TCM315 Fall 2022: Introduction to Open Quantum Systems

Lecture 1: Open Classical Systems

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

There are plenty of excellent textbooks on Hamiltonian dynamics. The textbook [3] has the merit to provide a rigorous introduction using nothing more than the notions of plain calculus. The main references for the derivation of derivation of "Nonlinear Generalized Langevin Equations" is Robert Zwanzig's 1973 paper [6]. Zwanzig discusses the similar results also in § 1.6 of [7].

ONE PARTICLE HAMILTONIAN MECHANICS

We consider an even, $2d$, dimensional vector space where we introduce a set of Euclidean coordinates. A generic point is denoted by a column vector, say $\boldsymbol{x}$. In these coordinates, a dynamical system is **Hamiltonian** if we can describe it by assigning

1. the Hamiltonian: a **scalar function** $H(\boldsymbol{x})$ of the vector space ;

2. a $2d \times 2d$ **symplectic** matrix $\mathbb{J}$ with properties

$$\mathbb{J}^\top \mathbb{J} = \mathbb{1}_{2d} \qquad \text{orthogonality} \tag{1a}$$

$$\mathbb{J}^\top = -\mathbb{J} \qquad \text{anti-symmetry} \tag{1b}$$

where $\mathbb{1}_{2d}$ denotes the $2d \times 2d$ identity matrix.

We use the Hamiltonian and the symplectic matrix to construct an **Hamiltonian vector field**

$$\boldsymbol{v}_H(\boldsymbol{x}) = \mathbb{J}(\boldsymbol{\partial} H)(\boldsymbol{x}) \tag{2}$$

Remark. *Notation: we write*

$$\partial_{\boldsymbol{x}} H(\boldsymbol{x}) = (\boldsymbol{\partial} H)(\boldsymbol{x}) \qquad \textit{or} \qquad \partial_{x_i} H(\boldsymbol{x}) = (\partial_i H)(\boldsymbol{x})$$

to denote the ***gradient*** *of a scalar function with respect to its argument. According to this notation, if we need to* ***first*** *compute the gradient and* ***then*** *evaluate the result for $x = f(\boldsymbol{y})$ we write*

$$\partial_{\boldsymbol{x}} H(\boldsymbol{x})|_{x=f(\boldsymbol{y})} = (\boldsymbol{\partial} H)(f(\boldsymbol{y}))$$

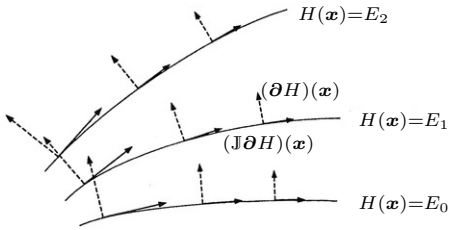
* *

Geometrically, an Hamiltonian vector field is always **tangent** to the any hyper-surface of the vector space where the Hamiltonian is constant. To justify this claim, we recall that the definition of the differential

$$\mathrm{d}H(\boldsymbol{x}) = H(\boldsymbol{x} + \mathrm{d}\boldsymbol{x}) - H(\boldsymbol{x}) = (\mathrm{d}\boldsymbol{x})^\top (\boldsymbol{\partial} H)(\boldsymbol{x})$$

implies that gradient be perpedicular to iso-surfaces of H. The definition of the Hamiltonian vector field (2) then immediately justifies the geometric interpretation

$$\left(\mathbb{J}(\boldsymbol{\partial} H)(\boldsymbol{x})\right)^\top (\boldsymbol{\partial} H)(\boldsymbol{x}) = (\boldsymbol{\partial} H)^\top(\boldsymbol{x}) \mathbb{J} (\boldsymbol{\partial} H)(\boldsymbol{x}) = 0$$



The latter expression in components means

$$(\boldsymbol{\partial} H)(\boldsymbol{x}_t)^\top \mathbb{J} (\boldsymbol{\partial} H)(\boldsymbol{x}_t) = \sum_{i\,j=1}^{2\,d} (\partial_i H)(\boldsymbol{x}_t) \mathbb{J}_{ij} (\partial_j H)(\boldsymbol{x}_t) = 0$$

and vanishes owing to the anti-symmetry of $\mathbb{J}$.

The **Hamilton equations of motion** for an initial data in $\boldsymbol{x}$ assigned at time $t = 0$ read

$$\dot{\boldsymbol{x}}_t = \boldsymbol{v}_H(\boldsymbol{x}_t) \tag{3a}$$

$$\boldsymbol{x}_0 = \boldsymbol{x} \tag{3b}$$

Example. *The explicit form (1) depends upon the choice of coordinates. For instance, the choice*

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{q} \\ \boldsymbol{p} \end{bmatrix} \qquad \& \qquad \boldsymbol{x}^\top = [q_1, \ldots, q_d, p_1, \ldots, p_d]$$

where the d-dimensional column vector $\boldsymbol{q}$ *(*$\boldsymbol{p}$*) with components $q_1, \ldots, q_d$ ($p_1, \ldots, p_d$) describes the* ***position*** *(****momentum****) of the system, implies*

$$\mathbb{J} = \begin{bmatrix} 0_d & \mathbb{1}_d \\ -\mathbb{1}_d & 0_d \end{bmatrix} \tag{4}$$

Here $\mathbb{1}_d$ and 0_d are, respectively the ***identity*** *and the* ***null*** *matrices acting on a d-dimensional vector space. The* ***Hamiltonian vector field*** *takes then the familiar form*

$$\boldsymbol{v}(\boldsymbol{q}, \boldsymbol{p}) = \begin{bmatrix} \partial_{\boldsymbol{p}} H(\boldsymbol{q}, \boldsymbol{p}) \\ -\partial_{\boldsymbol{q}} H(\boldsymbol{q}, \boldsymbol{p}) \end{bmatrix}$$

