



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA

YURI FALCÃO BASTOS

QUANTIZAÇÃO ESTOCÁSTICA DE NELSON E PROCESSOS ESTOCÁSTICOS

Recife

2023

YURI FALCÃO BASTOS

QUANTIZAÇÃO ESTOCÁSTICA DE NELSON E PROCESSOS ESTOCÁSTICOS

Dissertação apresentada ao Programa de Pós-graduação em Física do Centro de Ciências Exatas e da Natureza da Universidade Federal de Pernambuco, como requisito parcial para obtenção do grau de Mestre em Física em 29 de setembro de 2023. Área de concentração: Dinâmica Não-linear, Caos e Sistemas Complexos

Orientador: Antônio Murilo Santos Macêdo

Recife

2023

Catálogo na fonte
Bibliotecária Nataly Soares Leite Moro, CRB4-1722

B327q Bastos, Yuri Falcão
Quantização estocástica de Nelson e processos estocásticos / Yuri Falcão Bastos. – 2023.
93 f.: il., fig.

Orientador: Antônio Murilo Santos Macêdo.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CCEN, Física, Recife, 2023.
Inclui referências e apêndices.

1. Dinâmica não-linear, caos e sistemas complexos. 2. Probabilidade. 3. Estocástico. 4. Itô. 5. Berry. 6. Madelung. I. Macêdo, Antônio Murilo Santos (orientador). II. Título.

515.39 CDD (23. ed.) UFPE- CCEN 2023 - 188

YURI FALCÃO BASTOS

QUANTIZAÇÃO ESTOCÁSTICA DE NELSON E PROCESSOS ESTOCÁSTICOS

Dissertação apresentada ao Programa de Pós-Graduação em Física da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Física.

Aprovada em: 29/09/2023.

BANCA EXAMINADORA

Prof. Dr. Antônio Murilo Santos Macêdo
Orientador
Universidade Federal de Pernambuco

Prof. Dr. Fernando Roberto de Luna Parisio Filho
Examinador Interno
Universidade Federal de Pernambuco

Prof. Dr. Giovani Lopes Vasconcelos
Examinador Externo
Universidade Federal do Paraná

AGRADECIMENTOS

Gostaria de dedicar agradecimentos, primeiramente, a minha mãe Rita de Cássia Falcão Bastos e minha vó, Janilda Maria Falcão Bastos, por me darem todo suporte necessário na minha trajetória até então, eu não seria ninguém sem elas. Ao meu orientador, Antônio Murilo Santos Macêdo, pela paciência e por sua competência como orientador, que tornou a trajetória deste trabalho menos tortuosa. Agradeço aos meus professores da UFRPE Edgar Amorim, Pedro Hugo Figueirêdo e Antônio de Pádua Santos pelos aprendizados tão preciosos, que carregarei comigo para o resto da vida. Aos meus amigos de curso Eudes Gomes, Luan Martins, Vinícius Farah e Emanuel Pinheiro, por todo o suporte, prestatividade e companheirismo sempre que necessitei, nunca me senti só com vocês por perto. Aos meus amigos da graduação Guilherme Borges, João Victor Siqueira, Erike Simon e Maria Inês Arruda por estarem sempre me apoiando na trajetória até aqui. Ao meu amigo Leonardo Gueiros, hoje professor titular no departamento de letras da UFPB, por tanto me incentivar a adentrar a carreira acadêmica. Também aos meus amigos João Lucas Araújo, Fátima Luiza, Aline Olímpio, Ellen Bulamarque e tantos outros que, tenho certeza, sempre poderei contar.

RESUMO

Este trabalho tem como objetivo o estudo da formulação estocástica de Nelson da mecânica quântica, visando possíveis aplicações em áreas onde métodos híbridos oriundos da mecânica quântica e da física estocástica têm sido usados com sucesso, como econofísica, epidemiologia matemática e física de sistemas complexos hierárquicos. Inicialmente, realizamos uma revisão direcionada dos conceitos fundamentais da mecânica estatística fora do equilíbrio e processos estocásticos. O formalismo de Nelson é introduzido através da integral de Itô, juntamente com uma breve discussão de suas particularidades em relação à integral de Riemann, além da apresentação sucinta das condições de existência e unicidade de equações diferenciais estocásticas (EDE) e a aplicabilidade de EDE em sua forma "backward" em problemas de otimização. A partir disso, dividindo este estudo em duas partes, primeiro utilizamos o método variacional estocástico (MVE) para compreendermos como fazer a ponte entre mecânica estocástica, quântica e clássica e, em seguida, estudarmos como se dá a solução de EDEs via métodos numéricos, especialmente tratando do método de Euler-Maruyama. Por fim, na segunda parte, estudamos a obtenção e discutimos as propriedades da fase de Berry, a fim de propôr um método para a representação estocástica clássica a partir das correspondentes equações de Madelung e da equação dissipativa de Schrödinger.

Palavras-chave: probabilidade; estocástico; Itô; quantização; Berry; Madelung.

ABSTRACT

This work aims to study Nelson's stochastic formulation of quantum mechanics, with the aim of possible applications in areas where hybrid methods originating from quantum mechanics and stochastic physics have been successfully used, such as econophysics, mathematical epidemiology, and physics of complex hierarchical systems. Initially, we perform a targeted review of the fundamental concepts of non-equilibrium statistical mechanics and stochastic processes. Nelson's formalism is introduced through the Itô integral, along with a brief discussion of its peculiarities compared to the Riemann integral, as well as a succinct presentation of the conditions for the existence and uniqueness of stochastic differential equations (SDEs) and the applicability of SDEs in their "backward" form in optimization problems. Following this, dividing this study into two parts, we first use the stochastic variational method (SVM) to bridge the gap between stochastic, quantum, and classical mechanics, and then investigate how the solution of SDEs is achieved through numerical methods, particularly focusing on the Euler-Maruyama method. Finally, in the second part, we study the derivation and discuss the properties of the Berry phase in order to propose a method for classical stochastic representation based on the corresponding Madelung equations and the dissipative Schrödinger equation.

Keywords: probability; stochastic; Ito; quantization; Berry; Madelung.

SUMÁRIO

1	INTRODUÇÃO	9
2	INTRODUÇÃO AO FORMALISMO	14
2.1	MECÂNICA ESTATÍSTICA FORA DO EQUILÍBRIO	14
2.2	PROCESSOS ESTOCÁSTICOS	16
2.2.1	Cadeias de Markov	19
2.2.2	Martingais	22
2.2.3	Movimento browniano e processos de Wiener	25
2.3	CÁLCULO ESTOCÁSTICO DE ITÔ	27
2.3.1	Teorema da variação quadrática	27
2.3.2	Lema de Itô	29
2.3.3	Integral de Itô	32
2.4	EQUAÇÃO DE FOKKER PLANCK	38
2.5	EQUAÇÕES DIFERENCIAIS ESTOCÁSTICAS (EDE)	42
2.6	EQUAÇÕES DIFERENCIAIS ESTOCÁSTICAS "BACKWARD" (BSDE)	45
3	QUANTIZAÇÃO ESTOCÁSTICA	50
3.1	MECÂNICA DE NELSON	50
3.2	EQUAÇÕES DE MADELUNG	55
3.3	MÉTODO VARIACIONAL ESTOCÁSTICO	56
3.3.1	Dinâmica dos processos de difusão	57
3.3.2	Métodos numéricos	59
3.3.3	Aplicação do método: processo de Ornstein-Uhlenbeck	60
4	FASE DE BERRY	63
4.1	TEOREMA ADIABÁTICO	63
4.2	FASE DE BERRY	66
4.3	FASE DE BERRY A PARTIR DA FORMULAÇÃO POR INTEGRAIS DE CAMINHO	69
4.4	REPRESENTAÇÃO ESTOCÁSTICA DA FASE DE BERRY	72
5	CONCLUSÕES E PERSPECTIVAS	74
	REFERÊNCIAS	75

APÊNDICE A – REVISÃO DE CONCEITOS BÁSICOS DE PROBABILIDADE	78
APÊNDICE B – CONCEITOS DE GEOMETRIA DIFERENCIAL . .	83
APÊNDICE C – EQUAÇÕES DIFERENCIAIS PARCIAIS PARABÓLICAS (EDP PARABÓLICA)	91

1 INTRODUÇÃO

Nesta dissertação de mestrado, exploraremos os processos estocásticos e sua aplicação na quantização estocástica de Nelson. Nosso objetivo principal é investigar, analisar e compreender a natureza destes processos, suas propriedades e algumas aplicações. Temos dois objetivos principais: aprofundar nosso entendimento das soluções numéricas para equações diferenciais estocásticas, empregando o poderoso método variacional estocástico, e propor um método para quantização estocástica da fase de Berry.

Iniciamos este trabalho focados na compreensão dos sistemas mecânicos interagentes, particularmente aqueles que envolvem um grande número de partículas, que desafiam a nossa intuição. Quando nos deparamos com sistemas fora do equilíbrio, múltiplas forças atuam simultaneamente, fazendo com que o sistema evolua para exibir comportamentos complexos. Nesses casos, as equações de movimento fenomenológicas, como a equação de Langevin, revelam uma divisão entre uma parte determinística (o termo de fricção) e uma parte aleatória (a força flutuante). Isto torna a evolução do sistema imprevisível devido às influências da força aleatória. Além da descrição fenomenológica fornecida pela equação de Langevin, existem questões mais profundas a serem exploradas, como o papel da reversão temporal em sistemas que operam sob o regime de movimento browniano. Essa complexidade, originada da interação entre partículas submetidas a forças externas nos sistemas fora do equilíbrio fornece um campo fértil para estudos e pesquisas. Neste trabalho, pretendemos discutir e compreender estes fenômenos à luz dos processos estocásticos, explorando os conceitos fundamentais que regem tais processos.

O estudo do movimento browniano, cujo nome se deve às observações de Robert Brown em 1827, é um tema de profundo interesse na teoria de probabilidade. Esse fenômeno, que descreve o comportamento aparentemente aleatório de partículas em suspensão, é fundamentalmente um processo estocástico. O movimento de cada partícula em um fluido está associado a uma curva de probabilidade que descreve a probabilidade de diferentes eventos, como a partícula se mover de um ponto para outro. No entanto, a compreensão desse processo estocástico requer uma análise cuidadosa das características subjacentes, que envolvem incrementos independentes e distribui-

ções normais. Norbert Wiener, em 1918, desenvolveu uma formulação rigorosa do movimento browniano em termos de processos estocásticos de tempo contínuo, conhecidos como processos de Wiener. Essa dissertação explorará as propriedades e aplicações do movimento browniano, abordando a necessidade de um formalismo matemático, como o cálculo de Itô, para lidar com sua continuidade e integração.

O cálculo de Itô é uma ferramenta essencial na análise de processos estocásticos, desempenhando um papel fundamental na compreensão do movimento browniano. Ele nos permite lidar com a natureza não suave das trajetórias em processos de Wiener, graças a um dos pilares desse cálculo: o teorema da variação quadrática. Este demonstra que as funções continuamente diferenciáveis não acumulam variação quadrática ao longo de suas trajetórias, enquanto os processos de Wiener o fazem. Isso levanta a necessidade de uma nova abordagem para a integração, levando-nos à integral de Itô. Esta ferramenta poderosa permite a análise de flutuações locais em processos estocásticos, medindo a magnitude cumulativa das diferenças entre os valores sucessivos desses processos. No contexto da equação de Liouville, lidar analiticamente com sistemas que envolvem ruído é bastante desafiador, especialmente em casos não lineares. Neste ponto entra em cena a equação de Fokker-Planck, uma extensão da equação de Liouville que lida com o comportamento estatístico desses sistemas. Este trabalho explora a construção da equação de Fokker-Planck a partir da fórmula de Feynman-Kac, revelando como a fricção e o ruído afetam a evolução das funções de distribuição. Essa equação é um poderoso instrumento na modelagem de sistemas complexos, sendo essencial para compreender uma ampla gama de fenômenos físicos, químicos e biológicos.

As equações diferenciais estocásticas (EDE) desempenham um papel crucial na modelagem de sistemas dinâmicos, sujeitos a influências estocásticas. Este estudo visa explorar a resolução dessas equações, que muitas vezes não admitem soluções fechadas, tornando sua resolução uma tarefa desafiadora. Este estudo começa explorando a solução analítica de EDEs lineares, concluindo que seguem sempre a mesma estrutura. Contudo, EDEs mais complexas necessitam abordagens numéricas para serem solucionadas. Para entender as mais variadas formas que EDEs podem ser encontradas, adentramos numa discussão sobre o intrigante mundo das equações diferenciais estocásticas "backward" (BSDE). Essas equações desempenham um papel vital em várias aplicações, incluindo o controle estocástico, o mercado financeiro e a

solução de equações diferenciais parciais (EDP) por meio de métodos numéricos. Enquanto as BSDEs compartilham raízes teóricas com as equações diferenciais estocásticas (EDE), sua conexão não é direta, e este estudo explora suas nuances. Escolhemos começar expondo-as como uma ferramenta para caracterizar soluções de EDEs e EDPs, construindo uma ponte entre esses domínios com a equação de Kolmogorov, concluindo que as BSDEs são versões não lineares do teorema de representação de martingais. Em seguida, exploramos aplicações de controle ótimo, que desempenham um papel essencial na busca da melhor estratégia sob incerteza. Muitos problemas de controle ótimo são representados por BSDEs e solucionados via métodos numéricos, veremos um exemplo destes no capítulo 4. Esta jornada no formalismo estocástico forma a base teórica imprescindível para a introdução da quantização estocástica de Nelson.

Tradicionalmente, a quantização envolve a promoção de variáveis clássicas a operadores quânticos, mas a abordagem estocástica estende esse conceito, introduzindo variáveis estocásticas e variáveis aleatórias na evolução dos campos. A mecânica de Nelson, desenvolvida a partir das ideias de Imre Fényes, é uma teoria fundamental que explora a conexão entre a mecânica quântica e a estocasticidade. Este estudo apresenta a cinemática e a dinâmica dos processos de difusão, abrindo novas possibilidades de interpretação para os processos estocásticos e quânticos, fazendo a ligação entre o formalismo de Nelson e de Schrödinger. Ademais, abordamos a decomposição de Madelung das equações de Schrödinger e o método variacional estocástico (MVE) como uma maneira de unificar a mecânica clássica com a mecânica quântica. A decomposição de Madelung desmembra a equação de Schrödinger em duas equações, uma relacionada ao módulo da função de onda e outra relacionada à fase da função de onda. Essas são as chamadas "equações de Madelung". O MVE é uma abordagem que busca descrever tanto comportamentos clássicos quanto quânticos de maneira unificada. Isso é feito por meio da quantização estocástica, que permite descrever sistemas quânticos de forma precisa, mesmo sem recorrer à equação de Schrödinger, demonstrado que podemos obtê-la a partir da dinâmica dos processos de difusão, na mecânica de Nelson, considerando sua conexão com o princípio da mínima ação. A otimização da lagrangiana em processos de difusão desempenha um papel crucial nessa unificação. A ação estocástica é um conceito central, onde a variação estocástica é expressa por meio de derivadas estocásticas. A equação de Euler-Lagrange

é generalizada na forma estocástica e, através de cálculos detalhados, chega-se a uma forma que se assemelha à equação de Schrödinger. A aplicação dessas teorias e métodos numéricos, como o Método de Euler-Maruyama, se mostra essencial para explorar essas conexões e proporciona uma abordagem promissora para a resolução de problemas complexos da física quântica. Em (KOPPE; GRECKSCH; PAUL, 2017.), a autora expõe um exemplo deste tipo de problema, descrito como uma EDE e uma BSDE acopladas, isto é, uma equação diferencial estocástica "forward backward". A resolução numérica de FBSDEs pode ser implementada com técnicas de simulação de Monte Carlo. Esse enfoque permite não apenas obter soluções numéricas de sistemas simples, mas também explorar o potencial da quantização estocástica em contextos mais amplos, com ingredientes de geometria diferencial, como a fase de Berry, revelando aplicações inovadoras dessa abordagem na física contemporânea.

A mecânica quântica e a teoria dos vetores de estado desempenham um papel fundamental na descrição de sistemas físicos complexos. Quando esses vetores de estado percorrem trajetórias fechadas no espaço de parâmetros, sob um regime adiabático, uma notável transformação ocorre: eles adquirem uma fase puramente geométrica, denominada "fase de Berry". Esta dissertação investigando sua obtenção e formalização no contexto da mecânica de Nelson. O percurso se inicia com uma introdução ao teorema adiabático, estabelecendo os fundamentos necessários, seguido pela definição e propriedades da fase de Berry, que se revela como um observável intrinsecamente relacionado ao espaço de configuração. A análise aprofundada da fase de Berry e sua conexão com a geometria do espaço de parâmetros proporciona uma visão fundamental sobre as propriedades dos sistemas físicos, enriquecendo nosso entendimento das complexas dinâmicas quânticas. O objetivo principal desta parte é abordar a fase de Berry sob uma perspectiva inovadora, empregando o arcabouço da mecânica de Nelson. Um teorema fundamental, que conecta a solução da equação de Schrödinger à nova abordagem proposta, é apresentado. Além disso, equações de Madelung alternativas são obtidas, lançando luz sobre aspectos "invisíveis" do formalismo padrão. Concluímos que essa pesquisa promete revelar novas facetas na descrição da fase de Berry e enriquecer nosso entendimento das dinâmicas quânticas sob a perspectiva da mecânica estocástica.

Este trabalho pretende proporcionar uma visão abrangente do universo dos processos estocásticos, a partir da quantização estocástica de Nelson. Ao fazê-lo, es-

peramos contribuir para uma compreensão melhor das complexas interações entre a aleatoriedade, fenômenos clássicos e quânticos, bem como a importância do desenvolvimento de ferramentas numéricas eficazes na resolução de equações diferenciais estocásticas. Organizamos este trabalho partindo a introdução do formalismo da mecânica estocástica e do cálculo de Itô, seguido do capítulo 3 tratando da quantização estocástica de Nelson, fazendo a ponte entre o formalismo deste com o de Schrödinger e estudando métodos numéricos para solução de EDEs. Por fim, tratamos da fase geométrica de Berry, que é obtida através do formalismo quântico, mas buscamos obtê-la numa descrição estocástica, via equações de Madelung.

2 INTRODUÇÃO AO FORMALISMO

Este formalismo contém a "espinha dorsal" dos cálculos que envolvem processos estocásticos. Estes processos operam no âmbito da mecânica estatística fora do equilíbrio, a qual fazemos uma breve discussão qualitativa, tocando nos pontos de interesse deste trabalho. Feita esta contextualização, iniciamos a fundamentação teórica dos processos estocásticos, que segue com o estudo das cadeias de Markov, martingais e movimento browniano.

O cálculo de Itô emerge da necessidade de lidar especificamente com equações que possuam variáveis estocásticas, já que estas possuem características específicas e produzem efeitos não são acomodados pelo cálculo convencional, por consequência, isso muda como devemos pensar em integração e diferenciação destas variáveis. Iniciamos abordando o teorema da variação quadrática, importante para construirmos o lema de Itô e após isso, faço uma discussão comparativa com a integral de Riemann. Ao longo do restante do capítulo, abordamos a aplicabilidade da integral de Itô para solução de EDPs e faço uma conexão entre estas e EDEs (equações diferenciais estocásticas) e BSDEs (equações diferenciais estocásticas "backward").

O texto deste capítulo foi baseado nas seguintes referências: (ZWANZIG, 2001.), (REICHL, 1998.), (SCHILLING; PARTZSCH; BÖTTCHER, 2012.), (COHEN; ELLIOTT, 2015.), (FLEMING; RISHEL, 1975.a), (NELSON, 2012.), (NELSON, 1985.), (KARLIN; TAYLOR, 1974.), (BAUDOIN, 2012.), (MA; YONG, 2007.), (FITZSIMMONS, 2004.), (GUBINELLI, 2020.), (HELM, 1961.), e (WANG, 2022).

2.1 MECÂNICA ESTATÍSTICA FORA DO EQUILÍBRIO

Um sistema mecânico envolvendo duas partículas interagentes é algo fácil de resolver, se aumentarmos para cinco talvez se torne difícil, mas se aumentarmos para 10^{24} , fica fácil novamente. Quando lidamos com sistemas deste último tipo, não é interessante conhecer o comportamento de cada partícula envolvida no sistema, pois seria impossível e, principalmente, inútil. A perspectiva de análise que importa nesta situação é conhecer a qual configuração tende este sistema. Por exemplo: imagine que injetamos um gás qualquer em uma caixa (considerada um sistema isolado), com

uma barreira no meio que a divide em duas partes isoladas. Intuitivamente, sabemos que ao retirarmos a barreira, as moléculas logo se organizarão numa configuração aproximadamente uniforme, preenchendo toda a caixa. Moral da história: sistemas muito complicados podem se tornar simples, a depender da maneira como lidamos com tais problemas. Motivados por essa conclusão, nos voltamos para a estatística como ponto de partida para análise desse tipo de sistema.

Ao estudarmos qualquer livro tratando de mecânica estatística no equilíbrio, há dois aspectos que permeiam todas as discussões: grande número de partículas e noção de equilíbrio. No caso de um ensemble micro-canônico, por exemplo, ao analisarmos uma camada de energia isolada no espaço de fases, a noção de equilíbrio deste sistema vem da associação entre a integração de um observável (energia, magnetização, volume etc) cujos parâmetros dependem do tempo, ao longo de um grande intervalo de tempo, com seu valor médio

$$\lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^{\tau} A(q, p) dt \longleftrightarrow \langle A \rangle_{\text{micro-canônico}}. \quad (2.1)$$

Essa relação explicita algo fundamental sobre a estatística: Não é mais preciso resolver a dinâmica do sistema para compreender sua configuração, basta calcular a média em uma camada de energia. Neste referencial de análise, a noção de tempo é substituída por algo puramente estatístico.

Nos sistemas fora do equilíbrio, há várias forças atuando continuamente no sistema, ao mesmo tempo, fazendo com que este evolua a ponto de exibir um comportamento complexo. Se fizermos uma análise clássica do movimento de uma partícula num sistema fora do equilíbrio, a equação de movimento seria similar a uma equação de Langevin

$$m \frac{dv}{dt} = -\zeta v + \delta F(t), \quad (2.2)$$

onde o parâmetro ζ é o coeficiente de fricção e $\delta F(t)$ é a resultante de todas as forças atuando aleatoriamente sob a partícula, também chamada de força flutuante. Em outras palavras, a força total foi particionada em uma parte determinística (termo de fricção) e outra aleatória (termo de ruído). Nesse tipo de sistema, a relação (2.1) deixa de valer, já que evolui de forma imprevisível devido a $\delta F(t)$. O primeiro e segundo momentos, calculados como médias temporais sobre um intervalo de tempo infinitesimal,

sumarizam os efeitos deste termo

$$E[\delta F(t)] = 0, \quad (2.3)$$

$$E[\delta F(t)\delta F(t')] = 2\alpha\delta(t - t'), \quad (2.4)$$

onde α é a medida de intensidade da força flutuante. A função delta $\delta(t - t')$ indica que não há correlação entre impactos em quaisquer intervalos de tempo não nulos entre t e t' . A equação de Langevin fornece uma explicação meramente fenomenológica da mecânica estatística fora do equilíbrio. No entanto, existem outras questões mais profundas a serem tratadas, como veremos nos capítulos posteriores. Por exemplo, a dinâmica de sistemas que operam sob regime browniano carrega consigo uma noção de reversão temporal.

2.2 PROCESSOS ESTOCÁSTICOS

Podemos compreender um processo estocástico como uma coleção de variáveis aleatórias, indexadas pelo tempo, em que a cada instante de tempo, estas variáveis fornecem uma distribuição de probabilidade sobre um espaço de caminhos. Em outras palavras: um processo estocástico descreve uma variável aleatória se movimentando ao longo do tempo. Estes processos podem ser definidos (ver Apêndice A) em um espaço de probabilidade (Ω, \mathcal{F}, P) , em que a variável aleatória é um mapa do espaço amostral para os reais ($x: \Omega \rightarrow \mathcal{R}$) e um processo estocástico é o mapa $\{x_t\}_t: t \times \Omega \rightarrow \mathcal{R}$, sendo t o conjunto de instantes de tempo $\{1, 2, \dots, T, \dots, \infty\}$. Para compreender o comportamento dinâmico de tais processos, precisamos primeiro nos perguntar:

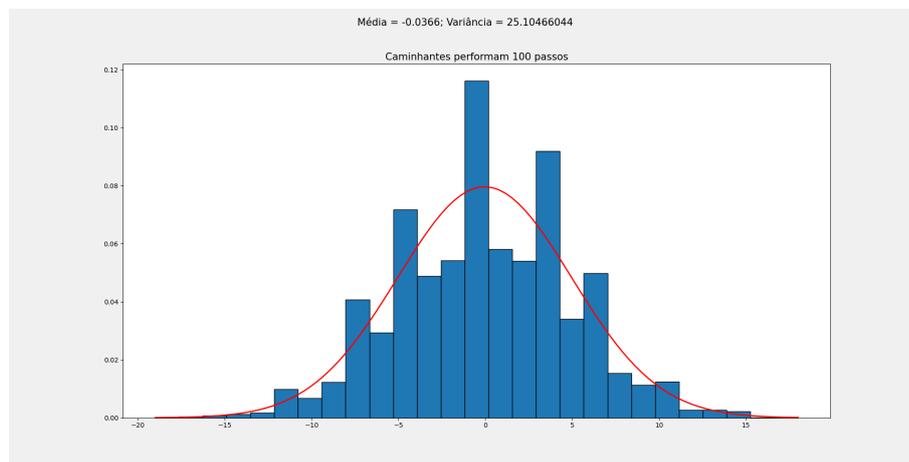
1. Quais são as interdependências entre os valores da sequência?
2. Qual é o comportamento da sequência a longo prazo?
3. Quais são as condições de contorno?

Como ponto de partida, vejamos um exemplo de processo estocástico amplamente estudado: a *caminhada aleatória unidimensional*. Um caminhante "bêbado" pode, aleatoriamente, dar um passo de mesmo tamanho para direita ou para esquerda e seguirá fazendo isso sucessivamente, a partir do ponto que ele estiver após o passo anterior. Supondo que temos um conjunto de estados ν , definidos no espaço amostral

das posições \mathcal{F} que define uma σ -álgebra, no qual cada caminhante percorre este espaço a partir de uma variável aleatória x_{t_i} , em que t indica o instante de tempo e i o caminhante em questão, que pode assumir os valores 1 (passo para direita) ou -1 (passo para esquerda), de maneira equiprovável. A probabilidade associada ao conjunto dos estados é igual a 1 ($P(v) = 1$).

