Mori-Zwanzig

Student: Lucas Amaral Taylor

Supervisor: Prof. Dr. Breno Raphaldini Ferreira da Silva

April de 2025

1 Introduction

The main objective of the article by Chekroun et al. (2021) is to simplify the Lorenz 80 model while preserving its behavior. To achieve this, we employ the Mori–Zwanzig method, a statistical-physical approach applicable to systems such as L80.

The Mori–Zwanzig method, developed by Robert Walter Zwanzig and Hajime Mori in the second half of the 20th century, is widely used in Hamiltonian systems. This method classifies the system's variables into two categories: "resolved" and "unresolved." Resolved variables are those whose behaviors and values are well known, whereas unresolved variables are those for which no direct information is available. To account for the unresolved variables, the method introduces stochastic terms, called noise, together with a damping term, also known as a memory term. This approach enables the behavior of the system of interest to be adequately preserved, even without complete knowledge of the unresolved variables.

Given the relevance of this method to the work of Chekroun et al. (2021), we chose to include an introduction to the Mori–Zwanzig formalism in order to provide a better understanding of its application in the context of the Lorenz 80 model and to establish a solid theoretical basis for further exploration.

2 Motivation

Consider the example from Chorin and Hald (2013, p. 173): a system with two particles in one spatial dimension, with Hamiltonian given by

$$H = \tfrac{1}{2}\left(q_1^2 + q_2^2 + q_1^2 q_2^2 + p_1^2 + p_2^2\right),$$

where q_i and p_i, $i = 1, 2$, represent positions and momenta. Once set in motion, harmonic oscillators continue oscillating indefinitely. The equations of motion are expressed as

$$\dot{q}_1 = p_1,$$
$$\dot{p}_1 = -q_1(1 + q_2^2),$$
$$\dot{q}_2 = p_2,$$
$$\dot{p}_2 = -q_2(1 + q_1^2). \tag{1}$$

Suppose that the initial values $q_1(0)$ and $p_1(0)$ are known, while $q_2(0)$ and $p_2(0)$ are sampled from the probability density function

$$W = \frac{e^{-H(q,p)}}{Z}.$$

This sampling can be performed, for example, using the Monte Carlo method via Markov chains.

Given a sample of $q_2(0)$ and $p_2(0)$, the system (1) can be solved. However, for each new sample of $q_2(0)$ and $p_2(0)$, a distinct trajectory is obtained for $q_1(t)$ and $p_1(t)$. In particular, one may wish to compute the expected values of q_1 and p_1 at time t, given their initial values, which represent the best estimates of $q_1(t)$ and $p_1(t)$:

$$\mathbb{E}[q_1(t) \mid q_1(0), p_1(0)], \quad \mathbb{E}[p_1(t) \mid q_1(0), p_1(0)].$$

Once q_2 and p_2 have been sampled, the complete system of four equations can be solved. This procedure can be repeated, allowing one to compute the averages of $q_1(t)$ and $p_1(t)$ over several runs.

However, there is a limitation: this approach is suitable only for the initial instants of the system. As time progresses, the expected value of $q_1(t)$ deviates from the actual value, compromising the accuracy of the approximation.

3 Prolegomena

3.1 Writing nonlinear ODEs as systems of linear PDEs

Consider the ordinary differential equation (ODE) system

$$\frac{d}{dt}\phi(x,t) = R(\phi(x,t)), \quad \phi(x,0) = x, \tag{2}$$

where R is a nonlinear function, ϕ is a time-dependent function, and R, ϕ, and x may be infinite-dimensional, formed by the vectors R_i, ϕ_i, and x_i, respectively.