Had we instead chosen coordinates

$$\tilde{\boldsymbol{x}}^\top = [q_1, p_1, \ldots, q_d, p_d]$$

the same mechanical system would have been described by an Hamiltonian $\tilde{H}(\tilde{\boldsymbol{x}})$ *satisfying*

$$H(\boldsymbol{x}) = \tilde{H}(\tilde{\boldsymbol{x}})$$

and $\mathbb{J}$ *would be replaced by*

$$\tilde{\mathbb{J}} = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ -1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & \ldots & 0 & 1 \\ 0 & \ldots & \ldots & -1 & 0 \end{bmatrix} = \overbrace{J_2 \oplus \cdots \oplus J_2}^{d-times} \qquad J_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

* *

Flow

The Hamilton equations of motion form a system of first order differential equations. If, as implied above, the Hamiltonian vector field $\boldsymbol{v}_H$ does not depend explicitly upon the time parameter, i.e. is **time-autonomous**, and enjoys sufficient mathematical regularity properties, the equations (3) induce a one-parameter group $\boldsymbol{\Phi}_t$, called the **flow**, such that

$$\boldsymbol{x}_t = \boldsymbol{\Phi}_t(\boldsymbol{x}) \tag{5}$$

The flow is in general a nonlinear function of the time t and of the initial conditions $\boldsymbol{x}$. The flow defines a **one parameter group** of transformations because for arbitrary $\boldsymbol{x}$

$$\boldsymbol{\Phi}_0(\boldsymbol{x}) = \mathbb{1}_{2d}\boldsymbol{x} = \boldsymbol{x} \tag{6a}$$
$$\boldsymbol{\Phi}_t(\boldsymbol{\Phi}_s(\boldsymbol{x})) = \boldsymbol{\Phi}_{t+s}(\boldsymbol{x}) \tag{6b}$$
$$\boldsymbol{\Phi}_{-t}(\boldsymbol{\Phi}_t(\boldsymbol{x})) = \boldsymbol{\Phi}_0(\boldsymbol{x}) = \boldsymbol{x} \tag{6c}$$

Te notion of flow is useful because it exhibits an important feature of classical mechanics. Given the evolution law (3a) if we now the state of the system in phase space at a certain instant of time, we are in position to reconstruct within limits posed only by the numerical integration accuracy we have available, its future and past states. In this sense Hamiltonian dynamics is **Markovian** and time reversible.

Example: the isotropic harmonic oscillator

We suppose that the symplectic matrix has the form (4) and the Hamilton function reads

$$H(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top \mathbb{M}\boldsymbol{x}$$

with $\mathbb{M}$ the $2\,d \times 2\,d$ matrix

$$\mathbb{M} = \begin{bmatrix} m\,\omega^2\,\mathbb{1}_d & 0_d \\ 0_d & \frac{1}{m}\mathbb{1}_d \end{bmatrix}$$

The equations of motion (3) then take the form

$$\begin{aligned} \dot{\boldsymbol{x}}_t &= \mathbb{J}\,\mathbb{M}\,\boldsymbol{x}_t \\ \boldsymbol{x}_0 &= \boldsymbol{x} \end{aligned} \tag{7}$$

As $\mathbb{J}\,\mathbb{M}$ is time independent the explicit expression of the flow is the exponential of $\mathbb{J}\,\mathbb{M}$

$$\boldsymbol{x}_t = \boldsymbol{\Phi}_t(\boldsymbol{x}) = \exp\left(\mathbb{J}\,\mathbb{M}\,t\right)\boldsymbol{x}$$

The meaning of writing the exponential of a matrix is given by the Taylor series

$$\exp\left(\mathbb{J}\,\mathbb{M}\,t\right) = \sum_{n=0}^{\infty} \frac{(\mathbb{J}\,\mathbb{M})^n}{n!} t^n$$

It is straightforward to verify that

$$(\mathbb{J}\,\mathbb{M})^2 = -\omega^2\,\mathbb{1}_{2d}$$

whence

$$\exp\left(\mathbb{J}\,\mathbb{M}\,t\right) = \mathbb{1}_{2d}\sum_{n=0}^{\infty}\frac{(-\omega^2)^n}{(2n)!}t^{2n} + \mathbb{J}\mathbb{M}\sum_{n=0}^{\infty}\frac{(-\omega^2)^n}{(2n+1)!}t^{2n+1} = \begin{bmatrix} \cos(\omega\,t)\mathbb{1}_d & \frac{1}{m\,\omega}\sin(\omega\,t)\mathbb{1}_d \\ -m\,\omega\,\sin(\omega\,t)\mathbb{1}_d & \cos(\omega\,t)\mathbb{1}_d \end{bmatrix}$$

Energy conservation

Physically a time-autonomous Hamiltonian describes an **isolated system**: a system not exchanging energy with the environment. Formally this means that if at time $t = 0$

$$E = H(\boldsymbol{x})$$

the value E of the energy is preserved by the flow

$$E = H(\boldsymbol{\Phi}_t(\boldsymbol{x}))$$

Namely, upon reaclling (5)

$$\frac{\mathrm{d}}{\mathrm{d}t}H(\boldsymbol{\Phi}_t(\boldsymbol{x})) \equiv \frac{\mathrm{d}}{\mathrm{d}t}H(\boldsymbol{x}_t) = \dot{\boldsymbol{x}}_t^{\top}(\boldsymbol{\partial}H)(\boldsymbol{x}_t) = (\mathbb{J}\boldsymbol{\partial}H)^{\top}(\boldsymbol{x}_t)(\boldsymbol{\partial}H)(\boldsymbol{x}_t) = -(\boldsymbol{\partial}H)^{\top}(\boldsymbol{x}_t)\mathbb{J}(\boldsymbol{\partial}H)(\boldsymbol{x}_t) = 0$$

Hence, for an isolated system the geometric property of the Hamiltonian vector field to be tangent to a surface of level of the Hamiltonian translates into the **conservative** character of the dynamics.