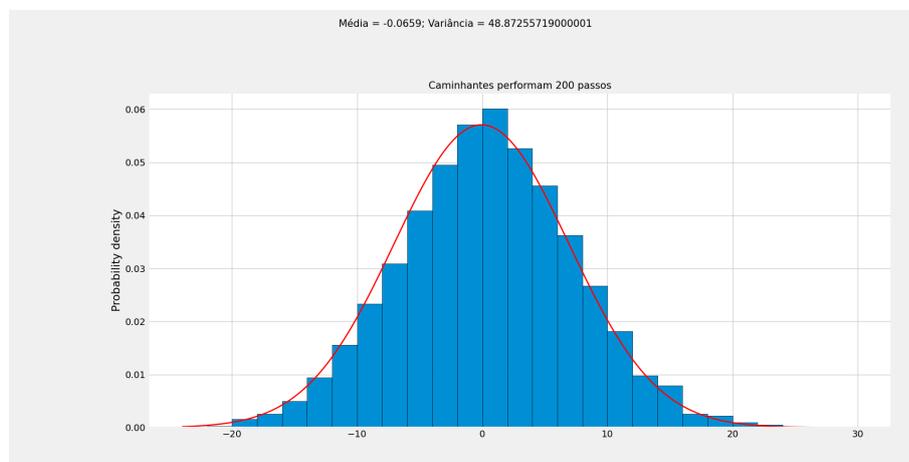
Podemos representar tal processo num algoritmo numérico. Estabelecendo a quantidade de passos e caminhantes, um contador de passos e associar uma variável aleatória a cada passo. Definindo 10.000 caminhantes aleatórios e, respectivamente, 100, 200 e 1000 passos, podemos obter os histogramas a seguir da densidade de probabilidade com relação a posição dos caminhantes:

Figura 1 – Caminhada aleatória de 100 passos



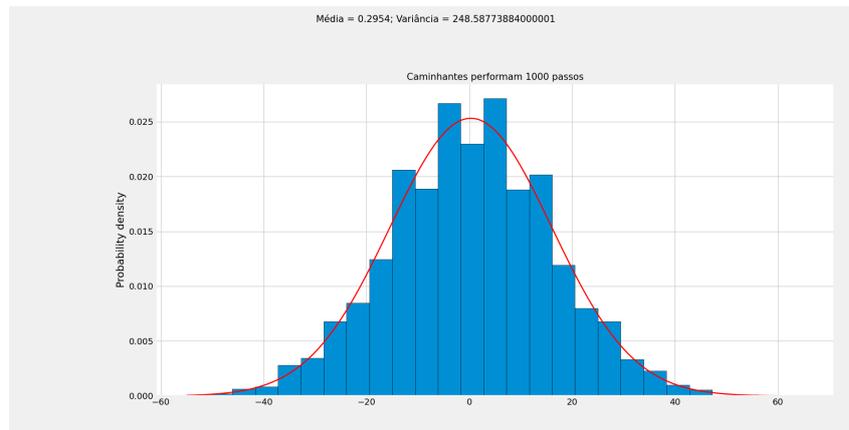
Fonte: O autor (2023).

Figura 2 – Caminhada aleatória de 200 passos



Fonte: O autor (2023).

Figura 3 – Caminhada aleatória de 1000 passos



Fonte: O autor (2023).

A partir da análise desses histogramas, podemos concluir algumas propriedades desse processo:

1. A soma das t variáveis aleatórias independentes x_t , com qualquer distribuição e variâncias semelhantes, é uma variável com distribuição que se aproxima da distribuição de Gauss (distribuição normal) a medida que t aumenta (teorema central do limite);
2. O valor esperado da variável aleatória é aproximadamente nulo: $E[x] \approx 0$; são mutuamente independentes;
3. O processo é estacionário.

Note que este processo é discreto no tempo. Processos estocásticos discretizados são solúveis, em geral, via métodos numéricos. Já com relação aos contínuos, é preciso cautela. Por conseguinte, dado que o objetivo principal deste trabalho é o estudo dos processos estocásticos, o dividimos em duas partes: a primeira, referente aos discretos no tempo, buscamos entender como solucionar uma equação estocástica discretizada via métodos numéricos. Por fim, iremos propor um método para a descrição estocástica de um objeto puramente geométrico, a fase de berry. Para cumprir tais objetivos, se faz necessário introduzir dois tipos de processos estocásticos, que formam a base teórica para o estudo de ambas as partes: martingais (partindo de cadeias de markov) e processos de Wiener.

2.2.1 Cadeias de Markov

Definição: Um processo estocástico discreto x_0, x_1, \dots é uma cadeia de Markov se

$$P(x_0, x_1, \dots, x_t | x_{t+1}) = P(x_t | x_{t+1}); \forall t \geq 0, \quad (2.5)$$

onde $P(x_t | x_{t+1})$ é a probabilidade condicional de x_{t+1} ocorrer, dado que x_t ocorreu.

Em suma, dado um instante de tempo t e uma variável estocástica x_t , o processo correspondente é uma cadeia de Markov se, para inferir algo sobre o futuro do processo, é possível ignorar tudo que veio antes de t e considerar apenas o presente. Para que tais processos ocorram, é preciso uma condição inicial e um espaço de estados, no qual há probabilidades de transição envolvidas em cada mudança entre estados. Um processo markoviano obedece a equação de Smoluchowski-Chapman-Kolmogorov ($q > t > s$):

$$P(x_q | x_s) = \int_{-\infty}^{\infty} P(x_q | x_t) P(x_t | x_s) dx_t. \quad (2.6)$$

Ambas as condições são necessárias para definir cadeias de Markov. Processos sob regime Markoviano são chamados de "não-antecipativos", isto é, seu comportamento futuro depende apenas do estado atual, independente dos anteriores. Esta propriedade, que caracteriza todos estes processos, é chamada *Propriedade de Markov*, tal que a equação (2.5) é uma representação discretizada desta propriedade e (2.6) a versão contínua.

Ademais, podemos descrever estes processos de maneira mais simples, num formato de somatório de probabilidades. Imagine que o processo estocástico x_i tem valores contidos num espaço de estados $S = [n]$ (conjunto finito), sendo assim:

$$P_{ij} = P(x_{t+1} = j | x_t = i); \forall i, j \in S, \quad (2.7)$$

$$\sum_j P_{ij} = 1. \quad (2.8)$$

Podemos então definir uma matriz de transição P_{ij} , que descreve todas as probabilidades do sistema transicionar de $i \rightarrow j$. Em geral, dada essa matriz de transição, o teorema de Perron-Frobenius (HORN; JOHNSON, 2013.) garante que: se todas as entradas da matriz forem positivas, existe um autovetor que caracteriza o comportamento a longo prazo da cadeia. Este vetor, chamado *distribuição estacionária* (representado

por π), introduz uma característica extraordinária das cadeias de Markov: a reversibilidade temporal.

Para uma cadeia de Markov positiva $\{x_t : t \in \mathbb{N}\}$, com matriz de transição P e distribuição estacionária π , seja $\{x_t^* : t \in \mathbb{N}\}$ a versão estacionária da cadeia, isto é, aquela que $x_0 \sim \pi$. Acontece que podemos estender esse processo para um tempo t que também assume valores negativos, $\{x_t^* : t \in \mathbb{Z}\}$, portanto para cada tempo $t \in \mathbb{Z}$ dado que a distribuição é estacionária, estendemos a compreensão que $x_0 \sim \pi$ para $P(x_t^* = j) = \pi_j$. Podemos interpretar isto como uma maneira de assumir que a cadeia inicia em $-\infty$ (no passado). A propriedade de Markov declara que o "futuro é independente do passado, dado o estado presente", mas para o caso de x_t^* , podemos re-interpretar esta declaração no sentido contrário, sem perda de significado: "o passado é independente do futuro, dado o estado presente". Isso significa que o processo $x_t = x_{-t}^* ; t \in \mathbb{N}$ ainda é uma cadeia de Markov estacionária, apenas compreendida com o tempo revertido.

Considerando P_{ij} a probabilidade de transição em tempo revertido, podemos computar cada uma destas em termos de π e P :

$$\begin{aligned} P_{ij} &= P(x_1 = j | x_0 = i) = P(x_0^* = j | x_1^* = i) \\ &= P(x_1^* = i | x_0^* = j) \frac{P(x_0^* = j)}{P(x_1^* = i)} \\ &= \frac{\pi_j}{\pi_i} P_{ji}. \end{aligned}$$

Então a cadeia de Markov de tempo revertido é uma cadeia de Markov com probabilidade de transição dada por

$$P_{ij} = \frac{\pi_j}{\pi_i} P_{ji}. \quad (2.9)$$

Isso nos leva a seguinte conclusão: Uma cadeia de Markov positiva, com matriz de transição P e distribuição estacionária π é reversível no tempo se a cadeia estacionária de tempo reverso $\{x_t : t \in \mathbb{N}\}$ possui a mesma distribuição que a de tempo normal $\{x_t^* : t \in \mathbb{N}\}$, isto é, satisfaz: $\pi_i P_{ij} = \pi_j P_{ji}$. Posto desta forma, parece que precisamos ter a distribuição estacionária a nossa disposição para checar se a cadeia é temporalmente reversível, mas não necessariamente. Fixando o estado j e somando sobre i , temos

que

$$\begin{aligned}\sum_i \pi_i P_{ij} &= \sum_i \pi_j P_{ji} \\ &= \pi_j \sum_i P_{ji}\end{aligned}$$

de (2.8), temos então

$$\sum_i \pi_i P_{ij} = \pi_j \quad (2.10)$$

$$\pi P = \pi. \quad (2.11)$$

Isso significa que se, para uma cadeia de Markov irredutível, existe um π para o conjunto de equações $\pi_i P_{ij} = \pi_j P_{ji}$, para todos os estados i, j , então a cadeia é positiva, temporalmente reversível e a solução π é única para a distribuição estacionária. A grande importância desta discussão é concluirmos que as equações temporalmente reversíveis são mais simples de serem solucionadas que as (2.11). Se pudermos obter ou já soubermos quais são as cadeias reversíveis, estas serão sempre mais fáceis de resolver.

Para finalizar este tópico, voltamos nossa atenção a uma classe de processos de Markov chamada de difusão. Sistemas sob regime markoviano geralmente estão associados a uma perturbação por flutuações independentes, que produzem o que é chamado processo de difusão. No capítulo 5 de (REICHL, 1998.), é demonstrado que a caminhada aleatória ao ser formulada como uma cadeia de Markov, considerando o limite em que o tamanho do passo e o tempo entre cada um tendem a zero, nos leva a equação de difusão

$$\frac{\partial P(x_t)}{\partial t} = D \frac{\partial^2 P(x_t)}{\partial x^2}. \quad (2.12)$$

Queremos resolver (2.12) para o caso $P(x_0) = \delta(x)$. Ao introduzirmos a transformada de Fourier

$$\tilde{P}(k_t) = \int_{-\infty}^{\infty} P(x_t) e^{ikx} dx$$

em que, neste caso: $\tilde{P}(k_0) = 1$. Utilizando este fato, a equação de difusão toma a forma

$$\frac{\partial \tilde{P}(k_t)}{\partial t} = -Dk^2 \tilde{P}(k_t),$$

cuja solução é trivial:

$$\tilde{P}(k_t) = e^{-Dk^2 t}.$$

Por fim, tomando a transformada inversa, obtemos

$$P(x_t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ikx} e^{-Dk^2 t} dk$$

$$P(x_t) = \frac{e^{-\frac{x^2}{4Dt}}}{\sqrt{4\pi Dt}}. \quad (2.13)$$

A equação acima nos permite uma conclusão importante: Se o movimento de um sistema pode ser descrito como o somatório de incrementos independentes (regime markoviano), no qual a estatística sobre curtos períodos de tempo possui variância bem definida, a caminhada aleatória resultante converge para um processo de difusão com distribuição normal. Processos de difusão, portanto, são processos markovianos nos quais as amostras de caminho $\{x_t\}$ são funções contínuas no tempo. Interpretamos o regime markoviano, neste caso, como uma garantia de que o sistema está completo no estado x_t , isto é: no caso de seu estado conter toda a informação sobre o passado que seja relevante para prever o futuro.

2.2.2 Martingais

Martingais possuem um papel fundamental no estudo de processos estocásticos por diversas razões, oferecendo um conjunto de ferramentas conceitualmente rico, permitindo a modelagem, caracterização e estudo de fenômenos aleatórios em diversas áreas das ciências naturais, matemática e principalmente finanças.

Um processo estocástico $F_t = x_0, x_1, \dots, x_t$ é um martingal se, partindo do tempo t no presente, ao tomar o valor esperado de $t + 1$, este deve ser exatamente igual a x_t , $\forall t \geq 0$

$$x_t = E[x_{t+1}|F_t]; \quad E[|x_{t+1}|] < \infty. \quad (2.14)$$

Em outras palavras: um martingal é todo processo estocástico no qual, se as probabilidades envolvidas nas transições são iguais, então em média tende-se a voltar para o ponto inicial. O martingal capta esta propriedade de "jogos justos", permitindo analisar processos aleatórios em que não podemos, consistentemente, prever resultados futuros, por exemplo: a caminhada aleatória unidimensional é um martingal, já o balanço de um jogador de caça-níquel não é um martingal. Podemos pensar na equação (2.14) da seguinte forma: o valor esperado do valor no futuro, condicionado às informações disponíveis no presente, é dado pelo valor atual.

Utilizamos martingais na filtragem de informação, ou sequência de informações. Um espaço de probabilidade que contenha, também, uma coleção crescente de sub- σ -álgebras de $\mathcal{F}: \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \mathcal{F}_3 \subseteq \dots \subseteq \mathcal{F}$ é chamado de espaço de filtração, em que

$$\mathcal{B} = \{F_t\}_{t \geq 0}, \quad (2.15)$$

$$\mathcal{P} = \{F_t\}_{t \leq 0} \quad (2.16)$$

são chamados, respectivamente, processo de sub-filtração e superfiltração. Intuitivamente, ambos dispõem uma sequência de processos estocásticos, $\mathbb{X} = \{x_t\}$, como a informação do sistema filtrada e acumulada ao longo do tempo, de forma discreta, sendo uma passagem de tempo direta e outra inversa. Ademais, a coleção de processos \mathbb{X} é dito \mathcal{P} -adaptado, ou \mathcal{B} -adaptado se, para todo $n \geq 0$, x_t é \mathcal{F} -mensurável.

Ao tratarmos de processos de longa duração, é imprescindível que relacionemos o martingal em questão a uma dada filtragem, para que os critérios de convergência a seguir sejam garantidos:

1. Existe um valor constante $a > 0$, tal que $|x_t| \leq a \forall t$;
2. Dada uma variável aleatória x , uma sequência x_t converge para x a medida que $t \rightarrow \infty$.

A condição 1 garante que o martingal não "exploda". Isso se provará fundamental para integração de processos estocásticos.

Suponha agora um jogo de apostas, não necessariamente honesto, em que um jogador possui um montante " $\xi(t)$ ", que funciona exatamente como um processo estocástico. Se o jogo for desfavorável, isto é, a chance de haver perda do montante é maior, temos:

$$E[\xi(t)|\mathcal{B}_s] \leq \xi(s); \quad s \leq t, \quad (2.17)$$

este processo é chamado de sub-martingal. O oposto

$$E[\xi(t)|\mathcal{P}_s] \geq \xi(s); \quad s \leq t, \quad (2.18)$$

é chamado de super-martingal. É importante notar que sub e super-martingais não são martingais e, mesmo sob condições gerais, não se tornam automaticamente martingais. Em geral, sub-martingais exibem tendência de crescer ao longo do tempo,

vice versa para super-martingais, o que indica um desvio nos valores esperados. A seguir, introduzimos duas categorias de martingais imprescindíveis para a construção do formalismo estocástico.

Definição: Um processo estocástico $\xi(t)$ é chamado de semi-martingal se este pode ser decomposto na soma de duas componentes: um martingal $M(t)$ cujo valor esperado pode ter saltos ocasionais no seu valor esperado, chamado de martingal local, e um processo estocástico $A(t)$ com variações finitas, limitadas por um intervalo de tempo finito, chamado processo de variação finita:

$$\xi(t) = M(t) + A(t). \quad (2.19)$$

Esta decomposição de um semi-martingal possibilita a análise das diferentes componentes do processo, dividido em parte determinística e estocástica. Muitos processos estocásticos são semi-martingais, incluindo o movimento browniano e o movimento browniano geométrico, e são extensamente utilizados em finanças para modelar preços de ativos, taxas de interesse e outras variáveis financeiras sujeitas a uma dinâmica regida por algo como (2.19). Por fim, cabe discorrer sobre tempo de parada.

Definição: Suponha um espaço filtrado $(\Omega, \mathcal{F}, \mathbb{P}, \{F_t\}_{t \in \mathbb{T}})$. A variável aleatória $\tau: \Omega \rightarrow \mathbb{T} \cup \{\infty\}$ é chamado tempo aleatório. Este tempo é chamado de tempo de parada com relação a filtração se

$$\{\tau \leq t\} = \mathcal{F}_t; \quad \forall t \in \mathbb{T}. \quad (2.20)$$

Intuitivamente, pense no tempo de parada como o evento τ que já ocorreu antes de t e, por conseguinte, é conhecido em t . Imagine que você lance uma moeda repetidamente, com filtração dada pelas observações da sequência de resultados, o tempo da terceira cara observada é um tempo de parada. O tempo da "última coroa antes da quinta coroa" não é um tempo de parada, pois você precisaria saber o valor do próximo lançamento. Por fim, a partir da definição de tempo de parada, temos o seguinte

Definição: Processos adaptados a filtração F_t , M_t , são chamados *martingais locais* se dada uma sequência τ_n de tempos de parada, os processos encerrados M_{τ_n} é também um martingal. Esta definição é similar para sub e super-martingais.

Todo martingal é um martingal local, mas a recíproca é falsa. Tempos de parada são usados, em geral, para truncar as trajetórias de um processo, para que se encerrem em algum ponto, ganhando mais "integrabilidade", enquanto se mantém adaptado

a filtração e garante sua continuidade. Veremos que, ao inserimos um termo de deriva em uma equação diferencial, a transformamos em uma equação diferencial estocástica. Solucionar estas EDs requer outro tipo de integração, que também está galgada na teoria de martingais.

2.2.3 Movimento browniano e processos de Wiener

O termo movimento browniano foi cunhado devido ao trabalho de Robert Brown, em 1827, ao observar o comportamento aleatório de partículas suspensas num fluido, devido a colisões entre as moléculas e partículas em suspensão. Associada ao movimento de cada uma, há uma curva de probabilidade (ou distribuição de probabilidade) que descreve a probabilidade de diferentes eventos (neste caso: a partícula se movimentar de um ponto a outro) ocorrerem. Ao lidarmos com este tipo de sistema, temos uma coleção de variáveis aleatórias indexadas pelo tempo, ou seja: um processo estocástico. Todavia, a informação sobre este processo está contida na função densidade de probabilidade associada ao sistema. Sendo assim, dada uma trajetória estocástica w , qual é a função densidade de probabilidade (FDP) $P(w)$ associada a essa curva? Antes de responder essa pergunta, é preciso dar um passo atrás nos debruçarmos sobre as características do problema imposto.

A teoria do movimento browniano foi estudada por diversos físicos-matemáticos, até que Norbert Wiener, em 1918, desenvolveu uma formulação matematicamente rigorosa, descrevendo este movimento a partir de processos estocásticos de tempo contínuo $W(t)$, chamados processos de Wiener.

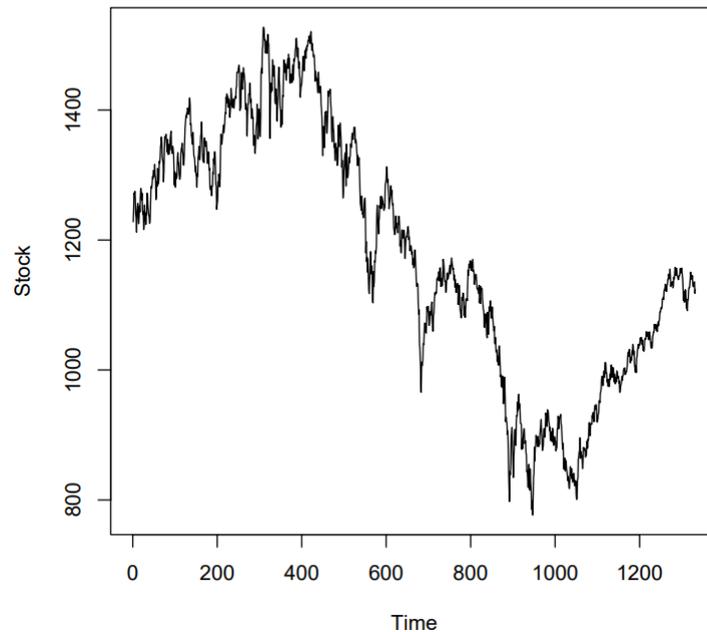
Definição: Um processo estocástico de tempo contínuo (processo de Wiener) $\{W_t\}_{t \geq 0}$, com $W_{t_0} = 0$, é um movimento browniano (MB) se, dados quaisquer instantes de tempo $0 = t_0 \leq t_1 \leq \dots \leq t_n$, os incrementos $W_{t_1} - W_{t_0}$, $W_{t_2} - W_{t_1}$, ..., $W_{t_n} - W_{t_{n-1}}$ são variáveis aleatórias independentes e o conjunto dessas variáveis segue uma distribuição normal:

$$W_{\{t_{i+1}\}} - W_{\{t_i\}} \sim N(0, t_{\{1+i\}} - t_i). \quad (2.21)$$

De maneira simples: A distribuição de probabilidade gaussiana, de incrementos independentes, cuja variância é a diferença entre os incrementos temporais, é chamada de movimento browniano. Isto implica que estamos lidando com um tipo específico de

cadeia de Markov. É possível observar este ocorrendo na prática ao analisarmos a variação do preço de uma ação no mercado financeiro:

Figura 4 – Variação diária da ação das empresas *Standard & Poor 500*, entre 1 de Janeiro de 1999 e 21 de Abril de 2004



Fonte: (FLORESCU; VIENS, 2005.).

Apesar da continuidade do MB, suas trajetórias são funções não-diferenciáveis. Isto pode ser observado empiricamente na figura acima, que não há curvas suaves na função. Esse fato é base fundadora para o cálculo de Itô. O MB possui as seguintes propriedades:

$$E[dW_t] = 0, \quad (2.22)$$

$$E[(dW_t^i)(dW_t^j)] = |dt|\delta^{ij}, \quad (2.23)$$

$$\text{Var}(\sigma dW_t^i) = \sigma^2 dt \quad (2.24)$$

em que σ é uma constante. Além disso, esses processos são martingais, isto é: os processos $\{F_t\}$ associados ao MB são uma coleção de σ -álgebras, tais que:

1. $0 \leq s < t \rightarrow F_s \subset F_t$;
2. $\{W_t\}$ deve ser adaptado à F_t ;
3. Se $0 \leq t < v$, então $W_v - W_t$ é independente de F_t .

Discorrer sobre continuidade de funções naturalmente nos leva a pensar sobre integração, contudo, integrar um processo estocástico via integral de Riemann é incorreto. Para compreendermos como se dará este processo de integração, é necessário uma introdução ao formalismo do cálculo de Itô.

2.3 CÁLCULO ESTOCÁSTICO DE ITÔ

O cálculo de Itô providencia um referencial rigoroso para análise de processos estocásticos, particularmente aqueles envolvendo processos de Wiener. Um dos desafios deste tipo de cálculo é considerar a natureza "não suave" das trajetórias nos processos de Wiener, caracterizados pela sua não diferenciabilidade e seus incrementos não integráveis, caso se utilize a integração convencional de Riemann. Portanto, é preciso compreender a origem da não-diferenciabilidade destes processos e construir uma nova abordagem para o que entendemos como integração, cujo integrando é um processo estocástico e o integrador um processo de Wiener ambos contínuos no tempo.

2.3.1 Teorema da variação quadrática

Munidos de arcabouço teórico suficiente, partimos para a formalização da não-diferenciabilidade dos processos de Wiener a partir do teorema da variação quadrática. Primeiro, suponha uma função continuamente diferenciável $f(t)$, em que num intervalo $[0, T]$ particionado n vezes, cujo tamanho de cada partição é dado por: $\pi_n = \max_{i=0,1,\dots,n-1} (t_{i+1} - t_i)$, tal que $t_0 = 0 < t_1 < t_2 < \dots < t_n = T$. A primeira coisa que podemos concluir é que o tamanho de cada partição vai a zero se o número de partições tende a infinito:

$$\lim_{n \rightarrow \infty} \pi_n = 0.$$

Queremos calcular a variação quadrática da função $f(t)$, dentro do intervalo em questão, escrita como:

$$\lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} [f(t_{i+1}) - f(t_i)]^2,$$

sendo $t_i^* \in [t_i, t_{i+1}]$. De acordo com o teorema do valor médio, podemos reescrever o lado direito da seguinte forma

$$\begin{aligned} \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} [f(t_{i+1}) - f(t_i)]^2 &= \lim_{n \rightarrow \infty} \sum_{i=0}^{\infty} |f'(t_i^*)|^2 (t_{i+1} - t_i)^2 \\ &\leq \lim_{n \rightarrow \infty} \pi_n \sum_{i=0}^{\infty} |f'(t_i^*)|^2 (t_{i+1} - t_i). \end{aligned}$$

dado que $f(t)$ é continuamente diferenciável. Então

$$\begin{aligned} \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} [f(t_{i+1}) - f(t_i)]^2 &= \lim_{n \rightarrow \infty} \pi_n \lim_{i \rightarrow \infty} \sum_{i=0}^{\infty} |f'(t_i^*)|^2 (t_{i+1} - t_i) \\ &= \lim_{n \rightarrow \infty} \pi_n \int_0^T |f'(t_i^*)|^2 dt. \end{aligned}$$

A integral está definida num intervalo finito, portanto não diverge. Finalmente, aplicando o limite, é possível concluir que

$$\lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} [f(t_{i+1}) - f(t_i)]^2 = 0. \quad (2.25)$$

Fica demonstrado, portanto, que funções continuamente diferenciáveis não acumulam variação quadrática ao longo de sua trajetória. Contudo, se aplicarmos o mesmo raciocínio a um processo estocástico, deve-se recorrer ao teorema da variação quadrática.

Teorema: Dado um caminho gerado por um processo estocástico X_t , definido no espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$, ocorrendo entre $[0, T]$, $\forall T > 0$, com incrementos temporais de tamanho $\frac{T}{n}$, temos que:

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \left\{ W\left(\frac{T}{n}i\right) - W\left[\frac{T}{n}(i-1)\right] \right\}^2 = T. \quad (2.26)$$

Prova: Reescrevendo o lado esquerdo de (2.26) da seguinte maneira:

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n [W(t_i) - W(t_{i-1})]^2$$

em que: $t_i = \frac{T}{n}i$. O termo entre colchetes têm distribuição normal $N(0, t_i - t_{i-1}) = N(0, \frac{T}{n})$.

Portanto

$$\lim_{n \rightarrow \infty} \sum_{t=1}^n [W(t_i) - W(t_{i-1})]^2 = \lim_{n \rightarrow \infty} \sum_{t=1}^n \left[N\left(0, \frac{T}{n}\right) \right]^2,$$

note que

$$\text{var}\left[N\left(0, \frac{T}{n}\right)\right] = \left[N\left(0, \frac{T}{n}\right)\right]^2 = \frac{T}{n},$$

como estamos somando n destes termos, ficamos com:

$$\lim_{n \rightarrow \infty} \sum_{t=1}^n \left[N\left(0, \frac{T}{n}\right) \right]^2 = n \left(\frac{T}{n} \right)$$

$$\lim_{n \rightarrow \infty} \sum_{t=1}^n \left[N\left(0, \frac{T}{n}\right) \right]^2 = T.$$

Esse teorema, portanto, garante a não-diferenciabilidade do processo de Wiener.

Ocorre que, esse processo oscila tanto, que mesmo se observarmos uma pequena escala da curva, a variação continuaria grande o suficiente para acumular e não desaparecer como uma função diferenciável. Podemos escrever esse teorema na forma diferencial

$$(dW)^2 = dt, \quad (2.27)$$

observando o lado direito de (2.26) como numa diferencial do processo de Wiener e o lado esquerdo como a diferencial da variável tempo. Esta discussão é, em suma, as provas das propriedades (2.22), (2.23) e (2.24).

