

Universidade Federal do Rio Grande do Sul
Instituto de Matemática
Programa de Pós-Graduação em Matemática

CADEIAS DE MARKOV CLÁSSICAS E QUÂNTICAS

por

CARLOS FELIPE LARDIZÁBAL RODRIGUES

Porto Alegre, Março de 2006

Dissertação submetida por Carlos Felipe Lardizábal Rodrigues¹ como requisito parcial para a obtenção do grau de Mestre em Matemática pelo Programa de Pós-Graduação em Matemática do Instituto de Matemática da Universidade Federal do Rio Grande do Sul.

Professor Orientador: Dr. Artur Oscar Lopes

Banca Examinadora:

Dr. Artur Oscar Lopes

Dr. Alexandre Tavares Baraviera

Dr. Jairo da Silva Bochi

Dr. Rafael Rigão Souza (UNISINOS)

Data de Defesa: 10 de março de 2006.

¹Bolsista do Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

Resumo: seguindo o trabalho de S. Gudder, fazemos uma construção de cadeias de Markov quânticas a partir de matrizes complexas, unitárias, estocásticas e analisamos o conceito de interferência nesse contexto, dando atenção para uma cadeia que chamamos de moeda quântica. Estamos interessados na entropia de cadeias de Markov reais, no princípio variacional para energia livre associado e em uma possível construção análoga no caso complexo. Este trabalho visa também dar uma introdução matematicamente rigorosa de certos aspectos de mecânica quântica.

Abstract: following the work of S. Gudder we make a construction of quantum Markov chains via complex, unitary, stochastic matrices and we analyse the concept of interference in that context, with emphasis on a certain chain we call the quantum coin. We are interested in the entropy of real Markov chains, in the associated variational principle for free energy and in analogous constructions in the complex case. This work also aims to be a mathematically rigorous introduction to certain aspects of quantum mechanics.

Agradecimentos

Esta dissertação de mestrado foi criada a partir de algumas notas simples que escrevi durante minha graduação e que serviram como um primeiro trabalho de iniciação científica. É claro que com o passar dos semestres, tais notas melhoraram em extensão e principalmente em profundidade. Aproveito para agradecer a meu orientador Artur Oscar Lopes pela orientação e pelas inúmeras sugestões para melhorar este trabalho, e mais geralmente pelo permanente incentivo, desde os tempos em que eu era um aluno da computação até os dias de hoje, ao concluir o bacharelado e mestrado em matemática.

Agradeço ainda aos professores Alexandre Baraviera, Rafael Rigão Souza e Jairo Bochi pelas diversas sugestões e apoio durante a elaboração deste trabalho. Agradeço aos meus colegas e à minha família pela generosa paciência na vida cotidiana.

Porto Alegre, março de 2006.

Sumário

1	Introdução	3
1.1	Sobre medições em um sistema quântico	4
2	Cadeias de Markov quânticas	7
2.1	Introdução	7
2.2	N-cadeias quânticas	8
2.3	Matrizes de Dirichlet	17
2.4	Distribuições de Probabilidade	23
2.5	Alguns cálculos	33
2.6	Apêndice	37
3	Integrais de Feynman	43
3.1	O formalismo das integrais de Feynman	43
3.2	Distribuição de Probabilidade	47
3.3	Sobre a amplitude de Feynman	50
4	Operadores densidade e mecânica quântica	54
4.1	Notação de Dirac	55
4.2	Produto tensorial	57
4.3	Operador densidade	62
4.4	Postulados da mecânica quântica	64
4.5	Estados emaranhados	66
5	Entropia	71
5.1	Entropia de Shannon	71
5.2	Entropia de von Neumann	75
5.3	Subaditividade forte	83
5.4	Cadeias quânticas: entropia de Shannon	85
5.5	Cadeias quânticas: entropia de von Neumann	88

6	Formalismo Termodinâmico	91
6.1	Introdução	91
6.2	Caracterização variacional	92
6.2.1	Introdução	92
6.2.2	Energia livre	93
6.2.3	Caracterização variacional	96
6.3	Observações sobre processos quânticos	104
6.4	Apêndice	108
6.4.1	Prova do lema de Wielandt	111
7	C^*-álgebras e cadeias de Markov quânticas	118
7.1	C^* -álgebras	118
7.2	Cadeias de Markov quânticas	119
7.3	CAR C^* -álgebras e estados KMS	124
7.4	Moeda quântica revisitada	128
8	Apêndice: Processos de Markov quânticos	130
8.1	Notações e definições	130
8.2	Exemplos	135
8.3	Processos de Markov quânticos quase-discretos	142

Capítulo 1

Introdução

Um princípio fundamental da mecânica quântica diz que a evolução temporal de um sistema físico isolado pode ser colocada em correspondência com um espaço vetorial de dimensão infinita de tal modo que os estados do sistema são representados por vetores e uma medição física qualquer corresponde a um operador linear. Com o intuito de simplificar o problema, vamos analisar um modelo onde o tempo é discreto, o espaço vetorial é de dimensão finita e a cadeia de Markov clássica é modificada pela introdução de uma matriz complexa.

As cadeias de Markov quânticas são úteis para obter um comportamento aproximado do problema em dimensão infinita da mecânica quântica (que é muito mais complexo). Esta aproximação (assim como outras abordagens discretas) é motivada não apenas pela necessidade de se estudar uma simplificação do problema maior, mas também pela possibilidade de que o próprio tempo seja discreto (questão complicada). É razoável também especular que o problema real com espaço de estados contínuo possa ser aproximado por um com espaço de estados discreto.

Inicialmente analisamos o trabalho feito por J. Marbeau e S. Gudder em [15], onde consideramos uma matriz de amplitude de transição que tem a forma de uma matriz de Dirichlet. Tais matrizes geram um análogo discreto da amplitude de Feynman contínua. Depois, calculamos a distribuição de probabilidade para estas cadeias.

A seguir, estaremos interessados em ferramentas que permitam descrever outros processos, ditos discretos ou quase-discretos, seguindo a descrição dada em [16]. Vamos considerar uma medida complexa sobre um espaço, onde também definiremos certos tipos de medições.

Como estamos interessados em estudar processos de natureza quântica, será útil ter em mente os postulados da mecânica quântica, que são discutidos com detalhe em uma formulação via operadores densidade. A notação

de Dirac, de uso corrente em mecânica quântica, também é introduzida. Estudaremos também as entropias de Shannon e von Neumann e analisamos a entropia de certos processos de Markov quânticos.

No capítulo 6, analisamos um problema variacional para cadeias de Markov finitas. Mais especificamente, se H denota a entropia, U é um potencial, e λ é um certo autovalor associado a U , temos o conhecido princípio variacional para energia livre:

$$H(\mu) - \int U d\mu \leq \log \lambda$$

onde vale a igualdade se e somente se μ for uma determinada medida especial $\nu = \nu^{(U)}$, que chamamos **medida de Markov**.

Dado este teorema para cadeias de Markov (reais), podemos nos perguntar se existe algum análogo para cadeias quânticas. Neste trabalho, não propomos uma solução para esse problema, mas fornecemos algumas ferramentas que podem ajudar na sua resolução.

Na literatura, uma cadeia de Markov quântica normalmente está associada a uma importante construção via álgebra de operadores, devida a L. Accardi ([1], veja também [30]). Nossa abordagem inicial difere bastante de tal contexto, mas no capítulo 7 faremos uma breve introdução às cadeias quânticas, fazendo uso de C^* -álgebras.

1.1 Sobre medições em um sistema quântico

Em processos estocásticos, estamos interessados nos resultados obtidos por uma família, ou seqüência, de funções mensuráveis. Podemos considerar, por exemplo, que o processo representa o resultado de se jogar uma moeda um número arbitrário de vezes, ou então a quantidade de chamadas telefônicas em um certo intervalo de tempo, ou mais geralmente, um fenômeno tal que sua evolução depende da probabilidade de ocorrência de um determinado evento. Tais problemas fazem parte de uma quantidade considerável de pesquisa em matemática.

Como a mecânica quântica possui um caráter probabilístico, é natural perguntar de que forma podemos descrever o movimento de uma partícula microscópica nos moldes de um processo estocástico. Veremos que a mecânica quântica impõe restrições severas quanto a evolução dos objetos que descreve. Dito de maneira geral, estaremos interessados em operadores que fornecem **amplitudes** de transição. No caso discreto, isso significa considerar matrizes unitárias agindo sobre os vetores estados que descrevem a evolução das partículas; no caso contínuo, isso significa que a evolução de tal vetor estado

satisfaz a equação de Schrödinger:

$$i\hbar \frac{d|\psi\rangle}{dt} = H|\psi\rangle$$

No entanto, descrever fenômenos quânticos é um problema substancialmente mais difícil do que modelar problemas clássicos. A idéia de realizar uma seqüência de medições sobre um elétron, como em um experimento comum com objetos macroscópicos, é no mínimo problemática. Surge então a necessidade de se construir um formalismo de **interferência** entre medições. Uma parte importante de nosso trabalho consistirá da construção matemática dessa idéia. Aqui X_t vai descrever a evolução temporal de amplitude quântica de um sistema. O processo X_t , $t \in \mathbb{N}$, vai tomar valores num conjunto finito $S = \{0, 1, \dots, n-1\}$.

É natural tentar encontrar um "setting quântico sobre um conjunto de estados $S = \{0, 1, \dots, n-1\}$ " que seja de alguma forma semelhante ao "setting probabilístico sobre um conjunto de estados $S = \{0, 1, \dots, n-1\}$ ". Neste último caso, se analisa $P(X_t = s) \in [0, 1]$, onde X_t é um processo estocástico e $s \in S$. Sendo assim, se espera que no primeiro "setting" intervenha de alguma forma uma medida complexa A e se deseja entender a evolução $A(X_t = s) \in \mathbb{C}$, onde X_t é um processo estocástico e $s \in S$. Vamos assumir que de alguma forma natural a hipótese markoviana apareça no problema. Desejamos que uma matriz \mathcal{A} , do tipo n por n , descreva a "amplitude de transição". Como veremos, neste formalismo nos deparamos com o conceito de interferência e com a questão do futuro (eventualmente poder) interferir no passado, o que nos parece indesejado (embora talvez fisicamente possível).

Dois resultados importantes que veremos nesse sentido são:

Teorema 1 *Em uma N -cadeia quântica $\{X_t\}$, $X_{t'}$ não interfere em X_t , para todo $0 \leq t \leq t' \leq N$.*

Teorema 2 *Se $(X_t)_{t \in T}$ é unitário então X_t não interfere em X_s para $s \leq t$.*

Em outras palavras, estes teoremas nos dizem que o futuro não interfere no passado. Esta é uma propriedade desejável da teoria que iremos desenvolver. O primeiro teorema se refere a um sistema discreto, e o outro a sistema contínuos ou quase-discretos. Provaremos estes dois teoremas no próximo capítulo. A mecânica quântica, como teoria física (ou matemática) da natureza, é incompleta e desafia muitas das características que consideramos senso comum no mundo clássico. Iremos explorar apenas algumas propriedades de tal teoria.

Em geral, não iremos discutir a validade experimental implicada pelo formalismo matemático que discutiremos aqui, mas em alguns casos particulares importantes, como os sistemas descritos pelas matrizes de Dirichlet que veremos no próximo capítulo, sabemos que existe tal confirmação com a realidade (as matrizes de Dirichlet são um análogo discreto das amplitudes de Feynman calculadas em eletrodinâmica quântica (QED)).

Capítulo 2

Cadeias de Markov quânticas

2.1 Introdução

Nesta seção consideramos apenas $t \in \mathbb{N}$ e $S = \{0, 1, 2, \dots, n-1\}$. O grupo das transformações lineares unitárias B , agindo sobre \mathbb{C}^n é dito **grupo unitário** de ordem n . A evolução de um sistema quântico é descrita pela ação de t no grupo unitário, denotada por $U(t)$, onde $t \in \mathbb{N}$ é um parâmetro temporal. Este semigrupo indexado por $t \in \mathbb{N}$ pode ser usado para calcular o estado do sistema em qualquer tempo dado, bem como encontrar amplitudes de transição.

Suponha que temos um operador unitário U em um espaço de Hilbert (sobre os corpos de escalares complexos) \mathcal{H} de dimensão n que gera um grupo unitário discreto $U(t) = U^t$, $t \in \mathbb{N}$. Dizemos que se o sistema encontra-se no estado ψ em um certo tempo então $\langle \phi, U(t)\psi \rangle$ é a amplitude de transição do estado ψ para o estado ϕ após t unidades de tempo. Se ψ_j , $j = 0, 1, \dots, n-1$ é uma base ortonormal para \mathcal{H} , definimos a $n \times n$ matriz de amplitude de transição \mathcal{A} relativa a esta base por $\mathcal{A}_{jk} = \langle \psi_j, U\psi_k \rangle$. É claro que \mathcal{A} é uma matriz unitária (e possivelmente com entradas complexas) \mathcal{A}_{jk} , onde $j, k \in S$. Então no nosso caso, identificamos $U = \mathcal{A}$ e $\mathcal{H} = \mathbb{C}^n$. Esta descrição simples da mecânica quântica será suficiente para o que estudaremos neste capítulo. Iremos formular mais precisamente o que entendemos por um sistema quântico no capítulo 4, quando considerarmos os postulados da mecânica quântica.

Como temos uma interpretação probabilística desta cadeia, podemos perguntar se ela está associada a um processo estocástico que gera estas amplitudes de transição. A resposta é afirmativa, e as funções X_j do processo podem ser interpretadas como observações quânticas, ou como chamaremos, **medições**. Como a matriz A é unitária, dizemos que X_j é um **processo**

unitário; queremos, além disso, que os processos considerados tenham a propriedade de Markov. Para que isso ocorra, A não deve ser apenas unitária, mas também estocástica. Diremos que X_j satisfazendo a isto é uma **cadeia quântica**.

2.2 N-cadeias quânticas

Um aspecto essencial da teoria quântica é o conceito de interferência. Um observador que realiza uma medição em um sistema irá em geral modificá-lo. Ainda, distintas possibilidades (e probabilidades) interferem uma nas outras. Começaremos descrevendo matematicamente o que significa interferência entre eventos e suas medições. Consideramos inicialmente uma medida complexa A denominada amplitude. Esta medida será usada posteriormente para medir probabilidades (do mesmo modo como uma medida em uma cadeia de Markov clássica).

Seja $\Omega = S^{\mathbb{N}}$ (ou, então $\Omega = S^N$, onde $N \in \mathbb{N}$ está fixo) que chamaremos espaço amostral. Os elementos w de Ω são ditos **pontos amostrais** e cada ponto representa uma determinada configuração de um sistema físico. Seja Λ uma σ -álgebra fixada de subconjuntos de Ω e seja $A: \Lambda \rightarrow \mathbb{C}$ uma medida complexa com $A(\Omega) = \int_{\Omega} dA = 1$. Quando $\Omega = S^{\mathbb{N}}$, então Λ é a sigma-algebra gerada pelos cilindros, ou seja, a sigma-algebra de Borel. Para $\Delta \in \Lambda$, o número complexo $A(\Delta) = \int_{\Delta} dA$ é dito **amplitude quântica** (ou amplitude de transição, ou amplitude de probabilidade) do evento $\Delta \in \Lambda$. Assumimos acima que A é σ -aditiva, e tal propriedade será necessária na prova de alguns teoremas, como por exemplo na proposição (2.2.3).

A probabilidade P de que um evento $\Delta \in \Lambda$ ocorra é definida por

$$P(\Delta) := |A(\Delta)|^2$$

Observação Em geral, $P = |A|^2$ não é uma probabilidade no sentido de teoria da medida se supormos que A é uma medida complexa. O motivo é que a σ -aditividade não vale em geral: para C_1, C_2 conjuntos disjuntos,

$$\begin{aligned} P(C_1 \cup C_2) &= |A(C_1 \cup C_2)|^2 = |A(C_1) + A(C_2)|^2 \\ &= |A(C_1)|^2 + |A(C_2)|^2 + 2\operatorname{Re}(A(C_1)\overline{A(C_2)}) \end{aligned}$$

Os dois primeiros termos da última igualdade fornecem o resultado clássico e o último é um termo de interferência, que contribui com uma soma ou subtração à parte clássica (veja também a seção 3.2). Um lema básico (e negativo) relacionado com esta questão é o seguinte (o leitor pode omitir a demonstração numa primeira leitura):

Lema 2.2.1 *Seja $\Omega = \{1, 2, 3\}^{\mathbb{N}}$, Λ uma σ -álgebra de subconjuntos de Ω , $A : \Lambda \rightarrow \mathbb{C}$, $A = \mu_1 + i\mu_2$ uma medida complexa, onde $\mu_1, \mu_2 : \Lambda \rightarrow \mathbb{R}$ são medidas com sinal. Então $P : \Lambda \rightarrow \mathbb{R}$, $P = |A|^2$ é uma medida real aditiva (sobre os cilindros de tamanho 1) se e somente se pelo menos um dos cilindros $\bar{1}, \bar{2}, \bar{3} \in \Lambda$ (mas não todos),*

$$\bar{1} := \{\omega \in \Omega : \omega = (\omega_1, \omega_2, \dots), \omega_1 = 1\}$$

$$\bar{2} := \{\omega \in \Omega : \omega = (\omega_1, \omega_2, \dots), \omega_1 = 2\}$$

$$\bar{3} := \{\omega \in \Omega : \omega = (\omega_1, \omega_2, \dots), \omega_1 = 3\}$$

tem medida nula (i.e., $A(\bar{1}) = 0$ ou $A(\bar{2}) = 0$ ou $A(\bar{3}) = 0$).

