TCM315 Fall 2022: Introduction to Open Quantum Systems

Lecture 2: Classical Master Equations

Course handouts are designed as a study aid and are not meant to replace the recommended textbooks. Handouts may contain typos and/or errors. The students are encouraged to verify the information contained within and to report any issue to the lecturer.

CONTENTS

INTRODUCTION

Introductions to stochastic calculus for a physicists' audience are e.g. [1, 2].

LANGEVIN DYNAMICS, MATHEMATICAL NOTATION

Stochastic differential equations with a noise source modulated by a local (ie. position dependent) amplitude term are called **multiplicative**. In the mathematical literature, it is customary to write Langevin equations in differential form to emphasize the inherent need of a **discretization prescription** in the presence of multiplicative noise

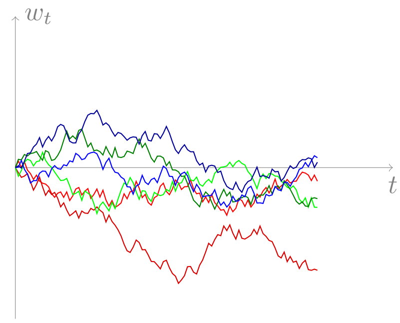
$$\mathrm{d}\boldsymbol{\chi}_t = \boldsymbol{b}(\boldsymbol{\chi}_t)\mathrm{d}t + \beta^{-1/2}\mathbb{D}^{1/2}(\boldsymbol{\chi}_t)\mathrm{d}\boldsymbol{w}_t \tag{1a}$$

$$\boldsymbol{\chi}_0 = \boldsymbol{x} \tag{1b}$$

Of the terms on the right hand side of (1a), we call the vector field proportional to $\mathrm{d}t$ the **drift** and the square, symmetric by hypothesis matrix $\mathbb{D}^{1/2}$ the **diffusion amplitude**. We define (1a) according to **the Itô prescription** and write the noise as the infinitesimal increment

$$\mathrm{d}\boldsymbol{w}_t = \boldsymbol{w}_{t+\mathrm{d}t} - \boldsymbol{w}_t$$

of a Gaussian process called the Wiener process.



The Wiener process starts by definition from the origin

$$\boldsymbol{w}_0 = 0$$

and has **independent increments** at any t obeying the Gaussian distribution

$$\mathrm{P}(\boldsymbol{x} \le \boldsymbol{w}_{t+s} - \boldsymbol{w}_t < \boldsymbol{x} + \mathrm{d}\boldsymbol{x}) = \frac{e^{-\frac{\|\boldsymbol{x}\|^2}{2\,s}}}{(2\,\pi\,s)^d}\mathrm{d}^{2d}\boldsymbol{x}$$

These properties imply that

$$\mathrm{E}\,\boldsymbol{w}_t = 0$$

$$\mathrm{E}\,\boldsymbol{w}_t\boldsymbol{w}_s^\top = \mathbb{1}_{2d} \times \begin{cases} s & \text{if}\, t > s \\ t & \text{if}\, t \le s \end{cases} = \mathbb{1}_{2d}\, t \wedge s$$

The white noise used in the physics notation is thus related to the Wiener process differential by the formal relation

$$\mathrm{d}\boldsymbol{w}_t = \boldsymbol{\eta}_t \mathrm{d}t$$

In terms of the Wiener process, the Itô property implies that for any fixed value of $\boldsymbol{\chi}_t$ the conditional expectation

$$\mathrm{E}\left(\mathbb{D}^{1/2}(\boldsymbol{\chi}_t)\mathrm{d}\boldsymbol{w}_t \middle| \boldsymbol{\chi}_t = \boldsymbol{x}\right) = \mathbb{D}^{1/2}(\boldsymbol{x})\,\mathrm{E}\,\mathrm{d}\boldsymbol{w}_t = 0$$

vanishes because the Wiener process and its increments vanish in mean.

In all practical manipulations we therefore need to recall

$$\mathrm{E}\,\mathrm{d}\boldsymbol{w}_t = 0$$

and that

$$\mathrm{E}\left(\mathrm{d}w_t \mathrm{d}w_s^\top\right) = \delta(t-s)\mathrm{d}t\mathrm{d}s\mathbb{1}_{2\,d}$$

	$\mathrm{d}t$	$\mathrm{d}w_{ti}$
$\mathrm{d}t$	0	0
$\mathrm{d}w_{tj}$	0	$\delta_{ij}\mathrm{d}t$

Wiener differential algebra.

In fact, squares of increments of the Wiener process are **self-averaging** so that for differential manipulations we can just apply the table on the right.

