) = \mathbb{1}_E$,
- iii) $\|T\| = 1$.

The *dynamical system* is actually the associated discrete-time semigroup $(\tilde{T}^n)_{n \in \mathbb{N}}$, when acting on points, or $(T^n)_{n \in \mathbb{N}}$, when acting on functions.

When the mapping \tilde{T} is invertible, then so is the associated operator T . The semigroups $(\tilde{T}^n)_{n \in \mathbb{N}}$ and $(T^n)_{n \in \mathbb{N}}$ can then be easily extended into one-parameter groups $(\tilde{T}^n)_{n \in \mathbb{Z}}$ and $(T^n)_{n \in \mathbb{Z}}$, respectively.

Let us now recall basic definitions concerning Markov chains. Let (E, \mathcal{E}) be a measurable space. A mapping ν from $E \times \mathcal{E}$ to $[0, 1]$ is a *Markov kernel* if

- i) $x \mapsto \nu(x, A)$ is a measurable function, for all $A \in \mathcal{E}$,
- ii) $A \mapsto \nu(x, A)$ is a probability measure, for all $x \in E$.

When E is a finite set, then ν is determined by the quantities

$$P(i, j) = \nu(i, \{j\})$$

which form a *stochastic matrix*, i.e. a square matrix with positive entries and sum of each line being equal to 1.

In any case, such a Markov kernel ν acts on $\mathcal{L}^\infty(E, \mathcal{E})$ as follows

$$\nu \circ f(x) = \int_E f(y) \nu(x, dy).$$

A linear operator T on $\mathcal{L}^\infty(E, \mathcal{E})$ which is of the form

$$Tf(x) = \int_E f(y) \nu(x, dy),$$

for some Markov kernel ν , is called a *Markov operator*.

In a dual way, a Markov kernel ν acts on probability measures on (E, \mathcal{E}) . Indeed, if \mathbb{P} is a probability measure on (E, \mathcal{E}) then so is the measure $\mathbb{P} \circ \nu$ defined by

$$\mathbb{P} \circ \nu(A) = \int_E \nu(x, A) \mathbb{P}(dx).$$

Finally, Markov kernels can be composed. If ν_1 and ν_2 are two Markov kernels on (E, \mathcal{E}) then so is

$$\nu_1 \circ \nu_2(x, A) = \int_E \nu_2(y, A) \nu_1(x, dy).$$

This kernel represents the Markov kernel resulting from making one first step following ν_1 and then another step following ν_2 .

A *Markov chain* is a discrete-time stochastic process $(X_n)_{n \in \mathbb{N}}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that

$$\mathbb{E}[f(X_{n+1}) | X_0, X_1, \dots, X_n] = \mathbb{E}[f(X_{n+1}) | X_n]$$

for all bounded function f , all $n \in \mathbb{N}$. In particular, if \mathcal{F}_n denotes the σ -algebra generated by X_0, X_1, \dots, X_n , then the above implies

$$\mathbb{E}[f(X_{n+1}) | \mathcal{F}_n] = L_n f(X_n)$$

for some function $L_n f$. The Markov chain is *homogeneous* if furthermore L_n does not depend on n . We shall be interested only in this case and we denote by L this unique value of L_n :

$$\mathbb{E}[f(X_{n+1}) | \mathcal{F}_n] = Lf(X_n) = L^n f(X_0). \quad (1)$$

If $\nu_n(x, dy)$ denotes the conditional law of X_{n+1} knowing $X_n = x$, which coincides with the conditional law of X_1 knowing $X_0 = x$, then ν is a Markov kernel and one can easily see that

$$Lf(x) = \int_E f(y) \nu(x, dy).$$

Hence L is the Markov operator associated to ν .

With our probabilistic interpretation we get easily that $\nu \circ f(x)$ is the expectation of $f(X_1)$ when $X_0 = x$ almost surely. The measure $\mathbb{P} \circ \nu$ is the distribution of X_1 if the distribution of X_0 is \mathbb{P} .

We end up this section with the following last definition. A Markov kernel ν is said to *deterministic* if for all $x \in E$ the measure $\nu(x, \cdot)$ is a Dirac mass. This is to say that there exists a measurable mapping $\tilde{T} : E \rightarrow E$ such that

$$\nu(x, dy) = \delta_{\tilde{T}(x)}(dy).$$

In other words, the Markov chain associated to ν is not random at all, it maps with probability 1, each point x to $\tilde{T}(x)$: it is a dynamical system.

2.2 Reduction of Dynamical Systems

Now consider two measurable spaces (E, \mathcal{E}) and (F, \mathcal{F}) , together with a dynamical system \tilde{T} on $E \times F$. As above, consider the lifted application T acting on $\mathcal{L}^\infty(E \times F)$.

For all bounded function f on E , we consider the bounded function $f \otimes \mathbb{1}$ on $E \times F$ defined by

$$(f \otimes \mathbb{1})(x, y) = f(x),$$

for all $x \in E, y \in F$.

