Introdução ao Formalismo Mori-Zwanzig

Aluno: Lucas Amaral Taylor

Orientador: Prof. Dr. Breno Raphaldini Ferreira da Silva

Abril de 2025

1 Introdução

O principal objetivo do artigo de Chekroun, Liu e McWilliams (2021) é simplificar o modelo de Lorenz 80, preservando seu comportamento. Para isso, utilizaremos o método de Mori-Zwanzig, que é uma abordagem física-estatística aplicável em sistemas como o L80.

O método de Mori-Zwanzig, desenvolvido por Robert Walter Zwanzig e Hajime Mori na segunda metade do século XX, é utilizado em sistemas hamiltonianos. Esse método consiste em classificar as variáveis do sistema em duas categorias: "resolvidas" e "não resolvidas". As variáveis resolvidas são aquelas cujos comportamentos e valores são bem conhecidos, enquanto as não resolvidas são aquelas para as quais não se possui informações diretas. Para substituir essas variáveis não resolvidas, o método introduz termos estocásticos, denominados ruídos (*noise*), além de um termo de amortecimento (*damping*), também conhecido como termo de memória (*memory term*). Essa abordagem permite que o comportamento do sistema de interesse seja preservado de maneira adequada, mesmo sem conhecer completamente as variáveis não resolvidas.

Dada a relevância deste método para o trabalho de Chekroun, Liu e McWilliams (2021), optamos por incluir uma introdução ao formalismo de Mori-Zwanzig, a fim de proporcionar uma melhor compreensão de sua aplicação no contexto do modelo de Lorenz 80 e permitir uma base teórica sólida para eventuais explorações.

2 Prolegômenos

2.1 Escrevendo sistemas de EDO não lineares como sistemas de EDPs lineares

Considere o sistema de equações diferenciais ordinárias (EDO) dado por:

$$\frac{d}{dt}\phi(x,t) = R(\phi(x,t)), \quad \phi(x,0) = x, \tag{1}$$

onde R é uma função não linear, ϕ é uma função dependente do tempo, e R, ϕ e x podem assumir dimensões infinitas, sendo formados pelos vetores R_i, ϕ_i e x_i, respectivamente.

A partir disso, podemos definir o *Operador de Liouville* associado à equação (1) como:

$$L = \sum_i R_i(x)\frac{\partial}{\partial x_i} \tag{2}$$

Utilizando o *Operador de Liouville*, podemos transformar o sistema de EDOs não lineares em um sistema de equações diferenciais parciais (EDPs) lineares da forma:

$$u_t = Lu, \quad u(x,0) = g(x) \tag{3}$$

A solução desse sistema existe, é única, e é dada por:

$$u(x,t) = g(\phi(x,t)) \tag{4}$$

Portanto, temos que a equação (3) é bem definida[1].

2.2 Notação de semigrupo

2.3 Definição de semigrupo

Tomemos X um conjunto não vazio, dotado de uma operação binária $*$, ou seja, $X \times X \to X$, que satisfaz a propriedade de associatividade:

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

[1]Detalhes da demonstração podem ser encontrados em Chorin e Hald (2013, p. 181-182)

2.3.1 Introdução à notação

A notação de semigrupo oferece uma forma compacta e eficiente de representar soluções para equações diferenciais, particularmente as parciais ou de evolução.

Considere o operador Δ definido por:

$$\Delta\psi = \psi_{xx}, \quad \text{onde } \psi \text{ é uma função suave.}$$

Agora, considere a equação diferencial:

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

cuja solução é bem conhecida: $v(t) = v_0 e^{kt}$.

De forma análoga, considere a equação do calor:

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

onde v_t é a derivada de v em relação ao tempo e $\phi(x)$ é a condição inicial. Em vez de resolver diretamente, expressamos a solução utilizando a notação de semigrupo:

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

Aqui, $e^{\frac{1}{2}t\Delta}$ é um operador semigrupo gerado pela operação de difusão (pelo operador Δ). Ele é aplicado à condição inicial $\phi(x)$, e a solução $v(t)$ descreve a evolução temporal de $v(x,t)$ ao longo do tempo t. Essa notação permite representar soluções de equações diferenciais de maneira compacta, explorando a estrutura associativa da operação de semigrupo. Especificamente, ela satisfaz a propriedade de composição:

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

2.4 Aplicação da notação

Dada a notação de semigrupo apresentada anteriormente, aplicamos esta notação à equação (4):

$$e^{tL}g(x) = g(\phi(x,t)) \tag{5}$$

Note que $e^{tL}x$ não representa uma avaliação direta de e^{tL}, mas sim a ação do operador e^{tL} sobre o vetor formado pelos componentes x_i. Além disso, a função g comuta com a variação temporal das condições iniciais de x_i.