From this, we define the Liouville operator associated with equation (2) as

$$L = \sum_i R_i(x) \frac{\partial}{\partial x_i}. \tag{3}$$

Using the Liouville operator, we can transform the nonlinear ODE system into a linear partial differential equation (PDE) system of the form

$$u_t = Lu, \quad u(x,0) = g(x). \tag{4}$$

The solution to this system exists, is unique, and is given by

$$u(x,t) = g(\phi(x,t)). \tag{5}$$

Therefore, equation (4) is well defined.[1]

3.2 Semigroup definition

Let X be a non-empty set equipped with a binary operation $*$, that is, $X \times X \to X$, which satisfies the associativity property:

$$(a * b) * c = a * (b * c), \quad \forall a, b, c \in X.$$

3.3 Introduction to notation

Semigroup notation offers a compact and efficient way to represent solutions to differential equations, particularly partial differential or evolution equations.

Consider the operator Δ defined by

$$\Delta\psi = \psi_{xx}, \quad \text{where } \psi \text{ is a smooth function.}$$

---

[1]Details of the proof can be found in Chorin and Hald (2013, p. 181–182).

Now, consider the differential equation

$$\frac{dv}{dt} - kv = 0, \quad v(0) = v_0,$$

whose solution is well known: $v(t) = v_0 e^{kt}$.

Similarly, consider the heat equation

$$v_t - \tfrac{1}{2}\Delta v = 0, \quad v(x,0) = \phi(x),$$

where v_t is the derivative of v with respect to time and $\phi(x)$ is the initial condition. Rather than solving directly, we express the solution using semigroup notation:

$$v(t) = e^{\frac{1}{2}t\Delta}\phi.$$

Here, $e^{\frac{1}{2}t\Delta}$ is a semigroup operator generated by the diffusion operator Δ. It acts on the initial condition $\phi(x)$, and the solution $v(t)$ describes the temporal evolution of $v(x,t)$ over time t. This notation allows us to represent solutions to differential equations in a compact form, exploiting the associative structure of the semigroup operation. Specifically, it satisfies the composition property:

$$e^{\frac{1}{2}(t+s)\Delta} = e^{\frac{1}{2}t\Delta}e^{\frac{1}{2}s\Delta}.$$

3.4 Application of notation

Given the semigroup notation presented above, we apply it to equation (5):

$$e^{tL}g(x) = g(\phi(x,t)), \tag{6}$$

Note that $e^{tL}x$ does not represent a direct evaluation of e^{tL}, but rather the action of the operator e^{tL} on the vector formed by the components x_i. Furthermore, the function g commutes with the time evolution of the initial conditions of x_i.

It is worth noting that g is a function independent of time with respect to the variables that describe the physical system, and its variation occurs exclusively due to changes in these variables over time. Thus, equation (4) can be expressed as

$$Le^{tL} = e^{tL}L. \tag{7}$$

An analogous relation holds for matrices: let A and B be two matrices. Then the following identity holds:

$$\exp(t(A+B)) = \exp(tA) + \int_0^t \exp((t-s)(A+B))\,B\exp(sA)\,ds, \tag{8}$$

This equation, known as the Duhamel formula or Dyson formula, is well defined.

3.5 Hermite polynomials

First, we define the inner product as

$$\langle u, v \rangle = \int_{-\infty}^{+\infty} \frac{e^{-x^2/2}}{\sqrt{2\pi}} \, u(x)v(x) \, dx, \tag{9}$$

The polynomials $p_n(x)$ and $p_m(x)$ are orthonormal with respect to this inner product (9) when they satisfy

$$\langle p_n, p_m \rangle = \int_{-\infty}^{\infty} e^{-x^2/2} p_n(x) p_m(x) \, dx = \delta_{nm},$$

where δ_{nm} is the Kronecker delta, which has the following properties:

1. **Orthogonality**: For $n \neq m$, the polynomials are orthogonal, i.e., the inner product between them is zero:

$$\langle p_n, p_m \rangle = 0 \quad \text{when} \quad n \neq m.$$

2. **Normalization**: For $n = m$, the polynomials are normalized, so that the inner product equals 1:

$$\langle p_n, p_n \rangle = 1.$$

In the n-dimensional case, the inner product generalizes to

$$\langle u, v \rangle = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} (2\pi)^{-n/2} \exp\left(-\sum_{i=1}^{n} \frac{x_i^2}{2} \right) u(x)v(x) \, dx_1 \cdots dx_n.$$

More generally, if $H(q, p)$ is a Hamiltonian, it is possible to define a family of polynomials in the variables q and p that are orthonormal with respect to the canonical distribution $Z^{-1}e^{-H/T}$. Polynomials that satisfy this condition are called *Hermite polynomials*.