Assume that (F, \mathcal{F}) is equipped with a probability measure μ . We shall be interested into the mapping L of $\mathcal{L}^\infty(E)$ defined by

$$Lf(x) = \int_F T(f \otimes \mathbb{1})(x, y) d\mu(y). \quad (2)$$

In other words, we have a deterministic dynamical system on a product space. We place ourselves from one component point of view only (we have access to E only). Starting from a point $x \in E$ and a function f on E we want to see how they evolve according to T , but seen from the E point of view. The function f on E is naturally lifted into a function $f \otimes \mathbb{1}$ on $E \times F$, that is, it still acts on E only, but it is now part of a “larger world”. We make f evolve according to the deterministic dynamical system T . Finally, in order to come back to E we project the result onto E , by taking the average on F according to a fixed measure μ on F . This is to say that, from the set E , what we see of the action of the “environment” F is just an average with respect to some measure μ .

Theorem 2.2 *The mapping L is a Markov operator on E .*

Proof: As \tilde{T} is a mapping from $E \times F$ to $E \times F$, there exist two measurable mappings:

$$X : E \times F \longrightarrow E \quad \text{and} \quad Y : E \times F \longrightarrow F,$$

such that

$$\tilde{T}(x, y) = (X(x, y), Y(x, y))$$

for all $(x, y) \in E \times F$.

Let us compute the quantity $Lf(x)$, with these notations. We have

$$\begin{aligned} Lf(x) &= \int_F T(f \otimes \mathbb{1})(x, y) d\mu(y) \\ &= \int_F (f \otimes \mathbb{1})(X(x, y), Y(x, y)) d\mu(y) \\ &= \int_F f(X(x, y)) d\mu(y). \end{aligned}$$

Denote by $\nu(x, dz)$ the (probability) measure on E , image of μ by the application $X(x, \cdot)$ (which goes from F to E , for each fixed x). By the Transfer Theorem we get

$$Lf(x) = \int_E f(z) \nu(x, dz).$$

Hence L acts on $\mathcal{L}^\infty(E)$ as the Markov transition kernel $\nu(x, dz)$. \square

Note the following important fact: the mapping Y played no role at all in the proof above.

Note that the Markov kernel ν is given by

$$\nu(x, A) = \mu(\{y \in F; X(x, y) \in A\}) .$$

In particular, when E is finite (or even countable), the transition kernel ν is associated to a Markovian matrix P whose coefficients are given by

$$P(i, j) = \nu(i, \{j\}) = \mu(\{k; X(i, k) = j\}) .$$

What we have obtained here is important and deserves more explanations. Mathematically, we have obtained a commuting diagram:

$$\begin{array}{ccc} & T & \\ \mathcal{L}^\infty(E \times F) & \longrightarrow & \mathcal{L}^\infty(E \times F) \\ & \otimes \mathbb{1} \uparrow & \downarrow \mu \\ \mathcal{L}^\infty(E) & \xrightarrow{L} & \mathcal{L}^\infty(E) . \end{array}$$

In more physical words, what we have obtained here can be interpreted in two different ways. If we think of the dynamical system \tilde{T} first, we have emphasized on the fact that loosing the information of a deterministic dynamics on one the components creates a random behavior on the other component. The randomness here appears only as a lack of knowledge on a deterministic behavior on a larger world. A part of the universe interacting with our system E is inaccessible to us (or at least we see a very small part of it: an average), then it results a random behavior on E .

In the converse direction, that is, seen from the Markov kernel point of view, what we have obtained is a *dilation* of a Markov transition kernel into a dynamical system. Consider the kernel L on the state space E . It does not represent the dynamics of a closed system, it is not a dynamical system. In order to see L as coming from a true dynamical system, we have enlarged the state space E with an additional state space F , which represents the environment. The dynamical system \tilde{T} represents the true dynamics of the closed system “ E +environment”. Equation (2) exactly says that the effective pseudo-dynamics L that we had observed on E is simply due to the fact that we are looking only at a subpart of a true dynamical system and an average of the F part of the dynamics.

These observations would be even more interesting if one could prove the converse: every Markov transition kernel can be obtained this way. This is what we prove now, with some small restriction on E .

We now assume that E is a Lusin space, that is, it is homeomorphic (as a measurable space) to a compact subset of a Polish space, equipped with its Borel σ -field. This condition is satisfied for example by all the spaces \mathbb{R}^n .

Let $\nu(x, dz)$ be a Markov kernel on (E, \mathcal{E}) . Let F be the set of applications from E to E . For every finite subset $\sigma = \{x_1, \dots, x_n\} \subset E$, every $A_1, \dots, A_n \in \mathcal{E}$ consider the set

$$F(x_1, \dots, x_n; A_1, \dots, A_n) = \{y \in F; y(x_1) \in A_1, \dots, y(x_n) \in A_n\}.$$

By Kolmogorov Consistency Theorem (it is here that the hypothesis on E is needed!) there exists a unique probability measure μ on F such that

$$\mu(F(x_1, \dots, x_n; A_1, \dots, A_n)) = \prod_{i=1}^n \nu(x_i, A_i).$$

Indeed, it is easy to check that the above formula defines a consistent family of probability measures on the finitely-based cylinders of F , then apply Kolmogorov's Theorem.

Theorem 2.3 *With the same notations as above, define the dynamical system*

$$\begin{aligned} \tilde{T} : E \times F &\longrightarrow E \times F \\ (x, y) &\longmapsto (y(x), y). \end{aligned}$$

Then the Markovian kernel associated to \tilde{T} restricted to E is equal to ν .