Itô differential

The differential algebra table implies that the differential of an arbitrary scalar indicator of a Langevin particle is

$$(\mathrm{d}f)(\boldsymbol{\chi}_t) = f(\boldsymbol{\chi}_t + \mathrm{d}\boldsymbol{\chi}_t) - f(\boldsymbol{\chi}_t) = (\mathrm{d}\boldsymbol{\chi}_t)^\top(\boldsymbol{\partial}f)(\boldsymbol{\chi}_t) + \frac{\beta^{-1}}{2}\,\mathrm{Tr}\left(\mathbb{D}(\boldsymbol{\chi}_t)(\boldsymbol{\partial}\otimes\boldsymbol{\partial}f)(\boldsymbol{\chi}_t)\right) \tag{2}$$

where the stochastic differential in the inner product

$$(\mathrm{d}\boldsymbol{\chi}_t)^\top(\boldsymbol{\partial}f)(\boldsymbol{\chi}_t) = \boldsymbol{b}^\top(\boldsymbol{\chi}_t)(\boldsymbol{\partial}f)(\boldsymbol{\chi})_t\mathrm{d}t + \beta^{-1/2}\left(\mathbb{D}^{1/2}(\boldsymbol{\chi}_t)\mathrm{d}\boldsymbol{w}_t\right)^\top(\boldsymbol{\partial}f)(\boldsymbol{\chi}_t)$$

must be interpreted in Itô sense.

THE STRATONOVICH PRESCRIPTION

In applications it is often convenient to change discretization prescription and adopt the **mid-point** or **Startonovich prescription**. This means evaluating the prefactor of the Wiener increment at a time

$$\bar{t} = t + \frac{\mathrm{d}t}{2}$$

In such a case, we can Taylor expand up to order $O(\mathrm{d}t)$ and obtain

$$\mathbb{D}^{1/2}(\boldsymbol{\chi}_t)\mathrm{d}\boldsymbol{w}_t = \mathbb{D}^{1/2}\Big(\boldsymbol{\chi}_{\bar{t}} - (\boldsymbol{\chi}_{\bar{t}} - \boldsymbol{\chi}_t)\Big)\mathrm{d}\boldsymbol{w}_t = \mathbb{D}^{1/2}(\boldsymbol{\chi}_{\bar{t}})\mathrm{d}\boldsymbol{w}_t - \frac{1}{2}((\mathrm{d}\boldsymbol{\chi}_t)^\top \boldsymbol{\partial})\mathbb{D}^{1/2}(\boldsymbol{\chi}_t)\mathrm{d}\boldsymbol{w}_t$$

The use of stochastic differential algebra table permit us to conclude

$$\begin{aligned}&((\mathrm{d}\boldsymbol{\chi}_t)^\top)\boldsymbol{\partial})\mathbb{D}^{1/2}(\boldsymbol{\chi}_t)\mathrm{d}\boldsymbol{w}_t = \\ &\quad \beta^{-1/2}\left(\left(\mathbb{D}^{1/2}(\boldsymbol{\chi}_t)\mathrm{d}\boldsymbol{w}_t\right)^\top \boldsymbol{\partial}\right)\mathbb{D}^{1/2}(\boldsymbol{\chi}_t)\mathrm{d}\boldsymbol{w}_t \equiv \beta^{-1/2}(\boldsymbol{\partial} \otimes \mathbb{D}^{1/2})(\boldsymbol{\chi}_t)\mathbb{D}^{1/2}(\boldsymbol{\chi}_t)\mathrm{d}t\end{aligned}$$

The stochastic increment

$$\mathbb{D}^{1/2}(\boldsymbol{\chi}_{\bar{t}})\mathrm{d}\boldsymbol{w}_t \equiv \mathbb{D}^{1/2}(\boldsymbol{\chi}_t)\diamond \mathrm{d}\boldsymbol{w}_t$$

is now evaluated according to the Stratonovich prescription. Accordningly, the increment of the Wiener process is **not anymore** independent of the state of the system at time $\bar{t} > t$:

$$\begin{aligned}&0 \neq \mathrm{E}\left(\mathbb{D}^{1/2}(\boldsymbol{\chi}_t)\diamond \mathrm{d}\boldsymbol{w}_t \big| \boldsymbol{\chi}_t = \boldsymbol{x}\right) = \\ &\mathrm{E}\left(\mathbb{D}^{1/2}(\boldsymbol{\chi}_t)\mathrm{d}\boldsymbol{w}_t \big| \boldsymbol{\chi}_t = \boldsymbol{x}\right) + \mathrm{E}\left(\frac{1}{2}(\boldsymbol{\partial} \otimes \mathbb{D}^{1/2})(\boldsymbol{\chi}_t)\mathbb{D}^{1/2}(\boldsymbol{\chi}_t)\big|\boldsymbol{\chi}_t = \boldsymbol{x}\right) = \frac{1}{2}(\boldsymbol{\partial} \otimes \mathbb{D}^{1/2})(\boldsymbol{x})\mathbb{D}^{1/2}(\boldsymbol{x})\end{aligned}$$