OPEN CLASSICAL SYSTEM

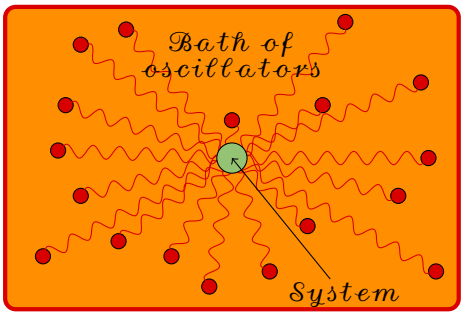
We consider a bipartite Hamiltonian dynamical system. The adjective "bipartite" means that we interpret the model as describing the interaction of two physically distinct constituents. The **central system**, or more simply from now on, the **system** described by a $2d$-dimensional coordinate vector $\boldsymbol{x}$ and the **environment** (or "**bath**") whose degrees of freedom are stored in a $2\,d\,\mathcal{N}$ dimensional coordinate vector $\boldsymbol{y}$. We also suppose $\mathcal{N} \gg 1$.

We surmise the Hamiltonian of the bipartite model to be of the form

$$H^{(tot)}(\boldsymbol{x},\boldsymbol{y}) = H(\boldsymbol{x}) + H^{(env)}(\boldsymbol{x},\boldsymbol{y})$$

For the moment, we do not make assumptions on H further than regularity in $\mathbf{x}$. We instead suppose that the environment Hamiltonian is quadratic in the coordinates $\boldsymbol{y}$

$$H^{(env)}(\boldsymbol{x},\boldsymbol{y}) = \frac{1}{2}\big(\boldsymbol{y} - \boldsymbol{A}(\boldsymbol{x})\big)^{\top}\mathbb{M}\big(\boldsymbol{y} - \boldsymbol{A}(\boldsymbol{x})\big)$$



The $2\,d\,\mathcal{N} \times 2\,d\,\mathcal{N}$ square matrix $\mathbb{M}$ is by hypothesis non singular and constant. The vector field with $2\,d\,\mathcal{N}$ components $\boldsymbol{A}$ is a smooth function of arbitrary form of the system coordinates $\boldsymbol{x}$ in the bath space. From the

physical point of view, the assumption of quadratic dependence in $\boldsymbol{y}$ of the total Hamiltonian deserves a justification. The idea (see also appendix C of [2] for a detailed discussion of the quantum couter-part of the model) is that we are stylizing a situation when the evolution of the environment is only **weakly perturbed** by the interaction with the system. The coordinates $\boldsymbol{y}$ describe the perturbation. In making this assumption **we completely neglect the internal dynamics of the bath**. Such a hypothesis can be quantitatively justified using multi-scale perturbation theory [5], if the parameters of the internal environment dynamics bring about a wide separation in the typical time scales of the motion. If this is the case, in a first approximation we can consider the internal degrees of freedom of the bath as frozen on a **stable fixed point** around which we consider small oscillations modeled by the $\boldsymbol{y}$ coordinates.

Hamiltonian dynamics implies the existence of a $2\,d \times 2\,d$ **constant, orthogonal and anti-symmetric** matrix $\mathbb{J}$ and of a $2\,d\mathcal{N} \times 2\,d\mathcal{N}$ **constant, orthogonal and anti-symmetric matrix** $\mathbb{J}_\Gamma$ such that

$$\dot{\boldsymbol{x}}_t = \mathbb{J}\boldsymbol{\partial}_{\boldsymbol{x}_t}\big(H(\boldsymbol{x}_t) + H^{(env)}(\boldsymbol{x}_t, \boldsymbol{y}_t)\big) \qquad \boldsymbol{x}_0 = \boldsymbol{x} \tag{8a}$$

$$\dot{\boldsymbol{y}}_t = \mathbb{J}_\Gamma\mathbb{M}\big(\boldsymbol{y}_t - \boldsymbol{A}(\boldsymbol{x}_t)\big) \tag{8b}$$

Note that at this stage we **do not** assign the value $\boldsymbol{y}_0$ of the initial data of the environment degrees of freedom.

Environment dynamics

The dynamics of the environment is linear and integrable (see appendix)

$$\boldsymbol{y}_t = e^{t\,\mathbb{J}_\Gamma\mathbb{M}}\boldsymbol{y}_0 - \int_0^t \mathrm{d}s\, e^{(t-s)\,\mathbb{J}_\Gamma\mathbb{M}}\mathbb{J}_\Gamma\mathbb{M}\boldsymbol{A}(\boldsymbol{x}_s) \tag{9}$$

We observe that the integrand is amenable to the form

$$\boldsymbol{y}_t = e^{t\,\mathbb{J}_\Gamma\mathbb{M}}\boldsymbol{y}_0 + \int_0^t \mathrm{d}s\, \left(\frac{\mathrm{d}}{\mathrm{d}s}\, e^{(t-s)\mathbb{J}_\Gamma\mathbb{M}}\right)\boldsymbol{A}(\boldsymbol{x}_s)$$

so that after an integration by parts we get into

$$\boldsymbol{y}_t = e^{t\,\mathbb{J}_\Gamma\mathbb{M}}\boldsymbol{y}_0 + \boldsymbol{A}(\boldsymbol{x}_t) - e^{t\mathbb{J}_\Gamma\mathbb{M}}\boldsymbol{A}(\boldsymbol{x}) - \int_0^t \mathrm{d}s\, e^{(t-s)\mathbb{J}_\Gamma\mathbb{M}}\,\frac{\mathrm{d}}{\mathrm{d}s}\boldsymbol{A}(\boldsymbol{x}_s) \tag{10}$$