Proof: With the same notations as in Theorem 2.2, we have here $X(x, y) = y(x)$ and hence

$$\mu(\{y \in F; X(x, y) \in A\}) = \mu(\{y \in F; y(x) \in A\}) = \nu(x, A).$$

□

Note that in this dilation of L , the dynamical system T has no reason to be invertible in general. It is worth noticing that one can construct a dilation where T is invertible.

Proposition 2.4 *Every Markov kernel ν admits a dilation \tilde{T} which is an invertible dynamical system.*

Proof: Consider the construction and notations of Theorem 2.3. Consider the space $F' = E \times F$. Let x_0 be a fixed element of E , define the mapping \tilde{T}' on $E \times F'$ by

$$\begin{cases} \tilde{T}'(x, (x_0, y)) = (y(x), (x, y)), \\ \tilde{T}'(x, (y(x), y)) = (x_0, (x, y)), \\ \tilde{T}'(x, (z, y)) = (z, (x, y)), & \text{if } z \neq x_0, z \neq y(x). \end{cases}$$

It is easy to check that \tilde{T}' is a bijection of $E \times F'$. Now extend the measure μ on F to the measure $\delta_{x_0} \otimes \mu$ on F' . Now, the dynamical system \tilde{T}' is invertible and dilates the same Markov kernel as \tilde{T} . □

2.3 Iterating the Dynamical System

We have shown that every dynamical system on a product set gives rise to a Markov kernel when restricted to one of the sets. We have seen that every Markov kernel can be obtained this way. But one has to notice that our construction allows the dynamical system \tilde{T} to dilate the Markov kernel L as a single mapping only. That is, iterations of the dynamical system T^n do not in general dilate the semigroup L^n associated to the Markov process. Let us check this with a simple counter-example.

Put $E = F = \{1, 2\}$. On F define the probability measure $P(1) = 1/4$ and $P(2) = 3/4$. Define the dynamical system \tilde{T} on $E \times F$ which is the ‘‘anticlockwise rotation’’:

$$\tilde{T}(1, 1) = (2, 1), \quad \tilde{T}(2, 1) = (2, 2), \quad \tilde{T}(2, 2) = (1, 2), \quad \tilde{T}(1, 2) = (1, 1).$$

With the same notations as in previous section, we have

$$X(1, 1) = 2, \quad X(2, 1) = 2, \quad X(2, 2) = 1, \quad X(1, 2) = 1.$$

Hence, we get

$$\begin{aligned} \mu(X(1, \cdot) = 1) &= \frac{3}{4}, & \mu(X(1, \cdot) = 2) &= \frac{1}{4}, \\ \mu(X(2, \cdot) = 1) &= \frac{3}{4}, & \mu(X(2, \cdot) = 2) &= \frac{1}{4}. \end{aligned}$$

Hence the Markovian matrix associated to the restriction of \tilde{T} to E is

$$L = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix}.$$

In particular

$$L^2 = L.$$

Let us compute \tilde{T}^2 . We get

$$\tilde{T}^2(1, 1) = (2, 2), \quad \tilde{T}^2(2, 1) = (1, 2), \quad \tilde{T}^2(2, 2) = (1, 1), \quad \tilde{T}^2(1, 2) = (2, 1).$$

Hence the associated X -mapping, which we shall denote by X_2 , is given by

$$X_2(1, 1) = 2, \quad X_2(2, 1) = 1, \quad X_2(2, 2) = 1, \quad X_2(1, 2) = 2.$$

This gives the Markovian matrix

$$L_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Which is clearly not equal to L^2 .

It would be very interesting if one could find a dilation of the Markov kernel L by a dynamical system \tilde{T} such that any power T^n would also dilate L^n . We would have realized the whole Markov chain as a restriction of iterations of a single dynamical system on a larger space.

This can be performed in the following way (note that this is not the only way, nor the more economical). Consider the set $\widehat{F} = F^{\mathbb{N}^*}$, equipped with the usual cylinder σ -field $\mathcal{F}^{\otimes \mathbb{N}^*}$ and the product measure $\widehat{\mu} = \mu^{\otimes \mathbb{N}^*}$. The elements of \widehat{F} are sequences $(y_n)_{n \in \mathbb{N}^*}$ in F . Let T be a dynamical system on $E \otimes F$ which dilates the Markovian kernel L . Put

$$\begin{aligned} \tilde{S} : E \times \widehat{F} &\longrightarrow E \times \widehat{F} \\ (x, y) &\longmapsto (X(x, y_1), \Theta(y)) \end{aligned}$$

where Θ is the usual *shift* on \widehat{F} : $\Theta(y) = (y_{n+1})_{n \in \mathbb{N}^*}$.

Then \tilde{S} can be lifted into a morphism S of $\mathcal{L}^\infty(E \times \widehat{F})$, as previously. Furthermore, any function f in $\mathcal{L}^\infty(E)$ can be lifted into $f \otimes \mathbb{1}$ on $\mathcal{L}^\infty(E \times \widehat{F})$, with $(f \otimes \mathbb{1})(x, y) = f(x)$.