Merit of the Stratonovich prescription

The reason for introducing the Stratonovich prescription is that Stratonovich differentials behave as ordinary differentials

$$\begin{aligned}&(\mathrm{d}f)(\boldsymbol{\chi}_t) = f(\boldsymbol{\chi}_t + \mathrm{d}\boldsymbol{\chi}_t) - f(\boldsymbol{\chi}_t) = f\left(\boldsymbol{\chi}_{\bar{t}} + \frac{\mathrm{d}\boldsymbol{\chi}_t}{2}\right) - f\left(\boldsymbol{\chi}_{\bar{t}} - \frac{\mathrm{d}\boldsymbol{\chi}_t}{2}\right) \\ &= \frac{1}{2}(\mathrm{d}\boldsymbol{\chi}_t)^\top \diamond (\boldsymbol{\partial} f)(\boldsymbol{\chi}_t) + \frac{\beta^{-1}}{4}\operatorname{Tr}\Big(\mathbb{D}(\boldsymbol{\chi}_t)(\boldsymbol{\partial} \otimes \boldsymbol{\partial} f)(\boldsymbol{\chi}_t)\Big) - \left(-\frac{1}{2}(\mathrm{d}\boldsymbol{\chi}_t)^\top \diamond (\boldsymbol{\partial} f)(\boldsymbol{\chi}_t) + \frac{\beta^{-1}}{4}\operatorname{Tr}\Big(\mathbb{D}(\boldsymbol{\chi}_t)(\boldsymbol{\partial} \otimes \boldsymbol{\partial} f)(\boldsymbol{\chi}_t)\Big)\right)\end{aligned}$$

Hence a mutual cancellation of terms insure that (2) admits the **equivalent** Stratonovich representation

$$(\mathrm{d}f)(\boldsymbol{\chi}_t) = (\mathrm{d}\boldsymbol{\chi}_t)^\top \diamond (\boldsymbol{\partial} f)(\boldsymbol{\chi}_t)$$

MARKOVIANITY

The Langevin process describes a **Markovian** process. If we assign the state of the process and the increment of the Wiener noise the evolution law completely determines the new state of the process

$$(\boldsymbol{\chi}_t\,, \mathrm{d}\boldsymbol{w}_t) \overset{evolution\ law}{\Rightarrow} \boldsymbol{\chi}_{t+\mathrm{d}t}$$

Suppose that at time zero the **initial state** of the system is specified by a probability density

$$\Pr(\boldsymbol{y} \leq \boldsymbol{\chi}_0 < \boldsymbol{y} + \mathrm{d}\boldsymbol{y}) = \rho(\boldsymbol{y})\mathrm{d}^{2\,d}\boldsymbol{y}$$

we then recover (1b) if

$$\rho(\boldsymbol{y}) = \delta^{2\,d}(\boldsymbol{x} - \boldsymbol{y})$$

Once the initial state is specified, we have at hand a complete description of the statistics of the system if we know the **transition probability**

$$\Pr\left(\boldsymbol{x} \leq \boldsymbol{\chi}_t < \boldsymbol{x} + \mathrm{d}\boldsymbol{x} \;\middle|\; \boldsymbol{\chi}_s = \boldsymbol{y}\right) = \mathrm{K}_{t,s}(\boldsymbol{x}|\boldsymbol{y})\mathrm{d}^{2\,d}\boldsymbol{x} \qquad \text{for all } t \geq s$$

If, as we are assuming in (1a), the Hamiltonian and the coupling embodied by the diffusion coefficient $\mathbb{D}$ do not depend explicitly upon time

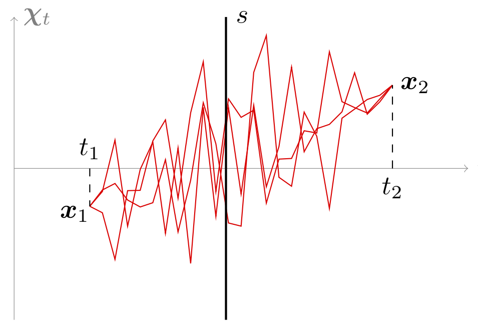
$$\mathrm{K}_{t,s}(\boldsymbol{x}|\boldsymbol{y}) = \mathrm{K}_{t-s,0}(\boldsymbol{x}|\boldsymbol{y})$$

In this case, the transition probability depends upon only upon the time elapsed between the instant t when we wish to know the state of the system $\boldsymbol{\chi}_t$ and the instant s at when we knew that the system was in $\boldsymbol{y}$ ($\boldsymbol{\chi}_s = \boldsymbol{y}$, conditioning event). We refer to this latter situation as **time autonomous** in the Langevin context. In order to neaten the notation, in the time autonomous case we denote the transition probability as depending upon only one time parameter.