O teorema da variação quadrática é o principal pilar do cálculo estocástico. A partir deste, é possível analisar flutuações locais de processos estocásticos, medindo a magnitude cumulativa do quadrado das diferenças entre os valores sucessivos do processo. Se a variação quadrática do processo é alta, indica que há muitas flutuações ou irregularidades entre os valores. Esse tipo de análise pode ser estendida para a compreensão de taxas de convergência, por exemplo: se a variação quadrática de um martingal tende a zero à uma taxa específica, isso indica que esta é a taxa na qual o martingal converge.

2.3.2 Lema de Itô

Início esta subseção com uma breve introdução a equações diferenciais estocásticas (EDEs), tópico este ao qual retornaremos mais adiante para aprofundamento. EDEs são equações diferenciais que envolvem um processo aleatório ou um termo de ruído, possuem forma geral dada pela *difusão de Itô de tempo homogêneo*, ou apenas equação de difusão

$$d\xi(t) = \mu(\xi(t), t)dt + \sigma(\xi(t), t)dW_t, \quad (2.28)$$

σ e μ são funções que especificam, respectivamente, a parte determinística e a parte estocástica da dinâmica e dW_t é um processo de Wiener. Funções $\xi(t)$ que satisfaçam

(2.28) são chamadas *processos de Itô*, o termo $\mu(\xi(t), t)dt$ é chamado *termo de deriva*, ou apenas *deriva*, $\sigma(\xi(t), t)$ é o termo de difusão, responsável pelo comportamento estocástico de $\xi(t)$. Ademais, há mais uma camada de interpretação sobre estas:

$$E[d\xi] = \mu(\xi, t)dt, \quad (2.29)$$

$$E[(d\xi)^2] = \sigma^2(\xi, t)dt \quad (2.30)$$

em que $\mu(\xi, t)dt$ e $\sigma^2(\xi, t)dt$ são, respectivamente, o valor esperado condicional (com relação ao tempo presente t) e a variância de $d\xi$. Intuitivamente, uma EDE descreve a variação instantânea no valor do processo estocástico $\xi(t)$ num tempo t . EDEs são utilizadas amplamente para modelar sistemas que envolvam processos aleatórios, como o mercado financeiro, reações químicas e sistemas biológicos. Estas equações diferenciais fornecem um poderoso referencial matemático para prever o comportamento de tais sistemas.

Antes de prosseguirmos, uma ressalva: ambas as equações (2.28) e (2.12) foram chamadas de "equação de difusão". Estas equações não são idênticas, contudo, cada uma envolve conceitos relacionados ao processo de difusão. Enquanto a primeira descreve a evolução da função probabilidade com relação ao tempo, a segunda descreve a evolução de uma variável estocástica com relação ao tempo, incorporando tanto efeitos determinísticos quanto estocásticos. A fim de evitar qualquer confusão, evocarei cada uma por sua respectiva numeração.

Para buscar soluções de (2.28), a regra da cadeia usual não é aplicável e deve ser substituída pelo lema de Itô, que rege a diferenciação de variáveis estocásticas. Esta fórmula é crucial para o desenvolvimento da teoria de equações diferenciais estocásticas. Como estamos tratando apenas da classe referente aos processos de difusão, as funções μ e σ , juntamente com o estado ξ dado um instante inicial s , são suficientes para especificar a distribuição de probabilidade de ξ em qualquer instante $t > s$. A solução para uma EDE é um processo estocástico, isso significa que não é apenas uma função, mas sim uma família de funções regidas por um processo aleatório. Há EDEs que podem ter sua solução exata encontrada por métodos analíticos, mas não é o caso da maioria delas, que requerem métodos numéricos.

Suponha agora que exista uma função suave do processo de Wiener e do tempo $f(t, W_t)$, em que $t \in [0, T]$ e $W_t = W$. Expandindo essa função em série de Taylor até

os termos de segunda ordem, temos:

$$f(t + dt, W + dW) = \left[f + \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial W} dW \right] + \frac{1}{2} \left[\frac{\partial^2 f}{\partial t^2} (dt)^2 + 2 \frac{\partial^2 f}{\partial t \partial W} (dt dW) + \frac{\partial^2 f}{\partial W^2} (dW)^2 \right] \quad (2.31)$$

$$f(t + dt, W + dW) - f(t, W) = \left[\frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial W} dW \right] + \frac{1}{2} \left[\frac{\partial^2 f}{\partial t^2} (dt)^2 + 2 \frac{\partial^2 f}{\partial t \partial W} (dt dW) + \frac{\partial^2 f}{\partial W^2} (dW)^2 \right]. \quad (2.32)$$

Note que o termo restante do lado esquerdo é apenas o trecho infinitesimal $f(dt, dW_t) = df$ e os dois primeiros termos, do segundo colchete do lado direito, são desprezíveis. Organizando o que restou de (2.32), obtemos

$$df = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial W} dW + \frac{1}{2} \frac{\partial^2 f}{\partial W^2} (dW)^2. \quad (2.33)$$

Essa equação é conhecida como *lema de Itô*. A equação funciona como uma contraparte estocástica para a regra da cadeia, pois dita o funcionamento da diferenciação de uma função estocástica e é imprescindível para resolução de equações diferenciais estocásticas. Permite calcular a evolução temporal de uma função que dependa tanto do tempo, quanto de um processo de Wiener, o que é fundamental para modelar o comportamento de sistemas dinâmicos complexos.

Como f é suave, podemos reescrever a equação (2.28) a partir de (2.33), considerando (2.27), e seguindo os mesmos passos para obtenção de (2.33)

$$\begin{aligned} df &= \frac{\partial f}{\partial t} dt + \frac{1}{2} \frac{\partial^2 f}{\partial t^2} dt^2 + \dots + \frac{\partial f}{\partial \xi} d\xi + \frac{1}{2} \frac{\partial^2 f}{\partial \xi^2} d\xi^2 + \dots \\ df &= \frac{\partial f}{\partial t} dt + \frac{1}{2} \frac{\partial^2 f}{\partial t^2} dt^2 + \dots + \frac{\partial f}{\partial \xi} (\mu dt + \sigma dW) + \frac{1}{2} \frac{\partial^2 f}{\partial \xi^2} (\mu dt + \sigma dW)^2 + \dots \\ df &= \left(\frac{\partial f}{\partial t} + \mu \frac{\partial f}{\partial \xi} + \frac{1}{2} \sigma^2 \frac{\partial^2 f}{\partial \xi^2} \right) dt + \sigma \frac{\partial f}{\partial \xi} dW. \end{aligned} \quad (2.34)$$

Esta forma (2.34) do lema de Itô é mais completa que (2.33) por envolver os termos de deriva e difusão μ e σ , aqui considerados constantes, fundamentais para expandir a discussão sobre equações diferenciais estocásticas. Dizemos que $f \in C^2$ por ser duplamente contínua e diferenciável, isto é, $\frac{df}{dt}$ e $\frac{d^2 f}{dt^2}$ existem e são contínuas para todo o domínio de f . Vale ressaltar que esta demonstração, ainda que válida, não é rigorosamente formal, pois tal prova requer que tomemos o limite de uma sequência de variáveis aleatórias, o que tomaria uma gama de detalhes matemáticos que não

cabem para as pretensões deste trabalho. O próximo passo é, enfim, compreender como integrar processos estocásticos.

2.3.3 Integral de Itô

Até agora tratei majoritariamente de diferenciação, agora é preciso compreender o que significa integrar uma variável estocástica. Quando pensamos em integral, logo é remetida a integral definida pela soma de Riemann. Contudo, essa descrição é estranha ao cálculo de Itô. A integral de Itô é um tipo de integral estocástica usada para solucionar EDEs governadas por processos de Wiener. Para os fins deste trabalho, trago uma descrição pouco rigorosa, porém mais intuitiva da integral de Itô, partindo do que já conhecemos da integração de Riemann.

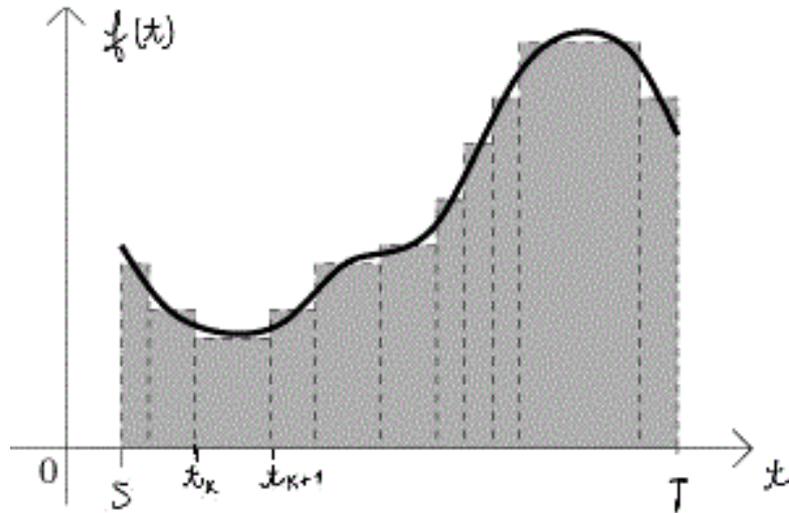
Na definição de Riemann-Stieltjes, consideramos os pontos t_k de uma curva $f(t)$ contida no intervalo $[s, T]$, tal que $s = t_0 < t_1 < t_2 < \dots < t_m = T$, com partições de tamanho $\Delta = t_{k+1} - t_k$. Escolhendo arbitrariamente t_0^*, \dots, t_{m-1}^* , tal que $t_{k+1} \geq t_k^* \geq t_k$ com $k = 0, \dots, m-1$ e tomando o limite $\Delta \rightarrow 0$, definimos esta integral como:

$$\int_s^T f(t) dW(t) = \lim_{\Delta \rightarrow 0} \sum_{k=0}^{m-1} f(t_k^*) [f(t_{k+1}) - f(t_k)].$$

Dada a existência do limite, tem-se que:

$$\Delta = \max_{1 \leq k \leq m-1} (t_{k+1} - t_k).$$

Figura 5 – Representação da integral de Riemann

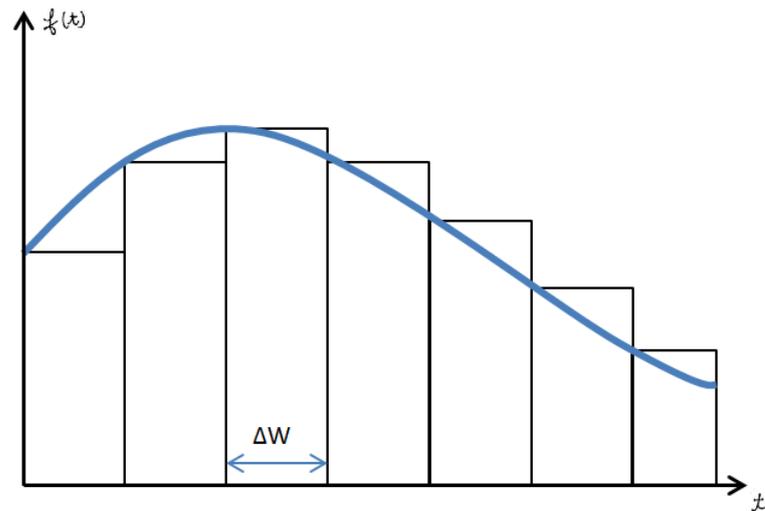


Fonte: Adaptada de (HABALA, 2023).

Como foi visto, funções contínuas no tempo como $f(t)$ acima não acumulam variação quadrática como os processo estocásticos $W(t)$. A razão disso é que $W(t)$ é definido num espaço de probabilidade filtrado $(\sigma, F_t, \mathbb{F}, \mathcal{P})$, tal que $W(t)$ é adaptado a $\{F_t\}$, isto é: é não antecipativo com relação a família crescente de σ -álgebras $\{F_t\}$ e com incrementos independentes dentro do dado intervalo.

Intuitivamente, devemos pensar na integral de Itô como o limite das somas de Riemann quando tomarmos sempre o ponto à esquerda do final de cada intervalo. Em outras palavras, divida o intervalo de tempo em pedaços infinitesimais ΔW . Para cada retângulo, pegue o ponto mais à esquerda e uso-o como partida para desenhar um retângulo e calcular sua área.

Figura 6 – Representação da integral de Itô



Fonte: Adaptada de (HALINA, 2011).

Tomando o limite desses intervalos, temos a integral de Itô. Apesar desta ser um caso particular da integral de Riemann, a questão de escolher o ponto mais à esquerda é profunda pois, quando estudamos processos estocásticos, não podemos ver o futuro, então partimos do fato presente para inferir algo sobre o futuro. Essa intuição está entranhada na teoria do cálculo de Itô, evidenciando ainda mais sua utilidade para descrever com eficácia processos estocásticos. Um processo é dito "adaptado" quando se refere apenas a informações no presente do sistema, não ao futuro.

Intuitivamente, a integral de Itô representa o efeito cumulativo de um processo de Wiener, em uma função, ao longo de um intervalo de tempo. A integral de um processo estocástico $f(t)$ com relação a um processo de Wiener $W(t)$ sobre um intervalo $[s, T]$ é definida como:

$$I(t) = \int_s^T f(t)dW(t) = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(t_i^*)[W(t_i) - W(t_{i-1})]. \quad (2.35)$$

Esses termos são definidos exatamente como na prova de (2.25). A integral de Itô satisfaz diversas propriedades importantes que a tornam útil para resolver EDEs, por exemplo: é um operador linear que obedece a propriedade distributiva e de multiplicação por escalar. Ademais, também é um martingal, o que significa que o seu valor esperado é igual ao seu valor inicial. Vejamos algumas propriedades desta integral:

1. O valor esperado da integral de um processo estocástico com relação ao tempo é finito:

$$E\left[\int_s^T |f(t)|^2 dt\right] < \infty. \quad (2.36)$$

2. Dado (1.23), temos:

$$E\left[\int_s^T f(t)dW(t)\right] = 0. \quad (2.37)$$

3. Dado (2.3), temos:

$$E\left[\left(\int_s^T f(t)dW(t)\right)\left(\int_s^T g(t)dW(t)\right)\right] = E\left[\int_s^T f(t)g(t)dt\right]. \quad (2.38)$$

4. Se $f(t) = g(t)$ na equação acima, temos a propriedade chamada *Isometria de Itô*:

$$E\left[\left(\int_s^T f(t)dW(t)\right)^2\right] = E\left[\int_s^T f(t)^2 dt\right]. \quad (2.39)$$

Dizemos que funções que obedecem a (2.36) são funções L^1 ($x \in L^1$). Intuitivamente, funções L^1 são aquelas cujo valor absoluto é integrável, significando que a área abaixo da curva do valor absoluto é finita. Isso implica que a função não oscila de maneira muito crítica ou possui alguma ponta que se estenda infinitamente. Tais funções são com frequência consideradas bem comportadas com crescimento controlado. A forma da integral estocástica em (2.35) pode ser alterada, graças a arbitrariedade do termo $x(t_i^*)$. Se fizermos $x(t_i^*) \rightarrow \frac{1}{2}(x_i + x_{i-1})$, o limite obtido é diferente da integral de Itô e, considerando sua existência, é chamado de integral de Stratonovich. Este método, ou forma alternativa da integral estocástica de Itô, em alguns casos é mais fácil de ser manipulado, pois é possível aplicar as regras convencionais da integração de Riemann tal como a regra da cadeia. Várias aproximações numéricas para solução de EDEs convergem para a integral de Stratonovich, contudo, o método de solução mais utilizado para a solução numérica de equações de Langevin, o método de Euler–Maruyama, exige que a equação esteja na forma de Itô. Integrais na forma de Stratonovich provam sua utilidade ao solucionarmos EDEs na forma de variedades diferenciáveis, ao invés de apenas em \mathbb{R}^n , pois aplicar o lema de Itô nestes casos pode acabar complicando as soluções.

Da mesma forma que a integral estocástica é construída com relação ao movimento browniano, há também com relação a um martingal quadrado-integrável (dize-

mos que é de classe L^2). A construção desta integral é similar ao caso anterior, mas aqui trago um pouco mais de rigor matemático.

Tomando como verdade os lemas a seguir:

1. Seja f um mapa $\mathbb{R}_{\geq 0} \mapsto \mathbb{R}$ quadrado-integrável ($f \in L^2$). O processo

$$\left[\exp\left(\int_0^T f(t) dW(t) - \frac{1}{2} \int_0^T f(t)^2 dt \right) \right]_{t \geq 0} \quad (2.40)$$

é um martingal quadrado-integrável.

2. Seja \mathbb{D} o conjunto de funções compactadas (isto é: contidas num conjunto fechado e limitadas), tal que este é o conjunto dos f que podem ser escritos como $f = \sum_{i=1}^n a_i [W(t_i) - W(t_{i-1})]$, a família de martingais

$$\left[\exp\left(\int_0^\infty f(t') dW(t') - \frac{1}{2} \int_0^\infty f(t')^2 dt' \right) \right]_{t \geq 0}; \quad (f \in \mathbb{D}) \quad (2.41)$$

está toda contida em L^2 .

A partir destes, é possível enunciar o teorema de representação de Itô.

Teorema: Para cada $f \in L^2$, há um único processo $h(t)_{t \geq 0}$, $E[\int_0^\infty h(t') dW(t')] < \infty$, tal que

$$f = E[f] + \int_0^\infty h(t) dW(t). \quad (2.42)$$

Isto é, para cada variável aleatória f em $L^2(\Omega, \mathcal{F}, \mathbb{P})$ pode ser escrita como uma integral estocástica de um processo quadrado-integrável $h(t)$. A unicidade de (2.42) pode ser demonstrada a partir da Isometria de Itô das integrais estocásticas. Suponha o seguinte:

$$f = E[f] + \int_0^\infty h_1 dW(t) = E[f] + \int_0^\infty h_2 dW(t),$$

$$\int_0^\infty (h_1 - h_2) dW(t) = 0.$$

de (2.39)

$$E\left[\int_0^\infty (h_1 - h_2) dW(t) \right]^2 = 0$$

$$\int_0^\infty E[(h_1 - h_2)^2] dt = 0,$$

portanto $h_1 = h_2$.

Para provar sua existência, consideramos os lemas já enunciados. Seja \mathbf{A} o conjunto de variáveis aleatórias $f \in \mathbb{L}^2$, tal que exista um processo progressivamente mensurável $(h_t)_{t \geq 0}$, tal que: $\int_0^\infty h_t^2 dt < \infty$ e $f = E[f] + \int_0^\infty h_t dt$. Desta forma, \mathbf{A} contém o conjunto do conjunto de variáveis aleatórias, dado por: $f = \exp[\int_0^T h(t) dW_t - \frac{1}{2} \int_0^T h(t)^2 dt]$; $f \in \mathbb{D}$. Considerando uma sequência de variáveis aleatórias $f_n \in \mathbf{A} \in \mathbb{L}^2$ e $f_n \mapsto f$ em \mathbb{L}^2 , tal que

$$f_n = E[f_n] + \int_0^\infty h_n dW(t). \quad (2.43)$$

Portanto, $E[f_n] \mapsto E[f] \iff E[f_n] = E[f] = 0$. Utilizando novamente a isometria de Itô, podemos concluir que

$$E[(f_n - f_m)^2] = \int_0^T E[(h_n - h_m)^2] dt$$

tende a zero, a medida que $n, m \rightarrow 0$. Portanto, h_n é uma sequência de Cauchy. Sendo f o limite desta sequência, a mesma converge para

$$f = E[f] + \int_0^T h dW(t). \quad (2.44)$$

Como uma consequência deste teorema de representação, podemos obter a seguinte descrição para martingais quadrado-integráveis com filtração $[F_t]_{t \geq 0}$.

Teorema: Sejam W os processos de Wiener, definidos num espaço de probabilidade filtrado $\mathbf{L}^2(\Omega, \mathcal{F}, \{F_t\}_{t \geq 0}, \mathbb{P})$ sendo $F_{t \geq 0}$ sua filtração. Então, todos os martingais locais M_t podem ser escritos como

$$M_t = E[M_t] + \int_0^T h_t dW(t). \quad (2.45)$$

Este é o chamado *teorema de representação de martingais*. Como a integração estocástica preserva a propriedade do martingal local, este resultado caracteriza o espaço de todos os martingais começando de 0, definido com relação a filtração, como sendo o conjunto de integrais estocásticas com relação ao processo de Wiener.

Finalmente, temos uma distribuição de probabilidade dada por $f(t, \xi(t))$ bem descrita em termos do tempo e de variáveis estocásticas. O próximo passo é encontrar a FDP associada à curva descrita pela função, agora que já sabemos como integrar sua parte estocástica.

2.4 EQUAÇÃO DE FOKKER PLANCK

Ao buscarmos uma equação que descreva uma função de distribuição ao longo do tempo, imediatamente voltamos nossa atenção para a equação de Liouville

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^n \left(\frac{\partial \rho}{\partial q_i} \dot{q}_i + \frac{\partial \rho}{\partial p_i} \dot{p}_i \right) = 0.$$

Esta equação é uma das mais fundamentais da mecânica estatística, construindo uma ponte da mecânica clássica para mecânica estatística via a equação de movimento para a função distribuição $\rho(q, p)$. A propriedade chave da equação de Liouville é sua reversibilidade temporal, consequência da natureza determinística da mecânica clássica, o que implica na conservação do volume do espaço de fase. À medida que o sistema evolui no tempo, a densidade dos pontos no espaço de fase permanece constante e o volume é conservado. Essa conservação é consequência do fato que a equação de Liouville é uma equação de continuidade, a preservação do volume no espaço de fase é propriedade fundamental da dinâmica hamiltoniana.

No apêndice A2, introduzimos a equação de Langevin, que carrega um termo de fricção e um de ruído. Posteriormente, a partir da equação (2.28), fica claro que a equação de Langevin é uma equação diferencial estocástica descrita classicamente. Na forma linear como a equação de Langevin foi apresentada em (2.2), esta pode ser facilmente resolvida e os efeitos do termo de ruído podem ser contornados. Contudo, casos não lineares, apesar de serem resolvidos satisfatoriamente via métodos numéricos, apresentam grandes dificuldades para um estudo analítico. É a partir desta dificuldade que construímos uma equação de Liouville que trate o comportamento estatístico da dinâmica do sistema com um termo de fricção Markoviano e um ruído branco Gaussiano, esta na qual será chamada de equação de Fokker-Planck. Iniciamos esta construção a partir da *fórmula de Feynman-Kac*: suponha que uma equação diferencial parcial (EDP) está sujeita a condição final $f(\xi_T, T) = h(\xi)$, dado que f é borel-mensurável com relação a filtração F_t e ξ um processo de Itô. Sob essas condições, a fórmula de Feynman-Kac determina que

$$f(\xi_t, t) = E[h(\xi_T)|F_t]. \quad (2.46)$$

tal que $E[\cdot|F_t]$ é o valor esperado com relação ao tempo presente. Todos os valores esperados que utilizarei nos cálculos a seguir serão com relação ao presente. Esta

fórmula é bastante poderosa pois, como veremos a seguir, dada uma EDP parabólica governada por um processo estocástico, (2.46) oferece um método de resolver estas EDPs simulando os caminhos aleatórios do processo estocástico.

Definição: O valor esperado de $f(t, \xi)$, com $\xi = \xi(t)$, é dado por

$$E[f(t, \xi)] = \int_{\mathbb{R}} f(\xi) \rho(\xi, t) d\xi, \quad (2.47)$$

onde $\rho(\xi, t)$ é a FDP. Derivando com relação ao tempo:

$$\frac{E[df(t, \xi)]}{dt} = \int_{\mathbb{R}} f(\xi) \frac{\partial \rho(\xi, t)}{\partial t} d\xi, \quad (2.48)$$

tomando o valor esperado de (2.34) e levando em conta as propriedades do processo de Wiener

$$E[df(t, \xi)] = \int_{\mathbb{R}} \left[\frac{\partial f(\xi)}{\partial t} dt + \mu \frac{df(\xi)}{d\xi} dt + \frac{1}{2} \sigma^2 \frac{d^2 f(\xi)}{d\xi^2} dt \right] \rho(\xi, t) d\xi \quad (2.49)$$

$$\frac{E[df(t, \xi)]}{dt} = \int_{\mathbb{R}} \left[\frac{\partial f(\xi)}{\partial t} + \mu \frac{df(\xi)}{d\xi} + \frac{1}{2} \sigma^2 \frac{d^2 f(\xi)}{d\xi^2} \right] \rho(\xi, t) d\xi, \quad (2.50)$$

por analogia entre (2.47) e (2.50):

$$\int_{\mathbb{R}} f(\xi) \frac{\partial \rho(\xi, t)}{\partial t} d\xi = \int_{\mathbb{R}} \left(\frac{\partial f(\xi)}{\partial t} + \mu \frac{df(\xi)}{d\xi} + \frac{\sigma^2}{2} \frac{d^2 f(\xi)}{d\xi^2} \right) \rho(\xi, t) d\xi \quad (2.51)$$

$$\int_{\mathbb{R}} f(\xi) \frac{\partial \rho(\xi, t)}{\partial t} d\xi = \mu \int_{\mathbb{R}} \frac{df(\xi)}{d\xi} \rho(\xi, t) d\xi + \frac{\sigma^2}{2} \int_{\mathbb{R}} \frac{d^2 f(\xi)}{d\xi^2} \rho(\xi, t) d\xi. \quad (2.52)$$

Nos limites 0 e T , a função f e a distribuição ρ tendem a zero. Dessa forma, a integração por partes nos dá

$$\int_{\mathbb{R}} f(\xi) \frac{\partial \rho(\xi, t)}{\partial t} d\xi = -\mu \int_{\mathbb{R}} f(\xi) \frac{\partial \rho(\xi, t)}{\partial \xi} d\xi + \frac{\sigma^2}{2} \int_{\mathbb{R}} f(\xi) \frac{\partial^2 \rho(\xi, t)}{\partial \xi^2} d\xi \quad (2.53)$$

$$\int_{\mathbb{R}} f(\xi) \frac{\partial \rho(\xi, t)}{\partial t} d\xi = \int_{\mathbb{R}} f(\xi) \left(-\mu \frac{\partial \rho(\xi, t)}{\partial \xi} + \frac{\sigma^2}{2} \frac{\partial^2 \rho(\xi, t)}{\partial \xi^2} \right) d\xi. \quad (2.54)$$

Obtemos, portanto:

$$\frac{\partial \rho(\xi, t)}{\partial t} = -\mu \frac{\partial \rho(\xi, t)}{\partial \xi} + \frac{\sigma^2}{2} \frac{\partial^2 \rho(\xi, t)}{\partial \xi^2}. \quad (2.55)$$

Essa é a *equação de Fokker-Planck* unidimensional. Ela descreve a evolução temporal da FDP em função de uma variável aleatória e do tempo. Escolhi demonstra-la a partir do lema de Itô, mas a mesma pode ser obtida também através da *expansão de Kramers-Moyal* (RISKEN, 1996.), via equação de Liouville (ZWANZIG, 2001.) ou diretamente através da *cadeia de Markov* para o caminhar aleatório unidimensional. Em

(2.55), o primeiro termo a direita é chamado *termo de deriva*, representando as variações determinísticas na FDP. O segundo é chamado de *termo de difusão*, que produz o efeito de difusão e introduz a característica estocástica a FDP.