Reduced system dynamics

In order to simplify the expression of the system equations of motion, it is expedient to define the Jacobian matrix of the vector field $\boldsymbol{A}$:

$$\mathbb{A}(\boldsymbol{x}) = (\boldsymbol{\partial} \otimes \boldsymbol{A})(\boldsymbol{x}) \tag{11}$$

which reads in components

$$\mathbb{A}_{i\,j}(\boldsymbol{x}) = (\partial_j\, A_i)(\boldsymbol{x})$$

We emphasize that $\mathbb{A}$ is by construction a **rectangular matrix** with $2\,d\mathcal{N} \times 2\,d$ components. For any fixed $\boldsymbol{x}$, $\mathbb{A}$ maps the system phase-space to the bath phase-space. Conversely, the transpose matrix $\mathbb{A}^\top$ maps column vectors in the environment space into column vectors in the system space. Once we are equipped with the definition (11) we couch (8a) into the form

$$\dot{\boldsymbol{x}}_t = \mathbb{J}(\boldsymbol{\partial}H)(\boldsymbol{x}_t) - \mathbb{J}\mathbb{A}^\top(\boldsymbol{x}_t)\mathbb{M}\big(\boldsymbol{y}_t - \boldsymbol{A}(\boldsymbol{x}_t)\big) \tag{12}$$

Furthermore we observe that

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(\boldsymbol{x}_t) = \mathbb{A}(\boldsymbol{x}_t)\dot{\boldsymbol{x}}_t$$

so that if we insert (10) into (12) we get into

$$\dot{\boldsymbol{x}}_t = \mathbb{J}(\boldsymbol{\partial}H)(\boldsymbol{x}_t) - \mathbb{J}\mathbb{A}^\top(\boldsymbol{x}_t)\mathbb{M}\, e^{t\,\mathbb{J}_\Gamma\,\mathbb{M}}\,(\boldsymbol{y}_0 - \boldsymbol{A}(\boldsymbol{x})) + \mathbb{J}\mathbb{A}^\top(\boldsymbol{x}_t)\mathbb{M}\int_0^t \mathrm{d}s\, e^{(t-s)\,\mathbb{J}_\Gamma\,\mathbb{M}}\,\mathbb{A}(\boldsymbol{x}_s)\dot{\boldsymbol{x}}_s$$

The system trajectory $\boldsymbol{x}_t$, satisfies an equation depending upon the system' s history from zero to t in which the bath variables enter only through their initial values.

STATISTICAL DESCRIPTION OF THE OPEN CLASSICAL SYSTEM

We now aim to use the exact reduced dynamics to model the action of the environment degrees of freedom on the system as a **random process** with well defined statistical properties. To this goal we define the noise source as

$$\boldsymbol{\omega}_t = -\mathbb{M}\, e^{t\,\mathbb{J}_\Gamma\,\mathbb{M}}\,(\boldsymbol{y}_0 - \boldsymbol{A}(\boldsymbol{x})) \tag{13}$$

This quantity depends parametrically upon the initial data of both system and environment. We surmise that

$$\Pr\left(\boldsymbol{y} \le \boldsymbol{y}_0 < \boldsymbol{y} + \mathrm{d}\boldsymbol{y} \middle| \boldsymbol{x}_0 = \boldsymbol{x}\right) \propto \exp\left(-\frac{\beta}{2}\big(\boldsymbol{y} - \boldsymbol{A}(\boldsymbol{x})\big)^\top \mathbb{M}\big(\boldsymbol{y} - \boldsymbol{A}(\boldsymbol{x})\big)\right) \tag{14}$$

In other words, we suppose that $\boldsymbol{y}_0$ is a Gaussian random variable with mean value

$$\mathrm{E}\,\boldsymbol{y}_0 = \boldsymbol{A}(\boldsymbol{x})$$

and co-variance

$$\mathrm{E}\,(\boldsymbol{y}_0 - \boldsymbol{A}(\boldsymbol{x}))\,(\boldsymbol{y}_0 - \boldsymbol{A}(\boldsymbol{x}))^\top = \beta^{-1}\mathbb{M}^{-1}$$

The statistical assignment of the initial data for the environment dynamics turns $\{\boldsymbol{\omega}_t\}_{t\ge 0}$ into a stochastic process (i.e. a one parameter t family of random variables) also Gaussian with vanishing mean value

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

and covariance

$$\mathrm{E}\left(\boldsymbol{\omega}_t\boldsymbol{\omega}_s^\top\right) = \mathbb{M}\, e^{t\,\mathbb{J}_\Gamma\,\mathbb{M}}\,\mathrm{E}\left(\,(\boldsymbol{y}_0 - \boldsymbol{A}(\boldsymbol{x}))\,(\boldsymbol{y}_0 - \boldsymbol{A}(\boldsymbol{x}))^\top\right)(e^{s\,\mathbb{J}_\Gamma\,\mathbb{M}})^\top\,\mathbb{M} = \beta^{-1}\mathbb{M}\, e^{t\,\mathbb{J}_\Gamma\,\mathbb{M}}\mathbb{M}^{-1}(e^{s\,\mathbb{J}_\Gamma\,\mathbb{M}})^\top\,\mathbb{M}$$

We simplify the covariance using the **symplectic property**

$$(e^{t\,\mathbb{J}_\Gamma\,\mathbb{M}})^\top\mathbb{J}_\Gamma = \mathbb{J}_\Gamma e^{-t\,\mathbb{J}_\Gamma\,\mathbb{M}}$$

hence

$$(e^{s\,\mathbb{J}_\Gamma\,\mathbb{M}})^\top\,\mathbb{M} = (e^{s\,\mathbb{J}_\Gamma\,\mathbb{M}})^\top\,\mathbb{J}_\Gamma\,\mathbb{J}_\Gamma^\top\,\mathbb{M} = \mathbb{J}_\Gamma e^{-s\,\mathbb{J}_\Gamma\,\mathbb{M}}\mathbb{J}_\Gamma^\top\,\mathbb{M} = \mathbb{J}_\Gamma\mathbb{J}_\Gamma^\top\,\mathbb{M}\,e^{-s\,\mathbb{J}_\Gamma\,\mathbb{M}}$$

We arrive in this way to the expression of the covariance we are after

$$\mathrm{E}\left(\boldsymbol{\omega}_t\boldsymbol{\omega}_s^\top\right) = \beta^{-1}\mathbb{M}\, e^{(t-s)\,\mathbb{J}_\Gamma\mathbb{M}}$$

The evolution of the system is now described by the **random differential equation**

$$\begin{aligned}
&\dot{\boldsymbol{x}}_t = \mathbb{J}(\boldsymbol{\partial}H)(\boldsymbol{x}_t) + \mathbb{J}\mathbb{A}^\top(\boldsymbol{x}_t)\boldsymbol{\omega}_t + \mathbb{J}\mathbb{A}^\top(\boldsymbol{x}_t)\mathbb{M}\int_0^t \mathrm{d}s\, e^{(t-s)\,\mathbb{J}_\Gamma\,\mathbb{M}}\,\mathbb{A}(\boldsymbol{x}_s)\dot{\boldsymbol{x}}_s\\
&\boldsymbol{x}_0 = \boldsymbol{x}
\end{aligned}$$

The equation is **non-Markovian** inasmuch it depends upon the full history of the system. As noted in [6] a special choice of the environment parameters can lead to an equation that becomes Markovian in the limit.