In all cases Markovianity imposes that the transition probability satisfies the **Chapman-Kolmogorov** equation

$$\mathrm{K}_{t_2,t_1}(\boldsymbol{x}_2|\boldsymbol{x}_1) = \int_{\mathscr{M}} \mathrm{d}^{2d}\boldsymbol{y}\ \mathrm{K}_{t_2,s}(\boldsymbol{x}_2|\boldsymbol{y})\,\mathrm{K}_{s,t_1}(\boldsymbol{y}|\boldsymbol{x}_1) \qquad \text{for all } t_1 \le s \le t_2$$

and arbitrary $\boldsymbol{x}_2$, $\boldsymbol{x}_1$. Geometrically this means that the probability to reach $\boldsymbol{x}_2$ at time t_2 from $\boldsymbol{x}_1$ at time t_1 is the sum of all the paths that at an intermediate time s pass through an arbitrary point of the accessible region of phase space $\mathscr{M}$. The accessible phase space may or may not coincide with $\mathbb{R}^{2d}$, depending upon the model.



A notable consequence of the Chapman–Kolmogorov equation is that

$$\lim_{t \to s} \mathrm{K}_{t,s}(\boldsymbol{y}|\boldsymbol{x}) = \delta^{(2d)}(\boldsymbol{y} - \boldsymbol{x})$$

The main consequence of the Chapman-Kolmogorov equation is that the transition probability also acts as **time propagator** of probability densities

$$\rho_t(\boldsymbol{x}) = \int_{\mathscr{M}} \mathrm{d}^{2d}\boldsymbol{y}\ \mathrm{K}_{t,s}(\boldsymbol{x}|\boldsymbol{y})\rho_s(\boldsymbol{y}) \qquad t \ge s \tag{3}$$

NON-EQUILIBRIUM EVOLUTION OF THE STATISTICAL ENSEMBLE

From the point of view of statistical mechanics the stochastic single particle evolution embodied by (1) is an example of a μ-space evolution obtained by coarse-graining (under suitable weak-coupling assumptions) an exact Liouville dynamics in Γ-space. We wish now to derive explicitly the equation governing the evolution of the probability density in μ space. In particular, we show that for the Langevin dynamics the **Fokker–Planck equation** governs the evolution of probability densities. In Γ-space, Koopmans' equation governs the evolution of averages of indicators of the state of a system of particles. The counter-part in μ-space for Langevin dynamics is the **backward Kolmogorov** equation which we also derive.

Fokker-Planck equation: evolution of density

Our goal is to prove that the probability density of a Langevin dynamics obeys the **Fokker–Planck equation**

$$\partial_t \rho_t(\boldsymbol{y}) + \boldsymbol{\partial}_{\boldsymbol{y}}^{\top} \boldsymbol{b}(\boldsymbol{y})\rho_t(\boldsymbol{y}) = \frac{\beta^{-1}}{2}\,\mathrm{Tr}\,\boldsymbol{\partial}_{\boldsymbol{y}} \otimes \boldsymbol{\partial}_{\boldsymbol{y}}\mathbb{D}(\boldsymbol{y})\rho_t(\boldsymbol{y}) \tag{4a}$$

$$\rho_0(\boldsymbol{y}) = \rho(\boldsymbol{y}) \tag{4b}$$

$$\rho_t(\boldsymbol{y})\big|_{\boldsymbol{y}\in\partial\mathscr{M}} = \text{proper boundary conditions} \tag{4c}$$

By "proper boundary conditions" we mean e.g. reflecting boundary conditions enforcing probability conservation in $\mathscr{M}$ if this is a finite region of phase space. If $\mathscr{M}$ coincides with $\mathbb{R}^{2d}$ then we require that $\rho_t(\boldsymbol{x})$ and its derivatives vanish at infinity sufficiently fast to justify our manipulations.