Note que, se fizermos $\sigma = 0$, temos uma equação de continuidade, implicando que a densidade de probabilidade é conservada. Além disso, fazendo $\mu = 0$ temos uma equação de difusão tradicional

$$\left(\frac{\partial}{\partial t} - \frac{\sigma^2}{2} \frac{\partial^2}{\partial \xi^2}\right) \rho(\xi, t) = 0. \quad (2.56)$$

Admitindo $\rho(\xi, t_0) = \delta(\xi - \xi_0)$, temos que a solução da equação de Fokker-Planck na forma (2.50) é, como deveria ser, uma distribuição gaussiana:

$$\rho(\xi, t) = \frac{e^{-\frac{(\xi - \xi_0)^2}{2\sigma^2(t-t_0)}}}{\sqrt{2\pi\sigma^2(t-t_0)}}. \quad (2.57)$$

A distribuição acima explicita dois fatos importantes para a obtenção da equação de Fokker-Planck neste caso específico: o ruído deve ser Gaussiano e está correlacionado a função delta. Veremos a seguir um caso mais geral, porém decidimos demonstrar apenas esta forma (2.55) pois o formalismo que precede outras mais gerais é desnecessariamente longo para nossas pretensões. Portanto, sem demonstrar, vejamos as condições para obtenção do caso mais geral.

Seja M uma variedade diferenciável, contida num intervalo fechado, e ξ_t^i um processo estocástico contido em M . Convencionando as coordenadas locais x^i com suporte compacto, dizemos que ξ_t^i é uma difusão suave se, dados os processos não antecipativos suaves, $b^i(\xi_t^i) = b^i$ e $\sigma^{ij}(\xi_t^i) = \sigma^{ij} \geq 0$, que obedecem as propriedades da integral de Itô, temos

$$d\xi_t^i = b^i dt + \sigma^{ij} dW^j(t), \quad (2.58)$$

tal que:

$$E[d\xi_t^i] = b^i dt, \quad (2.59)$$

$$d\xi_t^i d\xi_t^j = \sigma^2 dt. \quad (2.60)$$

Note que o lado esquerdo de (2.60) não contém o valor esperado, esse fato é crucial para excluir processos com descontinuidades. Ademais, assumindo que σ^{ij} é não-degenerado, então é um tensor métrico covariante, ou seja: uma métrica de Riemann

((NELSON, 1985.)). Este fato é importante, para conectar com o que veremos mais adiante, pois a partir desta descrição de um processo estocástico, vemos que a métrica de Riemann é parte intrínseca da estrutura probabilística, isso significa que a diferencial de um processo de Wiener é suficiente para o definirmos, de maneira geral, no contexto da geometria riemanniana. Discutirei mais sobre b^i no capítulo a seguir, mas já é perceptível que ele assume o papel do termo de deriva.

Integrando (2.58)

$$\xi_T^i = \xi_0^i + \int_0^T b^i dt + \int_0^T \sigma^{ij} dW^j(t); \quad 0 \leq t \leq T. \quad (2.61)$$

Sejam $f \in C_0^2(\mathbb{R}^n)$ e T um tempo de parada, com filtração $(F_t)_{t \geq 0}$ tal que $E_0[T] < \infty$ (E_0 é o valor esperado com relação ao valor inicial $\xi(0)$). Do lema de Itô (2.34) e levando em consideração (2.23):

$$f(\xi_T^i) = f(\xi_0^i) + \int_0^T \left[\frac{\partial f}{\partial x^i} b^i dt + \frac{1}{2} \frac{\partial^2 f}{\partial x^i \partial x^j} [\sigma^{ik} \sigma^{jn} (\delta^{kn} dt)] + \frac{\partial f}{\partial x^i} (\sigma^i dW^j) \right], \quad (2.62)$$

em que: $\sigma^{ik} \sigma^{jn} \delta^{kn} = (\sigma \sigma^*)^{ij}$ (este * sobrescrito significa "transposto"). Tomando o valor esperado com relação ao valor inicial $\xi(0)$, o termo dW^j desaparece e portanto ficamos com

$$E_0[f(\xi_T^i)] = f(\xi_0^i) + \int_0^T \left(b^i \frac{\partial f}{\partial x^i} + \frac{1}{2} (\sigma \sigma^*)^{ij} \frac{\partial^2 f}{\partial x^i \partial x^j} \right) dt. \quad (2.63)$$

Definição: Seja $(\xi_t)_{t \geq 0} \in \mathbb{R}^n$ um processo de difusão de Itô. Definimos o operador \mathcal{L} de f como

$$\mathcal{L}f(\xi_t) = \lim_{t \rightarrow 0} \frac{E_0[f(\xi_t)] - f(\xi_0)}{t} \quad (2.64)$$

substituindo (2.63) em (2.64) e utilizando a notação: $\nabla^i = \frac{\partial}{\partial x^i}$ e $\Delta = \frac{\partial^2}{\partial x^i \partial x^i}$:

$$\mathcal{L}f(\xi_t) = \left(b^i \nabla^i + \frac{1}{2} (\sigma \sigma^*)^{ij} \Delta \right) f(\xi_t). \quad (2.65)$$

Chamamos este termo \mathcal{L} de *operador de difusão*. Considerando $f \in C_0^2(\mathbb{R}^n)$ e ξ_t^i como definido em (2.58). Da equação (2.65) aplicada ao Lema de Itô

$$\frac{d}{dt} E[f(\xi_t^i)] = E[\mathcal{L}f(\xi_t^i)] \quad (2.66)$$

obtemos a chamada *equação de Kolmogorov "forward"*, que é apenas uma maneira de escrever a equação de Fokker-Planck, como demonstraremos a seguir. De (2.47), fazendo $\rho(\xi_t^i, t) = \rho^i$, temos

$$\frac{d}{dt} \int_{\mathbb{R}} f(\xi_t^i) \rho^i d\xi_t^i = \int_{\mathbb{R}} \mathcal{L}f(\xi_t^i) \rho^i d\xi_t^i = \int_{\mathbb{R}} f(\xi_t^i) \mathcal{L}^* \rho^i d\xi_t^i \quad (2.67)$$

em que \mathcal{L}^* é o L^2 -adjunto de \mathcal{L} e é dado por:

$$\mathcal{L}^* = \left(-b^i \nabla^i + \frac{1}{2} (\sigma \sigma^*)^{ij} \Delta \right) f(\xi_t^i). \quad (2.68)$$

Abrindo a última integral em (2.67)

$$\int_{\mathbb{R}} f(\xi_t^i) \mathcal{L}^* \rho^i d\xi_t^i = \int_{\mathbb{R}} f(\xi_t^i) \left(-b^i \nabla^i \rho^i + \frac{1}{2} (\sigma \sigma^*)^{ij} \Delta \rho^i \right) d\xi_t^i \quad (2.69)$$

$$= \int_{\mathbb{R}} -f(\xi_t^i) b^i \nabla^i \rho^i d\xi_t^i + \int_{\mathbb{R}} f(\xi_t^i) \frac{1}{2} (\sigma \sigma^*)^{ij} \Delta \rho^i d\xi_t^i \quad (2.70)$$

aqui, aplicamos integração por partes em cada uma das integrais. Porém, devido a f ter suporte compacto, o primeiro termo da integração por partes vai a zero por estar associado aos limites de integração, portanto ficamos apenas com o segundo termo. Como teremos que aplicar duas vezes na segunda integral, o sinal muda de volta para o positivo. Sendo assim, concluímos:

$$\int_{\mathbb{R}} f(\xi_t^i) \mathcal{L}^* \rho^i d\xi_t^i = \int_{\mathbb{R}} b^i \nabla^i f(\xi_t^i) \rho^i d\xi_t^i + \int_{\mathbb{R}} \frac{1}{2} (\sigma \sigma^*)^{ij} \Delta f(\xi_t^i) \rho^i d\xi_t^i \quad (2.71)$$

$$= \int_{\mathbb{R}} \left(b^i \nabla^i f(\xi_t^i) \rho^i + \frac{1}{2} (\sigma \sigma^*)^{ij} \Delta f(\xi_t^i) \rho^i \right) d\xi_t^i \quad (2.72)$$

$$= \int_{\mathbb{R}} \mathcal{L}f(\xi_t^i) \rho^i d\xi_t^i. \quad (2.73)$$

Portanto, podemos representar a equação de Fokker-Planck na forma:

$$\begin{aligned} \frac{\partial \rho}{\partial t} &= -\frac{\partial}{\partial x^i} (b^i \rho) + \frac{(\sigma \sigma^*)^{ij}}{2} \frac{\partial \rho}{\partial x^i \partial x^j} \\ \frac{\partial \rho}{\partial t} &= \mathcal{L}^* \cdot \rho \end{aligned} \quad (2.74)$$

Esta equação é uma *equação diferencial Parcial* (EDP) do tipo parabólica, a qual é feita uma breve introdução desta EDP no apêndice C.

Agora, já munidos de como integrar uma variável estocástica, precisamos compreender o tipo de problema que será preciso utilizar esta integração. Para isso, adentraremos na discussão sobre equações diferenciais estocásticas.

2.5 EQUAÇÕES DIFERENCIAIS ESTOCÁSTICAS (EDE)

A forma mais geral de uma EDE é dada pela equação (2.28). Após entender como integrar processos estocásticos, agora temos como objetivo encontrar o processo estocástico que satisfaça a EDE, ou seja: solucioná-la. A questão importante sobre esta

tarefa é que, geralmente, EDEs não apresentam soluções fechadas, podendo possuir múltiplas soluções que não significam nada relevante. Porém, há fatos importantes sobre EDEs que deve-se ter em mente antes de pensar sobre solucioná-las. Primeiramente, se μ e σ são funções "razoáveis", a solução existe e é única. Temos duas condições de unicidade:

1. Condição do crescimento linear:

$$|\mu(\xi, t) - \mu(\xi', t)|^2 + |\sigma(\xi, t) - \sigma(\xi', t)|^2 \leq K|\xi - \xi'|^2. \quad (2.75)$$

2. Condição de Lipschitz:

$$|\mu(\xi, t)|^2 + |\sigma(\xi, t)|^2 \leq K(1 + |\xi|^2). \quad (2.76)$$

Estas duas propriedades são chamadas Condições de Itô. A primeira condição diz que μ e σ estão limitados pela distância entre dois pontos vezes uma constante K , apenas uma relação de proporcionalidade. A segunda diz que, quando você cresce a variável espacial, neste caso representada pela variável estocástica ξ , esse crescimento não pode ser tão rápido a ponto da função "explodir".

Geralmente resolvemos EDEs mais simples por análise ou tentativa e erro. Vejamos um exemplo que pode ser resolvido analiticamente, o "movimento browniano geométrico":

$$d\xi(t) = \mu\xi(t)dt + \sigma\xi(t)dW(t), \quad (2.77)$$

onde μ e σ são constantes, com $\sigma > 0$ e condição inicial $\xi(0) = x_0$. Supondo o ansatz $\xi(t) = f(t, W(t)) = f(t)e^{\sigma W(t)}$, de acordo com o lema de Itô, temos:

$$d(f(t)e^{\sigma W(t)}) = \frac{df(t)}{dt}e^{\sigma W(t)} + \sigma f(t)e^{\sigma W(t)}dW(t) + \frac{1}{2}\sigma^2 f(t)e^{\sigma W(t)}dt,$$

comparando com (2.77), encontramos que

$$\frac{1}{f(t)} \frac{df(t)}{dt} + \frac{1}{2}\sigma^2 = \mu,$$

com solução

$$f(t) = f(0)e^{(\mu - \frac{1}{2}\sigma^2)t}.$$

Portanto $\xi(0) = f(0)e^{\sigma W(t)} = f(0)$ e a solução é

$$f = x_0 e^{\sigma\xi + (\mu - \frac{1}{2}\sigma^2)t}. \quad (2.78)$$

Para EDEs que satisfazem as condições de Itô, podemos solucionar utilizando apenas técnicas padrão aplicadas em equações diferenciais ordinárias. Estes tipos de EDEs são ditas lineares, possuem forma geral

$$d\xi_t = [a(t)\xi_t + c(t)]dt + [b(t)\xi_t + d(t)]dW(t), \quad (2.79)$$

tal que: $a(t)$, $b(t)$, $c(t)$ e $d(t)$ são funções contínuas do tempo. Nosso objetivo agora é encontrar uma solução geral para (2.79). Reescrevendo na forma

$$d\xi_t = [a(t)\xi_t + c(t) + b(t)d(t) - b(t)d(t)]dt + [b(t)\xi_t + d(t)]dW(t) \quad (2.80)$$

$$d\xi_t - \{[a(t)\xi_t + b(t)d(t)]dt + b(t)\xi_t dW(t)\} = [c(t) - b(t)d(t)]dt + d(t)dW(t). \quad (2.81)$$

Considere agora κ_t um processo de Itô, contido entre $[0, T]$, tal que $\kappa(0) = 1$.

$$d\kappa_t = f(\kappa_t, t)dt + g(\kappa_t, t)dW(t). \quad (2.82)$$

Multiplicando (2.81) por κ_t , obtemos:

$$\kappa_t d\xi_t - \kappa_t \{[a(t)\xi_t + b(t)d(t)]dt + b(t)\xi_t dW(t)\} = \kappa_t \{[c(t) - b(t)d(t)]dt + d(t)dW(t)\}. \quad (2.83)$$

Introduzindo a *regra do produto de Itô* para $d[\xi_t \kappa_t]$:

$$d(\xi_t \kappa_t) = \kappa_t d\xi_t + \xi_t d\kappa_t + d\xi_t d\kappa_t, \quad (2.84)$$

em comparação com (2.83), queremos os dois últimos termos de (2.84) iguais ao segundo termo de (2.83)

$$\xi_t d\kappa_t + d\xi_t d\kappa_t = -\kappa_t \{[a(t)\xi_t + b(t)d(t)]dt + b(t)\xi_t dW(t)\} \quad (2.85)$$

calculando apenas os termos do lado esquerdo:

$$\xi_t d\kappa_t = \xi_t [f(\kappa_t, t)dt + g(\kappa_t, t)dW(t)] \quad (2.86)$$

$$d\xi_t d\kappa_t = \{[a(t)\xi_t + c(t)]dt + [b(t)\xi_t + d(t)]dW(t)\} \{f(\kappa_t, t)dt + g(\kappa_t, t)dW(t)\},$$

das propriedades do processo de Wiener

$$d\xi_t d\kappa_t = [b(t)\xi_t + d(t)]g(\kappa_t, t)dt \quad (2.87)$$

substituindo (2.86) e (2.87) em (2.85)

$$\begin{aligned} \xi_t [f(\kappa_t, t)dt + g(\kappa_t, t)dW(t)] + [b(t)\xi_t + d(t)]g(\kappa_t, t)dt = \\ -\kappa_t \{[a(t)\xi_t + b(t)d(t)]dt + b(t)\xi_t dW(t)\}, \end{aligned} \quad (2.88)$$

percebe-se que $g(\kappa_t, t) = -b(t)\xi_t$. Então

$$\begin{aligned} \xi_t[f(\kappa_t, t)dt - b(t)\xi_t dW(t)] - [b(t)\xi_t + d(t)]b(t)\xi_t dt = \\ -\kappa_t\{[a(t)\xi_t + b(t)d(t)]dt + b(t)\xi_t dW(t)\} \end{aligned} \quad (2.89)$$

$$\xi_t f(\kappa_t, t)dt - b(t)^2 \kappa_t \xi_t dt = -a(t)\kappa_t \xi_t dt \quad (2.90)$$

$$\xi_t f(\kappa_t, t)dt = \xi_t[-a(t) + b(t)^2]\kappa_t dt, \quad (2.91)$$

concluimos que $f(\kappa_t, t) = [-a(t) + b(t)^2]\kappa_t$. Portanto:

$$d\kappa_t = [-a(t) + b(t)^2]\kappa_t dt - b(t)\kappa_t dW(t), \quad (2.92)$$

utilizando o lema de Itô, realizando processo similar a resolução de (2.77), é possível mostrar que $\phi(t) = (\phi_t) = (\kappa_t^{-1})$, tal que o processo de Itô ϕ_t satisfaça:

$$d\phi_t = a(t)\phi_t dt + b(t)\phi_t dW(t); \quad \phi_0 = 1, \quad (2.93)$$

com forma explícita

$$\phi_t = \phi_0 \exp\left\{\int_0^t \left[a(s) - \frac{b(s)^2}{2}\right] ds + \int_0^t b(s) dW(s)\right\}, \quad (2.94)$$

tal qual na solução de (2.77). Finalmente, temos que de (2.84)

$$d(\kappa_t \xi_t) = [c(t) - d(t)b(t)]\kappa_t dt + d(t)\kappa_t dW(t), \quad (2.95)$$

tal que o processo estocástico $\xi(t)$, solução geral da EDE linear, é escrito em termos de $\phi(t)$:

$$\xi_t = \phi_t \left\{ \xi_0 \kappa_0 + \int_0^t [a(s) - b(s)d(s)]\kappa_s ds + \int_0^t d(s)\kappa_s dW(s) \right\}. \quad (2.96)$$

É importante explicitar a solução analítica para esta classe de EDEs, contudo, a maior parte das EDEs não é linear. Para solucioná-las, é preciso buscar a solução numericamente. Antes de explorar este tipo de solução, é preciso expandir a discussão sobre um tipo específico de EDE, que também é peça chave nos estudos dos processos estocásticos: *equações diferenciais estocásticas "backward"*.

2.6 EQUAÇÕES DIFERENCIAIS ESTOCÁSTICAS "BACKWARD" (BSDE)

Estes tipos de equações diferenciais estocásticas foram introduzidas por (BISMUT, 1973.) para o caso contínuo, motivado pelas aplicações em controle estocástico, e por

(PARDOUX; PENG, 1990.) para o caso geral. As BSDEs são aplicáveis numa gama de diferentes campos, os mais notáveis sendo: mercado financeiro (via modelo de Black-Scholes), controle estocástico e métodos numéricos probabilísticos para solução de EDPs, este último sendo o caso de interesse neste tópico. Enquanto a EDE é uma extensão não-linear da integração estocástica, as BSDEs são uma versão não-linear do teorema de representação de martingais. Naturalmente, os argumentos para construção teórica das BSDEs são similares aos utilizados para EDEs, mas a conexão entre elas não é trivial e precisamos tratar disto com cautela.

Note que podemos compreender (2.74) como uma maneira de caracterizar a solução de uma EDE, solucionando uma EDP. A ideia aqui é realizar o caminho inverso e entender o que temos a ganhar com isso. Definindo $u(\xi_0, t) := E_0[f(\xi_t)]$ e assumindo que é uma função bem comportada, é solução da EDP:

$$u(\xi_0, t + h) = E_0[E_0[f(\xi_{t+h})|F_h]]. \quad (2.97)$$

Das propriedades de martingais, temos:

$$\begin{aligned} \frac{E_0[u(\xi_h, t)] - u(\xi_0, t)}{h} &= \frac{1}{h} E_0[E_\xi[f(\xi_t)] - E_0[f(\xi_t)]] \\ &= \frac{1}{h} E_0[E_0[f(\xi_{t+h})|\{F_s\}_{s \leq h}] - E_0[f(\xi_t)|\{F_s\}_{s \leq h}]] \\ &= \frac{1}{h} E_0[E_0[f(\xi_{t+h}) - f(\xi_t)|\{F_s\}_{s \leq h}]] \\ &= \frac{1}{h} E_0[f(\xi_{t+h}) - f(\xi_t)] \\ &= \frac{1}{h} E_0[f(\xi_{t+h})] - E_0[f(\xi_t)] \\ &= \frac{u(\xi_0, t + h) - u(\xi_0, t)}{h}. \end{aligned}$$

Tirando o limite $h \rightarrow 0$ dos dois lados, obtemos

$$\frac{\partial u(\xi_0, t)}{\partial t} = \mathcal{L}u(\xi_0, t). \quad (2.98)$$

A equação (2.98) é chamada *equação de Kolmogorov "backward"* e faz exatamente o pretendido: ao invés de resolver a EDP, podemos resolver a EDE e calcular $E_0[f(\xi_t)]$. Perceba que em (2.74) o operador \mathcal{L} perturba a posição final ξ_t , ao passo que na equação *backward* \mathcal{L} atua na condição inicial ξ_0 . As EDPs que nos interessam neste trabalho são as do tipo parabólica, portanto, queremos encontrar uma generalização da formula de Feynman-Kac do tipo:

$$\frac{\partial u(\xi_0, t)}{\partial t} + \mathcal{L}u(\xi_0, t) + f(\xi_0, t, u(\xi_0, t), \nabla_0 u(\xi_0, t)) = 0, \quad (2.99)$$

onde $\nabla_0 u(\xi_0, t)$ é a derivada covariante com relação a ξ_0 . Se u satisfaz (2.46), significa que o valor esperado com relação ao processo de filtragem pode ser obtido como uma solução da equação diferencial parcial associada (2.99), cuja condição terminal é $u(\xi_0, T) = f$. Assumindo que a equação acima possui solução $Y_t = u(\xi_t, t)$, então ao descrever sua dinâmica, podemos considerar uma versão de ξ_0 para cada par (t, ξ_0) . Do lema de Itô

$$dY_t = \left[\frac{\partial u(\xi_t, t)}{\partial t} + \mathcal{L}u(\xi_t, t) \right] dt + \nabla_0 u(\xi_t, t) \sigma(\xi_t) dW_t$$

$$dY_t = -f(\xi_0, t, u(\xi_0, t), \nabla_0 u(\xi_0, t)) + \nabla_0 u(\xi_t, t) \sigma(\xi_t) dW_t.$$

A equação acima sugere que devemos considerar uma versão de (2.99) que inclua $\sigma(\xi_t)$:

$$\frac{\partial u(\xi_0, t)}{\partial t} + \mathcal{L}u(\xi_0, t) + f(\xi_0, t, u(\xi_0, t), \nabla_0 u(\xi_t, t) \sigma(\xi_t)) = 0. \quad (2.100)$$

Se u resolve esta equação, fazendo $Z_t = \nabla_0 u(\xi_t, t) \sigma(\xi_t)$, obtemos

$$dY_t = -f(\xi_0, t, Y_t, Z_t) + Z_t dW_t. \quad (2.101)$$

Esta é a forma geral de uma equação diferencial estocástica "backward". Y_t é um martingal local, naturalmente induzido por ξ_0 , $Y_t : E[\xi_0 | F_t]$. Martingais locais estão intrinsicamente ligados ao conceito de tempo de parada, portanto, dizemos que ξ_0 é a condição terminal de (2.101), esta sendo uma equação diferencial estocástica *backward* cuja solução é dada pelo par (Y, Z) . De acordo com (2.45), há um único processo Z tal que

$$Y_0 = Y_T + \int_0^T f(\xi, t, Y_t, Z_t) - \int_0^T Z_t dW_t, \quad (2.102)$$

onde Y é contínuo e adaptado a filtração, f é chamada função geradora da BSDE e obedece a (2.76), Z é previsível e obedece (2.36). É importante compreender que uma BSDE **não** é uma EDE de tempo reverso. Num tempo t , (Y_t, Z_t) são F_t -mensuráveis, tal que o processo ainda não "conhece" sua condição terminal Y_T . Sendo assim, para encontrarmos a solução, precisamos dos parâmetros f e Y_T . Isto é o resumo de um teorema dado e provado por (KAROUI; PENG; QUENEZ, 1997.). A prova da unicidade de (2.101) pode ser encontrada em (PARDOUX; PENG, 1990.) teorema 1, a existência no teorema 3.2.

Métodos de controle ótimo são empregados quando queremos que alguma função atinja sua "melhor performance". Estratégias de controle, no contexto das BSDEs,

servem para otimizar os objetivos do processo enquanto minimiza os riscos em questão, para que possamos construir o melhor controle dentro dos limites impostos. Em finanças, por exemplo, o controle ótimo estocástico é utilizado para o gerenciamento de riscos das derivadas financeiras, minimizando os riscos de exposição. Considere a seguinte EDE controlada:

$$d\xi_t = [a\xi(t) + cu(t)]dt + dW(t); \quad \xi(0) = \xi_0, \quad t \in [0, T], \quad (2.103)$$

em que $u(t)$ é chamado *processo de controle*. O estado ξ_t e $u(t)$ precisam ser $\{F_t\}_{t \geq 0^-}$ adaptados quadrado-integráveis. Para simplificar, assumiremos a e c constantes e as funções todas unidimensionais. O *funcional de custo* é dado, de forma geral, por (FLEMING; RISHEL, 1975.b):

$$J(u) = E \left[\int_0^T f(t, \xi_t, u(t))dt + g(\xi_T) \right], \quad (2.104)$$

mas para o caso unidimensional, temos que

$$J(u) = \frac{1}{2} E \left[\int_0^T [\xi_t^2 + u(t)^2]dt + \xi_T^2 \right]. \quad (2.105)$$

O problema de otimização em questão tem como objetivo minimizar (2.105) sujeito a (2.103). Primeiro, precisamos determinar $u(t)$, para isso admitimos qualquer outro controle $v(t)$, tal que

$$0 \leq \lim_{\epsilon \rightarrow 0} \frac{J(u + \epsilon v) - J(u)}{\epsilon} \rightarrow E \left[\int_0^T [\xi'_t \xi_t + u(t)v(t)]dt + \xi'_T \xi_T \right], \quad (2.106)$$

em que ξ'_t satisfaz:

$$d\xi'_t = [a\xi'_t + cv(t)]dt; \quad \xi'_0 = 0, \quad t \in [0, T]. \quad (2.107)$$

Para obtermos mais informações sobre (2.106), introduzimos a equação adjunta:

$$\begin{cases} dY_t = -[aY_t + \xi_t]dt + Z_t dW(t), & t \in [0, T]; \\ Y_T = \xi_T. \end{cases} \quad (2.108)$$

esta equação é uma BSDE, tais que Y_t e Z_t obedecem às mesmas condições de (2.101) e neste caso $f(Y_t, \xi_t) = (aY_t + \xi_t)dt$. Sabemos, portanto, que (2.107) admite solução (Y_t, Z_t) . Aplicando o lema de Itô a $Y_t \xi'_t$, temos:

$$\begin{aligned} E[Y_T \xi'_T] &= E[\xi_T \xi'_T] \\ &= \int_0^T E \left\{ [-aY_t - \xi_t] \xi'_t + Y_t [a\xi'_t + cv(t)] \right\} dt \\ &= \int_0^T E \left[-\xi_t \xi'_t + cY_t v(t) \right] dt. \end{aligned} \quad (2.109)$$

Substituindo (2.109) em (2.106):

$$0 \leq E \int_0^T [cY_t + u(t)]v(t)dt, \quad (2.110)$$

como $v(t)$ é arbitrário e nosso objetivo é minimizar J , então

$$u(t) = -cY_t, \quad t \in [0, T]. \quad (2.111)$$

Ou seja: $u(t)$ só é um controle admissível se Y_t é $\{F_t\}_{t \geq 0}$. Substituindo (2.111) em (2.103), obtemos o sistema otimizado:

$$\begin{cases} d\xi_t = [a\xi(t) - c^2Y_t]dt + dW(t), \\ dY_t = -[aY_t + \xi_t]dt + Z_t dW(t); \quad t \in [0, T], \\ \xi(0) = \xi_0, \quad Y_T = \xi_T. \end{cases} \quad (2.112)$$

Ou seja, processos estocásticos otimizados são descritos como uma equação diferencial estocástica acoplada "forward & backward" (FBSDE). Se a solução adaptada (ξ, Y, Z) é única, o controle u também é ((MA; YONG, 2007.) teorema 2.2). Resolver uma FBSDE com controle ótimo envolve o uso de técnicas numéricas, tal como simulação de Monte Carlo. Estes métodos encontram soluções aproximadas para problemas de otimização complexos.