Finally, for the Mori–Zwanzig formalism, we consider an n-dimensional space Γ with a given probability density. We divide the coordinates into two types: $\hat{x}$ and $\tilde{x}$. Let g be a function of x; then $\mathcal{P}g = \mathbb{E}[g \mid \hat{x}]$ is the orthogonal projection onto the subspace of functions of $\hat{x}$. This projection generates a subspace of Hermite polynomials that depend only on $\hat{x}$, and the operator $\mathcal{P}$ acts by projecting functions onto this subspace.

4 Mori-Zwanzig

4.1 Construction

Let us take another look at the system (2), reproduced below:

$$\frac{d}{dt}\phi(x,t) = R(\phi(x,t)), \quad \phi(x,0) = x,$$

Recall that the equation consists of n components. Among these n components, we define the first m components of ϕ, with $m < n$, as the variables of interest. We then classify $\hat{\phi}$ as the "resolved" variables and $\tilde{\phi}$ as the "unresolved" variables:

$$\phi = (\hat{\phi}, \tilde{\phi}), \quad \hat{\phi} = (\phi_1, \ldots, \phi_m), \quad \tilde{\phi} = (\phi_{m+1}, \ldots, \phi_n).$$

The same applies to x and R: $x = (\hat{x}, \tilde{x})$ and $R = (\hat{R}, \tilde{R})$. From the resolved variables, we seek to create predictions for the model of interest, using the solutions from one part of the equation.

Based on the Liouville operator and semigroup notation, we can rewrite the components of $\hat{\phi}$ as[2]:

$$\hat{\phi}_j(x,t) = e^{tL}x_j, \quad 1 \leq j \leq m.$$

Still in semigroup notation, the equation for these components is given by:

$$\frac{\partial}{\partial t}e^{tL}x_j = Le^{tL}x_j = e^{tL}Lx_j, \tag{10}$$

From the orthogonal projection introduced in the previous section, we define $\mathcal{P}$ as the projection given by $\mathcal{P}g(x) = \mathbb{E}[g \mid \hat{x}]$. We assume that, at time $t = 0$, we know the joint density of all variables x, but only the initial data $\hat{x}$ are known. The density of the variables in $\tilde{x}$ is then the joint density of all variables x with $\hat{x}$ fixed. Thus, $\mathcal{P}$ is a projection onto a space of functions with fixed variables and is therefore independent of time.

The projections $\mathcal{P}\hat{\phi}(t) = \mathbb{E}[\hat{\phi}(t) \mid \hat{x}]$ are of greatest interest to us, as they estimate the behavior of the system from a reduced set of variables.

Defining $\mathcal{Q} = I - \mathcal{P}$ and considering that the following properties hold for any orthogonal projections:

1. $\mathcal{P}^2 = \mathcal{P}$;

---

[2]Note that each component depends on all values of x. Therefore, if $\tilde{x}$ is random, then $\hat{\phi}$ will also be random.

2. $\mathcal{Q}^2 = \mathcal{Q}$;

3. $\mathcal{P}\mathcal{Q} = 0$.

We can then rewrite equation (10) as:

$$\frac{\partial}{\partial t} e^{tL} x_j = e^{tL} \mathcal{P} L x_j + e^{tL} \mathcal{Q} L x_j, \tag{11}$$

Now, using the Dyson formula with $A = \mathcal{Q}L$ and $B = \mathcal{P}L$, we obtain:

$$e^{tL} = e^{t\mathcal{Q}L} + \int_0^t e^{(t-s)L} \mathcal{P} L \, e^{s\mathcal{Q}L} \, ds, \tag{12}$$

Due to the linearity of the Liouville equation and based on equations (11) and (12), we obtain:

$$\frac{\partial}{\partial t} e^{tL} x_j = e^{tL} \mathcal{P} L x_j + e^{t\mathcal{Q}L} \mathcal{Q} L x_j + \int_0^t e^{(t-s)L} \mathcal{P} L e^{s\mathcal{Q}L} \mathcal{Q} L x_j \, ds, \tag{13}$$

which is the Mori–Zwanzig equation.