Derivation of the Fokker–Planck equation

Our starting point to prove (4) is the expression of the average of an arbitrary smooth indicator f of the state of Langevin process (1) **subject to the condition** that the Langevin dynamics (1a) evolves from

$$\boldsymbol{\chi}_s = \boldsymbol{x}$$

for some **fixed** $s \leq t$. In terms of the transition probability, the conditional average reads

$$\mathrm{E}_{\boldsymbol{x},s}\, f(\boldsymbol{\chi}_t) = \int_{\mathscr{M}} \mathrm{d}^{2d}\boldsymbol{y}\, f(\boldsymbol{y})\, \mathrm{K}_{t,s}(\boldsymbol{y}|\boldsymbol{x})$$

The derivative with respect to the time when we perform the average

$$\partial_t\, \mathrm{E}_{\boldsymbol{x},s}\, f(\boldsymbol{\chi}_t) = \lim_{\mathrm{d}t \searrow 0} \int_{\mathscr{M}} \mathrm{d}^{2d}\boldsymbol{y}\, f(\boldsymbol{y}) \frac{\mathrm{K}_{t+\mathrm{d}t,s}(\boldsymbol{y}|\boldsymbol{x}) - \mathrm{K}_{t,s}(\boldsymbol{y}|\boldsymbol{x})}{\mathrm{d}t}$$

can be equivalently written in the form

$$\partial_t\, \mathrm{E}_{\boldsymbol{x},s}\, f(\boldsymbol{\chi}_t) = \lim_{\mathrm{d}t \searrow 0} \frac{\mathrm{E}_{\boldsymbol{x},s}\left(f(\boldsymbol{\chi}_{t+\mathrm{d}t}) - f(\boldsymbol{\chi}_t)\right)}{\mathrm{d}t}$$

We want to use the arbitrariness in the choice of the smooth indicator f to derive the equation governing the evolution of $\mathrm{K}_{t,s}(\boldsymbol{y}|\boldsymbol{x})$. Now we take advantage of the properties of the Ito differential to obtain the identity

$$\mathrm{E}_{\boldsymbol{x},s}\left(f(\boldsymbol{\chi}_{t+\mathrm{d}t}) - f(\boldsymbol{\chi}_t)\right) = \mathrm{E}_{\boldsymbol{x},s}\left(\boldsymbol{b}^{\top}(\boldsymbol{\chi}_t)(\boldsymbol{\partial} f)(\boldsymbol{\chi}_t) + \frac{\beta^{-1}}{2}\operatorname{Tr}\left(\mathbb{D}(\boldsymbol{\chi}_t)(\boldsymbol{\partial} \otimes \boldsymbol{\partial} f)(\boldsymbol{\chi}_t)\right)\right)\mathrm{d}t$$

which, once written explicitly in terms of the transition probability, allows us to arrive at

$$\int_{\mathscr{M}} \mathrm{d}^{2d}\boldsymbol{y}\, f(\boldsymbol{y})\partial_t\, \mathrm{K}_t(\boldsymbol{y}|\boldsymbol{x}) = \int_{\mathscr{M}} \mathrm{d}^{2d}\boldsymbol{y}\, \left(\boldsymbol{b}^{\top}(\boldsymbol{y})(\boldsymbol{\partial} f)(\boldsymbol{y}) + \frac{\beta^{-1}}{2}\operatorname{Tr}\left(\mathbb{D}(\boldsymbol{y})(\boldsymbol{\partial} \otimes \boldsymbol{\partial} f)(\boldsymbol{y})\right)\right)\mathrm{K}_t(\boldsymbol{y}|\boldsymbol{x})$$

Next we integrate by parts in order to free everywhere but in a boundary term, the test function f from the action of spatial derivatives

$$\begin{aligned}\int_{\mathscr{M}} \mathrm{d}^{2d}\boldsymbol{y}\, f(\boldsymbol{y})\partial_t\, \mathrm{K}_t(\boldsymbol{y}|\boldsymbol{x}) &= \int_{\mathscr{M}} \mathrm{d}^{2d}\boldsymbol{y}\boldsymbol{\partial}_{\boldsymbol{y}}^{\top}\left(f(\boldsymbol{y})\boldsymbol{B}(\boldsymbol{y},\boldsymbol{x}) + \frac{\beta^{-1}}{2}\mathrm{K}_t(\boldsymbol{y}|\boldsymbol{x})\mathbb{D}(\boldsymbol{y})(\boldsymbol{\partial} f)(\boldsymbol{y})\right)\\ &+ \int_{\mathscr{M}} \mathrm{d}^{2d}\boldsymbol{y}\, f(\boldsymbol{y})\left(-\boldsymbol{\partial}_{\boldsymbol{y}}^{\top}\boldsymbol{b}(\boldsymbol{y})\,\mathrm{K}_t(\boldsymbol{y}|\boldsymbol{x}) + \frac{\beta^{-1}}{2}\operatorname{Tr}\boldsymbol{\partial}_{\boldsymbol{y}} \otimes \boldsymbol{\partial}_{\boldsymbol{y}}\mathbb{D}(\boldsymbol{y})\,\mathrm{K}_t(\boldsymbol{y}|\boldsymbol{x})\right)\end{aligned} \tag{5}$$