O formalismo de Itô é bastante vasto e possui diversas ramificações que não foram contempladas por este trabalho. Para nossos objetivos (introduzir a quantização estocástica de Nelson, estudar soluções numéricas de EDEs e encontrar um método para a interpretação estocástica da fase de Berry), o amálgama das técnicas apresentadas até aqui formam o arcabouço teórico robusto para a introdução da quantização estocástica de Nelson, que por sua vez, irá fornecer significado físico ao formalismo de Itô, abrindo novas possibilidades de interpretação e solução das EDEs e EDPs.

3 QUANTIZAÇÃO ESTOCÁSTICA

A hipótese da quantização estocástica é uma abordagem à mecânica quântica que atribui a componente estocástica ao processo de quantização. Neste referencial estocástico, o conceito tradicional de quantização, que envolve promover as variáveis clássicas aos operadores quânticos, é estendido para incluir variáveis estocásticas, que introduz uma variável aleatória a evolução destes campos. Introduzimos o programa de quantização deste trabalho via equação de Langevin (2.2), mas a conexão mais intrínseca entre o comportamento quântico e estocástico envolve discretizar o espaço e o tempo, expressando a evolução do campo como a soma de diferentes caminhos. Isso nos leva, naturalmente, a utilização de integrais funcionais. A integral de Itô, abordada no capítulo anterior, é um tipo de integral funcional, contudo, a abordagem via mecânica de Nelson ((NELSON, 1988.), (NELSON, 2012.), (NELSON, 1966.), (NELSON, 1985.)), unida a conexão da mecânica clássica com quântica via equações de Madelung ((MADELUNG, 1926.), (REDDIGER; POIRIER, 2022.), (MOLLAI; FATHI, 2021.)), possibilitam aprofundar e ampliar o programa de quantização, além de dar fundamentos matemáticos sólidos para a abordagem de integral funcional (PAVON, 2000.).

3.1 MECÂNICA DE NELSON

Imre Fényes foi um físico húngaro, que em seu trabalho (FÉNYES, 1946.), foi o primeiro a propôr uma interpretação estocástica da mecânica quântica. Ele demonstra que a equação de Schrödinger pode ser entendida como uma equação de difusão para processos de Markov. A partir deste, uma gama de cientistas, como (GUERRA, 1981.), (PROTTER, 2007.), (YASUE, 1981.) etc buscaram desenvolver uma teoria mais fundamental do que a mecânica quântica, em que a interação entre partículas clássicas e um dito "éter" clássico causaria com que tais partículas experimentassem um processo de difusão, no espaço de configurações, que conservaria sua energia total. Graças a natureza dos processos de Markov, foi possível construir uma teoria na qual a cinemática das difusões é completamente simétrica com relação a direção do tempo, assim, destacou-se o trabalho de Edward Nelson com a mecânica estocástica.

Seja $dt > 0$ incremento temporal positivo e $f(t) \in \mathbb{R}$ definimos, respectivamente, os

incrementos temporais "forward" e "backward":

$$d_+f = f(t + dt) - f(t) \quad (3.1)$$

$$d_-f = f(t) - f(t - dt) \quad (3.2)$$

Os processos estocásticos L^1 , $\xi \in \mathbb{R}$, são (super)sub-martingais se obedecerem a ((2.18))(2.17). Para a maioria dos casos, as amostras de caminho " $\xi(t)$ " não são diferenciáveis, para quase todo " t ". Portanto, suas derivadas só existem no contexto de funções generalizadas, contudo, para evitar o uso de tais funções, Nelson propõe a definição abaixo.

Definição: Dado o instante de tempo " t ", $D_+(D_-)$ é a "mean forward(backward) derivative" da variável aleatória, $\xi(t) = \xi_t$, dada por:

$$D_+f = D_+f(\xi_t, t) = \lim_{dt \rightarrow 0} E_t \left[\frac{d_+f(t)}{dt} \right] \quad (3.3)$$

$$D_-f = D_-f(\xi_t, t) = \lim_{dt \rightarrow 0} E_t \left[\frac{d_-f(t)}{dt} \right] \quad (3.4)$$

tal que o limite existe no espaço de probabilidade $L^1(\Omega, \mathcal{F}, \mathbb{P})$, $E_t = E[\cdot | N_t]$ é o valor esperado com relação a σ -álgebra N_t representando o tempo presente " t ".

A necessidade das definições acima advém do fato que as derivadas dos processos de Itô não são bem definidas no limite $dt \rightarrow 0$. Isso pode ser constatado do fato que $dW(t)$ possui tamanho proporcional a \sqrt{dt} , portanto: $d\xi_t/dt \sim dW(t)/dt \sim 1/\sqrt{dt}$. Contudo, D_+ e D_- possuem um limite bem definido de dt .

Processos de difusão são descritos na mecânica de Nelson não apenas pelos incrementos $d_\pm f$, mas também pelos campos vetoriais b_+ e b_- , chamados respectivamente de "forward drift" e "backward drift":

$$b_\pm = E_t \left[\frac{d_\pm \xi(t)}{dt} \right], \quad (3.5)$$

que entram como os termos de deriva na equação de difusão. Supondo $\sigma = 1$ em (2.61), por exemplo, teríamos:

$$\xi(T) - \xi(s) = \int_s^T b_+ dt + W_+(t) - W_+(s), \quad (3.6)$$

$$\xi(T) - \xi(s) = \int_s^T b_- dt + W_-(t) - W_-(s), \quad (3.7)$$

onde o incremento $W_+(t) - W_+(s)$ independe da subfiltração \mathbb{P} e $W_-(t) - W_-(s)$ da superfiltração \mathbb{F} . Podemos concluir que ξ é um processo Markoviano, já que seus incrementos com relação ao passado e futuro são condicionalmente independentes do

presente. Munidos destas ferramentas, começamos a construção da cinemática dos processos estocásticos.

Definição: A velocidade de corrente v é a média de b_+ e b_- e a velocidade osmótica u é a metade da sua diferença

$$v \equiv \frac{b_+ + b_-}{2}, \quad (3.8)$$

$$u \equiv \frac{b_+ - b_-}{2}. \quad (3.9)$$

Nos processos estocásticos, " v " é um análogo direto da velocidade convencional de processos determinísticos. Já " u " mede o quão rápido a "aleatoriedade" do sistema cresce. Podemos representar a equação de Fokker-Planck para várias dimensões nas formas:

$$\frac{\partial \rho(\xi_t, t)}{\partial t} = -\nabla(\rho b_+) + \frac{1}{2} \Delta \rho, \quad (3.10)$$

$$\frac{\partial \rho(\xi_t, t)}{\partial t} = -\nabla(\rho b_-) - \frac{1}{2} \Delta \rho. \quad (3.11)$$

Estas são, respectivamente, a equação parabólica e anti-parabólica de Fokker-Planck. Quando somadas obtemos a equação de continuidade que relaciona ρ e v

$$\frac{\partial \rho(\xi_t, t)}{\partial t} = -\nabla(\rho v). \quad (3.12)$$

Para obtermos a velocidade osmótica u com relação a distribuição de probabilidade, é preciso aprofundar o formalismo. Definindo as difusões suaves $F_t = f(\xi_t)$ e $G_t = g(\xi_t)$, dividindo o intervalo $[s, T]$ em n partes iguais,

$$t_j = s + \frac{j(T - s)}{n}$$

com $j = 1, 2, \dots, n$. Então

$$E[F_T G_T] - E[F_s G_s] = \lim_{n \rightarrow \infty} \sum_{j=1}^{n-1} E[F(t_{j-1})G(t_j) - F(t_j)G(t_{j-1})],$$

reorganizando na forma:

$$E[F_T G_T] - E[F_s G_s] = \lim_{n \rightarrow \infty} \sum_{j=1}^{n-1} E \left\{ [F(t_{j+1}) - F(t_j)] \frac{G(t_j) + G(t_{j-1})}{2} + \frac{F(t_{j+1}) + F(t_j)}{2} [G(t_j) - G(t_{j-1})] \right\}.$$

Da definição dos incrementos temporais

$$\begin{aligned}
E[F_T G_T] - E[F_s G_s] &= \lim_{n \rightarrow \infty} \sum_{j=1}^{n-1} E[D_+ F(t_j) G(t_j) + F(t_j) D_- G(t_j)] \frac{T-s}{n} \\
&= \int_s^T E[(D_+ F_t) G_t + F_t (D_- G_t)] dt,
\end{aligned} \tag{3.13}$$

da regra do produto (2.84)

$$d(F_t G_t) = F_t dG_t + G_t dF_t + dG_t dF_t, \tag{3.14}$$

, e notando que

$$d(F_t G_t) \equiv \nabla^i F_t d\xi^i \nabla^j G_t d\xi^j \equiv E[\sigma^{ij} \nabla^i F_t \nabla^j G_t] dt, \tag{3.15}$$

a primeira integral de (3.13) toma a forma:

$$\begin{aligned}
\int_s^T E[(D_+ F_t) G_t] dt &= - \int_s^T E[F_t D_+ G_t] dt - \int_s^T E[F_t D_+ G_t] dt - \\
&\int_s^T E[\sigma^{ij} \nabla^i F_t \nabla^j G_t] dt + E[F_T G_T] - E[F_s G_s].
\end{aligned} \tag{3.16}$$

Comparando (3.16) com (3.13), temos que

$$\frac{d}{dt} E[f(\xi_t) g(\xi_t)] = E\{[D_+ f(\xi_t)] g(\xi_t)\} + E\{f(\xi_t) [D_- g(\xi_t)]\} \tag{3.17}$$

$$= E\{[D_+ f(\xi_t)] g(\xi_t)\} + E\{f(\xi_t) [D_+ g(\xi_t)]\} + E[\sigma^{ij} \nabla^i f(\xi_t) \nabla^j g(\xi_t)] dt, \tag{3.18}$$

tal que

$$E\{f(\xi_t) [D_- g(\xi_t)]\} = E\{f(\xi_t) [D_+ g(\xi_t)]\} + E[\sigma^{ij} \nabla^i f(\xi_t) \nabla^j g(\xi_t)]. \tag{3.19}$$

Das definições de (2.64) e (3.3), concluímos

$$D_+ g(\xi_t) = \left(b_+^i \nabla^i + \frac{1}{2} (\sigma \sigma^*)^{ij} \Delta \right) g(\xi_t) \tag{3.20}$$

$$D_- g(\xi_t) = \left(b_-^i \nabla^i - \frac{1}{2} (\sigma \sigma^*)^{ij} \Delta \right) g(\xi_t). \tag{3.21}$$

Substituindo ambas em (3.19) e considerando a definição (2.47):

$$\begin{aligned}
\int f(x) \left[\left(b_-^i \nabla^i - \frac{1}{2} (\sigma \sigma^*)^{ij} \Delta \right) g(x) \right] \rho(x, t) dx &= \int f(x) \left[\left(b_+^i \nabla^i + \frac{1}{2} (\sigma \sigma^*)^{ij} \Delta \right) g(x) \right] \rho(x, t) dx + \\
&\int f(x) \left[\sigma^{ij} \nabla^i f(\xi_t) \nabla^j g(x) \right] \rho(x, t) dx,
\end{aligned} \tag{3.22}$$

aplicando integração por partes na última integral, obtemos:

$$= - \int f(x) \delta g(x) \rho(x, t) dx - \int f(x) \nabla^j g(x) \frac{\nabla^i \rho(x, t)}{\rho(x, t)} \rho(x, t) dx.$$

Portanto, substituindo em (3.22), obtemos o operador de difusão "backward":

$$\begin{aligned} -\frac{1}{2}(\sigma\sigma^*)^{ij}\Delta + b_-^i \nabla^i &= -\frac{1}{2}(\sigma\sigma^*)^{ij}\Delta + [b_+^i - \nabla^i \log(\rho)] \nabla^i \\ b_-^i &= [b_+^i - \nabla^i \log(\rho)] \\ \frac{b_+^i - b_-^i}{2} &= \frac{1}{2} \nabla^i \log(\rho). \end{aligned}$$

Finalmente, da definição (3.9), temos que a velocidade osmótica com relação a distribuição de probabilidade é

$$u = \nabla \log(\sqrt{\rho}), \quad (3.23)$$

chamada *equação osmótica*. Escrevemos:

$$R = \frac{1}{2} \log(\rho) \quad (3.24)$$

tal que $\rho = e^{2R}$ e

$$\nabla R = u. \quad (3.25)$$

Voltando para (3.12), é importante notar que esta equação ainda é satisfeita se adicionarmos qualquer campo vetorial z independente do tempo a v , tal que para todo t , tenhamos

$$\begin{aligned} \nabla(\rho z) &= 0, \\ \int z^2 \rho dx &< \infty. \end{aligned}$$

Dentre todos estes vetores, haverá um v_0 , tal que sua integral é um mínimo. A condição para isso é

$$\frac{d}{d\lambda} \int_{\mathbb{R}^n} (v_0 + \lambda z)^2 \rho dx \Big|_{\lambda=0} = 0. \quad (3.26)$$

Para difusões suaves, ρ é sempre positivo. Definimos

$$\nabla S = v_0, \quad (3.27)$$

então S é unicamente determinado por uma função de t . Por definição, S satisfaz

$$-\nabla(\rho \nabla S) = \frac{\partial \rho}{\partial t}. \quad (3.28)$$

Dividindo por 2ρ , encontramos:

$$\frac{\partial R}{\partial t} = -\left(\frac{1}{2}\Delta + u\nabla\right)S, \quad (3.29)$$

com a hipótese de que S é suave para cada t . Ao trabalharmos na mecânica de Nelson, estaremos interessados majoritariamente em difusões de gradiente, por isso estas representações com R e S . Portanto, se v é um gradiente

$$\nabla S = v. \quad (3.30)$$

Seguindo nesta linha de representação por gradientes, R e S serão a porta de entrada para a ligação entre o formalismo de Nelson e Schrödinger.

3.2 EQUAÇÕES DE MADELUNG

A equação de Schrödinger unidimensional independente do tempo para partículas livres ($V = 0$) é dada por

$$-i\hbar \frac{\partial \Psi}{\partial t} = \frac{\hbar^2}{2m} \nabla^2 \Psi \quad (3.31)$$

consideremos agora o ansatz:

$$\Psi = R(x, t) \exp\left[\frac{i}{\hbar} S(x, t)\right], \quad (3.32)$$

em que $R(x, t)$ é o módulo da função de onda e $S(x, t)$ é a fase da função de onda. Substituindo na equação de Schrödinger, considerando $\frac{\partial f}{\partial t} = \dot{f}$ e $e^{\frac{i}{\hbar} S(x, t)} = e^{\kappa}$

$$-i\hbar \frac{\partial}{\partial t} \left\{ R(x, t) \exp\left[\frac{i}{\hbar} S(x, t)\right] \right\} = \frac{\hbar^2}{2m} \nabla^2 \left\{ R(x, t) \exp\left[\frac{i}{\hbar} S(x, t)\right] \right\} \quad (3.33)$$

$$(-i\hbar \dot{R} + R\dot{S})e^{\kappa} = \frac{\hbar^2}{2m} \nabla \left[\nabla R e^{\kappa} + R \frac{i}{\hbar} \nabla S e^{\kappa} \right] \quad (3.34)$$

$$(-i\hbar \dot{R} + R\dot{S})e^{\kappa} = \frac{\hbar}{2m} \left[\nabla^2 R + \frac{2i}{\hbar} \nabla R \nabla S + R \frac{i}{\hbar} \nabla^2 S + R \frac{i}{\hbar} \nabla S \frac{i}{\hbar} \nabla S \right] e^{\kappa} \quad (3.35)$$

$$-i\hbar \frac{\dot{R}}{R} + \dot{S} = \frac{\hbar}{2m} \frac{\nabla^2 R}{R} + \frac{i\hbar}{m} \frac{\nabla R \nabla S}{R} + \frac{i\hbar}{2m} \nabla^2 S - \frac{1}{2m} (\nabla S)^2. \quad (3.36)$$

Separando a parte imaginária da real e fazendo $\mathbb{R} = \mathbb{C} = 0$, respectivamente, nos dão:

$$-i\hbar \frac{1}{R} \frac{\partial R}{\partial t} = \frac{i\hbar}{m} \frac{\nabla R}{R} \nabla S + \frac{i\hbar}{2m} \nabla^2 S, \quad (3.37)$$

$$\frac{\partial S}{\partial t} = \frac{\hbar^2}{2m} \frac{\nabla^2 R}{R} - \frac{1}{2m} (\nabla S)^2. \quad (3.38)$$

Este passo é chamado "decomposição de Madelung". Tratando primeiro de (3.37), simplificando e multiplicando ambos os lados por R :

$$\begin{aligned} -\frac{\partial}{\partial t}R^2 &= R(\nabla R)\frac{\nabla S}{m} + \frac{R^2}{2} \frac{\nabla^2 S}{m} \\ -\frac{\partial}{\partial t}R^2 &= \nabla \left(\frac{R^2(\nabla S)}{2m} \right). \end{aligned}$$

Definindo: $\mathbf{J} = \frac{1}{2}R^2\mathbf{v}$, tal que $\mathbf{v} = \frac{\nabla S}{m}$ e $R^2 = \Psi\Psi^* = |\Psi|^2$, temos:

$$\frac{\partial}{\partial t}R^2 + \nabla \cdot \mathbf{J} = 0. \quad (3.39)$$

Agora sobre (3.38), multiplicando os dois lados por -1 :

$$-\frac{\partial R}{\partial t} = \frac{1}{2m}(\nabla S)^2 - \frac{\hbar}{2m} \frac{\nabla^2 R}{R}. \quad (3.40)$$

A equação (3.39) é uma forma da equação de continuidade para a densidade de probabilidade, a segunda é uma versão da equação de Hamilton-Jacobi da mecânica clássica. Estas são as chamadas *equações de Madelung*. Elas são uma maneira clássica de representar as equações de Schrodinger a partir do ansatz proposto. Vale ressaltar que, apesar da abordagem clássica, as equações de Madelung ainda não possuem uma definição matemática completamente desenvolvida, e portanto é necessário cautela ao tomarmos qualquer conclusão a partir delas. Esta discussão é abordada com mais profundidade em (REDDIGER; POIRIER, 2022.).

3.3 MÉTODO VARIACIONAL ESTOCÁSTICO

De acordo com Koide (KOIDE; AL., 2015): "O método variacional estocástico (MVE) pode descrever comportamentos clássicos e quânticos de maneira unificada". A aplicabilidade deste método é vasta, utilizamos o MVE a fim de encontrar a melhor base de funções, selecionando aleatoriamente seus parâmetros não-lineares e os otimizando. Utilizando o princípio variacional em conjunto com a mecânica estocástica, podemos descrever sistemas quanto-mecânicos de forma precisa por quantização estocástica, sem sequer utilizar a equação de Schrödinger. Essa quantização é importante, pois abre uma nova perspectiva aos cientistas dispostos a explorar as capacidades desse método, comparar suas vantagens e desvantagens com outros já propostos e como avançar a partir daqui. Na seção inicial, introduzo o método continuando o formalismo de Nelson, introduzindo sua dinâmica a partir de sua conexão com o

princípio da mínima ação. Em seguida, explicamos o que é a quantização estocástica e sua necessidade de utilizá-la em conjunto a métodos numéricos. Ao final, exponho um exemplo deste tipo de quantização para o oscilador harmônico, aproximado para a EDE de Ornstein-Uhlenbeck.

3.3.1 Dinâmica dos processos de difusão

O MVE busca uma representação unificada da mecânica clássica com a mecânica quântica. Nelson, em 1966, propôs uma interpretação totalmente clássica da equação de Schrödinger, a partir da demonstração que toda partícula massiva está sujeita a um movimento browniano com difusão igual a $\frac{\hbar}{2m}$, sob influência de um campo vetorial dado pela segunda lei de Newton (NELSON, 1966.). Em (NELSON, 1988.), obteve uma dinâmica conservativa para processos de difusão análoga a dinâmica determinística conservativa, unificando mecânica estocástica e princípio variacional, aprofundada posteriormente por (YASUE, 1981.). Tomando (KOIDE; AL., 2015) como referência principal para introduzir a dinâmica na mecânica de Nelson. Considere a otimização da lagrangiana de uma partícula

$$L = \frac{m}{2} \dot{\mathbf{r}}(t) - V(\mathbf{r}(t)), \quad (3.41)$$

onde m é a massa da partícula e V seu potencial. Para implementar a variação estocástica a esta lagrangiana, precisamos expressar cada termo pela sua respectiva quantia estocástica.

No nosso caso, a ação estocástica \mathcal{I} (utilizamos desta forma não convencional pois I já foi definido anteriormente como a integral de Itô) é escrita em termos das derivadas D_+ e D_- :

$$\mathcal{I}[\mathbf{r}(t)] = \int_{t_i}^{t_f} dt E[L(\mathbf{r}, D_+\mathbf{r}, D_-\mathbf{r})] \quad (3.42)$$

$$= \int_{t_i}^{t_f} dt E \left[\frac{m}{4} [(D_+\mathbf{r}(t))^2 + (D_-\mathbf{r}(t))^2 - V(\mathbf{r}(t))] \right], \quad (3.43)$$

dado $d\mathbf{r}(t) = \mathbf{u}(\mathbf{r}(t), t)dt + \sqrt{2\nu}d\mathbf{W}_t$, tal que ν caracteriza o termo de ruído. Da definição de "mean forward derivative" e sendo $f(\mathbf{x}, t)$ uma função arbitrária infinitesimal que

desaparece nos limites t_i e t_f , temos que a variação do termo cinético é dada por

$$\int_{t_i}^{t_f} dt E[(D_+ \mathbf{r}(t) + D_- \mathbf{f}(\mathbf{r}(t), t))^2] - \int_{t_i}^{t_f} dt E[(D_+ \mathbf{r}(t))^2] \quad (3.44)$$

$$= 2 \int_{t_i}^{t_f} dt E[\mathbf{b}_+(\mathbf{r}, t) \cdot D_+ \mathbf{f}(\mathbf{r}(t), t)] \quad (3.45)$$

$$= -2 \int_{t_i}^{t_f} dt E[\{D_- \mathbf{b}_+(\mathbf{r}, t)\} \cdot \mathbf{f}(\mathbf{r}(t), t)]. \quad (3.46)$$

A variação resultante é, portanto

$$\delta \mathcal{I} = \int_{t_i}^{t_f} dt E \left[\left\{ -\frac{m}{2} [D_- \mathbf{b}_+(\mathbf{r}, t) + D_+ \mathbf{b}_-(\mathbf{r}, t)] - V(\mathbf{r}(t)) \right\} \cdot \mathbf{f}(\mathbf{r}(t), t) \right]. \quad (3.47)$$

Pelo princípio da mínima ação, temos que $\delta \mathcal{I} = 0$, portanto:

$$\left[-\frac{m}{2} [D_- \mathbf{b}_+(\mathbf{r}, t) + D_+ \mathbf{b}_-(\mathbf{r}, t)] - V(\mathbf{r}(t)) \right]_{\mathbf{r}(t)=x} = 0, \quad (3.48)$$

que é formalmente expressa como:

$$\left[D_- \frac{\partial L}{\partial D \mathbf{r}(t)} + D_+ \frac{\partial L}{\partial D_- \mathbf{r}(t)} - \frac{\partial L}{\partial \mathbf{r}(t)} \right]_{\mathbf{r}(t)=x}. \quad (3.49)$$

Esta é a generalização da equação de Euler-Lagrange na forma estocástica. Portanto, temos que da definição de D_- e D_+ :

$$D_- \mathbf{b}_+(\mathbf{r}, t) = \left(\frac{\partial}{\partial t} + \mathbf{b}_-(\mathbf{r}(t), t) \nabla - v \nabla^2 \right) \mathbf{b}_+(\mathbf{r}(t), t) \quad (3.50)$$

$$D_+ \mathbf{b}_-(\mathbf{r}, t) = \left(\frac{\partial}{\partial t} + \mathbf{b}_+(\mathbf{r}(t), t) \nabla + v \nabla^2 \right) \mathbf{b}_-(\mathbf{r}(t), t). \quad (3.51)$$

Substituindo em (3.48), obtemos:

$$\left(\frac{\partial}{\partial t} + \mathbf{v} \nabla \right) \mathbf{v}(x, t) = -\frac{1}{m} \nabla V(x) + 2v^2 \nabla \sqrt{\rho^{-1}(x, t)} \nabla^2 \sqrt{\rho(x, t)}, \quad (3.52)$$

tal que v é dado por (3.8). As equações (3.12) e (3.52) determinam a dinâmica otimizada da ação. Estas equações acopladas podem ser trazidas a uma forma familiar, fazendo:

$$\psi(x, t) = \sqrt{\rho} e^{i\theta(x, t)}, \quad (3.53)$$

tal que

$$v(x, t) = 2v \nabla \theta(x, t) \quad (3.54)$$

e substituindo nas equações (3.12) e (3.52), podemos finalmente retornar até equação de Schrödinger.

Michele Pavon em (PAVON, 2000.) obteve as relações entre elementos quanto-mecânicos e clássicos de maneira coerente, incluindo as respectivas equações cinemáticas como restrições para o caminho ótimo $x(t)$ do funcional ação \mathcal{I} .

3.3.2 Métodos numéricos

Introduzimos aqui a noção de ensemble estatístico. Ao invés de lidar com densidade de probabilidade, trabalharemos com uma quantidade que é proporcional àquela, porém mais "transparente". Colhemos uma grande quantidade de ensembles idênticos e independentes, distribuídos de acordo com $W(t)$. Chamamos este conjunto de ensemble estatístico. Esse método consiste em, primeiro, determinar o tamanho h de cada iteração e uma condição inicial. Digamos que você queira achar a solução para uma ED como esta: $f'(t) = a f(t) + b$, sendo $a, b \in \mathbb{R}$, tal que $f(0) = 0$ e $h = 1/100$. Assim, precisamos iterar 100 vezes a equação para $f'(t)$, que terá a forma

$$f[(i+1)h] \approx f[ih] + hf'[ih],$$

sendo também aplicável para duas variáveis. Porém, o método não parece razoável para solucionar EDEs, pois se pretendemos computar um valor no presente, este pode depender de qualquer valor no passado. A única coisa que fornece informações sobre o processo estocástico é a distribuição de probabilidade. Se quisermos adaptar esse método, pegamos uma amostra de caminho do processo de Wiener para que saibamos como "retroceder" a instantes no passado.

Em posse dessa amostra, temos então:

$$\begin{aligned} x[(i+1)h] &\approx x[ih] + dx[ih], \\ x[(i+1)h] &\approx x[ih] + (b(x_t, t)dt + \sigma(x_t, t)dW_t). \end{aligned} \quad (3.55)$$

A partir dessa equação, podemos computar tudo. Assumindo um valor inicial $x[0]$, podemos computar b e σ a partir de suas coordenadas, dt aproximamos para h e dW_t pode ser computado a partir da amostra do caminho W_t , $W[(i+1)h] - W[ih]$. O problema é: por vezes, não temos como obter uma amostra de caminho. Para isso, recorreremos ao *método de Euler-Maruyama*, que é similar ao exposto acima, mas adaptado para EDE's.