4.2 Term-by-term analysis

4.2.1 First term

The first term is given by

$$e^{tL} \mathcal{P} L x_j, \tag{14}$$

Observe that

$$L x_j = \sum_i R_i \left(\frac{\partial}{\partial x_i} \right) x_j = R_j(x).$$

Therefore,

$$\mathcal{P} L x_j = \mathbb{E}[R_j(x) \mid \hat{x}], \quad \text{which is a function depending only on } \hat{x}.$$

From this, we can conclude that

$$e^{tL} \mathcal{P} L x_j = \bar{R}_j\big(\hat{\phi}(x,t)\big).$$

Moreover, the first term represents the system's intrinsic dynamics in the resolved variables. It is also a Markovian term, as it depends only on the current state of the system at time t.

4.2.2 Second term

For the second term, we define

$$w_j = e^{t\mathcal{Q}L}\mathcal{Q}Lx_j.$$

By definition, we have

$$\frac{\partial}{\partial t}w_j(x,t) = \mathcal{Q}Lw_j(x,t), \tag{15}$$

$$w_j(x,0) = \mathcal{Q}Lx_j = (I - \mathcal{P})R_j(x) = R_j(x) - \mathbb{E}[R_j(x) \mid \hat{x}]. \tag{16}$$

Note that $w_j(x,0) = \mathcal{Q}Lx_j = R_j(x) - \mathbb{E}[R_j(x) \mid \hat{x}]$ represents the *fluctuating part* of the variable $R_j(x)$, that is, the unpredictable component given $\hat{x}$. This part evolves according to the *orthogonal dynamics*, so that $\mathcal{P}w_j(x,t) = 0$ for all t, keeping the term as purely unresolved noise over time.

More specifically, the noise subspace is formed by the components of functions that are orthogonal to the functions of $\hat{x}$, typically corresponding to terms that depend on $\tilde{x}$.

4.2.3 Third term

The third term,

$$\int_0^t e^{(t-s)L}\mathcal{P}Le^{s\mathcal{Q}L}\mathcal{Q}Lx_j,$$

is classified as a memory term, since it involves the integration of quantities that depend on past states of the system.

Let $\mathcal{P}$ be the projection onto the span of Hermite polynomials $H_1, H_2, \ldots$ with arguments in $\hat{x}$. For a given function ψ, we have $\mathcal{P}\psi = \sum(\psi, H_k)H_k$, and thus

$$\mathcal{P}Le^{s\mathcal{Q}L}\mathcal{Q}Lx_j = \mathcal{P}L(\mathcal{P} + \mathcal{Q})e^{s\mathcal{Q}L}\mathcal{Q}Lx_j$$

$$= \mathcal{P}L\mathcal{Q}e^{s\mathcal{Q}L}\mathcal{Q}Lx_j$$

$$= \sum_k \langle L\mathcal{Q}e^{s\mathcal{Q}L}\mathcal{Q}Lx_j, H_k(\hat{x})\rangle H_k(\hat{x}).$$

The inner product is defined as an expectation with respect to the initial probability density. Assuming that L is antisymmetric, i.e., $(u, Lv) = -(Lu, v)$, we obtain

$$(L\mathcal{Q}e^{s\mathcal{Q}L}\mathcal{Q}Lx_j, H_k(\hat{x})) = -(\mathcal{Q}e^{s\mathcal{Q}L}\mathcal{Q}Lx_j, LH_k)$$

$$= -(e^{s\mathcal{Q}L}\mathcal{Q}Lx_j, \mathcal{Q}LH_k).$$

Both $\mathcal{Q}Lx_j$ and $\mathcal{Q}LH_k$ belong to the noise subspace, and $e^{s\mathcal{Q}L}\mathcal{Q}Lx_j$ is a solution at time s of the orthogonal dynamics equation with data in the noise subspace. Hence, $\mathcal{P}Le^{s\mathcal{Q}L}\mathcal{Q}Lx_j$ is a **sum of temporal covariances of noise**.