Markovian approximation

In order to derive the Markovian approximation we make two assumptions

1. We suppose that the Gaussian process $\{\boldsymbol{\omega}_t\}_{t\ge 0}$ can be replaced upon applying a **suitable limit procedure** by a **white noise** $\{\boldsymbol{\eta}_t\}_{t\ge 0}$ process: a Gaussian process with vanishing mean and δ-correlated in time:

$$\mathrm{E}\,\tilde{\boldsymbol{\omega}}_t = 0 \tag{15a}$$

$$\mathrm{E}\,\tilde{\boldsymbol{\omega}}_t\tilde{\boldsymbol{\omega}}_s^\top = \delta(t-s)\,\mathbb{1}_{2\,d\mathcal{N}} \tag{15b}$$

so that for any t

$$\boldsymbol{\omega}_t \to \beta^{-1/2}\mathbb{M}_\star^{1/2}\tilde{\boldsymbol{\omega}}_t$$

The under-script $\star$ here signify that we took a limit involving the parameters of the miscroscopic matrix $\mathbb{M}$ to produce the δ-function in (15b). We also emphasize that since $\mathbb{M}_\star$ is a square matrix both $\{\boldsymbol{\omega}_t\}_{t\geq 0}$ and $\{\tilde{\boldsymbol{\omega}}_t\}_{t\geq 0}$ are stochastic processes with $2\,d\,\mathcal{N}$ components.

2. We suppose that in the same limiting procedure powers of the rectangular matrix $\mathbb{A}$ or/and its transpose higher than the second become **negligible**

We refer e.g. to [6] or to § 1.6 of [7] for an explicit calculation showing how the two hypotheses may be derived from the choice of microscopic bath parameters.

If the two hypotheses are verified we formally arrive at

$$\dot{\boldsymbol{\chi}}_t = \tilde{\boldsymbol{b}}(\boldsymbol{\chi}_t) + \beta^{-1/2}\mathbb{B}(\boldsymbol{\chi}_t)\tilde{\boldsymbol{\omega}}_t \tag{16a}$$

$$\boldsymbol{\chi}_0 = \boldsymbol{x} \tag{16b}$$

with

$$\tilde{\boldsymbol{b}}(\boldsymbol{x}) = \mathbb{J}(\boldsymbol{\partial}H)(\boldsymbol{x}) - \mathbb{B}(\boldsymbol{x})\mathbb{B}^\top(\boldsymbol{x})(\boldsymbol{\partial}H)(\boldsymbol{x}) \tag{17}$$

and

$$\mathbb{B}(\boldsymbol{x}) = \mathbb{J}\mathbb{A}^\top(\boldsymbol{x})\,\mathbb{M}_\star^{1/2}$$

Even taking for granted the hypotheses which led us to (16), the result of our calculation is well defined only when the matrix $\mathbb{B}$ is constant i.e. independent of the state of the system.

WHITE NOISE AND THE DISCRETIZATION (CAUSALITY) PROBLEM

To understand the problem that (16) poses in general it is essential to analyze its meaning on a time lattice

$$0 \leq t_1 \leq t_2 \leq \ldots \leq t_k \leq \ldots$$

For simplicity we consider a time lattice with uniform mesh

$$t_{k+1} - t_k = \mathrm{d}t \qquad\qquad \text{for all } k$$

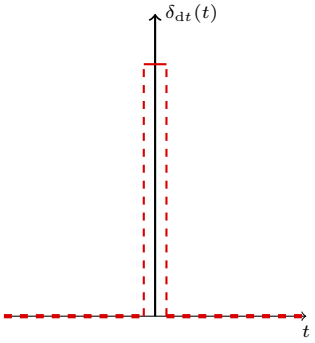
We then need to represent the δ-correlated random variables (15) on the lattice. To this goal we may introduce a sequence of **independent identically distributed** Gaussian $2\,d\,\mathcal{N}$-dimensional random vectors $\{\boldsymbol{\gamma}_k\}_{k\geq 1}$ with zero mean and unit variance

$$\mathrm{E}\,\boldsymbol{\gamma}_k = 0$$

$$\mathrm{E}(\boldsymbol{\gamma}_k\boldsymbol{\gamma}_l^\top) = \delta_{kl}\mathbb{1}_{2\,d\,\mathcal{N}}$$

Next we recall that we can approximate a Dirac δ with a step function of the form

$$\delta_{\mathrm{d}t}(t) = \begin{cases} \dfrac{1}{\mathrm{d}t} & t \in \left[-\dfrac{\mathrm{d}t}{2}, \dfrac{\mathrm{d}t}{2}\right] \\ 0 & t \notin \left[-\dfrac{\mathrm{d}t}{2}, \dfrac{\mathrm{d}t}{2}\right] \end{cases}$$