A solução numérica aproximada para uma EDE num intervalo de tempo $[0, T]$, com condição inicial x_0 , compreendendo os diferenciais como a diferença entre o valor do processo em dois instantes de tempo sucessivos, tal que esta diferença sirva como uma partição infinitesimal dos instantes de tempo. Graças às propriedades do processo de Wiener, $dW(t) \sim N(0, dt)$, o método de Euler-Maruyama pode ser aplicado iterando o seguinte algoritmo

$$x(t_{i+1}) = x(t_i) + b(x(t_i), t_i)dt + \sigma(x(t_i), t_i) \sqrt{dt}; \quad i = 0, 1, \dots, N - 1. \quad (3.56)$$

3.3.3 Aplicação do método: processo de Ornstein-Uhlenbeck

A autora (KOPPE; GRECKSCH; PAUL, 2017.) mostrou que aplicando o método de quantização de Nelson para o oscilador harmônico nos fornece as equações diferenciais estocásticas acopladas:

$$dx(t) = u(x(t))dt + dW_f(t), \quad (3.57)$$

$$du(x(t)) = x(t)dt + \tilde{q}dW_b(t). \quad (3.58)$$

Um algoritmo deste tipo pode ser descrito em duas etapas: (i) obter $u(x)$ resolvendo a EDO (KOPPE W. GRECKSCH, 2016.)

$$\begin{aligned} mu(x)\frac{du(x)}{dx} + \frac{\hbar}{2}\frac{d^2u(x)}{dx^2} - \frac{dV(x)}{dx} &= 0 \\ \Leftrightarrow \frac{m}{2}u^2(x) + \frac{\hbar}{2}\frac{du(x)}{dx} - V(x) &= -E \end{aligned} \quad (3.59)$$

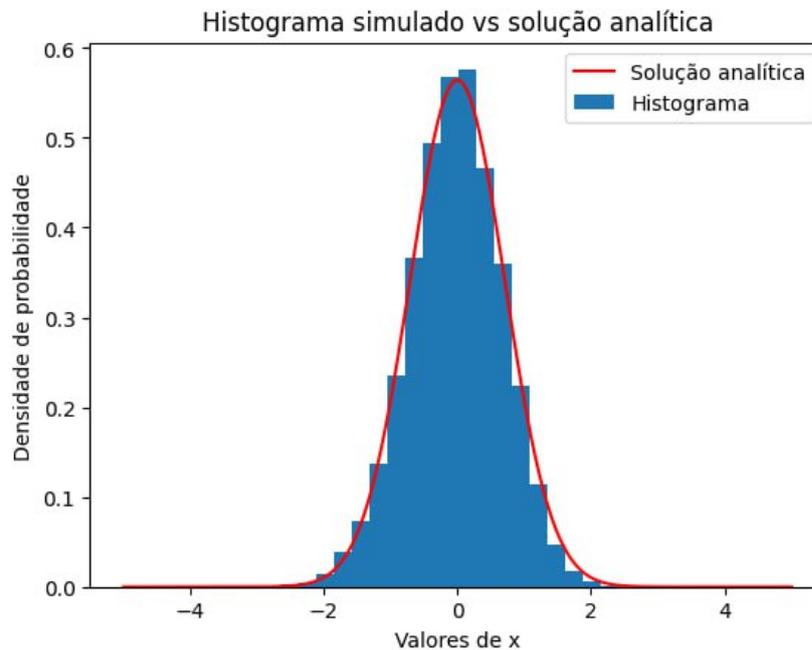
numericamente e (ii) utilizar a solução numérica de $u(x)$ para resolver a EDE (3.57) fazendo a aproximação $u(x(t)) = u(x)$. Esta tarefa soa elementar via método de Euler-Maruyama, mas em sua tese (KOPPE W. GRECKSCH, 2016.), Köppe estuda o caso do potencial do poço duplo unidimensional:

$$V(x) = \frac{V_0}{a^4}(x^2 - a^2)^2, \quad (3.60)$$

e explica que a medida que escolhermos diferentes potenciais, a obtenção de $u(x)$ pode ser uma tarefa bastante complexa. Portanto, a forma analítica de $u(x)$ nem sempre é conhecida, o que abre ainda mais possibilidade para desenvolvimento e estudos neste campo.

Ainda que haja esta dificuldade de resolver as equações diferenciais estocásticas acopladas, queremos ilustrar o método numérico mesmo que seja em um exemplo simples. Portanto, consideremos $\hbar = 1 = m$ em (3.59), e observemos que $u(x) = -x$ é solução para o potencial do oscilador harmônico $V(x) = \frac{x^2}{2}$, com energia $E = \frac{1}{2}$ e distribuição de probabilidade $P(x) = \frac{1}{\sqrt{\pi}}e^{-x^2}$. Este caso especial para $u(x)$ faz com que as EDEs (3.57) e (3.58) desacoplem, de forma que a primeira se torna um processo Ornstein-Uhlenbeck. Resolvendo (3.57) pelo método tradicional de Euler-Maruyama, obtemos o histograma mostrado na figura 7.

Figura 7 – Gráfico da FDP pela posição para o processo de Ornstein-Uhlenbeck.



Fonte: O autor (2023).

A solução numérica completa das EDEs acopladas da quantização de Nelson exige um esquema de iterações para produzir um ensemble de trajetórias acopladas, que é uma tarefa bastante complexa em geral. No entanto, esta abordagem pode em princípio ser usada para validar uma regularização discreta de uma integral funcional interpretada com o cálculo estocástico, em concordância com o que concluiu (PAVON, 2000.) e, portanto, a quantização de Nelson é uma das maneiras de regularizar integrais funcionais.

Poderíamos seguir com este trabalho apenas realizando aplicações deste algoritmo em outros potenciais, mas preferimos entender a aplicabilidade deste programa

de quantização para além de soluções numéricas, visto que uma das várias importâncias da quantização estocástica é trazer uma abordagem clássica a problemas que, aparentemente, não possuiriam tal construção. Motivados por isso, partimos para a quantização estocástica no contexto de geometria diferencial, aplicada a fase de Berry.

4 FASE DE BERRY

Vetores, ou no caso da mecânica quântica, vetores de estado, ao percorrerem um caminho fechado no espaço de parâmetros, sob um regime adiabático, adquirem uma fase puramente geométrica conhecida como fase de Berry. Nesta segunda parte da dissertação, tenho como objetivo entender o que é esta fase, como obtê-la e propor um método para descrevê-la no formalismo da mecânica de Nelson. Iniciamos com a introdução do teorema adiabático (WU; YANG, 2005.), para assim então definir a fase de Berry ((ZWANZIGER; KOENIG; PINES, 1990.), (PEDROSA; LIMA, 2014.), (CAROLLO et al., 2003.), (ALVES; THOMAZ, 2007.), (BERRY, 1983.))

4.1 TEOREMA ADIABÁTICO

Sistemas que evoluem seu estado gradualmente, a partir das mudanças de condições do sistema, modificam sua densidade de probabilidade ao longo do processo. Se esta evolução ocorre sob regime adiabático, o auto-estado do hamiltoniano inicial permanecerá constante até o final do processo. Em que pese esse fato, os estados final e inicial se diferem por uma fase. Este regime impõe uma extremização do tempo, chamamos estes estados extremizados de "estados instantâneos". Para encontrar auto-estados associados, é preciso obter as fases que satisfazem a equação de Schrödinger. A

gora, utilizando notação de Dirac, suponha um estado $|\psi(t)\rangle$, escrito como uma superposição de auto-estados instantâneos, que compõem uma base ortonormal e completa, na forma do *ansatz* geral:

$$|\psi(t)\rangle = \sum_n c_n(t)|\psi_n(t)\rangle. \quad (4.1)$$

A equação de Schrödinger dependente do tempo é:

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H|\psi(t)\rangle, \quad (4.2)$$

substituindo (4.1) em (4.2):

$$i\hbar \sum_n (\dot{c}_n(t)|\psi_n(t)\rangle + c_n(t)|\dot{\psi}_n(t)\rangle) = \sum_n c_n(t)E_n(t)|\psi_n(t)\rangle. \quad (4.3)$$

Projetando $\langle \psi_k(t) |$ sobre todos os vetores de estado, tendo em mente que $|\psi_n(t)\rangle$ forma uma base ortonormal, temos:

$$i\hbar\dot{c}_n(t) = c_k(t)E_k(t) - i\hbar \sum_n \langle \psi_k(t) | \dot{\psi}_n(t) \rangle c_n. \quad (4.4)$$

A equação acima fornece uma visão geral sobre o que ocorre com esses auto-estados instantâneos. Estamos interessados na aproximação em que, na transição de $n \rightarrow k$, o intervalo de tempo decorrido é extremizado. Portanto: $\langle \psi_k(t) | \dot{\psi}_n(t) \rangle \approx 0$, para $k \neq n$. Sendo assim:

$$i\hbar\dot{c}_k(t) = (E_k(t) - i\hbar\langle \psi_k | \dot{\psi}_k \rangle) c_k. \quad (4.5)$$

Esta E.D.O. têm solução geral:

$$c_k(t) = c(0) \exp \left[\frac{1}{i\hbar} \int_0^t \left(E_k(t') - i\hbar \langle \psi_k(t') | \dot{\psi}_n(t') \rangle \right) dt' \right]. \quad (4.6)$$

Para processos que tomam muitos intervalos de tempo para concluir, os termos $\langle \psi_k(t') | \dot{\psi}_n(t') \rangle$ (chamado "termo de acoplamento") começam a acumular, então precisamos fazer uma análise mais rigorosa desta aproximação. No regime adiabático, assume-se que o Hamiltoniano do sistema muda lentamente com relação ao tempo. Essa suposição nos permite tratar o sistema como, aproximadamente, independente do tempo em escalas curtas. A equação de Schrödinger independente do tempo para a aproximação adiabática é:

$$H(t)|\psi_n(t)\rangle = E_n(t)|\psi_n(t)\rangle, \quad (4.7)$$

A ideia central é que: se o Hamiltoniano varia lentamente, o sistema quântico permanecerá em um autovetor instantâneo do Hamiltoniano, desde que os níveis de energia não se cruzem ou se aproximem muito durante a variação. Diferenciando 4.7 com a relação ao tempo, temos

$$\dot{H}|\psi_n(t)\rangle + H|\dot{\psi}_n(t)\rangle = \dot{E}_n|\psi_n(t)\rangle + E_n|\dot{\psi}_n(t)\rangle. \quad (4.8)$$

Realizando a sobreposição pelo estado $\langle \psi_k(t) |$ (sabendo que: $\langle \psi_k(t) | H = \langle \psi_k(t) | E_k$)

$$\langle \psi_k(t) | \dot{H} | \psi_n(t) \rangle + E_k \langle \psi_k(t) | \dot{\psi}_n(t) \rangle = \dot{E}_n \langle \psi_k(t) | \psi_n(t) \rangle + E_n \langle \psi_k(t) | \dot{\psi}_n(t) \rangle, \quad (4.9)$$

em que: $\langle \psi_k(t) | \psi_n(t) \rangle = 0$ pois $k \neq n$. Desta forma:

$$\langle \psi_k(t) | \dot{\psi}_n(t) \rangle = \frac{\langle \psi_k(t) | \dot{H} | \psi_n(t) \rangle}{(E_n - E_k)} \quad (4.10)$$

$$\langle \psi_k(t) | \dot{\psi}_k(t) \rangle = \frac{\dot{H}_{kn}}{(E_n - E_k)}. \quad (4.11)$$

Concluimos, portanto, que a matriz dos termos $\langle \psi_k(t) | \dot{\psi}_n(t) \rangle$ está associada com a taxa de variação do hamiltoniano. Como estabelecemos a extremização do tempo, termos que dependam desta taxa tem efeito desprezível em (4.5).

Após todo esse panorama, podemos enunciar o teorema adiabático: Seja T o tempo decorrido durante uma mudança de configuração do sistema e definindo o módulo de ψ como sendo: $\|\psi\| \equiv \sqrt{\langle \psi | \psi \rangle}$, temos que:

$$\|\psi(t) - \psi_{adiabatico}(t)\| \leq \frac{C}{T}; \quad 0 \leq t \leq T. \quad (4.12)$$

Portanto, a aproximação adiabática se aproxima da real solução da equação de Schrödinger quanto maior for o T , tal que o erro associado seja desprezível. Voltando para (4.6), vamos nomear três termos:

$$\theta_k(t) \equiv -\frac{i}{\hbar} \int_0^t E_k(t') dt', \quad (4.13)$$

$$v_k(t) \equiv i \langle \psi_k(t) | \dot{\psi}_n(t) \rangle, \quad (4.14)$$

$$\gamma_k(t) \equiv \int_0^t v_k(t') dt', \quad (4.15)$$

tal que c_k possa ser reescrito na forma

$$c_k(t) = c(0) e^{i\theta_k(t)} e^{i\gamma_k(t)}. \quad (4.16)$$

O parâmetro $\theta_k(t)$ diz respeito a fase adquirida proporcional ao tempo que o sistema passa em um auto-estado de energia. A fase $\gamma_k(t)$ é chamada de **fase de Berry** e possui uma peculiaridade: apesar da dependência temporal, é de natureza puramente geométrica. Seu valor depende apenas do caminho percorrido no espaço de configuração, os parâmetros que definem o caminho estão todos contidos em H e, além disso, seu valor permanecerá constante ao longo do tempo. Podemos confirmar que (4.16) é consistente com a afirmação lembrando o vetor de estado para H independente do tempo:

$$\begin{aligned} |\psi_k(t)\rangle &= e^{-\frac{iE_k t}{\hbar}} |k\rangle \\ \langle \psi_k(t) | \dot{\psi}_k(t) \rangle &= \langle k | e^{\frac{iE_k t}{\hbar}} e^{-\frac{iE_k t}{\hbar}} \frac{-iE_k}{\hbar} |k\rangle \\ \langle \psi_k(t) | \dot{\psi}_k(t) \rangle &= \frac{-iE_k}{\hbar}. \end{aligned} \quad (4.17)$$

Substituindo (4.17) em (4.14) e tratando E_k como constante em (4.13):

$$\begin{cases} \theta_k(t) = \frac{-E_k t}{\hbar}, \\ \gamma_k(t) = \frac{E_k t}{\hbar}, \end{cases}$$

substituindo em $c_k(t)$: $c_k(t) = 1$. De (4.1), concluímos que:

$$|\psi(t)\rangle = \sum_n |\psi_n\rangle; \quad H \neq H(t). \quad (4.18)$$

Isto é: considerando a evolução temporal dos auto-estados, (4.18) demonstra que as fases não afetam estes estados evoluídos no tempo, como estabelece o regime adiabático.

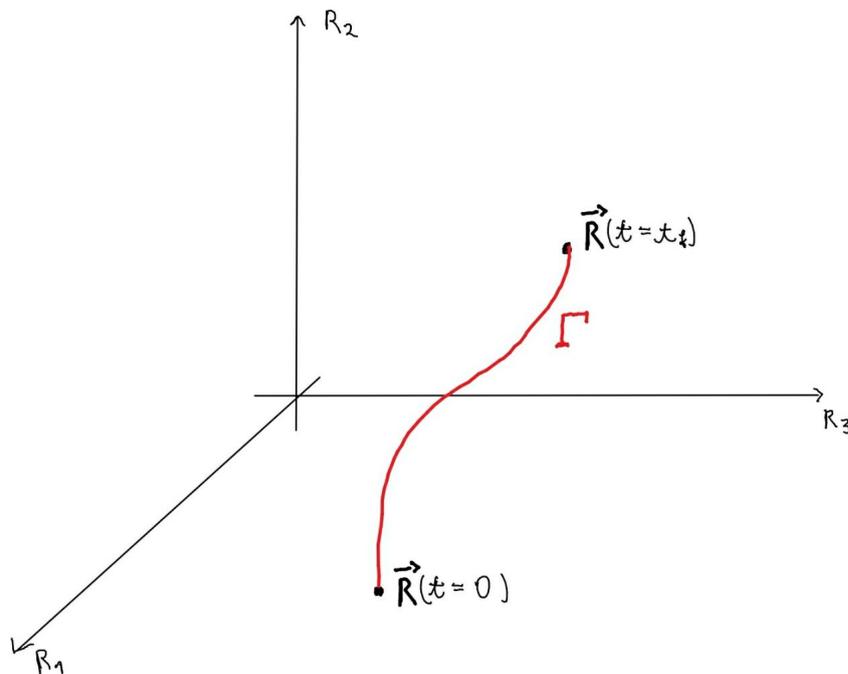
4.2 FASE DE BERRY

Suponha um hamiltoniano H dependente de N parâmetros \mathbf{R} , $H = H(\mathbf{R})$, tal que $\mathbf{R} = (R_1, R_2, R_3, \dots, R_N) \in \mathbb{R}^N$. Desta forma:

$$H(\mathbf{R})|\psi_n(\mathbf{R})\rangle = E_n(\mathbf{R})|\psi_n(\mathbf{R})\rangle, \quad (4.19)$$

em que $n = 1, 2, 3, \dots, N$, $|\psi_n(\mathbf{R})\rangle$ estão contidos no espaço de parâmetros, são ortogonais e todos são conhecidos. Supondo agora que os parâmetros \mathbf{R} dependam do tempo: $\mathbf{R}(t) = (R_1(t), R_2(t), \dots, R_N(t))$, teremos o seguinte esquema para a evolução temporal no espaço de configuração:

Figura 8 – Representação do caminho Γ que leva o estado $\psi_n(\mathbf{R}(0))$ para $\psi_n(\mathbf{R}(t))$.



Fonte: O autor (2023).

O novo hamiltoniano é dependente do tempo. Das propriedades de derivada covariante (apêndice B):

$$\frac{\partial}{\partial t} |\psi_n(\mathbf{R}(t))\rangle = \nabla_{\mathbf{R}} |\psi_n(\mathbf{R}(t))\rangle \cdot \frac{d\mathbf{R}}{dt}. \quad (4.20)$$

A partir disso, voltemos nossa atenção para as equações (4.14) e (4.15). Reescrevendo a primeira considerando 4.20, temos:

$$v_n(t) = i \langle \psi_k(\mathbf{R}(t)) | \nabla_{\mathbf{R}} |\psi_n(\mathbf{R}(t))\rangle \cdot \frac{d\mathbf{R}}{dt}. \quad (4.21)$$

Isto é: O gradiente de $|\psi_n(\mathbf{R}(t))\rangle$ nos dá N componentes de um vetor que realiza produto interno com $\hat{\mathbf{R}}$. Substituindo (4.21) em (4.15):

$$\begin{aligned} \gamma_n(t) &= \int_0^t \left(i \langle \psi_k(\mathbf{R}(t')) | \nabla_{\mathbf{R}} |\psi_n(\mathbf{R}(t'))\rangle \cdot \frac{d\mathbf{R}}{dt'} \right) dt' \\ \gamma_n(t) &= \int_0^t i \langle \psi_k(\mathbf{R}(t)) | \nabla_{\mathbf{R}} |\psi_n(\mathbf{R}(t))\rangle \cdot d\mathbf{R}. \end{aligned} \quad (4.22)$$

A equação acima confirma o que foi alegado anteriormente: γ depende apenas do espaço de configuração, integrando sobre o caminho Γ ,

$$\gamma_n(t) = \int_{\Gamma} i \langle \psi_k(\mathbf{R}(t)) | \nabla_{\mathbf{R}} |\psi_n(\mathbf{R}(t))\rangle \cdot d\mathbf{R}. \quad (4.23)$$

Portanto, independente de quanto tempo leve para realizarmos a integração sobre o caminho Γ , o valor da fase de Berry γ_n depende apenas do caminho de integração. É importante ressaltar que esta fase não se limita apenas ao regime adiabático, mas o formalismo tradicional (conforme dita (BERRY, 1983.)) é especificamente relacionado a sistemas quânticos em transformações adiabáticas. A fase de Berry em contextos dinâmicos, por exemplo, lida com situações em que a taxa de variação dos parâmetros não é necessariamente lenta, podendo ser aplicada em contextos mais gerais dependendo das características do sistema e da natureza das mudanças nos parâmetros, como em (MISAKI; MIYASHITA; NAGAOSA, 2021.), (SANBORN; SUSLOV; VINET, 2011.).

Definição: O vetor $\mathbf{A}_n(\mathbf{R})$ é chamado "conexão de Berry", dado por:

$$\mathbf{A}_n(\mathbf{R}) = i \langle \psi_k(\mathbf{R}(t)) | \nabla_{\mathbf{R}} |\psi_n(\mathbf{R}(t))\rangle. \quad (4.24)$$

Este vetor está definido no espaço de configurações. Há uma conexão de Berry para todos os auto-estados do sistema. Ademais, no espaço de configuração, podemos relacionar os vetores de estado da seguinte maneira: Seja $|\psi'_n\rangle$ um novo vetor de estado do espaço de configurações e $\beta(\mathbf{R})$ uma função real, temos que:

$$|\psi'_n\rangle = e^{-i\beta(\mathbf{R})} |\psi_n\rangle.$$

A conexão de Berry associada a $|\psi'_n\rangle$ é:

$$\mathbf{A}'_n(\mathbf{R}) = i\langle\psi'_k(R(t))|e^{i\beta(\mathbf{R})}\nabla_{\mathbf{R}}e^{-i\beta(\mathbf{R})}|\psi_n(R(t))\rangle$$

$$\mathbf{A}'_n(\mathbf{R}) = \mathbf{A}_n(\mathbf{R}) + \nabla_{\mathbf{R}}\beta(\mathbf{R}). \quad (4.25)$$

A equação acima nos mostra que $\mathbf{A}_n(\mathbf{R})$ obedece a transformação de calibre. Dessa forma, podemos imaginar $\mathbf{A}_n(\mathbf{R})$ como um potencial vetorial no espaço de configuração. Agora, substituindo (4.25) em (4.23):

$$\gamma_n = \int_{\Gamma} \mathbf{A}_n(\mathbf{R}) \cdot d\mathbf{R}. \quad (4.26)$$

Se quisermos calcular a fase de Berry com referência a (4.25), temos:

$$\begin{aligned} \gamma'_n &= \int_{\Gamma} \mathbf{A}'_n(\mathbf{R}) \cdot d\mathbf{R} \\ \gamma'_n &= \int_{\Gamma} \mathbf{A}_n(\mathbf{R}) \cdot d\mathbf{R} + \int_{\mathbf{R}_i}^{\mathbf{R}_f} \nabla_{\mathbf{R}}\beta(\mathbf{R}) \cdot d\mathbf{R}. \end{aligned} \quad (4.27)$$

Note que a fase de Berry somente obedece a transformação de calibre se, e somente se, o caminho Γ é fechado. Supondo dessa forma, temos:

$$\begin{aligned} \gamma'_n &= \int_{\Gamma} \mathbf{A}_n(\mathbf{R}) \cdot d\mathbf{R} + \int_{\mathbf{R}}^{\mathbf{R}} \nabla_{\mathbf{R}}\beta(\mathbf{R}) \cdot d\mathbf{R} \\ \gamma'_n &= \int_{\Gamma} \mathbf{A}_n(\mathbf{R}) \cdot d\mathbf{R} + \beta(\mathbf{R}) - \beta(\mathbf{R}) \\ \gamma'_n &= \gamma_n. \end{aligned} \quad (4.28)$$

A equação acima conclui que a fase γ_n é um observável, acumulada pelo sistema num movimento sob uma curva fechada no espaço de configuração. A fase geométrica de Berry é, portanto, invariante de calibre para um loop fechado no espaço de configuração.

Propriedades:

- Se $\psi_n(t)$ é real, $\gamma_n = 0$;
- Se o espaço de configuração é unidimensional, $\gamma_n = 0$;
- Em 3 dimensões, a fase de Berry é enunciada como um fluxo através de uma superfície:

$$\gamma_n = \oint_{\Gamma} \mathbf{A} \cdot d\mathbf{R} = \oiint_S (\nabla \times \mathbf{A}) \cdot d\mathbf{a}. \quad (4.29)$$

Definição: Curvatura de Berry

$$\mathbf{D} = \nabla_{\mathbf{R}} \times \mathbf{A}(\mathbf{R}). \quad (4.30)$$

Neste caso, podemos imaginar a fase de Berry análoga a um campo magnético.

É possível concluir, portanto, que a fase de Berry é o espaço de configuração (ou espaço de parâmetros) descrevendo uma variedade diferenciável, através de um ciclo adiabático fechado, resultado das propriedades dos parâmetros do hamiltoniano.

4.3 FASE DE BERRY A PARTIR DA FORMULAÇÃO POR INTEGRAIS DE CAMINHO

Seja $H_s(R(t))$ o hamiltoniano não degenerado num tempo t na representação de Schrödinger, $R(t)$ uma função vetorial, periodicamente lenta, d -dimensional, cujos componentes são funções clássicas dependentes do tempo. Seja T o tempo quando $R(0) = R(T)$, a dependência temporal do hamiltoniano é devido apenas a $R(t)$.

$$H_s(R(t))|\psi_n(R(t))\rangle_s = E_n(R(t))|\psi_n(R(t))\rangle_s, \quad (4.31)$$

tal que

$${}_s\langle\psi_n(R(t))|\psi_m(R(t))\rangle_s = \delta_{nm}. \quad (4.32)$$

Em cada tempo t , o hamiltoniano possui um conjunto completo de autovetores ortonormais, isto é $\sum_n |\psi_m(R(t))\rangle_s {}_s\langle\psi_m(R(t))| = I$ (operador identidade). Simplificando a notação:

$$\begin{cases} H_s(R(t)) \equiv H(t); \\ |\psi_n(R(t))\rangle_s \equiv |\psi(t)\rangle; \\ E_n(R(t)) \equiv E_n(t). \end{cases}$$

Daqui em diante, mesmo sem índice, estarei utilizando apenas a representação de Schrödinger.