5 Returning to the motivation

5.1 Introduction

First, let us recall the problem at hand: we have a Hamiltonian system, with the Hamiltonian defined by

$$H = \tfrac{1}{2}\left(q_1^2 + q_2^2 + q_1^2 q_2^2 + p_1^2 + p_2^2\right).$$

The associated equations of motion are

$$\dot{q}_1 = p_1,$$

$$\dot{p}_1 = -q_1(1 + q_2^2),$$

$$\dot{q}_2 = p_2,$$

$$\dot{p}_2 = -q_2(1 + q_1^2).$$

Applying the Liouville operator, we obtain

$$L = p_1 \frac{\partial}{\partial q_1} - q_1(1 + q_2^2)\frac{\partial}{\partial p_1} + p_2 \frac{\partial}{\partial q_2} - q_2(1 + q_1^2)\frac{\partial}{\partial p_2}.$$

As before, we assume that the initial values $q_1(0)$ and $p_1(0)$ are known, while q_2 and p_2 are distributed according to a probability density function given by

$$W(x) = \frac{1}{Z}\exp\left(-H(q_1, p_1, q_2, p_2)\right).$$

Our goal is to evaluate $q_1(t)$ and $p_1(t)$ using the Mori–Zwanzig equations (13).

5.2 First approximation

Given the nature of the system, it is necessary to make an approximation to introduce the memory term. We know that the orthogonal dynamics described in (16) are approximately equivalent to the complete dynamics, since they are not sensitive to the resolved variables. Based on this, the approximation consists of replacing $e^{t\mathcal{Q}L}$ with e^{tL} in the memory term. This substitution is valid when the influence of the resolved variables on the unresolved ones is small.[3]

That said, we can begin the algebraic manipulation. An operator commutes with any function of itself, so that

$$\mathcal{Q}Le^{s\mathcal{Q}L} = e^{s\mathcal{Q}L}\mathcal{Q}L.$$

---

[3]This substitution is justified when the coupling between the resolved and unresolved variables is weak.

Using this identity and substituting $e^{sQL} \to e^{sL}$ on the right-hand side, we obtain

$$\mathcal{P} L e^{sQL} \approx L e^{sL} - e^{sL} QL.$$

Hence,

$$e^{(t-s)L} \mathcal{P} L e^{sQL} \approx e^{(t-s)L} L e^{sL} - e^{(t-s)L} e^{sL} QL = e^{tL} \mathcal{P} L,$$

which makes the integrand in the memory term of the Mori–Zwanzig equation independent of s, resulting in

$$\int_0^t e^{tL} \mathcal{P} L \mathcal{Q} L x_j \, ds = t e^{tL} \mathcal{P} L \mathcal{Q} L x_j,$$

where $\hat{x}$ is the vector with components $x_1 = q_1$ and $x_2 = p_1$. The memory term is thus reduced to a differential operator multiplied by time t. Note that $t = 0$ corresponds to the instant when the initial values of $q_1(t)$ and $p_1(t)$ are assigned, i.e., when there is no uncertainty. This approximation can also be justified by assuming that the integrand in the memory term is approximately constant in s, and therefore can be evaluated at $s = 0$.

The equations with the simplified integral term constitute the so-called t-model. Combining the terms, we have

$$\frac{d}{dt} e^{tL} \hat{x} = e^{tL} \mathcal{P} L \hat{x} + t e^{tL} \mathcal{P} L \mathcal{Q} L \hat{x} + e^{tL} \mathcal{Q} L \mathcal{Q} L x_j.$$