Hence, we can approximate at every step

$$\boldsymbol{\omega}_{t_k} = \frac{\boldsymbol{\gamma}_k}{\sqrt{\mathrm{d}t}}$$

To neaten the notation we set

$$\boldsymbol{\chi}_{t_k} \equiv \boldsymbol{\chi}_k$$

If we now try to write the finite version of (16) we realize that writing

$$\boldsymbol{\chi}_{k+1} - \boldsymbol{\chi}_k = \tilde{\boldsymbol{b}}(\boldsymbol{\chi}_k)\mathrm{d}t + \beta^{-1/2}\mathbb{B}(\boldsymbol{\chi}_k)\frac{\boldsymbol{\gamma}_{k+1}}{\sqrt{\mathrm{d}t}}\mathrm{d}t = \tilde{\boldsymbol{b}}(\boldsymbol{\chi}_k)\mathrm{d}t + \beta^{-1/2}\mathbb{B}(\boldsymbol{\chi}_k)\boldsymbol{\gamma}_{k+1}\sqrt{\mathrm{d}t} \tag{18}$$

differs from

$$\boldsymbol{\chi}_{k+1} - \boldsymbol{\chi}_k = \tilde{\boldsymbol{b}}(\boldsymbol{\chi}_k)\mathrm{d}t + \beta^{-1/2}\mathbb{B}(\cos^2\varepsilon\,\boldsymbol{\chi}_k + \sin^2\varepsilon\,\boldsymbol{\chi}_{k+1})\boldsymbol{\gamma}_{k+1}\sqrt{\mathrm{d}t} \tag{19}$$

by a **finite** correction to the drift. Namely

$$\begin{aligned}
&\mathbb{B}(\cos^2\varepsilon\,\boldsymbol{\chi}_k + \sin^2\varepsilon\,\boldsymbol{\chi}_{k+1})\boldsymbol{\gamma}_{k+1}\sqrt{\mathrm{d}t} - \mathbb{B}(\boldsymbol{\chi}_k)\boldsymbol{\gamma}_{k+1}\sqrt{\mathrm{d}t}\\
&= \Big((\cos^2\varepsilon - 1)\,\boldsymbol{\chi}_k + \sin^2\varepsilon\,\boldsymbol{\chi}_{k+1}\Big)^\top(\boldsymbol{\partial}\otimes\mathbb{B})(\boldsymbol{\chi}_k)\boldsymbol{\gamma}_{k+1}\sqrt{\mathrm{d}t} + \textbf{h.o.t.}\\
&= \sin^2\varepsilon\,(\boldsymbol{\chi}_{k+1} - \boldsymbol{\chi}_k)^\top(\boldsymbol{\partial}\otimes\mathbb{B})(\boldsymbol{\chi}_k)\boldsymbol{\gamma}_{k+1}\sqrt{\mathrm{d}t} + \textbf{h.o.t.}
\end{aligned}$$

Since in the differentials (18), (19) the leading order term is $O(\sqrt{\mathrm{d}t})$ we can unambiguously estimate

$$\begin{aligned}
&\mathbb{B}(\cos^2\varepsilon\,\boldsymbol{\chi}_k + \sin^2\varepsilon\,\boldsymbol{\chi}_{k+1})\boldsymbol{\gamma}_{k+1}\sqrt{\mathrm{d}t} - \mathbb{B}(\boldsymbol{\chi}_k)\boldsymbol{\gamma}_{k+1}\sqrt{\mathrm{d}t}\\
&= \beta^{-1/2}\,\sin^2\varepsilon\,\mathbb{B}(\boldsymbol{\chi}_k)\,(\boldsymbol{\partial}\otimes\mathbb{B})(\boldsymbol{\chi}_k)\mathrm{d}t + \textbf{h.o.t.}
\end{aligned}$$

having also used

$$\mathrm{E}\,\boldsymbol{\gamma}_{k+1}^2 = 1$$

We thus arrive at the conclusion that (19) is equivalent to

$$\boldsymbol{\chi}_{k+1} - \boldsymbol{\chi}_k = \Big(\tilde{\boldsymbol{b}}(\boldsymbol{\chi}_k) + \sin^2\varepsilon\,\boldsymbol{v}_{\mathcal{J}}(\boldsymbol{\chi}_k)\Big)\mathrm{d}t + \beta^{-1/2}\mathbb{B}(\boldsymbol{\chi}_k)\boldsymbol{\gamma}_{k+1}\sqrt{\mathrm{d}t}$$

and therefore describes a different evolution law from (18).

The take-home message

The derivation of the stochastic differential equation governing the open system dynamics requires extra care in order to determine from first principles the discretization prescription which defines on a time lattice the dynamics.

We now assume to have performed such a refined analysis and that the result is

$$\dot{\boldsymbol{\chi}}_t = \boldsymbol{b}(\boldsymbol{\chi}_t) + \beta^{-1/2}\mathbb{B}(\boldsymbol{\chi}_t)\tilde{\boldsymbol{\omega}}_t$$

stands for the continuum limit of

$$\boldsymbol{\chi}_{k+1} - \boldsymbol{\chi}_k = \boldsymbol{b}(\boldsymbol{\chi}_k)\mathrm{d}t + \beta^{-1/2}\mathbb{B}(\boldsymbol{\chi}_k)\boldsymbol{\gamma}_{k+1}\sqrt{\mathrm{d}t}$$

A further important property of Gaussian random variables is that **the linear combination** of Gaussian random variable is a **Gaussian random variable**. Since at time t the random variable $\tilde{\boldsymbol{\omega}}_t$ is independent of $\boldsymbol{\chi}_t$ for **any fixed** value $\boldsymbol{x}$ of $\boldsymbol{\chi}_t$ we can construct a $2\,d$-dimensional white noise process $\{\boldsymbol{\eta}_t\}_{t\geq 0}$

$$\mathrm{E}\,\boldsymbol{\eta}_t = 0 \tag{20a}$$

$$\mathrm{E}\,\boldsymbol{\eta}_t\boldsymbol{\eta}_s^\top = \delta(t-s)\mathbb{1}_{2\,d} \tag{20b}$$

such that

$$\begin{aligned}
\mathbb{B}(\boldsymbol{x})\tilde{\boldsymbol{\omega}}_t &= \mathbb{D}^{1/2}(\boldsymbol{x})\boldsymbol{\eta}_t\\
\mathbb{B}(\boldsymbol{x})\mathbb{B}^\top(\boldsymbol{x}) &= \mathbb{D}(\boldsymbol{x})
\end{aligned}$$

for $\mathbb{D}(\boldsymbol{x})$ some positive definite symmetric $2\,d\,\times\,2\,d$ matrix.