Seja $|\psi(0)\rangle$ o vetor de estado inicial, queremos calcular sua evolução dada a condição adiabática. Da equação de Schrödinger, temos que

$$H(t)|\psi(t)\rangle = i\hbar \frac{d}{dt} |\psi(t)\rangle, \quad (4.33)$$

cuja solução é:

$$|\psi(t + \Delta t)\rangle = e^{\frac{-i}{\hbar}\Delta t H(t)}|\psi(t)\rangle; \quad \Delta t \rightarrow 0 \quad (4.34)$$

Definindo o operador de evolução temporal:

$$u(t) = \exp\frac{i}{\hbar}\Delta t H(t) \quad (4.35)$$

Dividindo o intervalo $[0, t]$ em "n" intervalos Δt , em que $n \rightarrow \infty$ e $\Delta t \rightarrow 0$, tal que $t = n\Delta t$. Da equação acima, para $t > 0$:

$$|\psi(t)\rangle = e^{\frac{-i}{\hbar}\Delta t H((n-1)\Delta t)} \dots e^{\frac{-i}{\hbar}H(t)\Delta t} e^{\frac{-i}{\hbar}H(0)\Delta t} |\psi(0)\rangle \quad (4.36)$$

Podemos expressar $|\psi(0)\rangle$ em termos dos autovetores de $H(0)$:

$$|\psi(0)\rangle = \sum_m c_m(0) |\psi_m(0)\rangle \quad (4.37)$$

introduzindo o operador identidade num tempo $t_j = j\Delta t$, $j = 1, 2, \dots, n - 1$ e tomando o limite $\Delta t \rightarrow 0$, então projetando o autovetor sobre o vetor posição, temos

$$\begin{aligned} \langle x|\psi(t)\rangle = \psi(x, t) = \sum_m c_m(0) \left\{ \sum_{n_0, n_1, \dots, n_{N-1}} \langle x|u((N-1)\Delta t)|\psi_{n_{N-1}}((N-1)\Delta t)\rangle \right. \\ \left. \langle \psi_{n_{N-1}}((N-1)\Delta t)|\psi_{n_{N-1}}((N-2)\Delta t)\rangle \dots \langle \psi_{n_1}(\Delta t)|u(0)|\psi_{n_0}(0)\rangle \langle \psi_0|\psi_m(0)\rangle \right\} \end{aligned} \quad (4.38)$$

Queremos analisar a evolução deste sistema para o hamiltoniano sob regime adiabático. De (4.31), considere um hamiltoniano degenerado, cuja degenerescência é representada por $\lambda_n = (1, 2, \dots, D_n)$. Em cada instante "t", o hamiltoniano possui um conjunto completo de auto-estados ortonormais $|\psi_n^{\lambda_n}(t)\rangle$

$$\langle \psi_n(t)|\psi_m(t)\rangle = \delta_{nm}\delta_{nm} \quad (4.39)$$

$$H(t)|\psi_n^{\lambda_n}(t)\rangle = E_n(t)|\psi_n^{\lambda_n}(t)\rangle; \quad (\lambda_n = 1, 2, \dots, D_n) \quad (4.40)$$

D_n é a degenerescência do autovalor $E_n(t)$. Temos $E_n(t) \neq E_m(t)$ se $n \neq m$. Da equação (4.1), temos que $u(t_i)|\psi_{n_i}(t_i)\rangle$ é a evolução do vetor de estado ao longo de um tempo Δt . Para o caso quando $D_n = 1$ (não degenerado), temos:

$$u(t_i)|\psi_{n_i}(t_i)\rangle \propto |\psi_{n_i}(t_i + \Delta t)\rangle \quad (4.41)$$

tal que

$$\langle \psi_{n_i+1}(t_i + \Delta t)|u(t_i)|\psi_{n_i}(t_i)\rangle \equiv \delta_{n_i, n_{i+1}} \quad (4.42)$$

Essas equações nos dizem que, se usarmos a base de auto-estados instantâneos $|\psi_n(t_i)\rangle$ para descrever uma integral de caminho com amplitude $\psi(x, t)$, apenas um caminho no espaço de Hilbert contribui para a função. Substituindo (4.31), (4.32) e (4.39) em (4.38):

$$\psi(x, t) = \sum_m c_m(0) \psi_m(x, t) \exp\left[-\frac{i}{\hbar} \int_0^t dt' E_m(t')\right] \langle \psi_m((N-1)\Delta t) | \psi_m((N-2)\Delta t) \rangle \quad (4.43)$$

$$\dots \langle \psi_m(\Delta t) | \psi_m(0) \rangle$$

em que $\psi_m(x, t) \equiv \langle x | \psi_m((N-1)\Delta t) \rangle$. Definimos, agora, a derivada do vetor de estado:

$$\frac{d}{dt} |\psi_m(t)\rangle \equiv \lim_{\Delta t \rightarrow 0} \frac{|\psi_m(t) + \Delta t\rangle - |\psi_m(t)\rangle}{\Delta t} \quad (4.44)$$

Seja: $l = 0, 1, 2, \dots, N-2$, temos que

$$\langle \psi_m((l+1)\Delta t) | \psi_m(l\Delta t) \rangle = 1 + \Delta t \left(\frac{d}{dt} \langle \psi_m(l\Delta t) | \right) | \psi_m(l\Delta t) \rangle \quad (4.45)$$

expandindo a equação acima em série de potência:

$$\langle \psi_m((l+1)\Delta t) | \psi_m(l\Delta t) \rangle = \exp\left\{ \frac{d}{dt} \langle \psi_m(l\Delta t) | \psi_m(l\Delta t) \rangle \right\} \quad (4.46)$$

tomando o limite $\Delta t \rightarrow 0$ em (4.43) e levando em consideração o obtido em (4.46):

$$\psi(x, t) = \sum_m c_m(0) \psi_m(x, t) \exp\left[-\frac{i}{\hbar} \int_0^t dt' E_m(t')\right] \exp\left\{ \int_0^t \frac{d}{dt'} \langle \psi_m(t') | \psi_m(t') \rangle dt' \right\} \quad (4.47)$$

De (4.32), temos:

$$\left(\frac{d}{dt} \langle \psi_m(t) | \right) | \psi_m(t) \rangle = -\langle \psi_m(t) | \left(\frac{d}{dt} | \psi_m(t) \rangle \right) \quad (4.48)$$

Assumindo $\mathbf{R}(t)$ tridimensional: $\mathbf{R}(t) = (X(t), Y(t), Z(t))$, então o lado direito de (4.48) pode ser escrito como

$$\langle \psi_m(t) | \left(\frac{d}{dt} | \psi_m(t) \rangle \right) = \langle \psi_m(\mathbf{R}(t)) | \left(\nabla_{\mathbf{R}} | \psi_m(\mathbf{R}(t)) \rangle \right) \cdot d\mathbf{R} \quad (4.49)$$

Finalmente, obtemos $\psi(x, t)$ escrito como

$$\psi(x, t) = \sum_m c_m(0) \psi_m(x, t) \exp\left[-\frac{i}{\hbar} \int_0^t dt' E_m(t')\right] \exp\left\{ - \int_{\mathbf{R}(0)}^{\mathbf{R}(t)} \langle \psi_m(\mathbf{R}(t)) | \left(\nabla_{\mathbf{R}} | \psi_m(\mathbf{R}(t)) \rangle \right) \cdot d\mathbf{R} \right\} \quad (4.50)$$

Em que a segunda exponencial é a fase geométrica, ou fase de Berry, adquirida ao projetarmos o vetor de estado ao longo de um caminho em x .

4.4 REPRESENTAÇÃO ESTOCÁSTICA DA FASE DE BERRY

Definida a fase de Berry, nosso objetivo agora é, finalmente, buscar sua descrição sob a ótica da mecânica de Nelson.

Teorema ((MOLLAI; FATHI, 2021.)): Se existe um ψ que satisfaça (3.31) e pode ser escrita na forma $\psi(x, t) = e^{-\beta(x, t)}\psi_D(x, t)$, $\beta(x, t) \in \mathbb{C}$, então ψ também é solução de

$$-i\hbar\frac{\partial\psi}{\partial t} = \frac{\hbar^2}{2m}(\nabla + \nabla\beta)^2\psi \quad (4.51)$$

no mesmo artigo, é demonstrado que este teorema nos leva até a equação

$$-i\hbar\frac{\partial\psi}{\partial t} = \frac{\hbar^2}{2m}(\nabla + \nabla\beta)^2\psi + i\hbar\frac{\partial\beta}{\partial t}\psi \quad (4.52)$$

esta é chamada equação dissipativa de Schrödinger (EDS). Utilizando o ansatz (3.32) nesta equação e aplicando a decomposição de Madelung, com cálculo muito similar, é possível obter equações de Madelung em formas alternativas:

$$-\frac{\partial}{\partial t}(S + \hbar\beta_i) = -\frac{\hbar^2}{2m}\frac{(\nabla + \nabla\beta_r)^2R}{R} + \frac{1}{2m}(\nabla S + \hbar\nabla\beta_i)^2 \quad (4.53)$$

$$\left(\frac{\partial}{\partial t} + 2\frac{\partial\beta_r}{\partial t}\right)\rho + (\nabla + 2\nabla\beta_r)\left[\frac{\rho(\nabla S + \hbar\nabla\beta_i)}{m}\right] = 0 \quad (4.54)$$

tal que $\beta = \beta_r + i\beta_i$, β_r sendo a parte real e β_i a imaginária. Portanto, como a fase de Berry $\gamma_n \in \mathbb{C}$, podemos trocar β por γ .

Nelson proveu uma conexão direta dos gradientes, via as equações propostas em (NELSON, 1966.):

$$\mathbf{v}(x, t) = \mathbf{v} = \frac{1}{m}\nabla S \quad (4.55)$$

$$\mathbf{u}(x, t) = \mathbf{u} = \frac{\hbar}{m}\nabla R \quad (4.56)$$

Começando de (4.53), tirando o gradiente de ambos os lados e sabendo que γ_n é puramente imaginário, portanto $\beta_r = 0$ e $\beta_i = \gamma_n$, temos

$$\frac{\partial}{\partial t}(m\mathbf{v} + \hbar\nabla\gamma_n) = -\frac{\hbar^2}{2m}\nabla\left[\frac{2m}{\hbar}\frac{\nabla u}{R}\right] + \frac{1}{2m}[m^2v^2 + \hbar m\mathbf{v} \cdot \nabla\gamma_n + (\nabla\gamma_n)^2], \quad (4.57)$$

que é o resultado final obtido nesta dissertação.

Apesar do cálculo não estar finalizado, esperávamos obter as EDPs (4.53) e (4.54) descritas em termos apenas de u , v e γ_n (como o que é feito em (KOPPE; GRECKSCH;

PAUL, 2017.), para obtenção das equações (8) (9)) e, a partir disso, interpretá-las de acordo com esta nova descrição. Para a equação de continuidade (4.54), é esperado que a evolução adiabática não produza efeitos adversos na equação (como discutido no tópico 3 em (MOLLAI; FATHI, 2021.)). Porém, a equação de Hamilton-Jacobi (4.53) no contexto da fase de Berry, temos

$$-\frac{\partial S}{\partial t} = \frac{1}{2m}(\nabla S)^2 - \frac{\hbar^2}{2m} \frac{\nabla^2 R}{R} + \hbar \frac{\partial \gamma_n(t)}{\partial t} \quad (4.58)$$

Uma descrição clássica para a equação acima, a partir da quantização de Nelson, poderia elucidar aspectos "invisíveis" desta forma usual, trazidos à tona ao introduzirmos os efeitos da difusão nesta interpretação.

5 CONCLUSÕES E PERSPECTIVAS

Nesta dissertação estudamos os processos estocásticos e a quantização estocástica, a primeira apresentada no formalismo de Itô e a segunda no de Nelson. A mecânica estocástica de Nelson preocupa-se primariamente em ser uma alternativa ao formalismo da mecânica quântica, tentando tratar de questões sobre a natureza da medição quântica, o papel de variáveis escondidas e a interpretação de funções de onda. Em seus trabalhos originais, Nelson se questiona diversas vezes se os desenvolvimentos em mecânica estocástica seriam úteis de fato. Mas trabalhos recentes como (BEYER; PAUL, 2021.), (PEÑA; CETTO; VALDÉS-HERNÁNDEZ, 2020.), (KUIPERS, 2021.), (KURIHARA, 2017.) etc provam que ainda há muito a se desenvolver neste campo.

Inicialmente, traçamos dois objetivos para este trabalho: estudar soluções numéricas de EDEs discretizadas e obter um método para descrição estocástica da fase de Berry. O processo da implementação do algoritmo foi bastante desafiador, tendo em vista a complexidade dos algoritmos de solução de EDEs acopladas *backward* e *forward* (KOPPE; GRECKSCH; PAUL, 2017.). No entanto, considero que cumprimos parcialmente o primeiro objetivo apesar das dificuldades. A principal dificuldade da implementação deveu-se ao fato que a primeira etapa do algoritmo (obtenção de $u(x)$) estava demorando muito para finalizar, por volta de 5 a 6 dias rodando ininterruptamente, seguindo o limite de convergência proposto por (KOPPE; GRECKSCH; PAUL, 2017.). Tentamos ainda ajustar o limite de convergência, mas o gráfico não convergiu para a solução esperada. Como discutido no texto, implementar um algoritmo para solucionar estas equações é possível através do método de Euler-Maruyama, mas não é uma tarefa trivial.

Com relação ao problema fase de Berry, infelizmente, não houve tempo de finalizar o cálculo e não podemos afirmar, conclusivamente, que o método funciona. Contudo, o caminho que traçamos via equações de Madelung e equação dissipativa de Schrödinger parece muito promissor para nosso objetivo de estabelecer um método para obtenção da fase de Berry numa descrição estocástica. Pretendemos continuá-lo no futuro.

REFERÊNCIAS

- ALVES, M. A.; THOMAZ, M. T. Berry's phase through the path integral formulation. *Am. J. Phys.* 75, 552, 2007.
- BAUDOIN, F. 2012. Disponível em: <<https://fabricebaudoin.wordpress.com/2012/09/23/lecture-23-itos-representation-theorem/>>.
- BERRY, M. V. Classical adiabatic angles and quantal adiabatic phase. *Journal of Physics A: Mathematical and General*, 1983.
- BEYER, M.; PAUL, W. On the stochastic mechanics foundation of quantum mechanics. 2021.
- BISMUT, J.-M. Theorie probabiliste du controle des diffusions. *Mem Amer Math. Soc.*, 1973.
- BOYCE, W. E.; DIPRIMA, R. C. *Elementary differential equations and boundary value problems*. [S.l.]: John Wiley & Sons, Inc., 2001.
- CAROLLO, A.; SANTOS, M. F.; GURIDI, I. F.; VEDRAL, V. Geometric phase in open systems. *Geometric Phase in Open Systems. Physical Review Letters*, 2003.
- COHEN, S. N.; ELLIOTT, R. J. *Stochastic Calculus and Applications*. [S.l.]: Birkhäuser Basel, 2015.
- FITZSIMMONS, P. J. *Martingale Representation Theorem*. 2004. Disponível em: <<https://mathweb.ucsd.edu/~pfitz/downloads/courses/fall04/math286/martrep.pdf>>.
- FLEMING, W.; RISHEL, R. *Applications of Mathematics 1 - Deterministic and Stochastic Optimal Control*. [S.l.]: Springer-Verlag New York, 1975.
- FLEMING, W.; RISHEL, R. *Applications of Mathematics 1 - Deterministic and Stochastic Optimal Control*. [S.l.]: Springer-Verlag New York, 1975.
- FLORESCU, I.; VIENS, F. A binomial tree approach to stochastic volatility driven model of the stock price. *Annals of the University of Craiova*, 2005.
- FRANKEL, T. *The Geometry of Physics: An Introduction*. [S.l.: s.n.], 2012.
- FÉNYES, I. A deduction of schrödinger equation. *Acta Bolyaiana*, 1946.
- GUBINELLI, M. *Brownian martingale representation theorem*. 2020. Disponível em: <https://www.iam.uni-bonn.de/fileadmin/user_upload/gubinelli/stochastic-analysis-ss20/sa-ss20-script-18.pdf>.
- GUERRA, F. Structural aspects of stochastic mechanics and stochastic field theory. *Physical Reports*, 1981.
- HABALA, P. 2023. Disponível em: <<https://math.fel.cvut.cz/cz/mt/txtd/1/txe3da1a.htm>>.
- HALINA, P. 2011. Disponível em: <<http://patrickhalina.com/posts/what-is-stochastic-calculus/>>.

- HELM, L. A representation of the infinitesimal generator of a diffusion process. 1961.
- HORN, R. A.; JOHNSON, C. R. *Matrix Analysis 2nd ed.* [S.I.]: Cambridge University Press, 2013. Chapter 8 p.
- KARLIN, S.; TAYLOR, H. M. *A First Course In Stochastic Processes 2nd ed.* [S.I.]: New York Academic Press, 1974.
- KAROUI, E.; PENG, S.; QUENEZ, M. C. *Backward stochastic differential equations in finance.* [S.I.: s.n.], 1997.
- KOIDE, T.; AL. et. Unified description of classical and quantum behaviours in a variational principle. *J. Phys.: Conf. Ser.* **626** 012055., 2015.
- KOPPE, J.; GRECKSCH, W.; PAUL, W. Derivation and application of quantum hamilton equations of motion. *Annalen der Physik* 529(3), 1600251, 2017.
- KOPPE W. GRECKSCH, W. P. J. Derivation and application of quantum hamilton equations of motion. *Annalen der, Physik* Volume 529, Issue 3, 2016.
- KUIPERS, F. Analytic continuation of stochastic mechanics. *J. Math. Phys.* 63, 04, 042301 (2022), 2021.
- KURIHARA, Y. Stochastic metric space and quantum mechanics. <https://doi.org/10.48550/arXiv.1612.04228>, 2017.
- MA, J.; YONG, J. *Forward-Backward Stochastic Differential Equations and Their Applications.* [S.I.]: Springer, 2007.
- MADELUNG, V. E. Quantentheorie in hydrodynamischer form. *Z. Phys.* 40, 322, 1926.
- MISAKI, K.; MIYASHITA, S.; NAGAOSA, N. Diffusive real-time dynamics of a particle with berry curvatures. *Phys. Rev. B* 97, 075122 (2018), 2021.
- MOLLAI, M.; FATHI, S. M. S. An application of the madelung formalism for dissipating and decaying systems. 2021.
- NEEDHAM, T. *Visual Differential Geometry and Forms: A Mathematical Drama in Five Acts.* [S.I.]: Princeton University Press, 2021.
- NELSON, E. Derivation of the schrödinger equation from newtonian mechanics. *Phys. Rev.* **150**, 1079, 1966.
- NELSON, E. *Quantum Fluctuations.* [S.I.]: Princeton University Press, 1985.
- NELSON, E. Stochastic mechanics and random fields. *Ecole d'Eté de Probabilités de Saint-Flour XV–XVII.* Springer, Berlin, p. 427–459, 1988.
- NELSON, E. Review of stochastic mechanics. *J. Phys.: Conf. Ser.* 361 012011, 2012.
- PAPOULIS, A. *Probability, Random Variables, and Stochastic Processes; 9th ed.* [S.I.]: Tokyo: McGraw–Hill, 1965.
- PARDOUX, E.; PENG, S. Adapted solution of a backward stochastic differential equation, systems & control. *Lett.* 14, 1990.

- PAVON, M. Stochastic mechanics and the feynman integral. *J. Math. Physics*, 41 (2000), 6060-6078., 2000.
- PEDROSA, I. A.; LIMA, D. A. P. Coherent states and geometric phases of a generalized damped harmonic oscillator with time-dependent mass and frequency. *International Journal of Modern Physics B Vol. 28*, 2014.
- PEÑA, L. de la; CETTO, A. M.; VALDÉS-HERNÁNDEZ, A. Connecting two stochastic theories that lead to quantum mechanics. *Front. Phys.* 8:162., 2020.
- PROTTER, P. The work of kyoshi itô. *Notices of the American Mathematical Society*, 2007.
- REDDIGER, M.; POIRIER, B. Towards a mathematical theory of the madelung equations. 2022.
- REICHL, L. E. *A Modern Course in Statistical Physics; 2nd ed.* [S.I.]: John Wiley & Sons, 1998.
- RISKEN, H. *The Fokker-Planck equation: methods of solution and application.* [S.I.]: Springer, Berlin, 1996. pp. 179–195. p.
- SANBORN, B.; SUSLOV, S. K.; VINET, L. Dynamical invariants and berry's phase for generalized driven harmonic oscillators. 2011.
- SCHILLING, R. L.; PARTZSCH, L.; BÖTTCHER, B. *Brownian motion: an introduction to stochastic processes.* [S.I.]: De Gruyter, 2012.
- TOBOGA, M. Parameter space, lectures on probability theory and mathematical statistics. Kindle Direct Publishing, 2021. Disponível em: <<https://www.statlect.com/glossary/parameter-space>>.
- TONON, L. *O Teorema de Cramér-Lundberg via martingais.* Tese (Doutorado) — PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO - PUC-RIO, 2005.
- WANG, J.-Y. 2022. Disponível em: <[https://homepage.ntu.edu.tw/~jryanwang/courses/Financial%20Computation%20or%20Financial%20Engineering%20\(graduate%20level\)/FE_Ch01%20Wiener%20Process.pdf](https://homepage.ntu.edu.tw/~jryanwang/courses/Financial%20Computation%20or%20Financial%20Engineering%20(graduate%20level)/FE_Ch01%20Wiener%20Process.pdf)>.
- WU, Z.; YANG, H. Validity of the quantum adiabatic theorem. *Physical Review A*, 72(1), 2005.
- YASUE, K. Stochastic calculus of variation. *J. Funct. Anal.* 41, p. 327, 1981.
- ZWANZIG, R. *Nonequilibrium Statistical Mechanics.* [S.I.]: Oxford University Press, 2001.
- ZWANZIGER, J. W.; KOENIG, M.; PINES, A. *Berry's Phase.* [S.I.]: Annual review of physical chemistry, 41, 1990.

APÊNDICE A – REVISÃO DE CONCEITOS BÁSICOS DE PROBABILIDADE

Neste apêndice, tratamos os conceitos imprescindíveis para os estudos deste trabalho. Tudo que envolve processos estocásticos está atrelado a compreensão do que é probabilidade e seu papel. Sendo assim, façamos uma sucinta exposição conceitual. Com a seção A.1 discorreremos sobre espaço de probabilidade (TONON, 2005.), em A.2 espaço de parâmetros (TOBOGA, 2021.) e em A.3 variáveis aleatórias (PAPOULIS, 1965.).

A.1 ESPAÇO DE PROBABILIDADE

Ao trabalharmos com teoria de probabilidade, o espaço de probabilidade definido pela tripla $(\Omega, \mathcal{F}, \mathbb{P})$ é o construto matemático que formaliza os elementos básicos necessários para modelar experimentos aleatórios e eventos, onde:

- $\Omega \neq \emptyset$ é o espaço amostral de eventos.

Ao lançarmos duas moedas, por exemplo, temos uma gama de possíveis eventos que estão contidos num espaço amostral: $\Omega = \{HH, HT, TH, TT\}$ (H: cara, T: coroa).

- O conjunto \mathcal{F} é uma σ -álgebra de subconjuntos de Ω , seus eventos.

Para que um conjunto \mathcal{F} seja uma σ -álgebra, é preciso que três requisitos sejam cumpridos:

1. $\emptyset \in \mathcal{F}$;
2. Se $A \in \mathcal{F}$, então $A^c \in \mathcal{F}$;
3. Se os conjuntos A_i são sub-conjuntos de Ω e cada um deles pertence a \mathcal{F} , então sua união também pertence a \mathcal{F} .

Seguindo o exemplo das moedas acima, fixando que o primeiro lançamento caiu cara, temos o seguinte conjunto de possibilidades: $A_H = \{HH, HT\}$. O conjunto de eventos que não pertencem a A_H é o seu complemento $A_H^c = \{TH, TT\}$. Sendo assim, encontramos o conjunto $\mathcal{F}_1 = \{\emptyset, \Omega, A_H, A_H^c\}$ que configura uma σ -álgebra.

Podemos considerar o corolário: σ -álgebras são fechadas sob interseções contáveis. Dada uma coleção \mathcal{B} de sub-conjuntos de \mathcal{F} , a σ -álgebra gerada por \mathcal{B} é definida como: as menores σ -álgebras de subconjuntos de \mathcal{F} contendo \mathcal{B} . Estes tipos são chamadas "borel-mensuráveis".

- \mathbb{P} é a medida de probabilidade associada aos elementos de \mathcal{F} .

Sejam A_i elementos de \mathcal{F} , tal que a probabilidade associada ao evento é $P(A)$, temos as propriedades:

1. $P(\Omega) = 1, P(\emptyset) = 0$;
2. Se $A_i \cap A_j = \emptyset$, para $i \neq j$ (disjuntos dois a dois), então:

$$P(A_i \cup A_j) = P(A_i) + P(A_j).$$

O espaço de probabilidade provê uma estrutura formal para modelar e analisar fenômenos aleatórios, que consistem numa amostra de espaço representando os possíveis resultados, um espaço de eventos que contém conjuntos de resultados e a medida de probabilidade que associada a cada evento. É desnecessário aprofundar a discussão deste tópico além do exposto acima, pois questões como finitude dos espaços amostrais, natureza de seus valores e distribuições de probabilidade são discutidas ao longo do texto. Portanto, prosseguimos este apêndice definindo outro espaço fundamental para a base teórica deste trabalho, o espaço de parâmetros.

A.2 ESPAÇO DE PARÂMETROS

Suponha uma variável aleatória x , com distribuição normal $N(\mu, \sigma^2)$, em que μ e σ^2 são, respectivamente, a média e a variância. O espaço de parâmetros no qual x está contido é dado por:

$$\Theta = \{\mu, \sigma^2; -\infty < \mu < \infty, \sigma^2 > 0\}. \quad (\text{A.1})$$

Marco Toboga em (TOBOGA, 2021.) define da seguinte forma: Sejam ρ o conjunto de todas as distribuições de probabilidade que produziram x e Θ um conjunto de vetores reais. O conjunto Θ é chamado espaço de parâmetros para ρ se, e somente se, o conjunto ρ está associado a pelo menos um parâmetro e este esteja associados a distribuição de probabilidade pertencente a ρ .

Intuitivamente, podemos pensar no espaço de parâmetros como o conjunto de todos os possíveis valores assumidos pelos parâmetros do sistema. Em um problema de estatística, utilizamos uma amostra de uma certa variável para obter a sua distribuição de probabilidade. Quando estas distribuições são postas em correspondência com um grupo de números reais e vetores, esse conjunto é chamado de espaço de parâmetros. Portanto, o espaço de parâmetros define o conjunto de possíveis valores que um parâmetro pode assumir num modelo estatístico. Este caracteriza a gama de propriedades dos parâmetros e guia o processo de estimativa e interferência na análise estatística.

Ao trabalhar neste espaço, focamos em selecionar valores apropriados para os parâmetros, considerando as devidas restrições, caso existam. Ao determinarmos parâmetros plausíveis, este espaço guia o processo de estimar os valores associados ao processo estatístico em questão. Ficará claro como ocorre na prática ao buscarmos a solução de um processo estocástico via solução numérica nos capítulos adiante.

Neste trabalho, o espaço de probabilidade e o de parâmetros se complementam. O primeiro se preocupa em fundamentar a aleatoriedade e a distribuição de probabilidade dos dados, definindo o conjunto de possíveis resultados, os eventos em questão e a assimilar probabilidades a cada evento, modelando o comportamento probabilístico que produz um processo. Já o segundo, tal como descrito acima, fornece os parâmetros apropriados para tirarmos conclusões ou produzirmos interferência em um processo estatístico. Trabalhar simultaneamente com esses espaços permite uma análise mais compreensível dos modelos estatísticos.

A.3 VARIÁVEIS ALEATÓRIAS

Uma variável aleatória é uma função na qual seus possíveis valores são determinados pelas realizações de um evento aleatório, tais como o rolar de um dado ou o lançamento de uma moeda. Estas variáveis podem ter valores discretos ou contínuos, a depender se seus possíveis valores são enumeráveis ou não. Por exemplo: o lançamento de uma moeda fornece valores discretizados, enquanto o peso de uma maçã, selecionada ao acaso num cesto com várias outras, é uma variável contínua. Quando tratamos deste tipo de variável, estas carregam intrinsecamente uma relação com o conceito de probabilidade, pois se tratando de eventos aleatórios, entender o papel

das probabilidades envolvidas é fundamental para compreender processos que dependam deste tipo de evento. Dois conceitos fundamentais envolvendo probabilidade, utilizados ao longo deste trabalho, são: valor esperado e variância.

Podemos entender, intuitivamente, o que é um valor esperado, levantando indagações do tipo: "Se estou diante de várias ocorrências de um evento aleatório, de natureza numérica, qual valor irei obter em média?". O valor esperado, portanto, nos dirá qual o valor mais provável dentre uma coleção de eventos aleatórios, sejam estes equiprováveis ou não, de forma que não necessariamente o valor deva pertencer ao espaço amostral dos eventos em questão.