Prova Como A é medida, vale a aditividade para os conjuntos disjuntos $\bar{1}$ e $\bar{2}$:

$$P(\bar{1} \cup \bar{2}) = |A(\bar{1} \cup \bar{2})|^2 = |A(\bar{1}) + A(\bar{2})|^2$$

Como $A(\bar{1}), A(\bar{2}) \in \mathbb{C}$,

$$|A(\bar{1}) + A(\bar{2})|^2 = |A(\bar{1})|^2 + |A(\bar{2})|^2 + 2\text{Re}(A(\bar{1})\overline{A(\bar{2})}) \quad (2.1)$$

e portanto,

$$P(\bar{1} \cup \bar{2}) = P(\bar{1}) + P(\bar{2}) \Leftrightarrow \text{Re}(A(\bar{1})\overline{A(\bar{2})}) = 0.$$

Mas

$$\text{Re}(A(\bar{1})\overline{A(\bar{2})}) = 0 \Leftrightarrow \mu_1(\bar{1})\mu_1(\bar{2}) + \mu_2(\bar{1})\mu_2(\bar{2}) = 0$$

e analogamente para os pares de cilindros $\bar{1}, \bar{3}$ e $\bar{2}, \bar{3}$:

$$\text{Re}(A(\bar{1})\overline{A(\bar{3})}) = 0 \Leftrightarrow \mu_1(\bar{1})\mu_1(\bar{3}) + \mu_2(\bar{1})\mu_2(\bar{3}) = 0$$

$$\text{Re}(A(\bar{2})\overline{A(\bar{3})}) = 0 \Leftrightarrow \mu_1(\bar{2})\mu_1(\bar{3}) + \mu_2(\bar{2})\mu_2(\bar{3}) = 0$$

Temos o sistema

$$\begin{cases} \mu_1(\bar{1})\mu_1(\bar{2}) + \mu_2(\bar{1})\mu_2(\bar{2}) = 0 \\ \mu_1(\bar{1})\mu_1(\bar{3}) + \mu_2(\bar{1})\mu_2(\bar{3}) = 0 \\ \mu_1(\bar{2})\mu_1(\bar{3}) + \mu_2(\bar{2})\mu_2(\bar{3}) = 0 \end{cases}$$

Suponha que a medida de um dos cilindros seja não nula, digamos a do cilindro $\bar{1}$. Então ao menos $\mu_1(\bar{1}) \neq 0$ ou $\mu_2(\bar{1}) \neq 0$. Suponha sem perda de generalidade que $\mu_2(\bar{1}) \neq 0$. Daí, a primeira e a segunda equação nos fornecem

$$\mu_2(\bar{2}) = -\frac{\mu_1(\bar{1})\mu_1(\bar{2})}{\mu_2(\bar{1})}, \quad \mu_2(\bar{3}) = -\frac{\mu_1(\bar{1})\mu_1(\bar{3})}{\mu_2(\bar{1})}.$$

Substituindo essas expressões na terceira equação, temos

$$\begin{aligned} \mu_1(\bar{2})\mu_1(\bar{3}) + \frac{\mu_1(\bar{1})\mu_1(\bar{2})}{\mu_2(\bar{1})} \frac{\mu_1(\bar{1})\mu_1(\bar{3})}{\mu_2(\bar{1})} &= 0 \Leftrightarrow \\ \Leftrightarrow \mu_1(\bar{2})\mu_1(\bar{3}) \left[1 + \frac{\mu_1(\bar{1})^2}{\mu_2(\bar{1})^2} \right] &= 0 \end{aligned}$$

Logo, $\mu_1(\bar{2})=0$ ou $\mu_1(\bar{3}) = 0$. Se $\mu_1(\bar{2}) = 0$ então $\mu_2(\bar{2}) = 0$ pois

$$\mu_2(\bar{2}) = -\frac{\mu_1(\bar{1})\mu_1(\bar{2})}{\mu_2(\bar{1})} = 0$$

Se $\mu_1(\bar{3}) = 0$ então $\mu_2(\bar{3}) = 0$ pois

$$\mu_2(\bar{3}) = -\frac{\mu_1(\bar{1})\mu_1(\bar{3})}{\mu_2(\bar{1})} = 0.$$

Portanto, em qualquer caso, um dos cilindros básicos deve ter medida nula. □

A demonstração deste lema vale para qualquer espaço mensurável que admita uma cobertura finita.

Vamos fazer agora a construção de N-cadeias quânticas. Começamos com o conceito de medição.

Definição Seja $S = \{s_0, s_1, \dots, s_n\}$ um conjunto finito e seja $X : \Omega \rightarrow S$. Dizemos que X é uma **medição** se $X^{-1}(s_j) \in \Lambda$, $j = 0, \dots, n - 1$, ou seja, que X é mensurável.

Note que se P é probabilidade sobre Ω então: $\sum_j P(X^{-1}(s_j)) = 1$ onde $X^{-1}(s_j)$ é o conjunto de eventos que produzem o resultado s_j .

Usaremos a notação $[X = s_j]$ para o conjunto $X^{-1}(s_j)$. Ainda, $P(X = s_j) = P(X^{-1}(s_j))$.

A menos que seja especificado, podemos supor que $\Omega = S^N$ para algum N ou então $\Omega = S^{\mathbb{N}}$.

Notação Escreveremos

$$\overline{a_0 a_1 \cdots a_t} := \{\omega \in \Omega | X_0(\omega) = a_0, X_1(\omega) = a_1, \dots, X_t(\omega) = a_t\},$$

onde $a_j \in S$ para todo $j \in \{0, 1, 2, \dots, t\}$. Com esta notação, podemos aplicar a σ -aditividade de A para escrever

$$A[X_t = s_k] = A\left[\bigcup_{i_0, i_1, \dots, i_{t-1}} \overline{s_{i_0} s_{i_1} \dots s_{i_{t-1}} s_k}\right] = \sum_{i_0, i_1, \dots, i_{t-1}} A[\overline{s_{i_0} s_{i_1} \dots s_{i_{t-1}} s_k}]$$

Outra notação que usaremos é $i(t) := i_t$.

Definição Dizemos que a seqüência $\{X_t\}$, $0 \leq t \leq N$, com $X_t : \Omega \rightarrow S$ e com todas as medições sobre o mesmo conjunto de resultados $S = \{s_0, \dots, s_{n-1}\}$, é uma **N-cadeia** se, para todo $t = 1, \dots, N$ e

$$s_{j(1)}, s_{j(2)}, \dots, s_{j(t-1)}, s_{j(t)},$$

temos

$$A[X_t = s_{j(t)}, X_{t-1} = s_{j(t-1)}, \dots, X_1 = s_{j(1)}] \neq 0$$

onde $s_{j(k)}$ são eventos que ocorrem no tempo k .

Observação Em [15], a definição de N -cadeia inclui ainda a seguinte condição:

$$[X_0 = s_0] = \Omega$$

ou seja, temos um certo estado inicial fixado s_0 . Em geral não vamos supor que essa condição vale.

A partir de agora as medições $X_t : \Omega = S^N \rightarrow S$ serão **sempre** da forma

$$X_t(w) = w_t, \text{ onde } w = (w_0, w_1, \dots, w_N) \in S^N.$$

Ou então, $X_t : \Omega = S^{\mathbb{N}} \rightarrow S$ serão **sempre** da forma

$$X_t(w) = w_t, \text{ onde } w = (w_0, w_1, \dots, w_t, \dots) \in S^{\mathbb{N}}.$$

Definição A amplitude do evento Δ_1 condicionada ao evento Δ_2 é

$$A(\Delta_1|\Delta_2) := \frac{A(\Delta_1 \cap \Delta_2)}{A(\Delta_2)}$$

se $A(\Delta_2) \neq 0$ e vale 0 caso contrário.

Definição Uma N -cadeia $\{X_t\}$ é **homogênea no tempo** se

$$A[X_{t+1} = s_j | X_t = s_k] = A[X_1 = s_j | X_0 = s_k]$$

para todo $j, k = 1, \dots, n-1$ e $t = 1, 2, \dots, N-1$.

Definição Fixada uma medida complexa A , uma N -cadeia $\{X_t\}$ homogênea no tempo é uma **cadeia de Markov** se

$$A[X_{t+1} = s_j | X_t = s_{j(t)}, X_{t-1} = s_{j(t-1)}, \dots, X_1 = s_{j(1)}] = A[X_{t+1} = s_j | X_t = s_{j(t)}]$$

para todo $t = 1, \dots, N - 1$ e $j, j(t), \dots, j(1) = 0, \dots, n - 1$.

A matriz \mathcal{A} da forma n por n associada a tal medida complexa A (que define uma cadeia de Markov homogênea no tempo X_t como definida acima) sobre $\Omega = S^N$ é dada por

$$\mathcal{A}_{ij} = A[X_1 = s_j | X_0 = s_i],$$

$i, j \in \{0, 1, 2, \dots, n - 1\}$.

Os conceitos análogos para $X_t : \Omega = S^{\mathbb{N}} \rightarrow S$, com $t \in \mathbb{N}$, são definidos de maneira semelhante.

Definição Sejam X e Y medições, $X : \Omega \rightarrow R$, $Y : \Omega \rightarrow S$, $R = \{r_0, \dots, r_{m-1}\}$, $S = \{s_0, \dots, s_{n-1}\}$. Dizemos que Y **não interfere** em X se, $\forall j, k, j = 0, \dots, m - 1$:

$$P[X = r_j] = \sum_{k=0}^{n-1} P[X = r_j, Y = s_k]$$

A probabilidade do evento Δ_1 condicionada a Δ_2 é dada por $P(\Delta_1 | \Delta_2) = P(\Delta_1 \cap \Delta_2) / P(\Delta_2)$, $P(\Delta_2) \neq 0$. Dizemos que $A[X_2 = s_j | X_1 = s_k]$ é a amplitude de transição do sistema do estado s_k para o estado s_j em uma unidade de tempo. Interpretamos o conjugado complexo $\overline{A}[X_2 = s_j | X_1 = s_k]$ como sendo a amplitude de transição de s_j para s_k em -1 unidades de tempo. Esta propriedade será útil para indicarmos que na cadeia em que estivermos trabalhando, o sistema não pode dar saltos em acréscimos de tempo $t = 0$, ou seja, se $j \neq k$ então ela satisfaz a seguinte condição:

$$\sum_{r=0}^{n-1} A[X_2 = s_r | X_1 = s_j] \overline{A}[X_2 = s_r | X_1 = s_k] = 0. \quad (2.2)$$

Diremos que

$$P_{jk} = |A[X_2 = s_j | X_1 = s_k]|^2$$

é a probabilidade de transição de s_k para s_j . Se X_2 não interfere em X_1 então

$$\sum_{j=0}^{n-1} P_{jk} = \frac{1}{P[X_1(s_k)]} \sum_{j=0}^{n-1} P[X_2(s_j), X_1(s_k)] = 1$$

ou seja, dado um estado fixo s_k em $t = 1$, as somas das probabilidades de transição para qualquer outro estado vale 1. Logo, dizer que X_2 não interfere em X_1 significa que para qualquer s_k fixado

$$\sum_{j=0}^{n-1} P[X_2 = s_j | X_1 = s_k] = 1 \quad (2.3)$$

Definição Uma N-cadeia é **unitária** se X_2 não interfere em X_1 (o que implica $\sum_r |A_{rj}|^2 = 1$) e se satisfaz a equação (2.2).

Definição Uma N-cadeia homogênea no tempo, unitária e Markov é chamada **N-cadeia quântica**.

Conceitos similares podem ser considerados para o caso de $\Omega = S^{\mathbb{N}}$ e $X_t : \Omega \rightarrow S$, medições com $t \in \mathbb{N}$. Neste caso teríamos uma \mathbb{N} -cadeia quântica, ou, cadeia quântica para simplificar.

A matriz \mathcal{A} da forma $n \times n$ com entradas

$$\mathcal{A}_{jk} = A[X_2 = s_k | X_1 = s_j]$$

é a **matriz de amplitude quântica de transição** para $\{X_t\}$. Desta maneira, a evolução do processo fica determinada por uma matriz \mathcal{A} com valores complexos. Associada a esta matriz temos uma matriz estocástica P real. Examinemos com mais detalhes as propriedades da matriz de amplitude.

Primeiro, A é uma **matriz unitária**, pois da equação (2.2), $\sum_r \bar{A}_{jr} A_{rk} = \sum_r \bar{A}_{rj} A_{rk} = 0$ se $j \neq k$; se $j = k$, temos pela equação (2.3) (que é válida neste caso pois estamos supondo que $\{X_t\}$ é N-cadeia quântica, e portanto X_2 não interfere em X_1), que

$$\sum_r \bar{A}_{jr} A_{rk} = \sum_r \bar{A}_{rj} A_{rj} = \sum_r |A_{rj}|^2 = \sum_r |A[X_2 = s_r | X_1 = s_k]|^2 = 1.$$

Segundo, pela definição de N-cadeia, temos que **os valores A_{jk} são diferentes de zero**, $j, k = 0, \dots, n - 1$.

Terceiro, A_{jk} é uma matriz **coluna-estocástica**, pois

$$\sum_j A_{jk} = \frac{1}{A[X_1 = s_k]} \sum_j A[X_2 = s_j, X_1 = s_k] = 1.$$

Observe que a primeira propriedade que citamos vale para qualquer N-cadeia unitária, e as outras duas valem para qualquer N-cadeia. Em resumo,

A descreve a evolução com probabilidades complexas (ondulatória) e P é a evolução correspondente com probabilidades reais.

Resumindo, mostramos que a cada N-cadeia quântica corresponde uma matriz estocástica, unitária e com entradas não nulas. Reciprocamente, toda matriz $n \times n$ com essas propriedades é a matriz de amplitude de transição de uma N-cadeia quântica [15]. De fato:

Observação O processo acima pode se revertido a partir de uma matriz complexa \mathcal{A} do tipo n por n que seja estocástica e unitária. Dada a matriz \mathcal{A} , temos uma maneira natural de definir a medida de amplitude complexa A sobre $S^{\mathbb{N}}$. Isso é análogo ao que é feito no caso real [36]. Seja $\vec{p} = (p_0, \dots, p_{n-1})$ uma distribuição de probabilidade inicial, isto é, um vetor tal que $p_0 + \dots + p_{n-1} = 1$. Podemos definir uma probabilidade sobre os cilindros de $S^{\mathbb{N}} = \{s_0, \dots, s_{n-1}\}^{\mathbb{N}}$ por

$$A(\overline{a_0 a_1 \dots a_t}) = \mathcal{A}_{a_t a_{t-1}} \mathcal{A}_{a_{t-1} a_{t-2}} \dots \mathcal{A}_{a_1 a_0} (\vec{p})_{a_0}$$

Isto define a medida complexa sobre toda a σ -álgebra gerada pelos cilindros, pelo teorema da extensão de Kolmogorov. Esta A é sigma aditiva. Chamamos A de **medida complexa de Markov**.

Seja agora $X_t : S^{\mathbb{N}} \rightarrow S$, onde $X_t(w) = w_t$, e $w = (w_0, w_1, \dots, w_t, \dots)$, com $t \in \mathbb{N}$.

É fácil ver que X_t , $t \in \mathbb{N}$, e A satisfazem a hipótese de ser uma cadeia quântica.

Se \vec{p} é tal que $\mathcal{A}\vec{p} = \vec{p}$, então o processo X_t será A - estacionário, ou seja,

$$A(X_{t_1} = a_1, X_{t_2} = a_2, \dots, X_{t_n} = a_n) = A(X_{t_1+t} = a_1, X_{t_2+t} = a_2, \dots, X_{t_n+t} = a_n)$$

para todos $t \geq 0$, e $0 \leq t_1 < t_2 < \dots < t_n$, e $a_1, a_2, \dots, a_n \in S$.

Teorema 2.2.2 *Em uma N-cadeia quântica $\{X_t\}$, $X_{t'}$ não interfere em X_t , para todo $0 \leq t \leq t' \leq N$. Em outras palavras, o futuro não interfere no passado.*

Prova Usaremos a notação $j(t) := j_t$. Podemos assumir que $t \geq 1$. Pela propriedade de Markov e pela estacionariedade, temos:

$$\begin{aligned} & A[X_t^{-1}(s_j) \cap X_{t'}^{-1}(s_k)] \\ &= \sum_{j(1), \dots, j(t'-1)} A[X_{t'}^{-1}(s_k) \cap X_{t'-1}^{-1}(s_{j(t'-1)}) \cap \dots \cap X_t^{-1}(s_j) \cap \dots \cap X_1^{-1}(s_{j(1)})] \end{aligned}$$

$$\begin{aligned}
&= \sum_{j(1), \dots, j(t'-1)} A[X_{t'-1}^{-1}(s_{j(t'-1)}) \cap \dots \cap X_1^{-1}(s_{j(1)})] A[X_{t'}^{-1}(s_k) | X_{t'-1}^{-1}(s_{j(t'-1)})] \\
&= \sum_{j(t'-1)} A[X_t^{-1}(s_j) \cap X_{t'-1}^{-1}(s_{j(t'-1)})] A_{kj(t'-1)} \tag{2.4}
\end{aligned}$$

Iterando esta equação obtemos

$$\begin{aligned}
&A[X_t^{-1}(s_j) \cap X_{t'}^{-1}(s_k)] \\
&= \sum_{j(t'-1)} \sum_{j(t'-2)} A[X_t^{-1}(s_j) \cap X_{t'-2}^{-1}(s_{j(t'-2)})] A_{kj(t'-1)} A_{j(t'-1)j(t'-2)} \\
&= \sum_{j(t'-2)} A[X_t^{-1}(s_j) \cap X_{t'-2}^{-1}(s_{j(t'-2)})] A_{kj(t'-2)}^2 = \dots \\
&= \dots = \sum_{j(t)} A[X_t^{-1}(s_j) \cap X_t^{-1}(s_{j(t)})] (A^{t'-t})_{kj(t)} = A[X_t^{-1}(s_j)] (A^{t'-t})_{kj}
\end{aligned}$$

Como A é unitária, $A^{t'-t}$ também é, e portanto

$$\sum_k P[X_t^{-1}(s_j) \cap X_{t'}^{-1}(s_k)] = P[X_t^{-1}(s_j)] \sum_k |A_{kj}^{t'-t}|^2 = P[X_t^{-1}(s_j)]$$

o que conclui a prova. □

Note que no último passo da demonstração fica clara a importância da matriz A ser unitária: esse fato é essencial para garantir a não interferência entre as medições da N -cadeia. Note que podemos analisar uma teoria de amplitudes associadas a matrizes complexas estocásticas, sem levar em conta a unitariedade das matrizes, mas nesse caso a demonstração que fizemos do teorema acima não é válida.