THE ITÔ STOCHASTIC DIFFERENTIAL EQUATION

We thus arrive at the central result of this lecture, the Itô representation of **Langevin-Kramers stochastic differential equation (SDE)**

$$\dot{\boldsymbol{\chi}}_t = \boldsymbol{b}(\boldsymbol{\chi}_t) + \beta^{-1/2}\mathbb{D}^{1/2}(\boldsymbol{\chi}_t)\boldsymbol{\eta}_t \tag{21a}$$

$$\boldsymbol{\chi}_0 = \boldsymbol{x} \tag{21b}$$

The drift in (21a)

$$\boldsymbol{b}(\boldsymbol{x}) = \mathbb{J}(\boldsymbol{\partial}H)(\boldsymbol{x}) - \mathbb{D}(\boldsymbol{x})(\boldsymbol{\partial}H)(\boldsymbol{x}) + \boldsymbol{v}_{\mathcal{J}}(\boldsymbol{x})$$

includes the two terms that we explicitly obtained in our derivation and **a last term** $\boldsymbol{v}_{\mathcal{J}}(\boldsymbol{x})$ stemming from a refined analysis that we only imply. This last term ensures that is indeed the continuum limit of the the finite difference equation

$$\boldsymbol{\chi}_{k+1} - \boldsymbol{\chi}_k = \boldsymbol{b}(\boldsymbol{\chi}_k)\mathrm{d}t + \beta^{-1/2}\mathbb{D}^{1/2}(\boldsymbol{\chi}_k)\,\boldsymbol{\nu}_{k+1}\sqrt{\mathrm{d}t} \tag{22a}$$

$$\boldsymbol{\chi}_0 = \boldsymbol{x} \tag{22b}$$

where now $\{\boldsymbol{\nu}_k\}_{k\geq 1}$ are a sequence of $2\,d$-dimensional Gaussian random variables with zero mean and unit variance

$$\mathrm{E}\,\boldsymbol{\nu}_k = 0$$
$$\mathrm{E}(\boldsymbol{\nu}_k\boldsymbol{\nu}_l^\top) = \delta_{kl}\mathbb{1}_{2\,d}$$

Example: kinetic plus potential Hamiltonian

Suppose that the system Hamiltonian has a kinetic plus potential form. We set $\boldsymbol{x}^\top = [\boldsymbol{q}^\top, \boldsymbol{p}^\top]$ and write

$$H(\boldsymbol{q},\boldsymbol{p}) = \frac{\|\boldsymbol{p}\|^2}{2\,m} + U(\boldsymbol{q})$$

Our choice of coordinates implies

$$\mathbb{J} = \begin{bmatrix} 0_d & \mathbb{1}_d \\ -\mathbb{1}_d & 0_d \end{bmatrix}$$

Finally we surmise that

$$\mathbb{D}(\boldsymbol{x}) = \frac{m}{\tau}\begin{bmatrix} 0_d & 0_d \\ 0_d & \mathbb{1}_d \end{bmatrix}$$

with $\tau > 0$. Physically this choice corresponds to a system environment coupling of the involving only **position coordinates**. The Langevin–Kramers equation reduces in such a case to the form

$$\dot{\boldsymbol{q}}_t = \frac{\boldsymbol{p}_t}{m}$$
$$\dot{\boldsymbol{p}}_t = -(\boldsymbol{\partial}U)(\boldsymbol{q}_t) - \frac{\boldsymbol{p}_t}{\tau} + \sqrt{\frac{m}{\beta\,\tau}}\boldsymbol{\eta}_t$$

The parameter τ is called the Stokes time. The example illustrates the fact that the positive definite matrix $\mathbb{D}$ describes **dissipative** effects from the system to the environment.

APPENDIX

Linear non homogeneous equation

A linear non-homogeneous system in N-dimensions of the form

$$\dot{\boldsymbol{y}}_t = \mathbb{C}\boldsymbol{y}_t + \boldsymbol{\ell}_t \tag{23a}$$

$$\boldsymbol{y}_0 = \boldsymbol{y} \tag{23b}$$

where $\mathbb{C}$ is a constant (i.e. time independent) square matrix and $\boldsymbol{f}_t$ is an assigned N dimensional column vector, eventually explicitly time dependent is always amenable to an exact integral expression.

We recall here two appraches

Separation of variables

We look for a solution of the form

$$\boldsymbol{y}_t = \mathbb{Y}_t \, \boldsymbol{g}_t \tag{24}$$

with $\mathbb{Y}_t$ a non-singular square matrix. Upon inserting (24) into (23a) we get into

$$\dot{\mathbb{Y}}_t \, \boldsymbol{g}_t + \mathbb{Y}_t \, \dot{\boldsymbol{g}}_t = \mathbb{C}\mathbb{Y}_t \, \boldsymbol{g}_t + \boldsymbol{f}_t$$

We then require that

$$\begin{aligned} \dot{\mathbb{Y}}_t &= \mathbb{C}\mathbb{Y}_t \\ \mathbb{Y}_0 &= \mathbb{1}_N \end{aligned} \tag{25}$$

i.e. that $\mathbb{Y}_t$ be the **fundamental solution** of the **linear homogeneous** equation associated to (23a). We get

$$\mathbb{Y}_t = e^{t\,\mathbb{C}}$$

It remains then to determine the vector $\boldsymbol{g}_t$. Since $\mathbb{Y}_t$ is non-singular by construction

$$\dot{\boldsymbol{g}}_t = \mathbb{Y}_t^{-1} \boldsymbol{f}_t$$

whence

$$\boldsymbol{g}_t = \boldsymbol{g}_0 + \int_0^t \mathrm{d}s \, \mathbb{Y}_s^{-1} \boldsymbol{f}_s$$