Theorem 2.5 *For all $n \in \mathbb{N}^*$, all $x \in E$, all $f \in \mathcal{L}^\infty(E)$ we have*

$$\int_{\widehat{F}} S^n(f \otimes \mathbb{1})(x, y) d\widehat{\mu}(y) = (L^n f)(x).$$

Proof: Recall that we noticed in the proof of Theorem 2.2, that the mapping Y associated to \tilde{T} played no role in the proof of this theorem, only the mapping X was of importance. In particular this implies that Theorem 2.5 is true for $n = 1$, for the dynamical systems \tilde{T} and \tilde{S} share the same X -mapping.

By induction, let us assume that the relation

$$\int_{\widehat{F}} S^k(f \otimes \mathbb{1})(x, y) d\widehat{\mu}(y) = (L^k f)(x)$$

holds true for all $f \in \mathcal{L}^\infty(E)$, all $x \in E$, all $k \leq n$. Put $\widehat{F}_{[2]}$ to be the set of sequences $(y_n)_{n \geq 2}$ in \widehat{F} and $\widehat{\mu}_{[2]}$ the restriction of $\widehat{\mu}$ to $\widehat{F}_{[2]}$. We have

$$\begin{aligned} \int_{\widehat{F}} S^{n+1}(f \otimes \mathbb{1})(x, y) d\widehat{\mu}(y) &= \\ &= \int_{\widehat{F}} S^n(f \otimes \mathbb{1})(X(x, y_1), \Theta(y)) d\widehat{\mu}(y) \\ &= \int_F \int_{\widehat{F}_{[2]}} S^n(f \otimes \mathbb{1})(X(x, y_1), y) d\widehat{\mu}_{[2]}(y) d\mu(y_1). \end{aligned}$$

Put $\tilde{x} = X(x, y_1)$, the above is equal to

$$\begin{aligned}
& \int_F \int_{\tilde{F}|_2} S^n(f \otimes \mathbb{1})(\tilde{x}, y) d\hat{\mu}|_2(y) d\mu(y_1) \\
&= \int_F L^n(f)(\tilde{x}) d\mu(y_1) \quad (\text{by induction hypothesis}) \\
&= \int_F L^n(f)(X(x, y_1)) d\mu(y_1) \\
&= L^{n+1}(f)(x).
\end{aligned}$$

□

With this theorem and with Theorem 2.3, we see that every Markov chain on E can be realized as the restriction on a E of the iterations of a deterministic dynamical system \tilde{T} acting on a larger set.

The physical interpretation of the construction above is very interesting. It represents a scheme of “repeated interactions”. That is, we know that the result of the deterministic dynamics associated to \tilde{T} on $E \times F$ gives rise to the Markov operator L on E . The idea of the construction above is that the environment is now made of a chain of copies of F , each of which is going to interact, one after the other, with E . After, the first interaction between E and the first copy of F has happened, following the dynamical system \tilde{T} , the first copy of F stops interacting with E and is replaced by the second copy of F . This copy now interacts with E following \tilde{T} . And so on, we repeat these interactions. The space E keeps the memory of the different interactions, while each copy of F arrives independently in front of E and induces one more step of evolution following \tilde{T} .

As resulting from this procedure, the successive evolutions restricted to E are the one of the iterations of the Markov operator L . This gives rise to announced behavior: a whole path of the homogeneous Markov chain with generator L .

2.4 Defect of Determinism and Loss of Invertibility

We end up this section with some algebraic characterizations of determinism for Markov chains. The point is to characterize what exactly is lost when going from the deterministic dynamics T on $E \times F$ to the Markov operator L on E .

Theorem 2.6 *Let (E, \mathcal{E}) be a Lusin space. Let (X_n) be a Markov chain with state space (E, \mathcal{E}) and with transition kernel ν . Let T be the Markov operator on $\mathcal{L}^\infty(E, \mathcal{E})$ associated to ν :*

$$Tf(x) = \int_E f(y) \nu(x, dy).$$

Then the Markov chain (X_n) is deterministic if and only if T is a $$ -homomorphism of the algebra $\mathcal{L}^\infty(E, \mathcal{E})$.*

Proof: One direction is immediate: if the Markov chain is deterministic, then T is associated to a dynamical system and hence it is a $*$ -homomorphism.

Conversely, suppose that T is a $*$ -homomorphism. We shall first consider the case where (E, \mathcal{E}) is a Polish space together with its Borel σ -field \mathcal{E} .

Take any $A \in \mathcal{E}$, any $x \in E$ and recall that we always have

$$\nu(x, A) = T(\mathbb{1}_A)(x).$$

The homomorphism property gives

$$T(\mathbb{1}_A)(x) = T(\mathbb{1}_A^2)(x) = T(\mathbb{1}_A)^2(x) = \nu(x, A)^2.$$

Hence $\nu(x, A)$ satisfies $\nu(x, A)^2 = \nu(x, A)$. This means that $\nu(x, A)$ is equal to 0 or 1, for all $x \in E$, all $A \in \mathcal{E}$.

Consider a covering of E with a countable family of balls $(B_i)_{i \in \mathbb{N}}$, each of which with diameter smaller than 2^{-n} (this always possible as E is separable). From this covering one can easily extract a partition $(S_i)_{i \in \mathbb{N}}$ of E with measurable sets, each of which with diameter smaller than 2^{-n} . We shall denote by \mathcal{S}_n this partition.