where $\boldsymbol{B}$ denotes the **probability current** produced by the integration by parts:

$$\boldsymbol{B}(\boldsymbol{y},\boldsymbol{x}) = \boldsymbol{b}(\boldsymbol{y})\,\mathrm{K}_t(\boldsymbol{y}|\boldsymbol{x}) - \frac{\beta^{-1}}{2}\boldsymbol{\partial}_{\boldsymbol{y}}\mathbb{D}(\boldsymbol{y})\,\mathrm{K}_t(\boldsymbol{y}|\boldsymbol{x})$$

The first integral on the right hand side of (5) is the divergence of a vector field in the integration variable $\boldsymbol{y}$. If $\mathscr{M}$ is a bounded region in $\mathbb{R}^{2d}$ the integral localizes on the boundary $\partial\mathscr{M}$ of $\mathscr{M}$. "Proper" probability conserving boundary conditions are those which ensure that the resulting boundary integral vanishes. For example, **reflecting boundary conditions** correspond to probability densities satisfying

$$\boldsymbol{B}(\boldsymbol{y},\boldsymbol{x})\Big|_{\boldsymbol{y}\in\partial\mathscr{M}} = 0$$

and which can be used to compute well defined averages on the class of system indicators satisfying

$$(\boldsymbol{\partial} f)(\boldsymbol{y})\Big|_{\boldsymbol{y}\in\partial\mathscr{M}} = 0$$

Assuming "proper" probability conserving boundary conditions, we have

$$\int_{\mathscr{M}} \mathrm{d}^{2d}\boldsymbol{y}\boldsymbol{\partial}_{\boldsymbol{y}}^{\top}\left(f(\boldsymbol{y})\boldsymbol{B}(\boldsymbol{y},\boldsymbol{x})+\frac{\beta^{-1}}{2}\,\mathrm{K}_t(\boldsymbol{y}|\boldsymbol{x})\mathbb{D}(\boldsymbol{y})(\boldsymbol{\partial}f)(\boldsymbol{y})\right)=0$$

As a consequence, we are left with the identity

$$\int_{\mathscr{M}} \mathrm{d}^{2d}\boldsymbol{y}\, f(\boldsymbol{y})\left(\partial_t\,\mathrm{K}_t(\boldsymbol{y}|\boldsymbol{x})+\boldsymbol{\partial}_{\boldsymbol{y}}^{\top}\boldsymbol{b}(\boldsymbol{y})\,\mathrm{K}_t(\boldsymbol{y}|\boldsymbol{x})-\frac{\beta^{-1}}{2}\,\mathrm{Tr}\,\boldsymbol{\partial}_{\boldsymbol{y}}\otimes\boldsymbol{\partial}_{\boldsymbol{y}}\mathbb{D}(\boldsymbol{y})\,\mathrm{K}_t(\boldsymbol{y}|\boldsymbol{x})\right)=0$$

The identity must hold independently of f. Therefore we conclude that the transition probability density must satisfy the Fokker–Planck equation

$$\begin{aligned}
&\partial_t\,\mathrm{K}_{t,s}(\boldsymbol{y}|\boldsymbol{x})+\boldsymbol{\partial}_{\boldsymbol{y}}^{\top}\boldsymbol{b}(\boldsymbol{y})\,\mathrm{K}_{t,s}(\boldsymbol{y}|\boldsymbol{x})=\frac{\beta^{-1}}{2}\,\mathrm{Tr}\,\boldsymbol{\partial}_{\boldsymbol{y}}\otimes\boldsymbol{\partial}_{\boldsymbol{y}}\mathbb{D}(\boldsymbol{y})\,\mathrm{K}_{t,s}(\boldsymbol{y}|\boldsymbol{x})\\
&\lim_{t\searrow s}\mathrm{K}_{t,s}(\boldsymbol{y}|\boldsymbol{x})=\delta^{(2d)}(\boldsymbol{y}-\boldsymbol{x})\\
&K_{t,s}(\boldsymbol{x})\big|_{\partial\mathscr{M}}=\text{proper boundary conditions}
\end{aligned}$$