We conclude that

$$\boldsymbol{y}_t = e^{t\,\mathbb{C}} \boldsymbol{y} + \int_0^t \mathrm{d}s \, e^{(t-s)\,\mathbb{C}} \boldsymbol{f}_s$$

having done the obvious identification

$$\boldsymbol{g}_0 = \boldsymbol{y}$$

and used the property of the one parameter flow

$$\mathbb{Y}_t \mathbb{Y}_s^{-1} = \mathbb{Y}_t \mathbb{Y}_{-s} = \mathbb{Y}_{t-s}$$

Remark. *The solution takes the form of a linear combination of the solution of the homogeneous equation carrying information about the initial data plus a non-homogeneous solution with* ***vanishing initial data****. In other words*

$$\partial_{\boldsymbol{y}} \otimes \boldsymbol{y}_t = e^{t\,\mathbb{C}}$$

We notice that the method can be extended to the case of ***time-dependent*** *matrices $\mathbb{C}_t$. In that case, however, the fundamental solution of the associated linear system becomes a* ***two-parameters*** *family flow. Physically this means that the dynamics depends not only upon the the time elapsed from the moment when initial data are assigned but also upon the time when the initial data are assigned.*

* *

Laplace transform

Another general method of solution comes from the resort to **integral transforms**. We set

$$\tilde{\boldsymbol{y}}_z = \int_0^\infty \mathrm{d}t\, e^{-z\,t}\, \boldsymbol{y}_t$$

with z a complex number with real part $\mathrm{Re}\, z$ sufficiently large to ensure the convergence of the integral (see below for details). If we then consider

$$\int_0^\infty \mathrm{d}t\, e^{-z\,t} \dot{\boldsymbol{y}}_t = \int_0^\infty \mathrm{d}t\, e^{-z\,t}\, (\mathbb{C}\boldsymbol{y}_t + \boldsymbol{f}_t)$$

we obtain

$$z\,\tilde{\boldsymbol{y}}_z - \boldsymbol{y} = \mathbb{C}\tilde{\boldsymbol{y}}_z + \tilde{\boldsymbol{f}}_z$$

where

$$\tilde{\boldsymbol{f}}_z = \int_0^\infty \mathrm{d}t\, e^{-z\,t} \boldsymbol{f}_t$$

The merit of the transformation is to turn the differential system into an algebraic system that can be readily solved

$$\tilde{\boldsymbol{y}}_z = (z\,\mathbb{1}_N - \mathbb{C})^{-1}\,(\boldsymbol{y} + \tilde{\boldsymbol{f}}_z) \tag{26}$$

The matrix

$$\mathbb{R}_z = (z\,\mathbb{1}_N - \mathbb{C})^{-1}$$

is called the resolvent of $\mathbb{C}$. The resolvent is well defined for all z **distinct** from eigenvalues of $\mathbb{C}$. In order to recover from (26) explicit form of the solution we need then to **invert** the Laplace transform. We refer to any good textbook of mathematical methods for physics for details (e.g. chapter 20 § 7 of [1]). Here we only sketch the main steps.

The inverse Laplace transform is an integral on the complex z plane on a contour, called the Bromwich path, parallel to the axis purely imaginary axis $\mathrm{Im}\, z$. The condition

$$x_\star = \mathrm{Re}\, z$$

sufficiently large acquires the definite meaning of **larger** than the largest eigenvalue of $\mathbb{C}$. Then

$$e^{t\,\mathbb{C}} = \int_{x_\star - \imath\,\infty}^{x_\star + \imath\,\infty} \frac{\mathrm{d}z}{2\,\pi\,\imath}\, e^{z\,t} \mathbb{R}_z$$

The integral can be evaluated by closing the contour at infinity and to the left of the Bromwich path. In this way we can apply Cauchy's residue theorem and evaluate the integral simply by computing the residues of the poles of the resolvent.

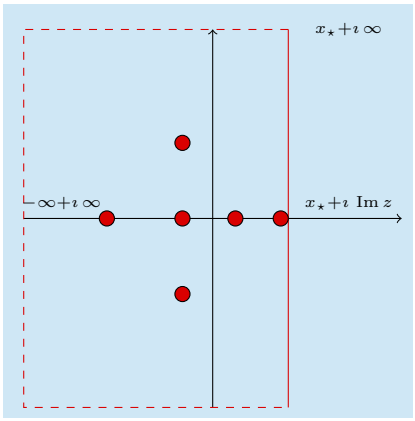


FIG. 1. The Bromwich path (red line), embedded in a contour closed at infinity (dashed red line). Red dots illustrate possible locations of the poles of the resolvent.

Finally we note that the properties of convolutions under Fourier-Laplace transform imply

$$\int_0^t \mathrm{d}s\, e^{(t-s)\,\mathbb{C}} \boldsymbol{f}_s = \int_{x_\star - \imath\,\infty}^{x_\star + \imath\,\infty} \frac{\mathrm{d}z}{2\,\pi\,\imath}\, e^{z\,t} \mathbb{R}_z\, \tilde{\boldsymbol{f}}_z$$

1 * * *

[1] G. B. Arfken, F. E. Harris, and H. J. Weber. *Mathematical Methods for Physicists*. Elsevier, 2013.
[2] A. O. Caldeira and A. J. Leggett. Quantum tunnelling in a dissipative system. *Annals of Physics*, 149(2):374–456, September 1983.
[3] A. Fasano and S. Marmi. *Analytical Mechanics : an Introduction*. Oxford Graduate Texts. Oxford Graduate Texts), 2006.
[4] K. Jacobs. *Stochastic Processes for Physicists. Understanding Noisy Systems*. Cambridge University Press, September 2010.
[5] G. A. Pavliotis and A. M. Stuart. *Multiscale methods: averaging and homogenization*, volume 53 of *Texts in applied mathematics*. Springer, 2008.
[6] R. Zwanzig. Nonlinear generalized Langevin equations. *Journal of Statistical Physics*, 9(3):215–220, November 1973.
[7] R. Zwanzig. *Nonequilibrium statistical mechanics*. Oxford University Press, 2001.