Let $x \in E$ be fixed. As we have $\sum_{E \in \mathcal{S}_n} \nu(x, E) = 1$ we must have $\nu(x, E) = 1$ for one and only one $E \in \mathcal{S}_n$. Let us denote by $E^{(n)}(x)$ this unique set. Clearly, the sequence $(E^{(n)}(x))_{n \in \mathbb{N}}$ is decreasing (for otherwise there will be more than one set $E \in \mathcal{S}_n$ such that $\nu(x, E) = 1$). Let $A = \bigcap_n E^{(n)}(x)$. The set A satisfies $\nu(x, A) = 1$, hence A is non-empty. But also, the diameter of A has to be 0, for it is smaller than 2^{-n} for all n . As a consequence A has to be a singleton $\{y\}$, for some $y(x) \in E$. Hence we have proved that for each $x \in E$ there exists a $y(x) \in E$ such that $n(x, \{y(x)\}) = 1$. This proves the deterministic character of our chain.

The case where E is a subset of a Polish space is obtained in the same way, by restricting the σ -field.

The case where E is only homeomorphic to a subset of a Polish space is obtained by transferring on E the partitions \mathcal{S}_n of the Polish space, via the homeomorphism (a bi-continuous bijection). \square

The result above is quite amazing. It gives such a clear and net characterization of the difference between a true Markov operator and a deterministic one! One can even think of several applications of that characterization, for example one may be able to measure the “level of randomness” of some Markov operator by evaluating for example

$$\sup\{\|T(f^2) - T(f)^2\|; f \in \mathcal{L}^\infty(E, \mathcal{E}), \|f\| = 1\}.$$

I do not know if such things have already been studied or not. It is not my purpose here to develop this idea, I just mention it.

Another strong result on determinism of Markov chains is the way it is related to non-invertibility.

Theorem 2.7 *Let (E, \mathcal{E}) be a Lusin space. Let T be a Markov operator on $\mathcal{L}^\infty(E, \mathcal{E})$ associated to a Markov chain (X_n) . If T is invertible in the category of Markov operators then (X_n) is deterministic.*

Proof: Recall that a Markov operator T maps positive functions to positive functions. Hence, in the same way as one proves Cauchy-Schwarz inequality, we always have

$$\overline{T(f)} = T(\bar{f}).$$

and

$$T(|f|^2) \geq T(\bar{f})T(f)$$

(hint: write the positivity of $T(\overline{(f + \lambda g)}(f + \lambda g))$ for all $\lambda \in \mathbb{C}$).

Let S be a Markov operator such that $ST = TS = I$. We have

$$|f|^2 = \bar{f}f = S \circ T(\bar{f}f) \geq S(T(\bar{f})T(f)) \geq S \circ T(\bar{f})S \circ T(f) = \bar{f}f = |f|^2.$$

Hence we have equalities everywhere above. In particular

$$S \circ T(\bar{f}f) = S(T(\bar{f})T(f)).$$

Applying T to this equality, gives

$$T(\bar{f}f) = T(\bar{f})T(f),$$

for all $f \in \mathcal{L}^\infty(E, \mathcal{E})$.

By polarization it is easy to prove now that T is a homomorphism. By Theorem 2.6 it is the Markov operator associated to a deterministic chain. \square

The result above is more intuitive than the one of Theorem 2.6, from the point of view of open systems. If the dynamical system \tilde{T} on the large space $E \times F$ is invertible, this invertibility is always lost when projecting on E . The fact we do not have access to one component of the coupled system makes that we loose all chance of invertibility.

3 Continuous Time

We now leave the discrete-time setup to concentrate on continuous-time dynamical systems. We aim to show that stochastic differential equations are actually some particular kind of continuous-time dynamical systems. In particular they are “deterministic”. The type of dynamical system we shall obtain this way is a continuous-time version of the construction of Theorem 2.5.

3.1 Preliminaries

Let us consider the d -dimensional Brownian motion W on its canonical space $(\Omega, \mathcal{F}, \mathbb{P})$. This is to say that $\Omega = C_0(\mathbb{R}^+; \mathbb{R}^d)$ is the space of continuous functions on \mathbb{R}^+ with values in \mathbb{R}^d and which vanish at 0, the σ -field \mathcal{F} is the Borel σ -field of Ω , the measure \mathbb{P} is the Wiener measure, that is, the law of a d -dimensional Brownian motion on Ω . The canonical Brownian motion (W_t) is defined by $W_t(\omega) = \omega(t)$, for all $\omega \in \Omega$, all $t \in \mathbb{R}^+$. This is to say, coordinate-wise: $W_t^i(\omega) = \omega_i(t)$, for $i = 1, \dots, d$.

We define for all $s \in \mathbb{R}^+$ the *shift* θ_s as an application from Ω to Ω by

$$\theta_s(\omega)(t) = \omega(t + s) - \omega(s).$$

We extend it as the *shift operator* Θ_s on $L^2(\Omega, \mathcal{F}, \mathbb{P})$ by

$$\Theta_s(F)(\omega) = F(\theta_s(\omega)).$$

In particular we have $\Theta_s(W_t) = W_{t+s} - W_s$.