Suponha uma N -cadeia $\{X_t\}$ com matriz de amplitude de transição $\mathcal{A} = (\mathcal{A})_{jk}$, $j, k \in S$. A **amplitude** e a **distribuição** dos estados $s_0, s_1, \dots, s_{n-1} \in S$ no tempo $t = 0, \dots, N$ são dados, respectivamente, por

$$\vec{X}_t = (A[X_t = s_0], \dots, A[X_t = s_{n-1}]) \in \mathbb{C}^n$$

que é um vetor unitário (ou seja, $\sum |A|^2 = 1$) e

$$P_t(k) = P[X_t = s_k] = |A[X_t = s_k]|^2 = |(\vec{X}_t)_k|^2.$$

Teorema 2.2.3 Em uma N -cadeia quântica $\{X_t\}$, $\vec{X}_t = A^t \vec{X}_0$ e

$$p(X_t = k) = P_t(k) = |(\mathcal{A}^t \vec{X}_0)_k|^2. \quad (2.5)$$

Prova Por ser uma cadeia de Markov e pela homogeneidade, temos

$$\begin{aligned} & A[X_t^{-1}(s_k) \cap X_{t-1}^{-1}(s_{j(t-1)}) \cap \cdots \cap X_0^{-1}(s_{j(0)})] \\ &= A[X_{t-1}^{-1}(s_{j(t-1)}) \cap \cdots \cap X_0^{-1}(s_{j(0)})] A[X_t^{-1}(s_k) | X_{t-1}^{-1}(s_{j(t-1)})] \\ &= \mathcal{A}_{kj(t-1)} A[X_{t-1}^{-1}(s_{j(t-1)}) \cap \cdots \cap X_0^{-1}(s_{j(0)})] \\ &= \mathcal{A}_{kj(t-1)} \mathcal{A}_{j(t-1)j(t-2)} \cdots \mathcal{A}_{j(2)j(1)} A[X_1^{-1}(s_{j(1)}) | X_0^{-1}(s_{j(0)})] \\ &= \mathcal{A}_{kj(t-1)} \mathcal{A}_{j(t-1)j(t-2)} \cdots \mathcal{A}_{j(1)j(0)} (\vec{X}_0)_{j(0)} \end{aligned}$$

Portanto,

$$\begin{aligned} (\vec{X}_t)_k &= A[X_t = s_k] \\ &= \sum_{j(t-1), \dots, j(0)} \mathcal{A}_{kj(t-1)} \mathcal{A}_{j(t-1)j(t-2)} \cdots \mathcal{A}_{j(1)j(0)} (\vec{X}_0)_{j(0)} = (\mathcal{A}^t \vec{X}_0)_k \end{aligned}$$

Segue que $\vec{X}_t = \mathcal{A}^t \vec{X}_0$ e que $P_t(k) = |(\mathcal{A}^t \vec{X}_0)_k|^2$.

□

Este teorema ilustra o fato de que temos dois sistemas evoluindo com o tempo em paralelo: um complexo dado por $\mathcal{A}^t \vec{X}_0$ e outro real dado por $P^t \vec{X}_0$. O sistema definido por P é a cadeia de Markov clássica e o definido por \mathcal{A} é a cadeia de Markov quântica.

Uma maneira mais simples de se obter a evolução da cadeia quântica é através dos autovalores e autovetores de \mathcal{A} . Desta forma obtemos uma expressão explícita para o cálculo de probabilidades $P(X_t = k)$.

Sejam $\lambda_0, \dots, \lambda_{n-1}$ autovalores (possivelmente repetidos) de uma matriz unitária \mathcal{A} e seja $\psi_0, \dots, \psi_{n-1}$ a base ortonormal de autovetores correspondente. Como é feito em [15], supondo que $\vec{X}_0 = (1, 0, \dots, 0)$, temos:

$$\mathcal{A}^t \vec{X}_0 = \mathcal{A}^t \sum_j \langle \vec{X}_0, \psi_j \rangle \psi_j = \sum_j (\overline{\psi_j})_0 \mathcal{A}^t \psi_j = \sum_j (\overline{\psi_j})_0 \lambda_j^t \psi_j$$

E da equação (2.5), obtemos

$$P_t(k) = \left| \sum_j (\overline{\psi_j})_0 \lambda_j^t (\psi_j)_k \right|^2. \quad (2.6)$$

Usaremos esta expressão para calcular probabilidades de transição, mas em alguns exemplos iremos considerar o caso geral em que \vec{X}_0 é qualquer.

2.3 Matrizes de Dirichlet

Nesta seção vamos exibir uma grande classe de matrizes \mathcal{A} satisfazendo a hipótese de ser matriz quântica.

Definição Sejam n, a inteiros positivos, $(n, a) = 1$. A matriz de Dirichlet $M(n, a)$ é a matriz $n \times n$ com entradas

$$\mathcal{B}_{jk} = \frac{1}{\sqrt{n}} e^{i\pi a(j-k)^2/n}, \quad j, k = 0, 1, \dots, n-1$$

É possível mostrar que $M(n, a)$ gera um análogo discreto da amplitude de Feynman para uma partícula livre, e que se $n \rightarrow \infty$, esse análogo se aproxima da amplitude de Feynman.

Queremos saber quando a matriz $M(n, a)$ é unitária. Precisamos do seguinte lema, cuja prova é imediata.

Lema Dois inteiros positivos n e a são relativamente primos se e somente se $al \neq nm$, para quaisquer inteiros l, m , com $0 < |l| < n$.

Teorema A matriz $M(n, a)$ é unitária se e somente se n e a são relativamente primos.

-**Prova** Como

$$\sum_{k=0}^{n-1} |\mathcal{B}_{jk}|^2 = 1$$

para $j = 0, \dots, n-1$, é claro que $M(n, a)$ é unitária se e somente se

$$\sum_{k=0}^{n-1} \mathcal{B}_{jk} \overline{\mathcal{B}_{j'k}} = 0$$

para $j \neq j'$. Nesse caso, temos

$$\begin{aligned} n \sum_{k=0}^{n-1} \mathcal{B}_{jk} \overline{\mathcal{B}_{j'k}} &= \sum_{k=0}^{n-1} \exp \left\{ \frac{i\pi a}{n} [(j-k)^2 - (j'-k)^2] \right\} \\ &= e^{i\pi a(j^2-j'^2)/n} \sum_{k=0}^{n-1} e^{i2\pi a(j-j)k/n} \end{aligned}$$

Se n e a são relativamente primos temos, aplicando o lema, que a soma geométrica na última expressão satisfaz

$$\sum_{k=0}^{n-1} \left[e^{i2\pi a(j'-j)/n} \right]^k = \frac{1 - e^{i2\pi a(j'-j)}}{1 - e^{i2\pi a(j'-j)/n}} = 0$$

e portanto $M(n, a)$ é unitária. Se n e a não são relativamente primos então pelo lema, existe $j \neq j'$ tal que $a(j - j') = nm$ para algum inteiro m . Nesse caso, a série geométrica tem soma n e portanto $M(n, a)$ não é unitária, o que conclui a prova. □

Definição A soma de Dirichlet é dada por

$$S(n, a) = \sum_{j=0}^{n-1} e^{i\pi a j^2/n}$$

A distribuição de probabilidades gerada pela matriz de Dirichlet depende da paridade de na . O seguinte lema ilustra o problema.

Lema 2.3.1 1. Se na é par então

$$\sum_{j=0}^{n-1} e^{i\pi a(j-k)^2/n} = S(n, a), \quad 0 \leq k \leq 2n - 2$$

2. Se na é ímpar então

$$\sum_{j=0}^{n-1} e^{i\pi a(j-k-1/2)^2/n} = \frac{1}{2} S(4n, a), \quad 0 \leq k \leq 2n - 2$$

Prova 1. Primeiro suponha que $j \leq n - 1$. Então

$$S := \sum_{k=0}^{n-1} e^{i\pi a(k-j)^2/n} = \sum_{k=0}^{j-1} e^{i\pi a(k-j)^2/n} + \sum_{k=j}^{n-1} e^{i\pi a(k-j)^2/n}$$

Fazendo $r = j - k$ no primeiro somatório e $r = n - k + j$ no segundo, obtemos

$$S = \sum_{r=1}^j e^{i\pi a r^2/n} + \sum_{r=j+1}^n e^{i\pi a(n-r)^2/n}$$

Como na é par, obtemos

$$S = \sum_{r=1}^n e^{i\pi a r^2/n} = \sum_{r=0}^{n-1} e^{i\pi a r^2/n} = S(n, a)$$

A seguir, suponha que $n \leq j \leq 2n - 2$. Então $j = n + r$ para algum inteiro $0 \leq r \leq n - 2$. Novamente, como na é par, temos

$$S = \sum_{k=0}^{n-1} e^{i\pi a(k-n-r)^2/n} = \sum_{k=0}^{n-1} e^{i\pi a(k-r)^2/n}$$

Mas este último somatório é igual a $S(n, a)$ pelo resultado anterior.

2. Considere a soma

$$T := \sum_{k=0}^{4n-1} e^{i\pi a(k-2j)^2/4n}$$

Vale que T é a soma de suas somas parciais $T = U + E$, onde

$$U := \sum_{k=1}^{2n} e^{i\pi a(2k-1-2j)^2/4n}$$

$$E := \sum_{k=0}^{2n-1} e^{i\pi a(2k-2j)^2/4n}$$

Podemos escrever E como

$$E = \sum_{k=0}^{n-1} e^{i\pi a(k-j)^2/n} + \sum_{k=n}^{2n-1} e^{i\pi a(k-j)^2/n}$$

Fazendo $r = k - n$ no segundo somatório e usando o fato de que na é ímpar, obtemos

$$\sum_{k=n}^{2n-1} e^{i\pi a(k-j)^2/n} = \sum_{r=0}^{n-1} e^{i\pi a(r+n-j)^2/n} = - \sum_{r=0}^{n-1} e^{i\pi a(r-j)^2/n}$$

Portanto, $E = 0$. Como

$$e^{i\pi a(4n-1-2j)^2/4n} = e^{i\pi a(1+2j)^2/4n},$$

temos

$$U = \sum_{k=0}^{n-1} e^{i\pi a(2k-1-2j)^2/4n} + \sum_{k=n}^{2n-1} e^{i\pi a(2k-1-2j)^2/4n}$$

Novamente, fazendo $r = k - n$ no segundo somatório e usando o fato de que na é ímpar, obtemos

$$\sum_{k=n}^{2n-1} e^{i\pi a(2k-1-2j)^2/4n} = \sum_{r=0}^{n-1} e^{i\pi a(2r+2n-1-2j)^2/4n} = \sum_{r=0}^{n-1} e^{i\pi a(2r-1-2j)^2/4n}$$

Segue que

$$T = U = 2 \sum_{k=0}^{n-1} e^{i\pi a(k-j-1/2)^2/n}$$

Mas como $4na$ é par, segue da parte 1 que $T = S(4n, a)$.

□

Uma consequência deste lema é que se na é par então a soma das linhas e das colunas de $M(n, a)$ vale $n^{-1/2}S(n, a)$; além disso, é mostrado em [15],[17] que $|S(n, a)| = n^{1/2}$. Podemos concluir daí que a matriz

$$\mathcal{A} = M'(n, a) = \frac{\overline{S(n, a)}}{\sqrt{n}} M(n, a)$$

é estocástica. Ela também tem entradas não nulas e é unitária.

Desta forma obtemos uma grande classe de matrizes \mathcal{A} satisfazendo o que foi exigido anteriormente. Por exemplo, vamos supor $n = 4$, $a = 1$. Então

$$\mathcal{B}_{jk} = \frac{1}{2} e^{i\pi(j-k)^2/4}, \quad j, k = 0, \dots, 3$$

e

$$S(4, 1) = 1 + e^{i\pi/4} - 1 + e^{i\pi/4} = 2e^{i\pi/4}$$

Daí, a matriz M' é

$$\mathcal{A} = M'(4, 1) = \frac{1}{2} e^{-i\pi/4} \begin{pmatrix} 1 & e^{i\pi/4} & -1 & e^{i\pi/4} \\ e^{i\pi/4} & 1 & e^{i\pi/4} & -1 \\ -1 & e^{i\pi/4} & 1 & e^{i\pi/4} \\ e^{i\pi/4} & -1 & e^{i\pi/4} & 1 \end{pmatrix}$$

Entretanto, se na é ímpar então a matriz de Dirichlet não pode se tornar estocástica desta forma. Por exemplo, se $n = 3$ e $a = 1$, temos:

$$M(3, 1) = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & e^{i\pi/3} & e^{i\pi 4/3} \\ e^{i\pi/3} & 1 & e^{i\pi/3} \\ e^{i\pi 4/3} & e^{i\pi/3} & 1 \end{pmatrix}$$

A soma da primeira e terceira colunas difere da segunda e então nenhum múltiplo de $M(3, 1)$ pode se tornar estocástica.

Para que possamos usar a equação (2.6) obtida na seção anterior, precisamos dos autovalores e autovetores de $M(n, a)$.

Teorema 2.3.2 1. Se na é par então, para $r = 0, \dots, n-1$, os autovalores de $M(n, a)$ são

$$\lambda_r = n^{-1/2} S(n, a) e^{-i\pi ar^2/n}$$

e uma base ortonormal de autovetores correspondente é ψ_r , onde

$$(\psi_r)_k = n^{-1/2} e^{-i2\pi ark/n}, \quad k = 0, \dots, n-1$$

2. Se na é ímpar então, para $r = 0, \dots, n-1$, os autovalores de $M(n, a)$ são

$$\lambda_j = \frac{1}{2n^{1/2}} S(4n, a) e^{-i\pi a(r+1/2)^2/n}$$

e uma base ortonormal de autovetores correspondente é ψ_r , onde

$$(\psi_r)_k = n^{-1/2} e^{-i2\pi a(r+(1/2))k/n}, \quad k = 0, \dots, n-1$$

Prova 1. A j -ésima coordenada de $M(n, a)\psi_r$ é

$$[M(n, a)\psi_r]_j = n^{-1/2} \sum_{k=0}^{n-1} e^{i\pi a(j-k)^2/n} e^{-i2\pi ark/n}$$

Mas

$$\begin{aligned} & e^{-i2\pi arj/n} \exp \left\{ i\pi a[(j-k)^2 - 2rk + 2rj]/n \right\} \\ &= (\psi_r)_j e^{-i\pi ar^2/n} e^{i\pi a(k-(j+r))^2/n} \end{aligned}$$

Aplicando o lema 2.3.1, item 1, obtemos

$$\begin{aligned} [M(n, a)\psi_r]_j &= (\psi_r)_j n^{-1/2} e^{-i\pi ar^2/n} \sum_{k=0}^{n-1} e^{i\pi a(k-(j+r))^2/n} \\ &= n^{-1/2} S(n, a) e^{-i\pi ar^2/n} (\psi_r)_j = \lambda_r (\psi_r)_j \end{aligned}$$

2. A j -ésima coordenada de $M(n, a)\psi_r$ é

$$[M(n, a)\psi_r]_j = n^{-1/2} \sum_{k=0}^{n-1} e^{i\pi a(j-k)^2/n} e^{-i\pi a(2r+1)k/n}$$

Mas

$$\begin{aligned} & e^{-i\pi a(2r+1)j/n} \exp \left\{ i\pi a[(j-k)^2 - (2r+1)k + (2r+1)j]/n \right\} \\ &= (\psi_r)_j e^{-i\pi a(r+1/2)^2/n} e^{i\pi a(k-(j+r+1/2))^2/n} \end{aligned}$$

Aplicando o lema 2.3.1, item 2, obtemos

$$\begin{aligned} [M(n, a)\psi_r]_j &= (\psi_r)_j n^{-1/2} e^{-i\pi a(r+1/2)^2/n} \sum_{k=0}^{n-1} e^{i\pi a(k-(j+r+1/2))^2/n} \\ &= \frac{n^{-1/2}}{2} S(4n, a) e^{-i\pi a(r+1/2)^2/n} (\psi_r)_j = \lambda_r(\psi_r)_j \end{aligned}$$

□

Aplicamos a equação acima usando o teorema anterior, e obtemos para na par que

$$\sum_j (\overline{\psi_j})_0 \lambda_j^t(\psi_j)_k = \frac{S(n, a)^t}{n^{t/2+1}} \sum_j e^{-i\pi a t j^2/n} e^{-i2\pi a j k/n}$$

e para na ímpar obtemos

$$\sum_j (\overline{\psi_j})_0 \lambda_j^t(\psi_j)_k = \frac{S(4n, a)^t}{2^t n^{3/2}} \sum_j e^{-i\pi a t (j+1/2)^2/n} e^{-i2\pi a (j+1/2)k/n}$$

Mas como $|S(n, a)| = \sqrt{n}$, obtemos da equação (2.6), para $t > 0$ e na par que

$$P(X_t = k) = P_t(k) = \frac{1}{n^2} \left| \sum_{j=0}^{n-1} e^{-i\pi a (tj+k)^2/nt} \right|^2, \quad (2.7)$$

e para na ímpar que

$$P(X_t = k) = \frac{1}{n^2} \left| \sum_{j=0}^{n-1} e^{-i\pi a (t(j+1/2)+k)^2/nt} \right|^2 \quad (2.8)$$

Acima assumimos que a a matriz de transição estocástica \mathcal{P} é tal que $\mathcal{P}_{ij} = |\mathcal{A}_{ij}|^2$, para todo i, j , e ainda que o vetor de probabilidade inicial é a delta de Dirac em s_0 , ou seja, $(1, 0, 0, \dots, 0)$.