Vale destacar que g é uma função independente do tempo em relação às variáveis que descrevem o sistema físico, e sua variação ocorre exclusivamente devido à mudança dessas variáveis ao longo do tempo. Assim, a equação (3) pode ser expressa como:

$$Le^{tL} = e^{tL}L \tag{6}$$

Essa mesma relação se aplica a matrizes: sejam A e B duas matrizes, então a seguinte identidade é válida:

$$\exp(t(A+B)) = \exp(tA) + \int_0^t \exp\left((t-s)(A+B)\right) B \exp(sA)\, ds \tag{7}$$

Esta equação, conhecida como *Fórmula de Duhamel* ou *fórmula de Dyson*, é bem definida.

2.5 Polinômios Hermitianos

Primeiramente, definimos o produto interno como:

$$\langle u, v\rangle = \int_{-\infty}^{+\infty} \frac{e^{-x^2/2}}{\sqrt{2\pi}} u(x)v(x)\, dx \tag{8}$$

Os polinômios $p_n(x)$ e $p_m(x)$ são ortonormais em relação a esse produto interno (8) quando satisfazem a seguinte condição:

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

em que δ_{nm} é o delta de Kronecker, que apresenta as propriedades:

1. **Ortogonalidade**: Para $n \neq m$, os polinômios são ortogonais, ou seja, o produto interno entre eles é nulo:

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

2. **Normalização**: Para $n = m$, os polinômios são normalizados, de modo que o produto interno é igual a 1:

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

No caso n-dimensional, o produto interno se generaliza para:

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

De forma mais geral, se $H(q,p)$ é um Hamiltoniano, é possível definir uma família de polinômios nas variáveis q e p que sejam ortonormais com respeito à densidade canônica $Z^{-1}e^{-H/T}$. Os polinômios que satisfazem essa condição ainda são chamados de *polinômios hermitianos*.

O Hamiltoniano $H(q,p)$ descreve a energia total de um sistema físico e, em muitos casos, é expresso como:

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

onde $\frac{p^2}{2m}$ representa a energia cinética e $V(q)$ a energia potencial.

A densidade canônica $Z^{-1}e^{-H/T}$ surge no contexto da mecânica estatística e descreve a distribuição de probabilidade dos estados do sistema em equilíbrio térmico à temperatura T. Nessa configuração, dizemos que dois polinômios $f(q,p)$ e $g(q,p)$ são ortonormais quando:

$$\langle f, g \rangle = \int f(q,p)g(q,p)\, Z^{-1}e^{-H(q,p)/T}\, dq\, dp = \delta_{fg}.$$

Esse produto interno ponderado define a ortogonalidade no espaço de fases segundo a medida induzida pela densidade canônica, e é fundamental na análise de sistemas de interesse.

Por fim, para o formalismo de Mori-Zwanzig, consideraremos um espaço n-dimensional Γ com uma densidade de probabilidade dada. Dividiremos as coordenadas em dois tipos: $\hat{x}$ e $\tilde{x}$. Seja g uma função de x; então $\mathbb{P}g = \mathbb{E}[g \mid \hat{x}]$ é uma projeção ortogonal sobre o subespaço das funções de $\hat{x}$.Temos que essa projeção gera um subespaço de polinômios hermitianos que são funções de $\hat{x}$ e projetando sobre esses polinômios.

3 Mori-Zwanzig

3.1 Construção

Tomemos novamente o sistema (1), reproduzido abaixo:

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

Lembremos que a equação é composta por componentes de dimensão n. Dentre essas n componentes, definimos as primeiras m componentes de ϕ, com $m < n$, como as variáveis de interesse. Em seguida, classificamos $\hat{\phi}$ como as variáveis "resolvidas" e $\tilde{\phi}$ como as variáveis "não resolvidas":

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

O mesmo vale para x e R: $x = (\hat{x}, \tilde{x})$ e $R = (\hat{R}, \tilde{R})$. A partir das variáveis resolvidas, buscamos criar predições para o modelo de interesse, utilizando as soluções de uma parte da equação.