As the process $Y_t = W_{t+s} - W_s$, $t \in \mathbb{R}^+$, is again a d -dimensional Brownian motion, this implies that the mapping θ_s preserves the measure \mathbb{P} . As a consequence Θ_s is an isometry of $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

Lemma 3.1 *If H is a predictable process in \mathbb{R}^d , then, for all fixed $s \in \mathbb{R}^+$, the process $K_t = \Theta_s(H_{t-s})$, $t \geq s$ is also predictable.*

Proof: The process K as a mapping from $\Omega \times \mathbb{R}^+$ to \mathbb{R}^d is the composition of H with the mapping $\phi(\omega, t) = (\theta_s(\omega), t - s)$ from $\Omega \times [s, +\infty[$ to $\Omega \times \mathbb{R}^+$. We just need to check that ϕ is measurable for the predictable σ -algebra \mathcal{P} .

Consider a basic predictable set $A \times]u, v]$, with $u < v$ and $A \in \mathcal{F}_u$, then

$$\phi^{-1}(A \times]u, v]) = \theta_s^{-1}(A) \times]u + s, v + s].$$

We just need to check that $\theta_s^{-1}(\mathcal{F}_u) \subset \mathcal{F}_{u+s}$. The σ -algebra \mathcal{F}_u is generated by events of the form $(W_t^i \in [a, b])$, for $t \leq u$ and $i = 1, \dots, d$. The set $\theta_s^{-1}(W_t^i \in [a, b])$ is equal to $(W_{t+s}^i - W_s^i \in [a, b])$, hence it belongs to \mathcal{F}_{u+s} . \square

In the following, the norm $\|\cdot\|_2$ is the $L^2((\Omega, \mathcal{F}, \mathbb{P}); \mathbb{R}^d)$ -norm.

Lemma 3.2 *Let H be a predictable process in \mathbb{R}^d such that $\int_0^{t+s} \|H_u\|_2^2 du < \infty$. Then we have*

$$\Theta_s \left(\int_0^t H_u \cdot dW_u \right) = \int_s^{t+s} \Theta_s(H_{u-s}) \cdot dW_u. \quad (3)$$

Proof: If H is an elementary predictable process then the identity (3) is obvious from the fact that $\Theta_s(FG) = \Theta_s(F) \Theta_s(G)$. A general stochastic integral $\int_0^t H_s \cdot dW_s$ is obtained as a limit in the norm

$$\left\| \int_0^t H_s \cdot dW_s \right\|_2^2 = \int_0^t \|H_s\|_2^2 ds,$$

of stochastic integrals of elementary predictable processes. As Θ_s is an isometry, it is clear that Equation (3) holds true for any stochastic integral. \square

Here is now the main result of this section.

Theorem 3.3 *Let W be a d -dimensional Brownian motion on its canonical space $(\Omega, \mathcal{F}, \mathbb{P})$. Let f be a locally bounded, Lipschitz function from \mathbb{R}^n to \mathbb{R}^n and let g be a locally bounded, Lipschitz function from \mathbb{R}^n to $M_{n \times d}(\mathbb{R})$. Denote by X^x the unique stochastic process (in \mathbb{R}^n) which is solution of the stochastic differential equation*

$$X_t^x = x + \int_0^t f(X_u^x) du + \int_0^t g(X_u^x) \cdot dW_u.$$

Then, for all $s, t \in \mathbb{R}^+$, we have, for almost all $\omega \in \Omega$,

$$X_t^{X_s^{X_s^x(\omega)}}(\theta_s(\omega)) = X_{s+t}^x(\omega).$$

Proof: Let s be fixed. Define

$$Y_u^x(\omega) = \begin{cases} X_u^x(\omega) & \text{if } u \leq s, \\ X_{u-s}^{X_s^{X_s^x(\omega)}}(\theta_s(\omega)) & \text{if } u > s. \end{cases}$$

Then Y_{s+t}^x is solution of

$$\begin{aligned}
Y_{s+t}^x(\omega) &= X_t^{X_s^x(\omega)}(\theta_s(\omega)) \\
&= X_s^x(\omega) + \left[\int_0^t f(X_u^{X_s^x}) du \right] (\theta_s(\omega)) + \left[\int_0^t g(X_u^{X_s^x}) \cdot dW_u \right] (\theta_s(\omega)) \\
&= x + \left[\int_0^s f(X_u^x) du \right] (\omega) + \left[\int_0^s g(X_u^x) \cdot dW_u \right] (\omega) + \int_0^t f(X_u^{X_s^x})(\theta_s(\omega)) du + \\
&\quad + \left[\int_s^{s+t} \Theta_s \left(g(X_{u-s}^{X_s^x}) \right) \cdot dW_u \right] (\omega) \quad (\text{by Lemma 3.2}) \\
&= x + \int_0^s f(Y_u^x)(\omega) du + \left[\int_0^s g(Y_u^x) \cdot dW_u \right] (\omega) + \int_0^t f(Y_{u+s}^x)(\omega) du + \\
&\quad + \left[\int_s^{s+t} g(Y_u^x) \cdot dW_u \right] (\omega) \\
&= x + \left[\int_0^s f(Y_u^x) du + \int_0^s g(Y_u^x) \cdot dW_u + \int_s^{s+t} f(Y_u^x) du + \right. \\
&\quad \left. + \int_s^{s+t} g(Y_u^x) \cdot dW_u \right] (\omega) \\
&= x + \left[\int_0^{s+t} f(Y_u^x) du + \int_0^{s+t} g(Y_u^x) \cdot dW_u \right] (\omega).
\end{aligned}$$

This shows that Y^x is solution of the same stochastic differential equation as X^x . We conclude easily by uniqueness of the solution. \square

3.2 Stochastic Differential Equations and Dynamical Systems

We are now ready to establish a parallel between stochastic differential equations and dynamical systems. Recall how we defined discrete time dynamical systems \tilde{T} in Section 2 and their associated semigroup (\tilde{T}^n) . In continuous time the definition extends in the following way.