Since drift and diffusion do not depend explicitly upon time we could also replace in the above equations

$$\mathrm{K}_{t,s}(\boldsymbol{y}|\boldsymbol{x})=\mathrm{K}_{t-s,0}(\boldsymbol{y}|\boldsymbol{x})\;\equiv\;\mathrm{K}_{t-s}(\boldsymbol{y}|\boldsymbol{x})$$

Finally, if we multiply the Fokker–Planck for the transition probability by $\rho_s(\boldsymbol{x})$ and integrate over $\boldsymbol{x}$ then the Champman-Kolmogorov identity

$$\rho_t(\boldsymbol{y})=\int_{\mathscr{M}}\mathrm{d}^{2d}\boldsymbol{x}\,\mathrm{K}_{t,s}(\boldsymbol{y}|\boldsymbol{x})\rho_s(\boldsymbol{x})$$

yields the Fokker–Planck equation for an individual density ρ_t. We thus recover the announced result (4).

In the mathematical literature the Fokker–Planck equation is called the **forward Kolmogorov equation**.

Remark. *The derivation of the Fokker–Planck equation goes through in the same way if the drift* $\boldsymbol{b}$ *and diffusion* $\mathbb{D}$ *carry an explicit time dependence.*

* *

Backward Kolmogorov equation: evolution of indicators

We now turn to the evolution of averages and look for the counterpart of Koopman's equation in the Langevin context. To this goal we **fix** the time t when we compute the average and study how the result changes versus the time s when we assign or know the state of the Langevin process. In order to emphasize the change of perspective we write

$$F_s(\boldsymbol{x})\;\equiv\;\mathrm{E}_{\boldsymbol{x},s}\,f(\boldsymbol{\chi}_t)=\int_{\mathscr{M}}\mathrm{d}^{2d}\boldsymbol{y}\,f(\boldsymbol{y})\,\mathrm{K}_{t,s}(\boldsymbol{y}|\boldsymbol{x})$$

The Chapman–Kolmogorov equation imply that the time boundary condition

$$\lim_{s\nearrow t}F_s(\boldsymbol{x})=f(\boldsymbol{x})$$

must hold true. Our aim is to now show that for any $s\leq t$, F_s satisfies the **backward Kolmogorov equation**

$$\partial_sF_s(\boldsymbol{x})+\boldsymbol{b}^{\top}(\boldsymbol{x})\boldsymbol{\partial}_{\boldsymbol{x}}F_s(\boldsymbol{x})+\frac{\beta^{-1}}{2}\,\mathrm{Tr}\,\mathbb{D}(\boldsymbol{x})\boldsymbol{\partial}_{\boldsymbol{x}}\otimes\boldsymbol{\partial}_{\boldsymbol{x}}F_s(\boldsymbol{x})=0 \tag{6a}$$

$$F_t(\boldsymbol{x})=f(\boldsymbol{x}) \tag{6b}$$

supplemented by suitable boundary conditions on $\partial\mathscr{M}$. We emphasize that the derivation of (6) at variance with that of the Fokker–Planck equation does not make use of integration by parts.

Derivation of the backward Kolmogorov equation

In order to derive (6), we look for the expression of the derivative with respect to the "initial" time:

$$\partial_s F_s(\boldsymbol{x}) = \lim_{\mathrm{d}s \searrow 0} \frac{F_{s+\mathrm{d}s}(\boldsymbol{x}) - F_s(\boldsymbol{x})}{\mathrm{d}s} = \lim_{\mathrm{d}s \searrow 0} \int_{\mathcal{M}} \mathrm{d}^{2d}\boldsymbol{y}\, f(\boldsymbol{y}) \frac{\mathrm{K}_{t,s+\mathrm{d}s}(\boldsymbol{y}|\boldsymbol{x}) - \mathrm{K}_{t,s}(\boldsymbol{y}|\boldsymbol{x})}{\mathrm{d}s}$$

We avail us of the Chapman–Kolmogorov equation to couch the right hand side into the form

$$\partial_s F_s(\boldsymbol{x}) = \lim_{\mathrm{d}s \searrow 0} \int_{\mathcal{M}} \mathrm{d}^{2d}\boldsymbol{y} \int_{\mathcal{M}} \mathrm{d}^{2d}\boldsymbol{w}\, f(\boldsymbol{y})\, \mathrm{K}_{t,s+\mathrm{d}s}(\boldsymbol{y}|\boldsymbol{w}) \frac{\delta^{(2\,d)}(\boldsymbol{w}-\boldsymbol{x}) - \mathrm{K}_{s+\mathrm{d}s,s}(\boldsymbol{w}|\boldsymbol{x})}{\mathrm{d}s}$$