Definição: Seja uma variável aleatória x , seu valor esperado $E[x]$ é um valor dado pela soma $x(s)$ sobre os eventos s aos quais se aplica a variável, vezes a probabilidade $P(s)$ de ocorrência do evento

$$E[x] = \sum_s x(s)P(s). \quad (\text{A.2})$$

Sejam x e y processos aleatórios independentes, c uma constante qualquer, valem as seguintes propriedades:

1. Linearidade:

$$E[cx + y] = cE[x] + E[y]. \quad (\text{A.3})$$

2. Independência:

$$E[xy] = E[x]E[y]. \quad (\text{A.4})$$

3. Variância:

$$\sigma^2 = \text{Var}(x) = E[(x - E[x])^2]. \quad (\text{A.5})$$

É possível reescrever (A.5) de outra forma

$$\begin{aligned} \text{Var}(x) &= (E[x - E[x]])^2 \\ \text{Var}(x) &= E[x^2 - 2E[x]x + (E[x])^2] \\ \text{Var}(x) &= E[x^2] - 2E[x]E[x] + (E[x])^2 \\ \text{Var}(x) &= E[x^2] - 2E[x]^2 + (E[x])^2 \\ \text{Var}(x) &= E[x^2] - E[x]^2. \end{aligned} \quad (\text{A.6})$$

Quando lidamos com coleções de variáveis aleatórias, calcular o valor esperado é a chave para nos aproximarmos de conclusões precisas sobre processos que dependam destas variáveis.

Muitos processos contínuos estão definidos em domínios compactos (fechados e limitados). No contexto de probabilidade, temos que se X é uma variável aleatória discreta, é definida como o conjunto: $R_X = \{x \in \mathbb{R} : P(X = x) > 0\}$. Para X contínua: $R_X = \{x \in \mathbb{R} : \rho(x) > 0\}$, tal que $\rho(x)$ é a função densidade de probabilidade de x .

As propriedades acima serão suficientes para as pretensões deste trabalho, cujos processos de interesse são, especificamente, os estocásticos.

APÊNDICE B – CONCEITOS DE GEOMETRIA DIFERENCIAL

Neste apêndice, tratamos sobre alguns conceitos de geometria diferencial, imprescindíveis para posteriormente definir a fase de Berry e para prover mais generalidade na teoria de processos estocásticos. Começamos discorrendo sobre variedade diferenciável (FRANKEL, 2012.), pressupondo familiaridade com o conceito de mapa e espaço topológico. Em seguida, trazemos as definições de transporte paralelo e holonomia (NEEDHAM, 2021.). Estes conceitos elucidam o tipo de derivada que é utilizada ao longo do texto, como transportar um vetor através da superfície, sem alterá-lo.

B.1 VARIEDADE DIFERENCIÁVEL

O espaço de configurações de um sistema mecânico é uma variedade diferenciável M . Definimos uma variedade diferenciável n -dimensional da seguinte forma.

Definição: Seja M um conjunto qualquer, preenchido por sub-conjuntos $M = U \cup V \cup \dots$, onde cada sub-conjunto U correspondente 1:1 com $\phi_U : U \rightarrow \mathbb{R}^n$ com um sub-conjunto aberto $\phi_U(U)$ de \mathbb{R}^n . O par (U, ϕ_U) é chamado de atlas, representa a vizinhança de coordenadas.

Nós requeremos que cada $\phi_U(U \cap V)$ seja um sub-conjunto aberto de \mathbb{R}^n . Os mapas sobrepostos (ou mapas de transição) são:

$$f_{UV} = \phi_V \circ \phi_U^{-1} : \phi_U(U \cap V) \rightarrow \mathbb{R}^n.$$

ou mais especificamente

$$\begin{cases} \phi_U \circ \phi_V^{-1} : \phi_V(U \cap V) \rightarrow \phi_U(U \cap V), \\ \phi_V \circ \phi_U^{-1} : \phi_U(U \cap V) \rightarrow \phi_V(U \cap V). \end{cases}$$

Cada par (U, ϕ_U) define um trecho coordenado em M ; para $x \in U \subset M$ nós devemos atribuir as n coordenadas do ponto $\phi_U(x)$ em \mathbb{R}^n , $x^1, \dots, x^k \in \phi_U$, que são chamadas coordenadas locais e portanto chamamos ϕ_U de mapa coordenado, similarmente para $\phi_V, x^{1'}, \dots, x^{k'}$. Dizemos que um mapa $F : \mathbb{R}^{x'} \rightarrow \mathbb{R}^x$ é de classe:

- C^k se todas as k derivadas parciais são contínuas;
- C^∞ se é de classe $C^k \forall k$.

Denotamos a matriz de derivadas parciais por:

$$\frac{\partial x^k}{\partial x^k} \quad (\text{B.1})$$

Dizemos que uma variedade M^n é de classe C^k se os mapas sobrepostos f_{UV} são de classe C^k , isto é, se f_{UV} é k vezes continuamente diferenciável. Da mesma forma, há a noção de uma variedade C^∞ , na qual a ideia chave por trás da variedade é que, localmente, esta se parece com o espaço Euclidiano. Em torno de qualquer ponto da variedade, há um atlas que provê um sistema de coordenadas "suave": se f_{UV} é de classe C^∞ , dizemos que é uma curva suave. Curvas suaves são infinitamente diferenciáveis, com o entendimento de que expressando f em termos do sistema de coordenadas locais $f = f_U(x^1, \dots, x^k)$, implicamos que F é diferenciável.

A partir daqui, temos arcabouço teórico suficiente para introduzir a cinemática dos sistemas mecânicos que trataremos, considerando que estes variem sua configuração com o tempo. Seja $\xi(t)$ uma curva suave sobre M^n , ou seja, um mapa de algum intervalo de \mathbb{R} para M^n . No sistema de coordenadas (U, x) , descrevemos um ponto $\xi(0) = p$ por k funções diferenciáveis $x^k = x^k(t)$. O "vetor velocidade" $\dot{\xi}(0)$ (o ponto significa derivada temporal) é classicamente descrito pela k -ésima tupla de números reais:

$$\left. \frac{dx^1}{dt} \right|_{t=0}, \dots, \left. \frac{dx^k}{dt} \right|_{t=0}.$$

Similarmente para (V, x') , se p também está presente. Relacionando ambos por meio da regra da cadeia aplicada a $f_U(x^1, \dots, x^k)$:

$$\left. \frac{dx^j}{dt} \right|_{t=0} = \sum_{i=1}^n \left(\frac{\partial x^j}{\partial x^i} \right) [p] \left(\left. \frac{\partial x^i}{dt} \right|_{t=0} \right), \quad (\text{B.2})$$

tal que i representa as componentes de U e j as de V . A equação acima nos permite definir, portanto, o vetor tangente.

Definição: O vetor tangente a $p \in M^n$, chamado \mathbf{X} , é uma classe de equivalência de caminhos suaves, partindo de um ponto p , tangentes entre si. Atribui a cada conjunto coordenado (U, x) contendo p uma n -tupla de números reais $(X^i) = (X^1, \dots, X^n)$, tal que $p \in U \cap V$, então

$$X^j = \sum_{i=1}^n \left[\frac{\partial x^j}{\partial x^i} [p] \right] X^i. \quad (\text{B.3})$$

O campo vetorial suave X , na variedade M , traz à tona um fluxo local único. Para cada ponto em M , há um caminho suave $t \rightarrow \xi(t)$, definido pelo menos por valores pequenos de t , tal que o vetor velocidade de $\xi(t)$ é $X(\xi(t))$, isto é:

$$\frac{d}{dt} \xi^i(t) = X^i(\xi(t)). \quad (\text{B.4})$$

O conjunto de todos os vetores tangentes formam um espaço tangente próprio, chamado de fibrado tangente TM . Um ponto em TM é o par $(\xi, X(\xi))$. Na cinemática, estamos interessados (na maioria das vezes) no fluxo sobre o espaço das velocidades, como TM é por si só uma variedade, este caso é sempre incluso. Este fluxo é sempre suave, fato este que é possível provar, mas isto foge do escopo deste trabalho.

Seja $(X) = (X^1, \dots, X^n)^T$ (T significa o transposto da matriz) a coluna das componentes vetoriais de X , podemos escrever a equação matricial

$$X' = \sigma^{ji} X, \quad (\text{B.5})$$

em que a função de transição σ^{ji} é a matriz jacobiana avaliada no ponto em questão.

Seja ξ um campo vetorial tangente a M , definido na vizinhança de um ponto p qualquer e seja X um vetor tangente neste mesmo ponto p . Se X é um vetor em p , definimos a derivada de f com relação ao vetor X por:

$$X_p f := \nabla_X f := \sum_{j=1}^n \left[\frac{\partial f}{\partial x^j} (p) \right] X^j, \quad (\text{B.6})$$

a qual chamamos de **derivada covariante**. Podemos facilmente verificar que a equação acima independe do sistema de coordenadas utilizado. Sejam (U, x) e (V, x') os dois sistemas coordenados, da regra da cadeia, temos:

$$\begin{aligned} \nabla_X^V f &= \sum_j \left(\frac{\partial f}{\partial x^j} \right) X^j \\ \nabla_X^V f &= \sum_j \left(\frac{\partial f}{\partial x^j} \right) \sum_i \left(\frac{\partial x^j}{\partial x'^i} \right) X^i \\ \nabla_X^V f &= \sum_i \left(\frac{\partial f}{\partial x'^i} \right) X^i \\ \nabla_X^V f &= \nabla_X^U f. \end{aligned}$$

Ou seja: sempre que definimos algo utilizando coordenadas locais, se desejarmos que a definição tenha significado intrínseco, deve-se verificar que significa a mesma coisa para qualquer sistema coordenado. Note que há uma correspondência 1:1 entre o vetor tangente X para M^n em p com o operador diferencial de primeira ordem, definindo próximo a p , dado por

$$X_p = \sum_j X^j \left. \frac{\partial}{\partial x^j} \right|_p. \quad (\text{B.7})$$

Numa forma mais geral, a derivada covariante é escrita considerando a métrica do sistema, resumida no *símbolo de Cristoffel* Γ :

$$(\nabla_X f)^k = \left(\frac{\partial f^k}{\partial x^i} + \Gamma_{ij}^k f^j \right) X^i. \quad (\text{B.8})$$

Definindo de maneira mais rigorosa: Seja M uma variedade, o operador ∇ é a derivada covariante, que assimila esta a cada par de, respectivamente, vetor X tangente a p e o campo vetorial \mathbf{v} definido na vizinhança de p .

B.2 GEODÉSICAS E TRANSPORTE PARALELO

Um vetor é paralelamente transportado através de uma superfície quando sua derivada covariante sobre a curva é zero ($\nabla_X = 0$). O transporte independe do caminho se a superfície é plana, mas caso seja curva, cada caminho escolhido resultará em vetores finais diferentes. Esta propriedade, chamada de holonomia, é capaz de medir a curvatura de superfícies riemannianas. Para transportar paralelamente um vetor tangente a uma superfície, devemos projetá-lo na superfície à medida que o transporte ocorre ou rotacioná-lo um pouco em direção a superfície a cada avanço infinitesimal. O comprimento do vetor se mantém constante. Ademais, se dois vetores são transportados paralelamente sob uma superfície, o ângulo entre eles permanece constante. Todas essas características partem, porém, de uma construção extrínseca ao conceito de transporte paralelo. Para formalizar esta construção, é preciso discorrer sobre o conceito de Geodésica.

Seja $p(t)$ a posição do vetor, $\mathbf{v}(t)$ a velocidade unitária e $\mathbf{w}(t)$ um vetor tangente a curva S , a taxa com que $\mathbf{w}(t)$ varia ao longo de um tempo infinitesimal ϵ , num intervalo

entre os pontos $q, p \in S$ é, por definição:

$$\epsilon \Delta_v \mathbf{w} = \epsilon w'(t) \asymp \mathbf{w}(q) - \mathbf{w}(p), \quad (\text{B.9})$$

em que o símbolo " \asymp " indica: $A \asymp B \iff \lim_{\epsilon \rightarrow 0} \frac{A(\epsilon)}{B(\epsilon)} = 1$. A diferença entre os vetores tangentes a S , $\mathbf{w}(q)$ e $\mathbf{w}(p)$, não resulta necessariamente em um vetor tangente a S , sendo portanto uma característica extrínseca a S . A partir das características do transporte paralelo, podemos introduzir um novo vetor posição: $\mathbf{w}_{\parallel}(p \mapsto q)$, que representa o vetor paralelamente transportado de p a q . Esse vetor significa que, a cada distância infinitesimal percorrida, estamos projetando \mathbf{w} sobre o plano S , tal que permaneça tangente a S e mantenha-se constante o ângulo entre a geodésica e o vetor transportado. Além disso, a derivada associada ao transporte paralelo é ∇_v (derivada covariante), portanto:

$$\epsilon \nabla_v \mathbf{w} \asymp \mathbf{w}(q) - \mathbf{w}_{\parallel}(p \mapsto q). \quad (\text{B.10})$$

Neste contexto, podemos chamar ∇_v de "conexão de Levi-Civita" ou "derivada intrínseca". Como $\nabla_v \mathbf{w}$ garante a tangência de \mathbf{w} a S , temos portanto uma característica intrínseca a S . O trabalho de ∇_v é "quantificar" a taxa com que \mathbf{w} aumenta seu comprimento e gira no sentido anti-horário e corrigir isso. Para obter $\nabla_v \mathbf{w}$, subtraímos da taxa de variação total $\Delta_v \mathbf{w}$ em \mathbb{R}^3 a parte que não é tangente:

$$\nabla_v \mathbf{w} = \Delta_v \mathbf{w} - (\mathbf{n} \cdot \Delta_v \mathbf{w}) \cdot \mathbf{n}. \quad (\text{B.11})$$

Definição: Operador forma é a derivada negativa do vetor unitário normal à superfície

$$S(\mathbf{v}) = -\Delta_v \mathbf{w}. \quad (\text{B.12})$$

Substituindo em (B.11):

$$\nabla_v \mathbf{w} = \Delta_v \mathbf{w} + (\mathbf{n} \cdot S(\mathbf{v})) \cdot \mathbf{n}. \quad (\text{B.13})$$

Propriedades da derivada intrínseca ∇_v :

$$\nabla_v [a\mathbf{x} + b\mathbf{y}] = a\nabla_v \mathbf{x} + b\nabla_v \mathbf{y}; \quad (\text{B.14})$$

$$\nabla_{[fx+gy]} \mathbf{z} = f\nabla_x \mathbf{z} + g\nabla_y \mathbf{z}; \quad (\text{B.15})$$

$$\nabla_v [f\mathbf{x}] = f\nabla_v \mathbf{x} + [\nabla_v f]\mathbf{x} = f\nabla_v \mathbf{x} + f'\mathbf{x}; \quad (\text{B.16})$$

$$\nabla_v [\mathbf{x} \cdot \mathbf{y}] = [\nabla_v \mathbf{x}] \cdot \mathbf{y} + \mathbf{x} \cdot [\nabla_v \mathbf{y}]. \quad (\text{B.17})$$

Definição: Taxa de rotação de \mathbf{v} sobre a superfície:

$$K_G = \nabla_{\mathbf{v}}\mathbf{v}. \quad (\text{B.18})$$

Se \mathbf{v} é transportado paralelamente sobre si mesmo, significa que deve parecer constante para um observador que esteja no caminho de \mathbf{v} . Isto é, se estivermos sobre uma geodésica, vale a equação da geodésica:

$$K_G = 0 \iff \nabla_{\mathbf{v}}\mathbf{v} = 0. \quad (\text{B.19})$$

A geodésica G é, portanto, uma curva para qual a derivada intrínseca do campo tangente \mathbf{v} não possui componentes tangenciais a superfície. Uma geodésica traçada a partir de um vetor velocidade unitário, lançado com velocidade inicial $\mathbf{v}(0)$, então $\mathbf{v}(t)$ é obtido pelo transporte paralelo de $\mathbf{v}(0)$ sobre G . Essa geodésica é única, pois depende das condições iniciais. Dessa forma, para transportar paralelamente um vetor \mathbf{w} , tangente à uma curva S , sob uma curva qualquer K , aproxime K por uma sequência K^* de segmentos geodésicos G_i , cada um com comprimento infinitesimal ϵ , então carregue \mathbf{w} sobre as G_i 's, mantendo um ângulo constante entre a direção das G_i 's e o \mathbf{w} paralelamente transportado: \mathbf{w}_{\parallel} . Finalmente, tome o limite $\epsilon \rightarrow 0$, tal que K^* torne-se K .

Note que, se observarmos a variação infinitesimal desse vetor no tempo, esta também será igual a zero.

$$\frac{\partial}{\partial t}(\nabla_{\mathbf{v}}\mathbf{v} = 0). \quad (\text{B.20})$$

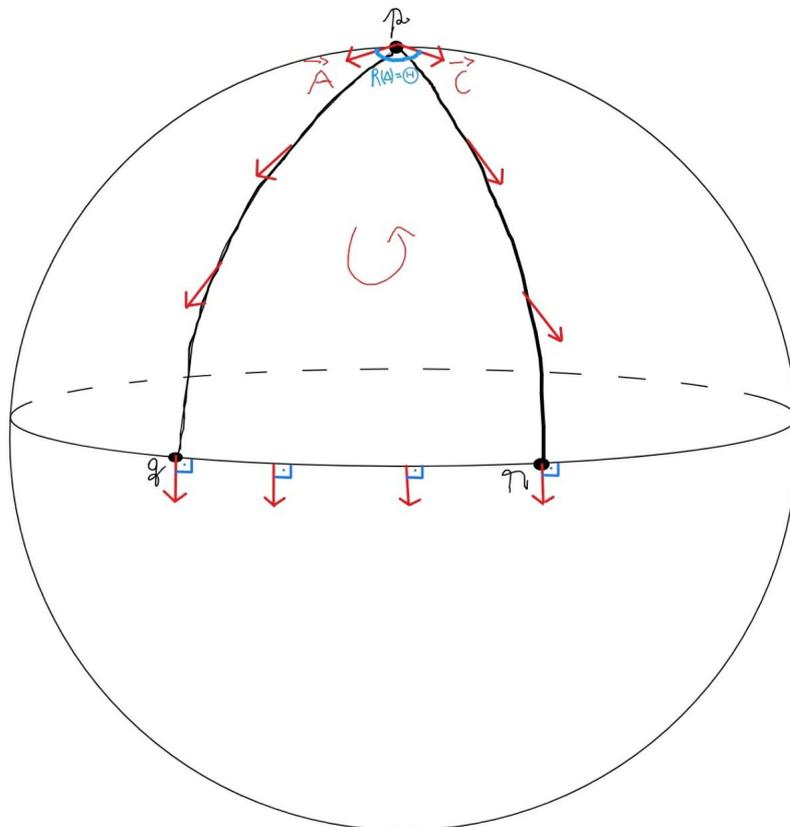
Portanto, \mathbf{v} é transportado sem mudança no seu comprimento ou direção. Isso nos leva diretamente ao conceito de holonomia que mencionamos anteriormente.

B.3 HOLONOMIA

Holonomia é o nome dado ao ato de transportar paralelamente um vetor através de uma curva geodésica. Tomando diferentes rotas que terminam em um mesmo ponto, obtemos vetores diferentes ao final. Suponha a seguinte situação: realizando o transporte paralelo de A através do triângulo geodésico $g\hat{p}r$ (9), obtemos o vetor C ao final. Dizemos que A foi rotacionado pela Holonomia de $g\hat{p}r$: $R(\Delta) = \Theta$. A partir dessa situação, podemos definir esse fenômeno, de maneira geral, de duas formas

- A holonomia $R(L)$ de um loop fechado L , sob uma superfície S , é a rotação em cadeia de um vetor tangente a S , paralelamente transportado ao redor de L ;
- Holonomia é a rotação do plano tangente, ao ser paralelamente transportado através do loop (independente do ponto inicial).

Figura 9 – Transporte paralelo do vetor A em torno da geodésica $g\hat{p}r$



Fonte: O autor (2023).

No caso em questão, a esfera é uma superfície de curvatura gaussiana constante $k = \frac{1}{R^2}$. A partir do teorema de Gauss-Bonnet, temos que a curvatura total do triângulo geodésico é:

$$K(\Delta) = \iint_{\Delta} k dA = \frac{1}{R^2} \iint_{\Delta} dA = \frac{1}{R^2} [R^2 \Theta] = \Theta. \quad (\text{B.21})$$

Portanto: $R(\Delta) = K(\Delta) = \Theta$. Observar o funcionamento deste teorema para uma esfera é bastante ilustrativo e suficiente para o escopo deste trabalho, contudo, no capítulo 17 de (FRANKEL, 2012.) o autor demonstra que essa é uma característica intrínseca de qualquer loop L em qualquer superfície S que possua dimensão maior que 2, independente da métrica. Esse fato é a chamada **condição de quantização topológica**. A derivada covariante, na forma como foi definida em (B.6), pode ser interpretada como a projeção da derivada em \mathbb{R}^n sobre um fibrado tangente. Ao projetarmos a derivada em \mathbb{C}^n , o teorema de Gauss-Bonnet fornece as condições de quantização topológicas para situações puramente geométricas, tal como o caso de interesse deste trabalho: a fase de Berry.

APÊNDICE C – EQUAÇÕES DIFERENCIAIS PARCIAIS PARABÓLICAS (EDP PARABÓLICA)

Uma EDP envolve a relação entre as derivadas parciais de uma função. O que a separa de uma EDO é que a função, e suas derivadas, dependem de várias variáveis, ao invés de apenas uma. Este tipo de equação é fundamental para compreender diversos fenômenos naturais, tais como: condução de calor (via equação de calor), movimento de um fluido (equação de Navier-Stokes), equação de Schrödinger, relatividade geral (equações de campo de Einstein) etc. Podemos classificar uma EDP das seguintes formas:

1. Quanto a ordem, referindo-se a ordem da maior derivada parcial na equação;
2. Quanto ao número de variáveis independentes;
3. Quanto a sua linearidade. Uma EDP linear é aquela cuja a variável que está sendo diferenciada surge apenas de forma linear;
4. Quanto a homogeneidade, na qual todos os termos da EDP envolvem a variável dependente ou suas derivadas;
5. Quanto aos tipos de coeficientes, se são constantes ou variáveis;
6. Se a EDP é parabólica, hiperbólica ou elíptica.

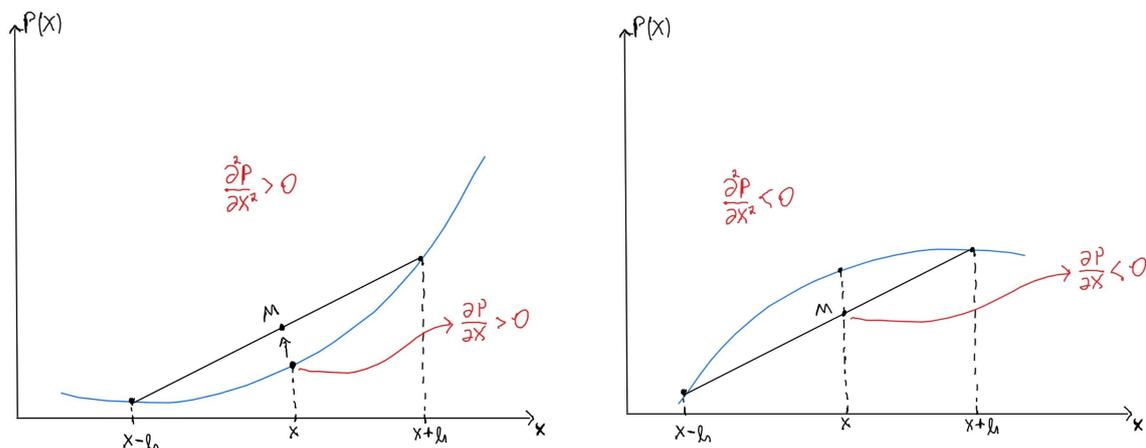
Tratando da equação de Fokker-Planck acima, é classificada como parabólica pois, dado que não há termo cruzado e nem uma derivada parcial de segunda ordem do tempo, a definição ((BOYCE; DIPRIMA, 2001.), capítulo 10) diz que é possível reorganizar os termos até que a equação assuma a forma parabólica usual. Ao buscarmos soluções para uma EDP, podemos lançar uma gama de ansatz diferentes e vários podem estar certos ao mesmo tempo. Este é um problema específico às EDPs, em EDOs podemos tentar um ansatz geral em termos de constantes arbitrárias desconhecidas, para EDPs não temos esta possibilidade. Portanto, para resolvê-las de maneira apropriada, precisamos das condições auxiliares, isto é: condições de contorno e condição inicial, estas ditam de fato como resolver uma EDP.

Por exemplo, para a distribuição de probabilidade $\rho(\xi_t, t)$, podemos especificar o estado espacial da função em um dado instante de tempo, bem como sua derivada

temporal neste mesmo instante e estas seriam as condições iniciais. O número de condições iniciais necessárias para a unicidade da solução é, tipicamente, igual a maior ordem da derivada temporal na EDP. As condições de contorno especificam propriedades da solução numa posição em particular, tipicamente nas bordas do domínio onde a EDP esta contida. Em geral, o número de condições de contorno necessárias é dado pela soma das maiores ordens das derivadas parciais espaciais. Para domínios não finitos ($-\infty$ a ∞), condições de contorno não podem ser definidas, mas existem outros tipos de condições que levam a soluções válidas e únicas, por exemplo: para solucionar a equação de Schrödinger há uma condição de normalização a ser cumprida, para que a função de onda seja fisicamente válida.

A EDP parabólica mais simples que podemos encontrar é a equação de difusão, na sua forma unidimensional em (2.12). Ao invés de entendermos como obter esta equação, é mais enriquecedor compreender sua interpretação física. $\frac{\partial P}{\partial x}$ nos informa o quão rápido a função P varia com relação a x , $\frac{\partial^2 P}{\partial x^2}$ informa o quão rápido a derivada varia. Sumarizamos isso no conceito de "concavidade". A seguir, exemplos de uma curva com concavidade para cima e outra para baixo, respectivamente:

Figura 10 – Gráficos da FDP pela posição



Fonte: O autor (2023).

Nestas curvas, temos três pontos adjacentes: $x - h$, x e $x + h$. Desenhando uma linha reta entre o ponto inicial e final, podemos estudar o valor médio M desta amostra da curva, dado por

$$M = \frac{1}{2}[P(x - h) + P(x + h)].$$

No caso da curva côncava para cima (para baixo), o valor M é maior (menor) que o valor de P no ponto x . Isso significa que a concentração de P em x é baixa (alta) com relação à média, então a difusão tende a ir em direção (contra a direção) de x . Portanto, a taxa com que P varia num dado ponto é proporcional a sua concavidade no mesmo ponto. Essa noção nos leva ao *Princípio de Máximo*.

Definição: Se P é descrita por uma EDP como a equação de difusão, sobre um domínio limitado $[s, T]$, para $t \geq 0$ o $\max[P(x, t)]$ ocorre em $t = 0$ ou nos limites $x = 0$ ou $x = L$. Porém, se P for constante, o "máximo" ocorre em todos os pontos. O princípio de máximo aplicado a $-P$ é o princípio de mínimo.

Em outras palavras, o que este princípio implica com relação ao comportamento da função é que a partir do momento que a função atinge seu máximo (ou mínimo), é impossível que a mesma ultrapasse esse valor à medida que evolui no tempo, tendendo sempre a cair (ou subir) a partir do ponto em questão. Percebe-se que este princípio está intrinsecamente relacionado ao ruído branco que temos tratado até então, dada sua forma e comportamento.