Estas equações nos fornecem expressões explícitas para P_t , mas elas não estão em uma forma fechada e portanto não nos mostram muita informação sobre a dinâmica do sistema. Faremos o trabalho técnico de calcular tais somas na próxima seção.

2.4 Distribuições de Probabilidade

Agora vamos calcular as probabilidades fornecidas pelas equações (2.7) e (2.8) dadas na seção anterior. Vamos analisar também alguns exemplos de outros tipos (não Dirichlet).

Notação Sejam a, b inteiros. Denotaremos por $a_{(2)}$ o número de vezes em que o fator 2 aparece na decomposição prima de a . Ainda, o máximo divisor comum entre a e b é dado por (a, b) .

Note que estamos usando as distribuições obtidas pela expressão (2.6) da seção anterior, que foi obtida tomando como distribuição inicial o vetor $(1, 0, \dots, 0)$.

Temos o seguinte teorema:

Teorema 2.4.1 *Sejam n e t inteiros, com $(n, t) = d, t > 0$.*

1. *Se n é par, então $P(X_t = k) = d/n$ se $t_{(2)} \neq n_{(2)}$ e $d|k$, ou, se $t_{(2)} = n_{(2)}$ e $2k/d$ é ímpar. Caso contrário, $P(X_t = k) = 0$.*
2. *Se n é ímpar então $P(X_t = k) = d/n$ se $d|k$. Caso contrário, $P(X_t = k) = 0$.*

Prova 1. Aplicando (2.7) e o lema 2.6.1, ítem 1¹, temos

$$P(X_t = k) = \frac{1}{n^2} \left| \sum_{m=0}^{d-1} (-1)^{amnt/d^2} e^{-i2\pi akm/d} \right|^2 \left| \sum_{j=0}^{n/d-1} e^{-ia\pi(tj+k)^2/nt} \right|^2$$

Suponha que $t_{(2)} \neq n_{(2)}$. Então nt/d^2 é par, e daí

$$\sum_{m=0}^{d-1} (-1)^{amnt/d^2} e^{-i2\pi akm/d} = \sum_{m=0}^{d-1} (e^{-i2\pi ak/d})^m$$

A soma geométrica vale d se $d|k$ e vale 0 caso contrário. Suponha que $t_{(2)} = n_{(2)}$. Então nt/d^2 é ímpar e daí

$$\sum_{m=0}^{d-1} (-1)^{amnt/d^2} e^{-i2\pi akm/d} = \sum_{m=0}^{d-1} (-e^{-i2\pi ak/d})^m$$

¹ver os lemas 2.6.1 e 2.6.2 no Apêndice deste capítulo

Se $2k/d$ é um inteiro ímpar então a soma geométrica vale d . Caso contrário, temos

$$\sum_{m=0}^{d-1} (-e^{-i2\pi ak/d})^m = \frac{1 - (-1)^d}{1 + e^{-i2\pi ak/d}}$$

Como $t_{(2)} = n_{(2)}$ e n é par, temos que d é par e portanto a última expressão vale zero. Concluimos que quando $P(X_t = k)$ não é zero, temos

$$P(X_t = k) = \frac{d^2}{n^2} \left| \sum_{j=0}^{n/d-1} e^{-i\pi a(tj+k)^2/nt} \right|^2 = \frac{d^2}{n^2} \left| \sum_{j=0}^{n'-1} e^{-i\pi a(t'j+1/2)^2/nt} \right|^2$$

onde $n' = n/d$ e $t' = t/d$. Segue que $(n', t') = 1$. Se $t_{(2)} \neq n_{(2)}$ e $d|k$, vale que nt/d^2 é par e $n'at' = ant/d^2$ é par. Aplicando o lema 2.6.2, parte 1 obtemos

$$\left| \sum_{j=0}^{n'-1} e^{-i\pi a(t'j+k/d)/n't'} \right|^2 = |S(n', at')|^2 = n'$$

Logo,

$$P(X_t = k) = \frac{d^2}{n^2} n' = \frac{d}{n}$$

Se $t_{(2)} = n_{(2)}$ e $2k/d$ é ímpar então nt/d^2 é ímpar e $n'at' = ant/d^2$ é ímpar. Fazendo $2k/d = 2\mu + 1$ e aplicando o lema 2.6.1, ítem 2, obtemos

$$\left| \sum_{j=0}^{n'-1} e^{-i\pi a(t'j+\mu+1/2)^2/n't'} \right|^2 = \frac{1}{2} |S(4n', at')|^2 = n'$$

Novamente,

$$P(X_t = k) = \frac{d^2}{n^2} n' = \frac{d}{n}$$

2. Seja n ímpar e a par. Aplicando (2.7) e o lema 2.6.1, parte 1, obtemos $P(X_t = k)$ como no item 1 deste teorema. Se $t_{(2)} \neq n_{(2)}$ então nt/d^2 é par e então como no item 1, a série geométrica tem soma d se $d|k$ e vale 0 caso contrário. Se $t_{(2)} = n_{(2)}$ então nt/d^2 é ímpar. Como a é par, temos

$$\sum_{m=0}^{d-1} (-1)^{amnt/d^2} e^{-i2\pi akm/d} = \sum_{m=0}^{d-1} (e^{-i2\pi ak/d})^m$$

Como antes, a série geométrica tem soma d se $d|k$ e vale 0 caso contrário. Como no item 1, $P(X_t = k) = d/n$ quando tal valor não se anula. Finalmente, seja na ímpar. Aplicando (2.8), e o lema 2.6.1, parte 2, temos

$$P(X_t = k) = \frac{1}{n^2} \left| \sum_{m=0}^{d-1} e^{-i2\pi akm/d} \right|^2 \left| \sum_{j=0}^{n/d-1} e^{-i\pi a(t(j+1/2)+k)^2/nt} \right|^2$$

A soma geométrica tem soma d se $d|k$ e vale zero caso contrário. Quando $P(X_t = k)$ não se anula, temos

$$P(X_t = k) = \frac{d^2}{n^2} \left| \sum_{j=0}^{n'-1} e^{-i\pi a(t'(j+1/2)+k/d)^2/n't'} \right|^2$$

onde $n' = n/d$ e $t' = t/d$. Novamente, temos $(n', t') = 1$. Aplicando o lema 2.6.2, partes 3 e 4, concluímos que

$$P(X_t = k) = \frac{d^2}{n^2} n' = \frac{d}{n}.$$

□

Este teorema nos fornece o fato interessante de que $P(X_t = k)$ independe de a . Alguns corolários seguem a partir deste teorema.

Corolário 2.4.2 *Temos os seguintes casos:*

1. *Suponha n par. Então $P(X_t = k) = P(X_s = k)$ se e somente se $(n, t) = (n, s)$ e $t_{(2)}, s_{(2)} \neq n_{(2)}$ ou $t_{(2)} = s_{(2)} = n_{(2)}$.*
2. *Suponha n ímpar. Então $P(X_t = k) = P(X_s = k)$ se e somente se $(n, t) = (n, s)$.*

Prova Seja $d_t = (n, t)$ e $d_s = (n, s)$.

1. É claro que a condição é suficiente. Para ver que é necessária, vemos que existe k tal que $P(X_t = k) \neq 0$. Então

$$\frac{d_t}{n} = P(X_t = k) = P(X_s = k) = \frac{d_s}{n}$$

Então $d_t = d_s$.

2. É claro que a condição é suficiente. Para ver que é necessária, suponha que $P(X_t = k) = P(X_s = k)$, $t_{(2)} \neq n_{(2)}$ e $d_t|k$. Então, como antes,

$$\frac{d_t}{n} = P(X_t = k) = P(X_s = k) = \frac{d_s}{n}$$

e então $d_t = d_s$. Se $s_{(2)} = n_{(2)}$, então como $P(X_s = k) \neq 0$, temos que $2k/d$ é ímpar. Mas daí $2(k/d_t)$ é ímpar, o que é uma contradição. Portanto, $s_{(2)} \neq n_{(2)}$, e portanto $t_{(2)}, s_{(2)} \neq n_{(2)}$. Suponha que $t_{(2)} = n_{(2)}$ e que $2k/d_t$ é ímpar. Então, como antes, $d_t = d_s$. Se $s_{(2)} \neq n_{(2)}$, então como $P(X_s = k) \neq 0$, temos que $d_s|k$. Mas daí $2(k/d_s)$ é ímpar, o que é uma contradição. Portanto, $t_{(2)} = s_{(2)} = n_{(2)}$.

□

Dizemos que z é o **período de** $x(t) \in \mathbb{R}^n$, $t \geq 0$, se z é o menor inteiro positivo tal que $x(t+z) = x(t)$ para todo $t \geq 0$. Vamos considerar assim $x(t) = (P(X_t = s_0), P(X_t = s_1), \dots, P(X_t = s_{n-1}))$ e observar que estamos assumindo neste momento que $x(0) = (1, 0, 0, \dots, 0)$. Então z é o período de $x(t)$ se $P(X_{t+z} = k) = P(X_t = k)$, para todo t e para todo k .

Corolário 2.4.3 *Temos os seguintes casos:*

1. Se n é ímpar o período de $x(t)$ (acima definido) é n .
2. Se n é par então o período de $x(t)$ (acima definido) é $2n$.

Prova 1. Como $(n, t) = (n, n+t)$, pelo corolário 2.4.2, vale que $P(X_t = k) = P(X_{t+n} = k)$ para cada t . O menor inteiro positivo tal que $n = (n, 0) = (n, p)$ é $p = n$. Portanto, n é o menor inteiro positivo tal que $P(X_n = k) = P(X_0 = k)$. Então n é o período de probabilidade.

2. Primeiro, $(n, t) = (n, 2n+t)$. Suponha que $t_{(2)} = n_{(2)} = m$. Então $t = 2^m p$, $n = 2^m q$, onde p e q são ímpares. Portanto,

$$2n + t = 2^{m+1}q + 2^m p = 2^m(2q + p)$$

Como $2q+p$ é ímpar, $(2n+t)_{(2)} = m$. A seguir, suponha que $(2n+t)_{(2)} = m$. Então $2n+t = 2^m p$, $n = 2^m q$, onde p e q são ímpares. Portanto,

$$t = 2^m p = 2^{m+1}q = 2^m(p - 2q)$$

Como $p-2q$ é ímpar, $t_{(2)} = m$. Pelo corolário 2.4.2, $P(X_t = k) = P(X_{t+2n} = k)$ para todo t . Agora suponha que $p > 0$ e que $P(X_t = k) = P(X_0 = k)$. Pelo corolário 2.4.2, $n = (n, 0) = (n, p)$. Portanto, $n|p$ e então $p = rn$ para algum inteiro positivo r . Se $r = 1$ então $p_{(2)} = n_{(2)}$, mas $0_{(2)} \neq n_{(2)}$ o que contradiz o corolário 2.4.2. Logo, $r \neq 1$. Segue que $2n$ é o menor inteiro positivo que satisfaz $P(X_{2n} = k) = P(X_0 = k)$. Logo, $2n$ é o período de probabilidade.

□

Isso mostra que não precisamos calcular $P(X_t = k)$ para $t \geq 2n$, n par e para $t \geq n$, n ímpar. O próximo corolário nos diz que para n par também não é necessário calcular $P(X_t = k)$, $t \geq n$.

Corolário 2.4.4 *Temos:*

1. $P(X_0 = k) = \delta_{0,k}$
2. Se n é par então $P(X_n = k) = \delta_{n/2,k}$ e $P(X_t = k) = P(X_{2n-t} = k)$, $0 \leq t \leq 2n$.

Prova A prova de 1 é evidente. Provemos o item 2. Pelo teorema 2.4.1, $P(X_n = k) = 1$ se e somente se $2k/n$ é ímpar. Mas $2k/n$ ímpar é equivalente a $2k = nr$, onde r é ímpar, e então $k = (n/2)r$. Como $0 \leq k \leq n-1$, isso vale se e somente se $k = n/2$. Para a segunda parte, $(n, t) = (n, 2n-t)$. Suponha que $t_{(2)} = n_{(2)} = m$. Então $t = 2^m p$, $n = 2^m q$, onde p e q são ímpares. Logo,

$$2n - t = 2^m(2q - p)$$

Como $2q - p$ é ímpar, $(2n - t)_{(2)} = m$. A seguir, suponha que $(2n - t)_{(2)} = n_{(2)} = m$. Por um argumento semelhante, $t_{(2)} = m$. Pelo corolário 2.4.2, $P(X_t = k) = P(X_{2n-t} = k)$.

□

Este último corolário significa que o sistema inicia em s_0 e que a seqüência de probabilidades para $t = 0, 1, 2, \dots, n$ é a mesma para $2n, 2n - 1, 2n - 2, \dots, n$, nesta ordem.

Lembre que assumimos até agora nesta seção que o vetor de probabilidade inicial é $p = (1, 0, 0, \dots, 0)$ e não um p tal que $\mathcal{A}p = p$. Assim, com não poderia deixar de ser, a matriz de Dirichlet (com tal vetor inicial) não determina um Processo Estocástico Estacionário, mas sim uma seqüência periódica de probabilidades de estado.

Abaixo temos a tabela de $P_t(k) = P(X_t = k)$, para $n = 12$.

t/k	0	1	2	3	4	5	6	7	8	9	10	11
0	1	0	0	0	0	0	0	0	0	0	0	0
1	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12
2	1/6	0	1/6	0	1/6	0	1/6	0	1/6	0	1/6	0
3	1/4	0	0	1/4	0	0	1/4	0	0	1/4	0	0
4	0	0	1/3	0	0	0	1/3	0	0	0	1/3	0
5	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12
6	1/2	0	0	0	0	0	1/2	0	0	0	0	0
7	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12
8	1/3	0	0	0	1/3	0	0	0	1/3	0	0	0
9	1/4	0	0	1/4	0	0	1/4	0	0	1/4	0	0
10	1/6	0	1/6	0	1/6	0	1/6	0	1/6	0	1/6	0
11	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12
12	0	0	0	0	0	0	1	0	0	0	0	0
13	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12
14	1/6	0	1/6	0	1/6	0	1/6	0	1/6	0	1/6	0
15	1/4	0	0	1/4	0	0	1/4	0	0	1/4	0	0
16	1/3	0	0	0	1/3	0	0	0	1/3	0	0	0
17	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12
18	1/2	0	0	0	0	0	1/2	0	0	0	0	0
19	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12
20	0	0	1/3	0	0	0	1/3	0	0	0	1/3	0
21	1/4	0	0	1/4	0	0	1/4	0	0	1/4	0	0
22	1/6	0	1/6	0	1/6	0	1/6	0	1/6	0	1/6	0
23	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12
24	1	0	0	0	0	0	0	0	0	0	0	0

Pelo corolário acima, para t de 13 a 24, a tabela é a mesma, mas em ordem inversa. Obtemos uma seqüência de probabilidades de período $2n$. Podemos dizer que em um tempo t , existem certas transições impossíveis (i.e., $P_t(k) = 0$ para certos k) e as que podem ocorrer, tem todas a mesma probabilidade. Como temos um número finito de estados (ou alternativamente, um número finito de vetores unitários que diferem de um ângulo constante), o número de caminhos possíveis para uma partícula é finito, e são todos igualmente prováveis no início de um processo.

Outra tabela, para $n=4$ (ver [17]):

t/k	0	1	2	3
0	1	0	0	0
1	1/4	1/4	1/4	1/4
2	1/2	0	1/2	0
3	1/4	1/4	1/4	1/4
4	0	0	1	0
5	1/4	1/4	1/4	1/4
6	1/2	0	1/2	0
7	1/4	1/4	1/4	1/4
8	1	0	0	0

Vamos mostrar abaixo que existem certos $p \in \mathbb{C}^n$ tais que para certas matrizes de Dirichlet \mathcal{A} vale $\mathcal{A}p = p$.

O seguinte exemplo é inspirado pela moeda quântica que é descrita no exemplo 1 da seção 8.2.

Exemplo Matrizes Quânticas \mathcal{A} dois por dois não necessariamente do tipo Dirichlet. Sejam $u, v \in \mathbb{C}$ tais que $u + v = 1$ e $|u|^2 + |v|^2 = 1$. Uma condição necessária e suficiente para que essas propriedades sejam satisfeitas é que $u = 1 - v$, $0 \leq \text{Re}(u) \leq 1$ e $\text{Im}(u) = \pm(\text{Re}(u))^{1/2}(1 - \text{Re}(u))^{1/2}$. Então, vemos que o número real $0 \leq \text{Re}(u) \leq 1$ determina u e v a menos de conjugação. Por exemplo, se $\text{Re}(u) = 1/2$, temos $u = (1 \pm i)/2$, $v = (1 \mp i)/2$.

Vamos construir uma matriz quântica complexa \mathcal{A} a partir do seguinte. Sejam u_1, v_1 e u_2, v_2 dois pares de números complexos tais que

$$u_1 + v_1 = 1, \quad |u_1|^2 + |v_1|^2 = 1$$

$$u_2 + v_2 = 1, \quad |u_2|^2 + |v_2|^2 = 1$$

Pelo que vimos acima, tais números tem a forma

$$u_1 = x_1 \pm i\sqrt{x_1(1-x_1)}, \quad v_1 = 1 - x_1 \mp i\sqrt{x_1(1-x_1)}$$

$$u_2 = x_2 \pm i\sqrt{x_2(1-x_2)}, \quad v_2 = 1 - x_2 \mp i\sqrt{x_2(1-x_2)}$$

onde x_1 e $x_2 \in [0, 1]$.