Next, we invert the order of integration and observe that by definition

$$F_{s+\mathrm{d}s}(\boldsymbol{w}) = \int_{\mathcal{M}} \mathrm{d}^{2d}\boldsymbol{y} f(\boldsymbol{y})\, \mathrm{K}_{t,s+\mathrm{d}s}(\boldsymbol{y}|\boldsymbol{w})$$

Thus we can rewrite the derivative as

$$\partial_s F_s(\boldsymbol{x}) = \lim_{\mathrm{d}s \searrow 0} \int_{\mathcal{M}} \mathrm{d}^{2d}\boldsymbol{w}\, F_{s+\mathrm{d}s}(\boldsymbol{w}) \frac{\delta^{(2\,d)}(\boldsymbol{w}-\boldsymbol{x}) - \mathrm{K}_{s+\mathrm{d}s,s}(\boldsymbol{w}|\boldsymbol{x})}{\mathrm{d}s} = - \lim_{\mathrm{d}s \searrow 0} \frac{\mathrm{E}_{\boldsymbol{x},s}\left(F_{s+\mathrm{d}s}(\boldsymbol{\chi}_{s+\mathrm{d}s}) - F_{s+\mathrm{d}s}(\boldsymbol{\chi}_s)\right)}{\mathrm{d}s} \tag{7}$$

In the last step we used

$$\mathrm{E}_{\boldsymbol{x},s}\, F_{s+\mathrm{d}s}(\boldsymbol{\chi}_s) = \int_{\mathcal{M}} \mathrm{d}^{2d}\boldsymbol{w}\, F_{s+\mathrm{d}s}(\boldsymbol{w}) \delta^{(2\,d)}(\boldsymbol{w}-\boldsymbol{x}) = F_{s+\mathrm{d}s}(\boldsymbol{x})$$

What we achieved in writing (7) is that the numerator is depends upon the differential along a path of the Langevin process of the function $F_{s+\mathrm{d}s}$. Hence we can express the differential in Itô form and obtain

$$\begin{aligned}
&\lim_{\mathrm{d}s \searrow 0} \frac{\mathrm{E}_{\boldsymbol{x},s}\left(F_{s+\mathrm{d}s}(\boldsymbol{\chi}_{s+\mathrm{d}s}) - F_{s+\mathrm{d}s}(\boldsymbol{\chi}_s)\right)}{\mathrm{d}s} \\
&= \lim_{\mathrm{d}s \searrow 0} \mathrm{E}_{\boldsymbol{x},s} \left(\boldsymbol{b}^{\top}(\boldsymbol{\chi}_s)(\boldsymbol{\partial} F_{s+\mathrm{d}s})(\boldsymbol{\chi}_s) + \frac{\beta^{-1}}{2} \operatorname{Tr} \mathbb{D}(\boldsymbol{\chi}_s)(\boldsymbol{\partial} \otimes \boldsymbol{\partial} F_{s+\mathrm{d}s})(\boldsymbol{\chi}_s) \right) \\
&= \lim_{\mathrm{d}s \searrow 0} \int_{\mathcal{M}} \mathrm{d}^{2d}\boldsymbol{w}\, \mathrm{K}_{s+\mathrm{d}s,s}(\boldsymbol{y}|\boldsymbol{x}) \left(\boldsymbol{b}^{\top}(\boldsymbol{y})(\boldsymbol{\partial} F_{s+\mathrm{d}s})(\boldsymbol{y}) + \frac{\beta^{-1}}{2} \operatorname{Tr} \mathbb{D}(\boldsymbol{y})(\boldsymbol{\partial} \otimes \boldsymbol{\partial} F_{s+\mathrm{d}s})(\boldsymbol{y}) \right)
\end{aligned}$$

Finally, the limit behavior of the transition probability

$$\lim_{\mathrm{d}s \searrow 0} \mathrm{K}_{s+\mathrm{d}s,s}(\boldsymbol{y}|\boldsymbol{x}) = \delta^{(2\,d)}(\boldsymbol{y}-\boldsymbol{x})$$

allows us to establish the result we are after.

Remark. *The derivation of the backward Kolmogorov equation goes through in the same way if the drift* $\boldsymbol{b}$ *and diffusion* $\mathbb{D}$ *carry an explicit time dependence.*

* *

* * *

[1] C. W. Gardiner. *Stochastic Methods: an Handbook for the Natural and Social Sciences*, volume 13 of *Springer series in synergetics*. Springer, 4 edition, 2009.

[2] K. Jacobs. *Stochastic Processes for Physicists. Understanding Noisy Systems.* Cambridge University Press, September 2010.