In the particular case considered, with components $x_j = q_1$ and $x_j = p_1$, we obtain

$$L q_1 = p_1,$$
$$\mathcal{P} L q_1 = p_1,$$
$$\mathcal{Q} L q_1 = 0,$$
$$L \mathcal{Q} L q_1 = 0,$$
$$\mathcal{P} L \mathcal{Q} L q_1 = 0,$$

and

$$L p_1 = -q_1 (1 + q_2^2),$$
$$\mathcal{P} L p_1 = -q_1 \left(1 + \frac{1}{1 + q_1^2} \right),$$
$$\mathcal{Q} L p_1 = -q_1 (1 + q_2^2) + q_1 \left(1 + \frac{1}{1 + q_1^2} \right),$$
$$L \mathcal{Q} L p_1 = p_1 \left(-(1 + q_2^2) + \left(1 + \frac{1}{1 + q_1^2} \right) \right) - \frac{2 q_1^2}{(1 + q_1^2)^2} - 2 q_1 q_2 p_2,$$
$$\mathcal{P} L \mathcal{Q} L p_1 = -\frac{2 q_1^2 p_1}{(1 + q_1^2)^2}.$$

Thus, the approximate equations of motion for q_1 and p_1 become

$$\frac{d}{dt}q_1 = p_1,$$

$$\frac{d}{dt}p_1 = -q_1\left(1 + \frac{1}{1+q_1^2}\right) - 2t\frac{q_1^2 p_1}{(1+q_1^2)^2} + e^{tL}\mathcal{Q}L\mathcal{Q}Lp_1. \tag{17}$$

Now suppose we are interested only in the quantities

$$\mathbb{E}[q_1(t) \mid q_1(0), p_1(0)], \quad \mathbb{E}[p_1(t) \mid q_1(0), p_1(0)],$$

that is, the conditional expectations of $q_1(t)$ and $p_1(t)$ given $q_1(0), p_1(0)$.

An equation for these quantities can be obtained by applying the operator $\mathcal{P}$ to equations (17), recalling that, by definition, $\mathcal{P}q_1(t) = \mathbb{E}[q_1(t) \mid q_1(0), p_1(0)]$, and similarly for $p_1(t)$. In this process, the noise term disappears.

However, a problem arises: the mean of a function of a variable is not, in general, equal to the function of the mean. To overcome this difficulty, further simplification is necessary. This complication can be avoided by considering specific trajectories of the resolved variables, where such a distinction becomes irrelevant, allowing the approximate equations to be applied directly without the difficulties associated with orthogonal dynamics.

5.3   Second approximation

Assume that, for the functions on the right-hand side of equations (17), the averaging operation and the evaluation of the function commute. This means that, for example,

$$\mathbb{E}\left[(1 + q_1^2(t))^{-1} \mid q_1(0), p_1(0)\right] \approx \left(1 + \mathbb{E}[q_1(t) \mid q_1(0), p_1(0)]^2\right)^{-1}.$$

This mean-field approximation is valid when the noise is sufficiently small. In the limiting case where the noise vanishes, the approximation becomes exact. In the specific problem considered, this assumption is reasonable if the initial data are sampled from a canonical distribution at low temperature. Here, we use an initial temperature $T = 1$.

We then define

$$Q_1(t) = \mathbb{E}[q_1(t) \mid q_1(0), p_1(0)], \quad P_1(t) = \mathbb{E}[p_1(t) \mid q_1(0), p_1(0)].$$

The approximate equations of motion become

$$\frac{d}{dt}Q_1 = P_1,$$

$$\frac{d}{dt}P_1 = -Q_1\left(1 + \frac{1}{1+Q_1^2}\right) - t \cdot \frac{2Q_1^2 P_1}{(1+Q_1^2)^2}.$$

12

The system above can be solved numerically, yielding the Mori–Zwanzig approximation for our problem of interest.

References

Chekroun, M. D., Liu, H., & McWilliams, J. C. (2021). Stochastic rectification of fast oscillations on slow manifold closures. *Proceedings of the National Academy of Sciences*, *118*(48). https://doi.org/10.1073/pnas.2113650118

Chorin, A. J., & Hald, O. H. (2013). *Stochastic tools in mathematics and science*. Springer New York. https://doi.org/10.1007/978-1-4614-6980-3