Seja a matriz

$$\begin{aligned} \mathcal{A}_{\pm} &= \begin{pmatrix} u_1 & u_2 \\ 1 - u_1 & 1 - u_2 \end{pmatrix} \\ &= \begin{pmatrix} x_1 \pm i\sqrt{x_1(1-x_1)} & x_2 \pm i\sqrt{x_2(1-x_2)} \\ 1 - x_1 \mp i\sqrt{x_1(1-x_1)} & 1 - x_2 \mp i\sqrt{x_2(1-x_2)} \end{pmatrix} \end{aligned}$$

Observe que, com respeito aos sinais da parte imaginária, devemos fazer duas escolhas independentes, ou seja, uma para a primeira coluna e outra para a segunda. Portanto, se queremos determinar sob quais condições a matriz \mathcal{A} é unitária, devemos resolver o sistema $\mathcal{A}\mathcal{A}^* = I$ em quatro casos:

1. Caso $(+, +)$, ou seja,

$$\mathcal{A} = \begin{pmatrix} x_1 + i\sqrt{x_1(1-x_1)} & x_2 + i\sqrt{x_2(1-x_2)} \\ 1 - x_1 - i\sqrt{x_1(1-x_1)} & 1 - x_2 - i\sqrt{x_2(1-x_2)} \end{pmatrix}$$

Resolvendo uma das equações de $AA^* = I$ (que é um sistema de duas equações independentes), concluímos primeiro que $x_1 + x_2 = 1$, e portanto A tem a forma

$$\mathcal{A} = \begin{pmatrix} x_1 + i\sqrt{x_1(1-x_1)} & 1 - x_1 + i\sqrt{x_1(1-x_1)} \\ 1 - x_1 - i\sqrt{x_1(1-x_1)} & x_1 - i\sqrt{x_1(1-x_1)} \end{pmatrix}$$

E resolvendo a outra equação, obtemos que $x_1 = 0$ ou 1 , e portanto, neste caso,

$$\mathcal{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ ou } \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Para a primeira matriz (identidade), o autovalor é 1 , com autovetores $(1, 0)$ e $(0, 1)$. Para a segunda, um autovalor é 1 com autovetor associado $(1, 1)$ e o outro é -1 com autovetor $1/\sqrt{2}(-1, 1)$.

2. Caso $(-, -)$, ou seja,

$$\mathcal{A} = \begin{pmatrix} x_1 - i\sqrt{x_1(1-x_1)} & x_2 - i\sqrt{x_2(1-x_2)} \\ 1 - x_1 + i\sqrt{x_1(1-x_1)} & 1 - x_2 + i\sqrt{x_2(1-x_2)} \end{pmatrix}$$

Aqui obtemos as mesmas soluções do caso 1.

3. Caso $(+, -)$,

$$\mathcal{A} = \begin{pmatrix} x_1 + i\sqrt{x_1(1-x_1)} & x_2 - i\sqrt{x_2(1-x_2)} \\ 1 - x_1 - i\sqrt{x_1(1-x_1)} & 1 - x_2 + i\sqrt{x_2(1-x_2)} \end{pmatrix}$$

Resolvendo uma das equações de $AA^* = I$, concluímos que $x_1 + x_2 = 1$, e portanto A é simétrica da forma

$$A = \begin{pmatrix} x_1 + i\sqrt{x_1(1-x_1)} & 1 - x_1 - i\sqrt{x_1(1-x_1)} \\ 1 - x_1 - i\sqrt{x_1(1-x_1)} & x_1 + i\sqrt{x_1(1-x_1)} \end{pmatrix}$$

A outra equação é uma identidade e não impõe restrições sobre x_1 ou x_2 . Um autovalor é $2x_1 - 1 + 2i\sqrt{x_1(1-x_1)}$ com autovetor associado $1/\sqrt{2}(1, -1)$ e o outro autovalor é 1 com autovetor $1/\sqrt{2}(1, 1)$.

4. O caso $(-, +)$ é semelhante ao caso $(+, -)$:

$$A = \begin{pmatrix} x_1 - i\sqrt{x_1(1-x_1)} & x_2 + i\sqrt{x_2(1-x_2)} \\ 1 - x_1 + i\sqrt{x_1(1-x_1)} & 1 - x_2 - i\sqrt{x_2(1-x_2)} \end{pmatrix}$$

Resolvendo $AA^* = I$, temos que $x_1 + x_2 = 1$, e portanto

$$A = \begin{pmatrix} x_1 - i\sqrt{x_1(1-x_1)} & 1 - x_1 + i\sqrt{x_1(1-x_1)} \\ 1 - x_1 + i\sqrt{x_1(1-x_1)} & x_1 - i\sqrt{x_1(1-x_1)} \end{pmatrix}$$

Um autovalor é $2x_1 - 1 - 2i\sqrt{x_1(1-x_1)}$ com autovetor associado $1/\sqrt{2}(-1, 1)$ e o outro autovalor é 1 com autovetor $1/\sqrt{2}(1, 1)$.

Assim obtemos condições para que as matrizes sejam unitárias. Se $x \neq 0, 1$, temos que possuem entradas não-nulas e portanto A_{\pm} é a matriz de amplitude de transição de uma N -cadeia quântica. Como fizemos na seção de cadeias quânticas, vamos calcular as probabilidades de transição. Para isso, precisamos dos autovalores e autovetores de A_{\pm} . Pelo que vimos nos casos acima, podemos construir uma medida de Markov estacionária (vista na seção 2.2) escolhendo como vetor estacionário o vetor associado ao autovalor 1 (que sempre é um autovalor para a moeda quântica, ao menos no caso de ordem 2).

Observação Supondo que temos um vetor de distribuição inicial $\vec{X}_0 = (1, 0, \dots, 0)$, podemos usar a expressão para probabilidades

$$P(X_t = k) = P_t(k) = \left| \sum_j (\overline{\psi_j})_0 \lambda_j^t (\psi_j)_k \right|^2$$

obtido na seção 2.2 para cadeias quânticas. Fazendo $S = \{s_0, s_1\} = \{0, 1\}$ e escrevendo $P_t(k) = P(f_t^{-1}(s_k))$, temos

$$P_t(0) = \left| (\overline{\psi_0})_0 \lambda_0^t (\psi_0)_0 + (\overline{\psi_1})_0 \lambda_1^t (\psi_1)_0 \right|^2 = \left| \frac{1 + \lambda_1^t}{2} \right|^2$$

$$P_t(1) = \left| (\overline{\psi_0})_0 \lambda_0^t (\psi_0)_1 + (\overline{\psi_1})_0 \lambda_1^t (\psi_1)_1 \right|^2 = \left| \frac{1 - \lambda_1^t}{2} \right|^2$$

Para mostrar um exemplo, observe o caso particular em que $x = \text{Re}(u) = 1/2$. Então temos a matriz

$$\mathcal{A} = \frac{1}{2} \begin{pmatrix} 1 + i & 1 - i \\ 1 - i & 1 + i \end{pmatrix}$$

Os autovalores são

$$\lambda_0 = i, \quad \lambda_1 = 1$$

e os respectivos autovetores são

$$\psi_0 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \quad \psi_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

Como antes, $S = \{s_0, s_1\} = \{0, 1\}$. Temos

$$P_t(0) = \left| \overline{(\psi_0)_0} \lambda_0^t (\psi_0)_0 + \overline{(\psi_1)_0} \lambda_1^t (\psi_1)_0 \right|^2 = \left| \frac{1 + i^t}{2} \right|^2$$

$$P_t(1) = \left| \overline{(\psi_0)_1} \lambda_0^t (\psi_0)_1 + \overline{(\psi_1)_1} \lambda_1^t (\psi_1)_1 \right|^2 = \left| \frac{1 - i^t}{2} \right|^2$$

Então obtemos uma seqüência periódica de probabilidades:

t/k	0	1
0	1	0
1	1/2	1/2
2	0	1
3	1/2	1/2
4	1	0

Observação Considere o teorema clássico de convergência de cadeias de Markov reais:

Teorema 2.4.5 *Seja \mathcal{P} uma matriz estocástica (real) e regular e seja $S = \{1, 2, 3, \dots, d\}$. Então:*

1. \mathcal{P} tem um único vetor de probabilidade fixo π e os componentes de π são todos positivos, ou seja é único o vetor de probabilidade π tal que $\mathcal{P}\pi = \pi$ e $\sum_{s \in S} \pi_s = 1$ e $\pi_s > 0, \forall s \in S$.
2. Se p é qualquer vetor de probabilidade, então a seqüência de vetores $\mathcal{P}p, \mathcal{P}^2p, \mathcal{P}^3p, \dots$ converge para o ponto fixo π , isto é,

$$\lim_{n \rightarrow \infty} \mathcal{P}^n p = \pi.$$

Como conseqüência, as entradas das matrizes $\mathcal{P}, \mathcal{P}^2, \mathcal{P}^3, \dots$ obtidas a partir de \mathcal{P} convergem para as entradas correspondentes da matriz cujas colunas são iguais ao vetor fixo π . Ou seja,

$$\lim_{n \rightarrow \infty} \mathcal{P}^n = \begin{pmatrix} \pi_1 & \pi_1 & \pi_1 & \dots & \pi_1 \\ \pi_2 & \pi_2 & \pi_2 & \dots & \pi_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \pi_d & \pi_d & \pi_d & \dots & \pi_d \end{pmatrix}$$

Lembramos que uma matriz estocástica real \mathcal{P} é regular se existir alguma potência de \mathcal{P} tal que todas as suas entradas sejam positivas.

Note que este teorema não vale, em geral, para cadeias de Markov quânticas (considerando que a condição de entradas positivas é substituída pela condição de possuir entradas não nulas). Por exemplo, considere o item 2 do teorema. Um cálculo simples mostra que se

$$\mathcal{A} = \frac{1}{2} \begin{pmatrix} 1+i & 1-i \\ 1-i & 1+i \end{pmatrix}$$

então a sequência de matrizes $(\mathcal{A}^n)_{n \in \mathbb{N}}$ é uma sequência periódica, e portanto não ocorre convergência para a matriz cujas colunas são o vetor estacionário $(1/\sqrt{2}, 1/\sqrt{2})$ de \mathcal{A} .

2.5 Alguns cálculos

Exemplo 1: Matriz de Dirichlet Seja $M(n, a)$ a matriz de Dirichlet, dada por

$$M_{jk} = \frac{1}{\sqrt{n}} e^{i\pi a(j-k)^2/n},$$

onde $n \in \mathbb{N}$ e $a \in \mathbb{Z}$, $j, k \in \{0, \dots, n-1\}$. Por (1) sabemos que $M(n, a)$ é unitária se e somente se n e a são relativamente primos (i.e., $(n, a) = 1$) e que se o produto na for par então tal matriz pode ser tornada estocástica. Denotamos o vetor de probabilidade inicial por $\vec{X}_0 = (a_0, \dots, a_{n-1})$.

Escrevemos abaixo alguns exemplos de matrizes onde $(n, a) = 1$ e na par, com os seus respectivos autovalores, autovetores e processo estocástico real P associado. O processo P é dado pela matriz $|M|^2$,

$$(P)_{ij} = (|M|^2)_{ij} = |M_{ij}|^2$$

e o vetor de probabilidade inicial real associado é

$$p = |\vec{X}_0|^2 = (|a_0|^2, \dots, |a_{n-1}|^2)$$

Se na é par, a renormalização

$$M'(n, a) = \frac{\overline{S(n, a)}}{\sqrt{n}} M(n, a)$$

é estocástica, onde $S(n, a)$ é a soma de Dirichlet:

$$S(n, a) = \sum_{j=0}^{n-1} e^{i\pi a j^2/n}$$

Vamos analisar os casos $M(2, 1)$, $M(3, 2)$ e $M(4, 1)$ e tentar determinar se existem autovetores complexos associados ao autovalor 1. Temos:

1. Matriz $M(2, 1)$.

$$M(2, 1) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$$

No trabalho de J. Kempe [23], tal matriz é vista como a de uma moeda quântica balanceada (pág. 312), onde cara e coroa são tratados da mesma forma e o passeio não é influenciado pelo seu estado inicial [23].

Os autovalores e autovetores de $M(2, 1)$ são

$$\frac{\sqrt{2}}{2}(1 - i), \text{ com autovetor } (-1, 1)$$

$$\frac{\sqrt{2}}{2}(1 + i), \text{ com autovetor } (1, 1)$$

Ainda,

$$S(2, 1) = 1 + i$$

e portanto,

$$M'(2, 1) = \frac{1}{2} \begin{pmatrix} 1 - i & 1 + i \\ 1 + i & 1 - i \end{pmatrix}$$

Os autovalores e autovetores de $M'(2, 1)$ são

$$-i, \text{ com autovetor } (-1, 1)$$

$$1, \text{ com autovetor } (1, 1)$$

O processo estocástico real associado é

$$P(2, 1) = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Os autovalores e autovetores de $P(2, 1)$ são

$$0, \text{ com autovetor } (-1, 1)$$

$$1, \text{ com autovetor } (1, 1)$$

2. Matriz $M(3, 2)$.

$$M(3, 2) = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & -\frac{1}{2} + i\frac{\sqrt{3}}{2} & -\frac{1}{2} + i\frac{\sqrt{3}}{2} \\ -\frac{1}{2} + i\frac{\sqrt{3}}{2} & 1 & -\frac{1}{2} + i\frac{\sqrt{3}}{2} \\ -\frac{1}{2} + i\frac{\sqrt{3}}{2} & -\frac{1}{2} + i\frac{\sqrt{3}}{2} & 1 \end{pmatrix}$$

Os autovalores e autovetores de $M(3, 2)$ são

$$\frac{\sqrt{3} - i}{2}, \text{ com autovetores } (-1, 1, 0), (-1, 0, 1)$$

$$i, \text{ com autovetor } (1, 1, 1)$$

Ainda,

$$S(3, 2) = i\sqrt{3}$$

e portanto,

$$M'(3, 2) = -\frac{i\sqrt{3}}{2} \begin{pmatrix} 1 & -\frac{1}{2} + i\frac{\sqrt{3}}{2} & -\frac{1}{2} + i\frac{\sqrt{3}}{2} \\ -\frac{1}{2} + i\frac{\sqrt{3}}{2} & 1 & -\frac{1}{2} + i\frac{\sqrt{3}}{2} \\ -\frac{1}{2} + i\frac{\sqrt{3}}{2} & -\frac{1}{2} + i\frac{\sqrt{3}}{2} & 1 \end{pmatrix}$$

Os autovalores e autovetores de $M'(3, 2)$ são

$$-\frac{1}{2} - i\frac{\sqrt{3}}{2}, \text{ com autovetores } (-1, 1, 0), (-1, 0, 1)$$

$$1, \text{ com autovetor } (1, 1, 1)$$

O processo estocástico real associado é

$$P(3, 2) = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Os autovalores e autovetores de $P(3, 2)$ são

$$0, \text{ com autovetores } (-1, 1, 0), (-1, 0, 1)$$

$$1, \text{ com autovetor } (1, 1, 1)$$

3. Matriz $M(4, 1)$.

$$M(4, 1) = \frac{1}{2} \begin{pmatrix} 1 & \frac{\sqrt{2}}{2}(1+i) & -1 & \frac{\sqrt{2}}{2}(1+i) \\ \frac{\sqrt{2}}{2}(1+i) & 1 & \frac{\sqrt{2}}{2}(1+i) & -1 \\ -1 & \frac{\sqrt{2}}{2}(1+i) & 1 & \frac{\sqrt{2}}{2}(1+i) \\ \frac{\sqrt{2}}{2}(1+i) & -1 & \frac{\sqrt{2}}{2}(1+i) & 1 \end{pmatrix}$$

Os autovalores e autovetores de $M(4, 1)$ são

1, com autovetores $(-1, 0, 1, 0), (0, 1, 0, -1)$

$-\frac{\sqrt{2}}{2}(1+i)$, com autovetor $(-1, 1, -1, 1)$

$\frac{\sqrt{2}}{2}(1+i)$, com autovetor $(1, 1, 1, 1)$

Ainda,

$$S(4, 1) = \sqrt{2}(1+i)$$

e portanto,

$$M'(4, 1) = \frac{\sqrt{2}}{4}(1-i) \begin{pmatrix} 1 & \frac{\sqrt{2}}{2}(1+i) & -1 & \frac{\sqrt{2}}{2}(1+i) \\ \frac{\sqrt{2}}{2}(1+i) & 1 & \frac{\sqrt{2}}{2}(1+i) & -1 \\ -1 & \frac{\sqrt{2}}{2}(1+i) & 1 & \frac{\sqrt{2}}{2}(1+i) \\ \frac{\sqrt{2}}{2}(1+i) & -1 & \frac{\sqrt{2}}{2}(1+i) & 1 \end{pmatrix}$$

Os autovalores e autovetores de $M'(4, 1)$ são

$\frac{\sqrt{2}}{2}(1-i)$, com autovetores $(-1, 0, 1, 0), (0, 1, 0, -1)$

-1 , com autovetor $(-1, 1, -1, 1)$

1, com autovetor $(1, 1, 1, 1)$

O processo estocástico real associado é

$$P(4, 1) = \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Os autovalores e autovetores de $P(4, 1)$ são

0, com autovetores $(0, 1, 0, -1), (0, 0, 1, -1), (1, 0, 0, -1)$

1, com autovetor $(1, 1, 1, 1)$

Exemplo 2 Considere a matriz dada por

$$\begin{pmatrix} x + i\sqrt{x(1-x)} & 1 - x - i\sqrt{x(1-x)} \\ 1 - x - i\sqrt{x(1-x)} & x + i\sqrt{x(1-x)} \end{pmatrix}$$

para $x \in (0, 1)$. Os autovalores e autovetores são

$$2x - 1 + i(2i\sqrt{x(1-x)}), \text{ com autovetor } (-1, 1)$$

$$1, \text{ com autovetor } (1, 1)$$

O processo estocástico real associado é

$$\begin{pmatrix} x & 1-x \\ 1-x & x \end{pmatrix}$$

Os autovalores e autovetores são

$$2x - 1, \text{ com autovetor } (-1, 1)$$

$$1, \text{ com autovetor } (1, 1)$$

2.6 Apêndice

Neste apêndice provamos dois lemas técnicos usados na prova do teorema 2.4.1. Esta seção segue [15]. Como antes, supomos que $(n, a) = 1$ e $k = 0, \dots, n-1$.