Com base no *Operador de Liouville* e na *notação de semigrupo*, podemos reescrever as componentes de $\hat{\phi}$ como[2]:

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

Ainda na notação de semigrupo, a equação dessas componentes é dada por:

$$\frac{\partial}{\partial t}e^{tL}x_j = Le^{tL}x_j = e^{tL}Lx \tag{9}$$

A partir da *projeção ortogonal* introduzida na seção anterior, definimos $\mathbb{P}$ como a projeção dada por: $\mathbb{P}g(x) = \mathbb{E}[g|\hat{x}]$. Assumimos que, no instante $t = 0$, conhecemos a densidade conjunta de todas as variáveis x, mas apenas os dados iniciais $\hat{x}$ são conhecidos. A densidade das variáveis em $\tilde{x}$ é, então, a densidade conjunta de todas as variáveis x com $\hat{x}$ fixado. Assim, $\mathbb{P}$ é uma projeção sobre um espaço de funções com variáveis fixas e, portanto, independente do tempo.

As projeções $\mathbb{P}\hat{\phi}(t) = \mathbb{E}[\hat{\phi}(t)|\hat{x}]$ são de nosso maior interesse, pois estimam o comportamento do sistema a partir de um conjunto reduzido de variáveis.

Definindo $\mathbb{Q} = I - \mathbb{P}$ e considerando que as seguintes propriedades são válidas para quaisquer projeções ortogonais:

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

[2] Note que cada componente depende de **todos** os valores de x. Portanto, se $\tilde{x}$ for aleatório, então $\hat{\phi}$ também será.

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

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

Podemos reescrever a equação (9) como:

$$\frac{\partial}{\partial t} e^{tL} x_j = e^{tL}\mathbb{P}Lx_j + e^{tL}\mathbb{Q}Lx_j \tag{10}$$

Utilizando agora a *fórmula de Dyson*, com $A = \mathbb{Q}L$ e $B = \mathbb{Q}L$, obtemos:

$$e^{tL} = e^{t\mathbb{Q}L} + \int_0^t e^{(t-s)L}\mathbb{P}Le^{s\mathbb{Q}L}\, ds \tag{11}$$

Pela linearidade da equação de Liouville e a partir das equações (10) e (11), obtemos:

$$\frac{\partial}{\partial t} e^{tL} x_j = e^{tL}\mathbb{P}Lx_j + e^{t\mathbb{Q}L}\mathbb{Q}Lx_j + \int_0^t e^{(t-s)L}\mathbb{P}Le^{s\mathbb{Q}L}\mathbb{Q}Lx_j\, ds \tag{12}$$

A equação acima expressa a *equação de Mori-Zwanzig*.

3.2 Análise termo a termo

3.2.1 Primeiro termo

O primeiro termo é dado por:

$$e^{tL}\mathbb{P}Lx_j \tag{13}$$

Note que:

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

Assim,

$$\mathbb{P}Lx_j = \mathbb{E}[R_j(x)|\hat{x}] \quad \text{Note que é uma função exclusiva de } \hat{x}$$

A Sistemas Hamiltonianos

B Aproximações

Referências

CHEKROUN, Mickaël D.; LIU, Honghu; MCWILLIAMS, James C. Stochastic rectification of fast oscillations on slow manifold closures. **Proceedings of the National Academy of Sciences**, Proceedings of the National Academy of Sciences, v. 118, n. 48, nov. 2021. ISSN 1091-6490. DOI: 10.1073/pnas.2113650118. Disponível em: ¡http://dx.doi.org/10.1073/pnas.2113650118¿.

CHORIN, Alexandre J.; HALD, Ole H. **Stochastic Tools in Mathematics and Science**. [S.l.]: Springer New York, 2013. ISBN 9781461469803. DOI: 10.1007/978-1-4614-6980-3. Disponível em: ¡http://dx.doi.org/10.1007/978-1-4614-6980-3¿.