A *continuous-time dynamical system* on a set E is a one-parameter family of applications $(\tilde{T}_t)_{t \in \mathbb{R}^+}$ on E such that $\tilde{T}_s \circ \tilde{T}_t = \tilde{T}_{s+t}$ for all s, t . That is, \tilde{T} is a semigroup of applications on E .

Each of the mappings \tilde{T}_t can be lifted into an operator on $\mathcal{L}^\infty(E)$, denoted by T_t and defined by

$$T_t f(x) = f(\tilde{T}_t x).$$

The following result is now a direct application of Theorem 3.3.

Corollary 3.4 *Let W be a d -dimensional Brownian motion on its canonical space $(\Omega, \mathcal{F}, \mathbb{P})$. Let f be a locally bounded Lipschitz functions from \mathbb{R}^n to \mathbb{R}^n and let g be a locally bounded*

Lipschitz function from \mathbb{R}^n to $M_{n \times d}(\mathbb{R})$. Consider the stochastic differential equation (on \mathbb{R}^n)

$$X_t^x = x + \int_0^t f(X_u^x) du + \int_0^t g(X_u^x) \cdot dW_u.$$

Then the mappings \tilde{T}_t on $\mathbb{R}^n \times \Omega$ defined by

$$\tilde{T}_t(x, \omega) = (X_t^x(\omega), \theta_t(\omega))$$

define a continuous time dynamical system.

This is to say that a stochastic differential equation is nothing more than a deterministic dynamical system on a product set $\mathbb{R}^n \times \Omega$, that is, it is a semigroup of point transformations of this product set.

We now have a result analogous to the one of Theorem 2.5 when this dynamical system is restricted to the \mathbb{R}^n -component. But before establishing this result, we need few technical lemmas. In the following $\Omega_{[t]}$ denotes the space of continuous functions from $[0, t]$ to \mathbb{R}^d . For all $\omega \in \Omega$ we denote by $\omega_{[t]}$ the restriction of ω to $[0, t]$. Finally $P_{[t]}$ denotes the restriction of the measure P to $(\Omega_{[t]}, \mathcal{F}_t)$.

Lemma 3.5 *The image of the measure P under the mapping*

$$\begin{aligned} \Omega &\rightarrow \Omega_{[t]} \times \Omega \\ \omega &\mapsto (\omega_{[t]}, \theta_t(\omega)) \end{aligned}$$

is the measure $P_{[t]} \otimes P$.

Proof: Recall that $\omega(s) = W_s(\omega)$ and $\theta_t(\omega)(s) = W_{t+s}(\omega) - W_t(\omega)$. If A is a finite cylinder of $\Omega_{[t]}$ and B a finite cylinder of Ω , then the set

$$\{\omega \in \Omega; (\omega_{[t]}, \theta_t(\omega)) \in A \times B\}$$

is of the form

$$\{\omega \in \Omega; W_{t_1}(\omega) \in A_1, \dots, W_{t_n}(\omega) \in A_n, (W_{s_1} - W_t)(\omega) \in B_1, \dots, (W_{s_k} - W_t)(\omega) \in B_k\}$$

for some $t_1, \dots, t_n \leq t$ and some $s_1, \dots, s_k > t$. By the independence of the Brownian motion increments, the probability of the above event is equal to

$$\begin{aligned} &\mathbb{P}(\{\omega \in \Omega_{[t]}; W_{t_1}(\omega) \in A_1, \dots, W_{t_n}(\omega) \in A_n\}) \times \\ &\quad \times \mathbb{P}(\{\omega \in \Omega; (W_{s_1} - W_t)(\omega) \in B_1, \dots, (W_{s_k} - W_t)(\omega) \in B_k\}). \end{aligned}$$

This is to say,

$$\mathbb{P}(\{\omega \in \Omega; (\omega_{[t]}, \theta_t(\omega)) \in A \times B\}) = \mathbb{P}_{[t]}(\{\omega_{[t]} \in A\}) \mathbb{P}(\{\omega \in \Omega; \theta_t(\omega) \in B\}).$$

This is exactly the claim of the lemma for the cylinder sets. As the measures \mathbb{P} and $\mathbb{P}_{[t]} \otimes \mathbb{P}$ are determined by their values on the cylinder sets, we conclude easily. \square

Lemma 3.6 *Let g be a bounded measurable function on $\Omega_{t_j} \times \Omega$. Then we have*

$$\int_{\Omega_{t_j}} \int_{\Omega} g(\omega, \omega') dP_{t_j}(\omega) dP(\omega') = \int_{\Omega} g(\omega_{t_j}, \theta_t(\omega)) dP(\omega).$$