Lema 2.6.1 *Seja $(n, t) = d$.*

1. *Então*

$$\begin{aligned} S_1 &:= \sum_{j=0}^{n-1} e^{-i\pi a(tj+k)^2/nt} \\ &= \sum_{m=0}^{d-1} (-1)^{amnt/d^2} e^{-i2\pi akm/d} \sum_{j=0}^{n/d-1} e^{-i\pi a(tj+k)^2/nt} \end{aligned}$$

2. *Se n é ímpar então*

$$\begin{aligned} S_2 &:= \sum_{j=0}^{n-1} e^{-i\pi a(t(j+1/2)+k)^2/nt} \\ &= \sum_{m=0}^{d-1} e^{-i2\pi akm/d} \sum_{j=0}^{n/d-1} e^{-i\pi a(t(j+1/2)+k)^2/nt} \end{aligned}$$

Prova 1. Divida a soma S_1 em d partes para obter

$$S_1 = \sum_{m=0}^{d-1} \sum_{j=mn/d}^{(m+1)n/d-1} e^{-i\pi a(tj+k)^2/nt}$$

Definindo $j = s + mn/d$, obtemos

$$S_1 = \sum_{m=0}^{d-1} \sum_{s=0}^{n/d-1} e^{-i\pi a(t(s+mn/d)+k)^2/nt}$$

Como $d|t$ e $d^2|nt$, temos

$$e^{-i\pi a(t(s+mn/d)+k)^2/nt} = e^{-i\pi a(ts+k)^2/nt} (-1)^{amnt/d^2} e^{-i2\pi akm/d}$$

e o resultado segue.

2. Novamente, separe a soma em d partes para obter

$$\begin{aligned} S_2 &= \sum_{m=0}^{d-1} \sum_{s=0}^{n/d-1} e^{-i\pi a(t(s+1/2)+k+mnt/d)^2/nt} \\ &= \sum_{m=0}^{d-1} e^{-i\pi am^2nt/d} \sum_{s=0}^{n/d-1} e^{-i\pi a(t(s+1/2)+k)^2/nt} e^{-i2\pi am(t(s+1/2)+k)/d} \end{aligned}$$

Como $d|t$, a última exponencial é igual a

$$e^{-i\pi amt/d} e^{-i2\pi akm/d} = (-1)^{amt/d} e^{-i2\pi akm/d}$$

Como na é ímpar, $(-1)^{amt/d} = (-1)^{mt/d}$ e

$$e^{-i\pi am^2nt/d} = (-1)^{mt/d}$$

e portanto o produto destes dois termos vale 1. O resultado segue. □

Lema 2.6.2 *Seja $(n, t) = 1$, $k = 0, \dots, n-1$.*

1. *Se nat é par então*

$$\sum_{j=0}^{n-1} e^{-i\pi a(tj+k)^2/nt} = e^{-i\pi ak^2(\alpha t-1)^2/nt} \overline{S(n, at)}$$

onde α é um inteiro definido por $\alpha t = 1 \pmod{n}$, $0 \leq \alpha \leq n-1$.

2. Se nat é ímpar então

$$\sum_{j=0}^{n-1} e^{-i\pi a(tj+k+1/2)^2/nt} = \frac{1}{2} e^{-i\pi a(2k+1)^2(\alpha t-1)^2/4nt} \overline{S(4n, at)}$$

onde α satisfaz $\alpha t = 1 \pmod{4n}$, $0 \leq \alpha \leq 4n - 1$.

3. Se na é ímpar e t é par então

$$\sum_{j=0}^{n-1} e^{-i\pi a(t(j+1/2)+k)^2/nt} = e^{-i\pi a((t/2+k)^2-\rho^2)/nt} e^{-i\pi a(\alpha t-1)^2\rho^2/nt} \overline{S(n, at)}$$

onde ρ é o resto de $1/2 + k \pmod{n}$.

4. Se nat é ímpar então

$$\sum_{j=0}^{n-1} e^{-i\pi a(t(j+1/2)+k)^2/nt} = \frac{1}{2} e^{-i\pi a(\alpha t-1)^2k^2/nt} e^{-i\pi a(\alpha t-1)k/n} \overline{S(4n, at)}$$

onde α satisfaz $\alpha t = 1 \pmod{n}$, $0 \leq \alpha \leq n - 1$.

Prova 1. Como $(n, t) = 1$, pelo algoritmo euclideano existem e são únicos os inteiros q e α tais que $\alpha t = 1 + qn$, $0 \leq \alpha \leq n - 1$. Então $k = (\alpha t - qn)k$ e daí temos

$$\begin{aligned} S &= \sum_{j=0}^{n-1} e^{-i\pi a(tj+k)^2/nt} = \sum_{j=0}^{n-1} e^{-i\pi a((j+\alpha k)t - qnk)^2/nt} \\ &= \sum_{j=0}^{n-1} e^{-i\pi at(j+\alpha k)^2/n} e^{-i\pi a(qn)^2k^2/nt} \end{aligned}$$

Mas $(qn)^2 = (\alpha t - 1)^2$ e daí

$$S = e^{-i\pi a(\alpha t-1)^2k^2/nt} \sum_{j=0}^{n-1} e^{-i\pi at(j+\alpha k)^2/n}$$

Seja m o inteiro que satisfaz

$$ak = -m \pmod{n}, \quad 0 \leq m \leq n - 1$$

Como nat é par,

$$e^{-i\pi at(j+\alpha k)^2/n} = e^{-i\pi at(j-m)^2/n}$$

Mas pelo lema 2.3.1, temos

$$\sum_{j=0}^{n-1} e^{i\pi at(j-m)^2/n} = \sum_{j=0}^{n-1} e^{i\pi atj^2/n} = S(n, at)$$

para $0 \leq m \leq 2n - 1$, o que prova o item 1.

2. A soma

$$S = \sum_{j=0}^{n-1} e^{-i\pi a(tj+k+1/2)^2/nt} = \sum_{j=0}^{n-1} e^{-i\pi a(2tj+2k+1)^2/4nt}$$

é uma soma parcial de

$$T = \sum_{j=0}^{4n-1} e^{-i\pi a(tj+2k+1)^2/4nt}$$

Vale que $4nat$ é par e $(4n, t) = 1$ pois $(n, t) = 1$ e t é ímpar. Além disso, $0 \leq 2k + 1 \leq 4n - 1$ e α satisfaz $\alpha t = 1 \pmod{4n}$, $0 \leq \alpha \leq 4n - 1$. Segue do item 1 que

$$T = e^{-i\pi a(2k+1)^2(\alpha t-1)^2} \overline{S(4n, at)}$$

Agora decompos T em uma parte par e outra ímpar $T = E + U$ onde

$$\begin{aligned} E &= \sum_{j=0}^{2n-1} e^{-i\pi a(tj+k+1/2)^2/nt} \\ &= \sum_{j=0}^{n-1} e^{-i\pi a(tj+k+1/2)^2/nt} + \sum_{j=n}^{2n-1} e^{-i\pi a(tj+k+1/2)^2/nt} \\ &= \sum_{j=0}^{n-1} e^{-i\pi a(tj+k+1/2)^2/nt} + \sum_{j=0}^{n-1} e^{-i\pi a(tj+k+1/2+nt)^2/nt} \end{aligned}$$

Como nat é ímpar, o termo geral da última soma se torna

$$e^{-i\pi a(tj+k+1/2)^2/nt}$$

Logo,

$$E = 2 \sum_{j=0}^{n-1} e^{-i\pi a(tj+k+1/2)^2/nt}$$

A parte ímpar é dada por

$$U = \sum_{j=0}^{2n-1} e^{-i\pi a((2j+1)t+2k+1)^2/4nt}$$

Como nat é ímpar, temos que $(2j+1)t + (2k+1)$ é par. Portanto,

$$\begin{aligned} \sum_{j=0}^{2n-1} e^{-i\pi a((2j+1)t+2k+1)^2/4nt} &= \sum_{j=0}^{n-1} e^{-i\pi a((2j+1)t+2k+1+2nt)^2/4nt} \\ &= \sum_{j=0}^{n-1} e^{-i\pi a((2j+1)t+2k+1)^2/4nt} e^{-i\pi ant} \\ &= - \sum_{j=0}^{n-1} e^{-i\pi a((2j+1)t+2k+1)^2/4nt} \end{aligned}$$

Portanto $U = 0$ e

$$\sum_{j=0}^{n-1} e^{-i\pi a(tj+k+1/2)^2/nt} = \frac{1}{2}T$$

O resultado segue.

3. Defina a soma

$$S = \sum_{j=0}^{n-1} e^{-i\pi a(t(j+1/2)+k)^2/nt} = \sum_{j=0}^{n-1} e^{-i\pi a(tj+1/2+k)^2/nt}$$

Como t é par, $t/2+k$ é um inteiro e nat é par então podemos usar o resultado do item 1. Seja ρ o resto (mod n) de $t/2+k$. Então $t/2+k = \rho + \mu$. Segue que

$$\begin{aligned} S &= e^{-i\pi a\mu^2/nt} e^{-i2\pi a\mu\rho/nt} \sum_{j=0}^{n-1} e^{-i\pi a(tj+\rho)^2/nt} \\ &= e^{-i\pi a(\mu^2+2\mu\rho)/nt} e^{-i\pi a\rho^2(at-1)^2/nt} \overline{S(n, at)} \\ &= e^{-i\pi a((1/2+k)^2-\rho^2)/nt} e^{-i\pi a\rho^2(at-1)^2/nt} \overline{S(4n, at)} \end{aligned}$$

4. Defina S como no item 3 e troque k por $(\alpha t - qn)k$ como no item 1 para obter

$$S = \sum_{j=0}^{n-1} e^{-i\pi a(t(j+1/2)+\alpha k)-qnk)^2/nt}$$

$$= e^{-i\pi a(\alpha t-1)^2 k^2/nt} e^{-i\pi a(\alpha t-1)k/n} \sum_{j=0}^{n-1} e^{-i\pi at(j+1/2)+\alpha k)^2/n}$$

Resta calcular a soma

$$T = \sum_{j=0}^{n-1} e^{-i\pi at(j+\alpha k+1/2)^2/n}$$

Seja m o inteiro definido por

$$\alpha k + 1 = -m + sn \quad , \quad 0 \leq m \leq n - 1$$

Então

$$j + \alpha k + \frac{1}{2} = j - m + sn - \frac{1}{2}$$

e

$$\begin{aligned} e^{-i\pi at(j+\alpha k+1/2)^2/n} &= e^{-i\pi at(j-(m+1/2)+sn)^2/n} \\ &= e^{-i\pi at(j-m-1/2)^2/n} e^{-i\pi ats^2 n} e^{i\pi ats} \end{aligned}$$

Como nat é ímpar temos $e^{i\pi ats} = (-1)^s$ e

$$e^{-i\pi ats^2 n} = (-1)^{s^2} = (-1)^s$$

Portanto,

$$T = \sum_{j=0}^{n-1} e^{-i\pi at(j-m-1/2)^2/n}$$

Aplicando o lema 2.3.1 temos

$$T = \frac{1}{2} \overline{S(4n, at)}$$

e o resultado segue. □

Capítulo 3

Integrais de Feynman

Algumas referências básicas para integrais de Feynman são [12],[20]. Uma introdução informal interessante é [13].

3.1 O formalismo das integrais de Feynman

A mecânica quântica tradicional teve bastante êxito tanto no que se refere a compreensão dos fenômenos quânticos como também nas previsões numéricas de experimentos. Ela foi substanciada por muitos experimentos feitos em laboratórios e por observações da natureza. Entretanto, a teoria não tem respostas para certos problemas, como por exemplo, a descrição de partículas elementares e suas interações. Isto pode ser devido ao fato de o conceito de amplitude de transição não estar na sua base axiomática (embora possa ser deduzido a partir dos postulados, este conceito não está na base da teoria). Esta opinião está de acordo com [18].

Uma tentativa de resolver este problema é o formalismo das integrais de caminhos, proposto por R. Feynman (ver referências em [18]). Embora este formalismo seja uma ferramenta útil, traz consigo alguns problemas devido a sua natureza matemática não rigorosa.

Para entender este formalismo, usaremos a formulação lagrangiana da mecânica clássica, que descreveremos brevemente. Suponha que temos um sistema mecânico com n graus de liberdade. Diremos que $C = \mathbb{R}^n$ é o espaço de configurações. Um ponto $q = (q_1, \dots, q_n) \in C$ descreve a configuração do sistema e é chamado **estado lagrangiano**. A dinâmica de um sistema é descrita por um **caminho**, ou **trajetória** $q(t) = (q_1(t), \dots, q_n(t))$, $-\infty < t < \infty$ em C . Na formulação hamiltoniana da mecânica clássica, um estado $s(t)$ é determinado para o tempo t desde que a condição inicial $s(t_0)$ seja dada. Como veremos, o estado lagrangiano $q(t)$, $t_a \leq t \leq t_b$ é

determinado desde que as condições de contorno $q(t_a)$, $q(t_b)$ sejam dadas. Então definimos as funções de velocidade $v_i = dq_i/dt$, $i = 1, \dots, n$ e fazemos $v(t) = (v_1(t), \dots, v_n(t))$. O **lagrangiano** do sistema é definido por

$$L(v, q) = \sum_i \left(\frac{M_i}{2} \right) v_i^2 - V(q).$$

(M_i é a massa associada a i -ésima coordenada). Logo, L é a energia cinética menos a energia potencial.

A **ação** sobre um caminho $q(t)$ entre os tempos t_a , t_b , $t_a < t_b$ é definida por

$$S[q(t)] = \int_{t_a}^{t_b} L[v(t), q(t)] dt.$$

Suponha que o sistema se encontra no estado lagrangiano q_a no tempo t_a e q_b no tempo t_b , isto é, $q_a = q(t_a)$, $q_b = q(t_b)$. Na formulação lagrangiana a trajetória que o sistema percorre, chamada **trajetória clássica**, é determinada pelo **princípio da mínima ação**. Este princípio nos diz que dentre todos os caminhos possíveis de q_a a q_b , a trajetória clássica $\hat{q}(t)$ é aquela onde $S[q(t)]$ é um extremo. Isto é, o valor de $S[\hat{q}(t)]$ não se modifica se o caminho $\hat{q}(t)$ é alterado ligeiramente.

Voltemos agora à discussão da amplitude de Feynman. Veremos que o formalismo destas amplitudes não leva em consideração apenas a trajetória clássica $\hat{q}(t)$, mas sim todas as trajetórias. Suponha que um elétron encontre-se na posição x_a no tempo t_a e então se move, sob a influência de alguma força. Queremos calcular a probabilidade $P(a, b)$ de que a partícula esteja em um ponto x_b no tempo t_b . O elétron irá se mover por um dos muitos caminhos possíveis, mas não seremos capazes de distinguir uma trajetória de outra sem interferir no sistema. Se fizermos uma medição, sabemos que seu movimento irá se modificar.

Na mecânica quântica, temos que **todos** os caminhos possíveis devem ser considerados. Cada caminho $x(t)$ contribui com uma amplitude $\phi[x(t)]$ para a amplitude total $K(a, b)$ e $P(a, b) = |K(a, b)|^2$. Em símbolos:

$$K(a, b) = \int \phi[x(t)].$$

onde $x(t)$ é todo caminho com $x(t_a) = x_a$ e $x(t_b) = x_b$. Com isso, temos que o axioma básico do formalismo das integrais de caminho é:

$$(P) \quad \phi[x(t)] = A e^{(i/\hbar)S[x(t)]}$$

onde

$$S[x(t)] = \int_{t_a}^{t_b} L[v(t), x(t)] dt \quad (3.1)$$

é a ação para o caminho $x(t)$, A é uma constante de normalização e L é o lagrangiano do sistema.

Resumindo, cada caminho possível contribui com uma determinada amplitude ϕ e soma destas amplitudes nos fornece uma amplitude total K . Esta amplitude K , que é a amplitude de Feynman que discretizaremos nas seções seguintes, nos fornece a probabilidade de ocorrência de um evento, fazendo $P(a, b) = |K(a, b)|^2$.

O axioma (P) nos diz que cada caminho contribui igualmente em módulo, embora as suas fases variem. Então não é claro se um caminho em particular é mais importante. Entretanto, na aproximação clássica, S é grande comparado com \hbar . Se um caminho arbitrário $x(t)$ é modificado por uma pequena quantidade $\delta x(t)$, embora δS seja pequeno na escala clássica, não é pequeno comparado com \hbar . Estas pequenas variações no caminho irão, em geral, variar bastante a fase, que irá oscilar rapidamente. Então, se os caminhos vizinhos de $x(t)$ possuem uma ação diferente, suas contribuições para $K(a, b)$ irão se cancelar, sem fornecer nenhuma contribuição. Mas para o caminho especial \hat{x} , onde S é um extremo, uma pequena mudança no caminho fornecerá praticamente nenhuma variação em S . Assim, caminhos na vizinhança de \hat{x} não fornecem fatores que se cancelam; desta forma, \hat{x} se destaca e as leis clássicas do movimento surgem a partir das leis quânticas. No nível atômico, quando S é comparável com \hbar , nenhum caminho se destaca. Neste caso, todos os caminhos (mais precisamente, caminhos contínuos) devem ser considerados ao se calcular $K(a, b)$.

Existem algumas dificuldades matemáticas ao se tentar construir uma formulação rigorosa para a amplitude total K . O problema principal é encontrar uma medida adequada no espaço de caminhos a fim de que a soma K possa ser formulada em termos de uma integral sobre este espaço (ver [18] para mais detalhes). Alguns casos especiais foram resolvidos, mas uma solução geral ainda não foi descoberta. Iremos seguir um argumento heurístico.

Seja $t_a = t_0 < t_1 < \dots < t_n = t_b$ uma partição de $[t_a, t_b]$ onde

$$t_{i+1} - t_i = \epsilon = \frac{(t_b - t_a)}{n}, \quad i = 0, \dots, n - 1.$$

Seja P_n o conjunto dos caminhos poligonais da forma $x(t)$ onde

$$x(t_a) = x_a, \quad x(t_b) = x_b, \quad x(t_i) = x_i, \quad i = 1, \dots, n - 1$$

e para $t_i < t < t_{i+1}$, $(t, x(t))$ é o segmento de reta de (t_i, x_i) para (t_{i+1}, x_{i+1}) . Tomamos uma constante de normalização A_n como sendo

$$A_n = \left(\frac{m}{2\pi i \hbar \epsilon} \right)^{n/2}$$

onde m é a massa da partícula e defina

$$K_n(a, b) = \sum \{ \phi[x(t)] : x(t) \in P_n \}.$$

Se escrevermos esta integral sobre os valores de x_i , $i = 1, \dots, n-1$, obtemos

$$K_n(a, b) = A_n \int \dots \int e^{(i/\hbar)S[x(t)]} dx_1, \dots, dx_{n-1}. \quad (3.2)$$

E daí definimos $K(a, b)$ como sendo

$$K(a, b) = \lim_{n \rightarrow \infty} K_n(a, b)$$

Não é garantido que este limite exista (e em muitos casos não existe). Quando o limite existe, não depende crucialmente da natureza poligonal dos caminhos. Usaremos a seguinte notação intuitiva:

$$K(a, b) = \int_a^b e^{(i/\hbar)S[x(t)]} D[x(t)]$$

e diremos que a equação acima é uma **integral de caminho**.

Apesar de a integral acima não possuir um sentido matemático rigoroso, ela possui importante significado físico. Além disso, ela pode ser usada para descobrir certas propriedades que $K(a, b)$ deve ter se uma formulação rigorosa for possível. Por exemplo, seja $t_c = t_i$ um certo tempo fixado para um $i \in \{1, \dots, n-1\}$ fixo. Escrevendo a ação para os intervalos de tempo $[t_a, t_c]$, $[t_c, t_b]$ como $S_a^c[x(t)]$, $S_c^b[x(t)]$, respectivamente, obtemos a partir da equação 3.1 que

$$S_b^a[x(t)] = S_a^c[x(t)] + S_c^b[x(t)].$$

Logo, a equação 3.2 nos fornece

$$K_n(a, b) = \int K_{i-1}(a, c) K_{n-i+1}(c, b) dx_c$$

Fazendo n arbitrariamente grande obtemos a seguinte fórmula

$$K(a, b) = \int K(a, c) K(c, b) dx_c. \quad (3.3)$$

Agora, vamos calcular a integral para um caso simples. Para uma partícula livre, o lagrangiano é $L = mv^2/2$. Para $x(t) \in P_n$ temos

$$S[x(t)] = \int_{t_a}^{t_b} L[v(t), x(t)] dt = \frac{m}{2\epsilon} \sum_{i=1}^n (x_i - x_{i-1})^2.$$

Então,

$$K_n(a, b) = A_n \int \cdots \int \exp\left[\frac{im}{2\hbar n \epsilon} \sum_{i=1}^n (x_i - x_{i-1})^2\right] dx_1, \dots, dx_{n-1}.$$

Aqui temos um produto de integrais Gaussianas que podem ser resolvidas individualmente para obter

$$K_n(a, b) = \left(\frac{m}{2\pi i \hbar n \epsilon}\right)^{1/2} \exp\left[\frac{im}{2\hbar n \epsilon} (x_b - x_a)^2\right].$$

Como $n\epsilon = t_b - t_a$, os K_n 's são iguais para todo n . Segue que

$$K(a, b) = \left[\frac{m}{2\pi i \hbar (t_b - t_a)}\right]^{1/2} \exp\left[\frac{im(x_b - x_a)^2}{2\hbar(t_b - t_a)}\right].$$

É possível também mostrar que a equação de Schrödinger pode ser obtida (novamente, sem o rigor matemático exigido normalmente) a partir do formalismo das integrais de caminho (ver [18]).

3.2 Distribuição de Probabilidade

Os conceitos mais importantes tratados neste trabalho estão baseados na **função da onda**, que representa o aspecto ondulatório de uma partícula, à medida que a posição e o tempo variam. A função da onda desempenha um papel central no cálculo de probabilidades em experimentos que envolvem partículas atômicas.

Para ilustrar o problema de tais cálculos, consideremos o seguinte experimento: de um lado, temos uma fonte emissora de partículas (elétrons, por exemplo). Cada partícula deve passar por um anteparo, que possui duas fendas, e atingir um detector no outro lado. O detector pode ser deslocado como quisermos, e assim podemos obter uma distribuição de probabilidade, realizando o experimento sucessivas vezes (figura 1). Se fecharmos a fenda 2, podemos calcular a probabilidade de o detector ser acionado por um elétron que passou pela fenda 1 (analogamente para a fenda 2). A soma destas duas distribuições nos fornece a probabilidade de o detector ser acionado

por um elétron que passou pela fenda 1 ou 2. Podemos concluir a partir daí o seguinte: se analisarmos classicamente o problema, concluiremos que a probabilidade de que uma partícula atinja o detector é

$$P(x) = P_1(x) + P_2(x)$$

onde P_1 e P_2 são as probabilidades de a partícula atingir o detector passando pela fenda 1 ou 2, respectivamente.

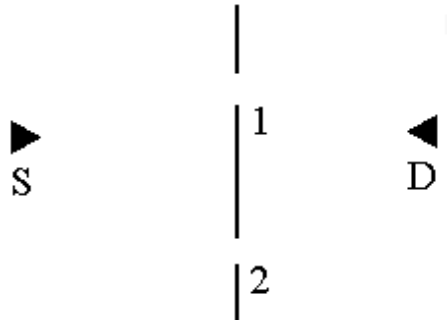


Figura 3.1: Versão simplificada do experimento de Young. S é a fonte de elétrons e D é um detector que se movimenta livremente na vertical.

A distribuição de probabilidade clássica seria uma curva gerada pela superposição das probabilidades via fenda 1 e 2 (figuras 3.2a-c). Entretanto, a experiência nos mostra que este cálculo não pode estar correto¹. Devemos ainda levar em conta a **interferência** entre caminhos alternativos. A distribuição real terá, na verdade, o aspecto aproximado de uma senóide com amplitude decrescente à medida que nos aproximamos dos extremos do anteparo (figura 3.2d). Desta forma, obtemos os padrões de interferência e difração, tal como no experimento de Young, revelando o caráter ondulatório de partículas como elétrons.

Um resultado importante da mecânica quântica, sobre o qual estão baseados as definições neste texto é que a probabilidade de se encontrar uma partícula em qualquer ponto é proporcional ao quadrado do valor absoluto da amplitude da onda de matéria nesse ponto. Este postulado baseia-se no fato de que o cálculo correto de probabilidades envolve uma função de onda (i.e., uma função complexa).

Sendo assim, temos que o problema anterior pode ser resolvido da seguinte maneira. Os padrões de interferência gerados no experimento sugerem que

¹O cálculo estará correto no caso em que colocarmos um detector para se determinar por qual fenda a partícula passou. Ou seja, ao introduzir tal detector estaremos destruindo o caráter ondulatório do experimento.

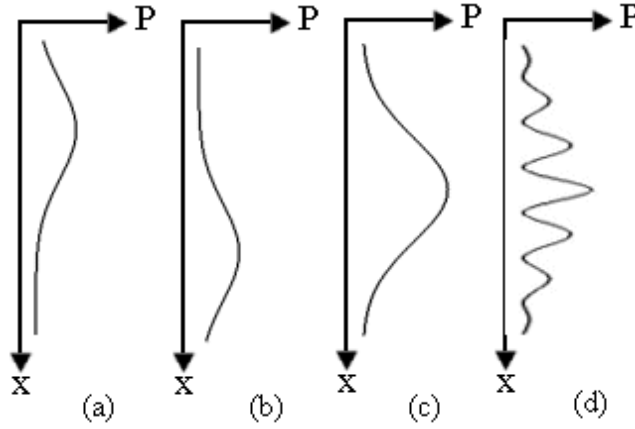


Figura 3.2: (a) $P_1(x)$. (b) $P_2(x)$. (c) $P_1(x)+P_2(x)$ (previsão clássica). (d) A distribuição de $P_1(x) + P_2(x)$ que realmente ocorre.

representemos a nossa função de distribuição, com tempo fixado, por uma função complexa. Então podemos postular que $P(x)$ é proporcional ao quadrado do valor absoluto de uma certa quantidade $\phi(x)$ e este valor é a amplitude de probabilidade, ou como chamamos nas outras seções, **amplitude quântica**. Além disso, temos o seguinte: $\phi(x)$ será a soma de duas contribuições, $\psi_1(x)$ e $\psi_2(x)$, as amplitudes de chegada via fenda 1 ou 2, respectivamente. Então, segue que

$$P_1(x) = |\psi_1(x)|^2, \quad P_2(x) = |\psi_2(x)|^2$$

e

$$P(x) = |\psi_1(x) + \psi_2(x)|^2 = P_1(x) + P_2(x) + 2\text{Re}\psi_1(x)\bar{\psi}_2(x).$$

Os dois primeiros termos do lado direito fornecem o resultado clássico; o último termo é o correspondente à interferência quântica, que contribui como uma soma ou subtração ao termo clássico.

Na física clássica, os modelos matemáticos nos fornecem informações sobre os fenômenos que observamos, enquanto que na mecânica quântica, os modelos não são observados sem que causemos alguma interferência. Uma abordagem dos fenômenos microscópicos consiste em descrever a evolução da função da onda, mas esta função nos permite apenas calcular a probabilidade de que certos eventos ocorram. Para relacioná-la com os experimentos, a função de onda deve ser interpretada de uma maneira apropriada.

3.3 Sobre a amplitude de Feynman

O formalismo de integrais de caminhos, sobre o qual as amplitudes de Feynman estão baseadas, nos permite imaginar algumas situações interessantes. Seja o seguinte experimento: temos um emissor de fótons, um detector, e abaixo um espelho (figura 3.3). Vamos supor válida a lei de incidência e reflexão da luz. Supondo que temos uma barreira entre o emissor e o detector (assim a luz emitida não pode ir diretamente para o detector) concluímos, após uma análise clássica, que todo fóton que atingiu o detector foi refletido pelo espelho, e além disso, deve ter sido no seu centro, pois o ângulo de incidência é igual ao ângulo de reflexão, e tanto o emissor como o detector estão a uma mesma distância da barreira. Isso é o que observamos quando fazemos experimentos com um feixe de luz. Entretanto, quando temos apenas uma partícula, o formalismo das integrais nos mostra um outro aspecto do fenômeno: devemos considerar todas as trajetórias “possíveis” para a luz; em outras palavras, não devemos supor que a luz anda apenas em linha reta, e muito menos que ela sabe qual o caminho mais curto! Mas para fins de simplificação, iremos supor que a luz percorre apenas linhas retas e a situação que iremos considerar é: um fóton, incidindo e refletindo em qualquer ângulo e em qualquer ponto do espelho.

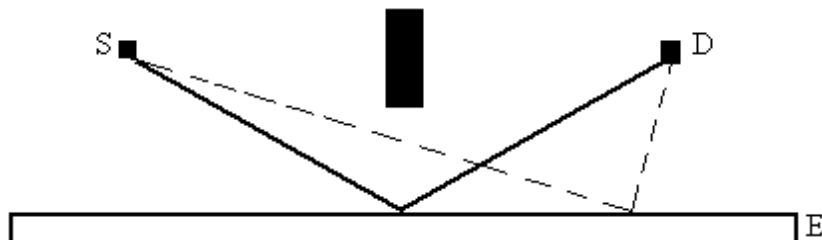


Figura 3.3: Dois caminhos para o fóton. Sabe-se que o caminho em que o fóton atinge o espelho no centro nos fornece o menor tempo, mas é errado dizer que os outros pontos são proibidos, ou que nunca irão ocorrer. Teoricamente, é possível que um fóton atinja outro ponto do espelho (embora a probabilidade seja pequena, e, em grande escala, descobrimos que as amplitudes de probabilidade destas alternativas se cancelam).

A princípio pode parecer estranho supor que um fóton atinge uma parte qualquer do espelho. Isso nos leva a uma pergunta: como poderia o fóton saber qual é o caminho mais curto e por que ele escolheria tal caminho? A figura 3.4 nos fornece um gráfico que indica os tempos necessários para um fóton sair da fonte e chegar até o detector.

Daremos apenas uma explicação intuitiva do que acontece. A reflexão da luz é um fenômeno que envolve uma certa porção do espelho. O fato impor-

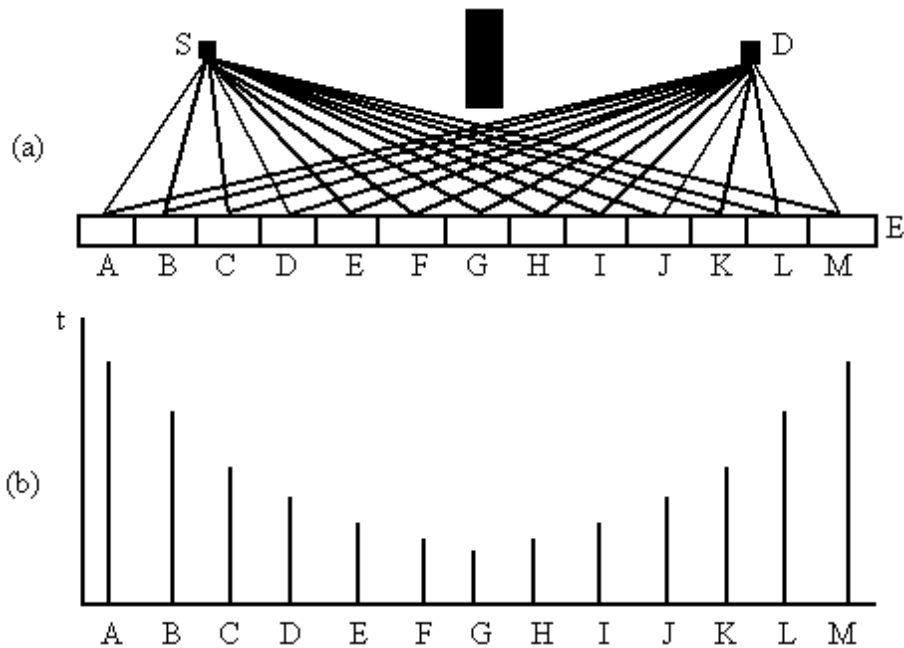


Figura 3.4: Vamos dividir o espelho em partições de mesmo tamanho. (a) Todos os caminhos possíveis. (b) Tempos associados aos caminhos. O importante aqui é notar que a diferença entre os tempos associados a dois pontos do centro, digamos F e G , é **menor** do que a diferença entre os tempos associados a dois pontos dos extremos, digamos A e B . É exatamente isso que determina o fato de que apenas observamos a luz gerando ângulos de incidência e reflexão iguais.

tante a ser observado é o seguinte: a diferença entre os tempos associados a pontos próximos do centro é menor do que a diferença entre os tempos associados a pontos dos extremos do espelho. Isso significa que a diferença de fase entre as exponenciais associadas à amplitude de Feynman (ver equações na seção 1) é pequena para pontos próximos ao centro, e logo temos amplitudes de probabilidade que contribuem para uma amplitude maior e logo, uma probabilidade maior; nos extremos, a variação de fase é maior, e existe uma contribuição menor de amplitudes, o que confere a esses pontos uma menor probabilidade. Logo, o centro é a região onde ocorrem pequenas variações de fase e onde é possível obter uma amplitude considerável. É por isso que podemos dizer, aproximadamente, que a luz percorre a trajetória em que o tempo é mínimo (também, não é difícil provar que no caminho onde o tempo é mínimo, o ângulo de incidência é igual ao ângulo de reflexão).

Vamos examinar um exemplo mais simples de fenômeno que pode ser descrito por uma cadeia quântica. Imagine uma partícula livre se movimen-

tando no espaço bidimensional. Queremos calcular a probabilidade de que esta partícula, saindo do ponto s_a , chegue ao ponto s_b . Como discutimos anteriormente, se estivéssemos considerando o fenômeno contínuo, teríamos de calcular a integral sobre todos os caminhos possíveis no espaço que tem como origem o ponto s_a e destino o ponto s_b (figura 3.5). Dizemos que a integral sobre cada caminho possível é uma amplitude de probabilidade e a soma destas amplitudes é a amplitude total. O quadrado do módulo desta amplitude total nos fornece a probabilidade de que a partícula saia de s_a e chegue a s_b . Isto é o que o formalismo das integrais de caminhos nos diz.

Lembre que fizemos $n = 12$ e na matriz de Dirichlet (de acordo com [2], o parâmetro a pode ser considerado a massa da partícula, e como já vimos, ela não influi na probabilidade). Como temos uma matriz com um parâmetro n finito, estamos descrevendo uma aproximação do fenômeno contínuo. Então podemos supor que as trajetórias possíveis para a partícula são poligonais. Mas vamos tornar as coisas ainda mais simples e supor que as frações de caminhos possíveis são as arestas de um reticulado.

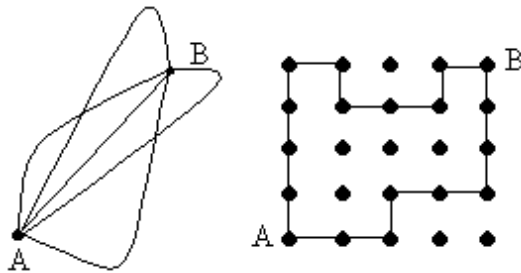


Figura 3.5: Quando consideramos sistemas discretos, o análogo de se calcular a integral sobre todos os caminhos de a a b é somar todas as amplitudes de transição de um estado a outro. Em particular, como estávamos interessados na *posição* de uma partícula, a matriz de transição associada a cadeia quântica que usamos foi a matriz de Dirichlet, que gera um análogo discreto para o movimento de uma partícula livre.