Proof: This is just the Transfer Theorem for the mapping of Lemma 3.5. \square

Theorem 3.7 *Let (Ω, \mathcal{F}, P) be the canonical space of a d -dimensional Brownian motion W . Let f be a locally bounded Lipschitz functions from \mathbb{R}^n to \mathbb{R}^n and let g be a locally bounded Lipschitz function from \mathbb{R}^n to $M_{n \times d}(\mathbb{R})$. Consider the stochastic differential equation (on \mathbb{R}^n)*

$$X_t^x = x + \int_0^t f(X_u^x) du + \int_0^t g(X_u^x) \cdot dW_u$$

and the associated dynamical system

$$\tilde{T}_t(x, \omega) = (X_t^x(\omega), \theta_t(\omega)).$$

For any bounded function h on \mathbb{R}^n consider the mapping

$$P_t h(x) = \mathbb{E} [T_t(h \otimes \mathbb{1})(x, \cdot)].$$

Then $(P_t)_{t \in \mathbb{R}^+}$ is a Markov semigroup on \mathbb{R}^n with generator

$$A = \sum_{i=1}^n f_i(x) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^n \sum_{\alpha=1}^d g_{\alpha}^i(x) g_{\alpha}^j(x) \frac{\partial^2}{\partial x_i \partial x_j}.$$

Proof: The fact that each P_t is a Markov operator is a consequence of Theorem 2.2. Let us check that they form a semigroup. By definition we have

$$\begin{aligned} P_t(P_s h)(x) &= \mathbb{E} \left[\tilde{T}_t(P_s h \otimes \mathbb{1})(x, \cdot) \right] \\ &= \int_{\Omega} P_s h(X_t^x(\omega)) dP(\omega) \\ &= \int_{\Omega} \int_{\Omega} h(X_s^{X_t^x(\omega)}(\omega')) dP(\omega') dP(\omega) \\ &= \int_{\Omega} h(X_s^{X_t^x(\omega_{t_j})}(\theta_t(\omega))) dP(\omega) \quad (\text{by Lemma 3.6}) \\ &= \int_{\Omega} h(X_{s+t}^x(\omega)) dP(\omega) \quad (\text{by Theorem 3.3}) \\ &= P_{s+t} h(x). \end{aligned}$$

We have proved the semigroup property.

The rest of the proof comes from the usual theory of Markov semigroups and their associated generator. \square

We have proved the continuous time analog of Theorem 2.5. Every Markov semigroup, with a generator of the form above, can be dilated on a larger set (a product set) into a deterministic dynamical system. What is maybe more surprising is that the deterministic dynamical system in question is a stochastic differential equation. Theorem 3.7 and Corollary 3.4 show that a stochastic differential equation can actually be seen as a particular deterministic dynamical system.

The theorem above means again an open system point of view on Markov processes: Markov processes are obtained by restriction of certain types of dynamical systems on a product space, when one is averaging over one inaccessible component. The role of the environment is now played by the Wiener space and the role of the global dynamics on the product space is played by the stochastic differential equation.

In this section we have developed the Brownian case only. But it is clear that all this discussion extends exactly in the same way to the case of the Poisson process. Indeed, the arguments developed above are mostly only based on the independent increment property.

We said that stochastic differential equations are particular dynamical systems which are continuous analogues of those of Section 2.3: repeated interactions. In the article [4], the convergence of discrete-time repeated interactions models to stochastic differential equations is proved.

References

- [1] S. Attal, Y. Pautrat: *From Repeated to Continuous Quantum Interactions*, Annales Henri Poincaré (Theoretical Physics) 7 (2006), p. 59-104.
- [2] C. Bernardin, S. Olla: *Fourier's law for a microscopic heat conduction model*, J. Stat. Phys 121 , p. 271-289 (2005).
- [3] T. Bodineau, G. Giacomin: *From dynamic to static large deviations in boundary driven exclusion particle systems*, Stoch. Proc. Appl. 110, 67-81 (2004).
- [4] J. Deschamps, *Continuous Time Limit of Classical Repeated Interaction Systems*, preprint.
- [5] F. Bonetto, J. L. Lebowitz, L. Rey-Bellet: *Fourier Law: A challenge to theorists*, In: Mathematical Physics 2000, A. Fokas, A. Grigoryan, T. Kibble, and B. Zegarlinski (Eds.) Imp. Coll. Press, London 2000.
- [6] D. Ruelle: *A departure from equilibrium*, Nature 414 no. 6861, 263-264 (2001).
- [7] F. Haake, D. Spehner: *Quantum measurements without macroscopic superpositions*, Phys. Rev. A 77 (2008).
- [8] J. Preskill: *Quantum Information and Quantum Computation*, Course Files on the Web: <http://www.theory.caltech.edu/people/preskill/ph229>
- [9] V. Jaksic, Y. Ogata, C.-A. Pillet: *The Green-Kubo formula and the Onsager reciprocity relations in quantum statistical mechanics*, Communications in Mathematical Physics 265, 3 (2006), p. 721-738.
- [10] D. Karevski, T. Platini: *Quantum Non-Equilibrium Steady States Induced by Repeated Interactions*, Phys Rev. Letter, to appear.