Uma observação importante. Lembre que em cadeias quânticas estamos sempre supondo que o conjunto S de estados possíveis do sistema é finito. Então, temos uma partícula em um plano movendo-se em um reticulado finito. Isso quer dizer a partícula livre que consideramos pode se mover apenas no reticulado, obviamente, e está de acordo com o cálculo de probabilidades que fizemos antes, pois lembre que tínhamos uma lista periódica de probabilidades (ou seja, isso deve-se também ao fato de que a partícula tem apenas um número finito de estados que podem ser assumidos. Se o conjunto de estados não fosse finito, não saberíamos, a priori, se existe uma

periodicidade nas transições do sistema).

Dizemos que f_t é a medição no tempo t e neste exemplo f representa a posição da partícula. Por exemplo, se o sistema está em um estado s_j no tempo t , realizando outra medição, obtemos $f(s_j) = s_k$, onde s_k é uma posição adjacente no reticulado. Inversamente, se estamos em um estado s_k no tempo t , temos que $f_t^{-1}(s_k)$ é o conjunto de estados que o sistema poderia estar no tempo $t - 1$, sabendo que passamos para s_k no tempo t (supomos que cada elemento deste conjunto deve necessariamente ser um vértice adjacente). É claro que neste exemplo, $s_j \in f_t^{-1}(s_k)$. Temos, também, o vetor de amplitude \hat{f}_t . Este vetor é unitário ou seja, $|\sum_j A[f_t^{-1}(s_j)]| = 1$, o que reflete a certeza de a partícula estar no reticulado.

Queremos calcular a probabilidade de que a partícula, saindo do estado s_a chegue ao estado s_b em t unidades de tempo. Note que desta forma, estamos considerando apenas as amplitudes de caminhos que levam t unidades de tempo. (Mesmo assim, conseguimos uma boa aproximação; usando o formalismo das integrais de caminhos, deveríamos considerar todos os caminhos, e não apenas os que duram t unidades de tempo). Logo, considerando caminhos de s_a a s_b em um tempo t , vemos que o formalismo das integrais (ou das somas de amplitudes sobre caminhos) se encaixa perfeitamente com a cadeia quântica que estamos considerando: dado um estado inicial $s_0 = s_a$, quando queremos determinar $P[f_t^{-1}(s_b)]$, procedemos da seguinte maneira: aplicando o estado inicial \hat{f}_0 á matriz de transição \hat{A} , podemos determinar como a amplitude varia com o tempo; somamos *todas* as amplitudes de caminhos que levam de s_a a s_b , calculamos o módulo e elevamos ao quadrado. Esse é o propósito da cadeia: verificar todas as probabilidades sobre o espaço de caminhos, exatamente como no formalismo das amplitudes de Feynman.

Capítulo 4

Operadores densidade e mecânica quântica

Computadores clássicos operam com estados construídos a partir de um número finito n de **bits**. Cada bit pode existir em um de dois estados, 0 ou 1. O estado do sistema é determinado ao se especificar os valores de cada um dos bits. Portanto, o conjunto de estados $\mathbb{B}^n = \{0, 1\}^n$ é finito e tem cardinalidade 2^n .

Um computador quântico trabalha com um conjunto finito de objetos chamados **q-bits**. Cada q-bit possui dois estados distintos, também denotados por 0 e 1. As 2^n combinações de estados para cada q-bit não consistem de todos os estados possíveis para o sistema, mas formam uma **base** para o espaço de estados. Denotaremos os estados base por $|x_1, \dots, x_n\rangle$, onde $x_j \in \mathbb{B}$ ou então por $|x\rangle$ onde $x \in \mathbb{B}^n$. Um estado arbitrário do sistema pode ser representado na forma

$$|\psi\rangle = \sum_{(x_1, \dots, x_n) \in \mathbb{B}^n} c_{x_1, \dots, x_n} |x_1, \dots, x_n\rangle, \text{ onde } \sum_{(x_1, \dots, x_n) \in \mathbb{B}^n} |c_{x_1, \dots, x_n}|^2 = 1$$

Ainda, se multiplicamos o vetor $|\psi\rangle = \sum_x c_x |x\rangle$ por um **fator de fase** $e^{i\varphi}$, $\varphi \in \mathbb{R}$, obtemos um estado fisicamente indistinguível e portanto o estado em um computador quântico é um vetor unitário definido a menos de um fator de fase.

Em um computador clássico, qualquer função $f : \mathbb{B}^n \rightarrow \mathbb{B}^n$ é permitida, ou seja composições e seqüências quaisquer de tais funções é o que entendemos como sendo computações clássicas; em um computador quântico as transformações permitidas são os operadores unitários, ou seja, operadores que preservam o comprimento $\sum_{x \in \mathbb{B}^n} |c_x|^2$ de cada vetor $|\psi\rangle = \sum_x c_x |x\rangle$.

Neste capítulo estamos interessados em estudar alguns fundamentos de mecânica quântica. Além disso, queremos entender a entropia de von Neu-

mann, o análogo quântico da entropia de Shannon da teoria clássica da informação. No próximo capítulo, calcularemos tal entropia para as cadeias de Markov quânticas que vimos no capítulo 2.

4.1 Notação de Dirac

Estamos interessados em obter algum análogo quântico para a entropia usada na teoria de informação clássica (entropia de Shannon). Veremos que algumas idéias da teoria da informação serão úteis para descrever os sistemas que nos interessam aqui. Parte desta seção é baseada em [32] e [41].

Começamos introduzindo a notação de Dirac, que é usual em mecânica quântica. Um vetor ψ de um espaço vetorial com produto interno será denotado por

$$|\psi\rangle$$

Dados dois vetores $|\psi\rangle$ e $|\varphi\rangle$, o produto interno é denotado por

$$(|\psi\rangle, |\varphi\rangle)$$

Escreveremos

$$\langle\psi|$$

para denotar o vetor dual do vetor $|\psi\rangle$. O vetor dual é um operador linear definido em um espaço vetorial com produto interno, toma valores em \mathbb{C} , e é definido por

$$\langle\psi|(|\varphi\rangle) := \langle\psi|\varphi\rangle := (|\psi\rangle, |\varphi\rangle),$$

e também denotaremos o produto interno entre dois vetores $|\psi\rangle$ e $|\varphi\rangle$ por $\langle\psi|\varphi\rangle$. Por convenção, se $|v\rangle$ é um vetor, definimos $|v\rangle^\dagger := \langle v|$. Suponha que A é um operador linear em um espaço de Hilbert V de dimensão finita. Então existe um único operador linear A^\dagger em V , dito o **adjunto** de A tal que para quaisquer $|v\rangle, |w\rangle \in V$,

$$(|v\rangle, A|w\rangle) = (A^\dagger|v\rangle, |w\rangle)$$

É fácil ver que $(A|v\rangle)^\dagger = \langle v|A^\dagger$, pois

$$(A|v\rangle)^\dagger|w\rangle = (A|v\rangle, |w\rangle) = (|v\rangle, A^\dagger|w\rangle) = \langle v|A^\dagger|w\rangle,$$

onde na última passagem acima, escrevemos o produto interno entre $|v\rangle$ e $|A^\dagger w\rangle$ como

$$\langle v|A^\dagger|w\rangle$$

que, equivalentemente, denota o produto interno entre $A|v\rangle$ e $|w\rangle$.

A elegância e praticidade da notação de Dirac fica evidente na seguinte definição. Seja $|v\rangle$ um vetor num espaço vetorial com produto interno V e $|w\rangle$ um vetor num espaço vetorial com produto interno W . Defina

$$|w\rangle\langle v| : V \rightarrow W$$

como sendo o operador linear cuja ação é definida por

$$(|w\rangle\langle v|)(|v'\rangle) := |w\rangle\langle v|v'\rangle = \langle v|v'\rangle|w\rangle$$

Tal operador é dito **produto exterior**. Note que a expressão $|w\rangle\langle v|v'\rangle$ pode ter as seguintes interpretações: (1) o resultado obtido ao se aplicar o operador $|w\rangle\langle v|$ no vetor $|v'\rangle$ ou (2) o resultado de se multiplicar $|w\rangle$ pelo número complexo $\langle v|v'\rangle$. As definições dadas acima são tais que essas duas interpretações coincidem. De fato, o que fizemos aqui foi simplesmente definir a primeira interpretação em termos da segunda.

Seja $|i\rangle$ uma base ortonormal qualquer para o espaço vetorial V . Um vetor $|v\rangle$ pode ser escrito como $|v\rangle = \sum_i v_i|i\rangle$, onde $v_i = \langle i|v\rangle$. Então temos

$$\left(\sum_i |i\rangle\langle i|\right)|v\rangle = \sum_i |i\rangle\langle i|v\rangle = \sum_i v_i|i\rangle = |v\rangle$$

Logo,

$$\sum_i |i\rangle\langle i| = I$$

onde I é o operador identidade. Esta é a **relação de completude para vetores ortonormais**. Esta relação nos permite mostrar que todo operador linear pode ser escrito como uma combinação linear de produtos exteriores. Suponha que $A : V \rightarrow W$ é um operador linear, $|v_i\rangle$ é uma base ortonormal para V e $|w_j\rangle$ é uma base ortonormal para W . Usando a relação de completude, obtemos

$$A = I_W A I_V = \sum_{ij} |w_j\rangle\langle w_j|A|v_i\rangle\langle v_i| = \sum_{ij} \langle w_j|A|v_i\rangle |w_j\rangle\langle v_i|$$

que é a representação em produto exterior de A . Note que a partir dessa expressão, vemos também que a matriz A_{ij} associada ao operador A é tal que $A_{ji} = \langle w_j|A|v_i\rangle$ (a matriz A_{ij} sendo tomada com respeito as bases $|v_i\rangle$ e $|w_j\rangle$).

Seja ψ um vetor unitário e $A : V \rightarrow V$ um operador qualquer. A partir de ψ , forme uma base ortonormal ordenada de modo que ψ seja o primeiro elemento. Então, pela expressão para A_{ij} que vimos acima,

$$tr(A|\psi\rangle\langle\psi|) = \sum_i \langle i|A|\psi\rangle\langle\psi|i\rangle = \langle\psi|A|\psi\rangle$$

que é uma expressão útil para se calcular o traço de operadores.

Suponha que W é um subespaço vetorial de dimensão k do espaço vetorial V de dimensão d . Usando o método de Gram-Schmidt, é possível construir uma base ortonormal $|1\rangle, \dots, |d\rangle$ para V tal que $|1\rangle, \dots, |k\rangle$ é uma base ortonormal para W . Definimos o **projetor** sobre o subespaço W como sendo

$$P := \sum_{i=1}^k |i\rangle\langle i|$$

4.2 Produto tensorial

Parte desta seção é baseada em [21]. Vamos considerar um sistema quântico que consiste de uma partícula de spin $1/2$ e outra de spin 1 . Se considerarmos apenas as propriedades de spin (ou seja, ignorando os graus de liberdade associados com as propriedades espaciais das partículas), os espaços de estados dos subsistemas são \mathbb{C}^2 e \mathbb{C}^3 , respectivamente, com bases típicas dadas pelos autoestados $\{|+\frac{1}{2}\rangle, |-\frac{1}{2}\rangle\}$ e $\{|+1\rangle, |0\rangle, |-1\rangle\}$. Uma suposição física natural é assumir que o sistema composto inclui estados que assumem cada um dos valores descritos acima. Então existem seis estados, que serão denotados por

$$|+\frac{1}{2}, +1\rangle, |+\frac{1}{2}, 0\rangle, |+\frac{1}{2}, -1\rangle, |-\frac{1}{2}, +1\rangle, |-\frac{1}{2}, 0\rangle, |-\frac{1}{2}, -1\rangle \quad (4.1)$$

Como estamos considerando uma teoria quântica, esperamos que combinações lineares dos estados acima sejam permitidos. Então parece natural supor que o estado mais geral pode ser escrito na forma

$$|\psi\rangle = c_1|+\frac{1}{2}, +1\rangle + c_2|+\frac{1}{2}, 0\rangle + c_3|+\frac{1}{2}, -1\rangle + c_4|-\frac{1}{2}, +1\rangle + c_5|-\frac{1}{2}, 0\rangle + c_6|-\frac{1}{2}, -1\rangle$$

e portanto os vetores na equação (4.1) formam uma base para o espaço de Hilbert do sistema composto. Como temos seis vetores, o espaço é isomorfo a \mathbb{C}^6 .

Para um sistema composto em geral, a operação matemática relevante toma vetores ψ_1 e ψ_2 nos espaços de Hilbert \mathcal{H}_1 e \mathcal{H}_2 , de dimensão m e n respectivamente, e os transforma em um vetor $\psi_1 \otimes \psi_2$ em um novo espaço de Hilbert $\mathcal{H}_1 \otimes \mathcal{H}_2$, chamado **produto tensorial** de \mathcal{H}_1 e \mathcal{H}_2 , que tem dimensão mn . O espaço de estados quântico do sistema composto é, portanto, o produto tensorial dos espaços de estados quânticos dos subsistemas constituintes. No exemplo acima, o vetor $|+\frac{1}{2}, +1\rangle$ denota o produto tensorial

$|+\frac{1}{2}\rangle \otimes |+1\rangle$. Em particular, temos que o produto tensorial $\mathbb{C}^2 \otimes \mathbb{C}^3$ é isomorfo a \mathbb{C}^6 .

Se $|v_i\rangle$ e $|w_j\rangle$ são bases ortonormais para \mathcal{H}_1 e \mathcal{H}_2 , respectivamente, então $|v_i\rangle \otimes |w_j\rangle$ é uma base para $\mathcal{H}_1 \otimes \mathcal{H}_2$. Usaremos também as notações

$$|v\rangle|w\rangle, |v, w\rangle \text{ e } |vw\rangle$$

para denotar o produto $|v\rangle \otimes |w\rangle$. Por exemplo, se \mathcal{H} é um espaço de Hilbert de dimensão 2 com vetores base $|0\rangle$ e $|1\rangle$, então $|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle \in \mathcal{H} \otimes \mathcal{H}$.

Agora definiremos formalmente o produto tensorial. Sejam V e W são espaços vetoriais, que iremos supor sempre sobre \mathbb{R} , e ainda que são de dimensão finita. A construção mais geral para módulos sobre anéis comutativos, que não precisaremos aqui, pode ser vista em [29] ou [5]. Denote por $\mathbb{R}[V \times W]$ o conjunto de elementos que são combinações lineares formais de elementos de $V \times W$ com coeficientes em \mathbb{R} , ou seja, expressões do tipo

$$\sum_{i=1}^n a_i \cdot (v_i, w_i), \quad a_i \in \mathbb{R}, v_i \in V, w_i \in W$$

Seja $\mathbb{R}(V, W)$ o subespaço gerado por todos os elementos de $\mathbb{R}[V \times W]$ da forma

$$\begin{aligned} (v_1 + v_2, w) - (v_1, w) - (v_2, w) \\ (v, w_1 + w_2) - (v, w_1) - (v, w_2) \\ (rv, w) - r(v, w) \\ (v, rw) - r(v, w) \end{aligned} \tag{4.2}$$

onde $v_i \in V$, $w_i \in W$ e $r \in \mathbb{R}$.

Definição O produto tensorial de espaços vetoriais V e W , denotado por $V \otimes W$ é o espaço

$$V \otimes W := \frac{\mathbb{R}[V \times W]}{\mathbb{R}(V, W)}$$

Seja $\pi : \mathbb{R}[V \times W] \rightarrow V \otimes W$ a projeção canônica e denote por $v \otimes w$ a imagem de $(v, w) \in \mathbb{R}[V \times W]$, ou seja,

$$v \otimes w := \pi(v, w)$$

Segue de (4.2) que $\pi : V \times W \rightarrow V \otimes W$ é bilinear. Além disso tal aplicação é universal no seguinte sentido:

Lema 4.2.1 *Sejam U, V e W espaços vetoriais, e seja $f : V \times W \rightarrow U$ uma aplicação bilinear. Então existe uma única aplicação linear $\bar{f} : V \otimes W \rightarrow U$ tal que $f = \bar{f} \circ \pi$.*

Prova Como o conjunto $V \times W$ é uma base para $R[V \times W]$, f se estende a uma aplicação linear $\hat{f} : R[V \times W] \rightarrow U$. A bilinearidade de f implica que $\hat{f}(R(V, W)) = 0$, e portanto \hat{f} induz uma aplicação \bar{f} do quociente $V \otimes W$ em U . Por construção, $f = \bar{f} \circ \pi$, \bar{f} é linear e como $\pi(V \times W)$ gera $V \otimes W$, \bar{f} é determinada de maneira única por f .

□

Segue do lema anterior que o produto tensorial é um funtor. De fato, se $\varphi : V \rightarrow V'$ e $\psi : W \rightarrow W'$ são aplicações lineares então a composição

$$V \times W \xrightarrow{\varphi \times \psi} V' \times W' \xrightarrow{\pi'} V' \otimes W'$$

é bilinear, portanto induz uma única aplicação

$$\varphi \otimes \psi : V \otimes W \rightarrow V' \otimes W'$$

A unicidade garante que $(\varphi' \otimes \psi') \circ (\varphi \otimes \psi) = \varphi' \circ \varphi \otimes \psi' \circ \psi$ quando $\varphi' : V' \rightarrow V''$, $\psi' : W' \rightarrow W''$.

Lema 4.2.2 *Sejam V e V' espaços vetoriais com bases B e B' , respectivamente. Então $V \otimes V'$ é um espaço vetorial com base $\{b \otimes b' | b \in B, b' \in B'\}$.*

Prova A bilinearidade de $\pi : V \times V' \rightarrow V \otimes V'$ mostra que o conjunto enunciado acima gera $V \otimes V'$. Suponha que

$$\sum r_{ij} b_i \otimes b'_j = 0 \tag{4.3}$$

Sejam $\varphi_0 : V \rightarrow R$ e $\varphi'_0 : V' \rightarrow R$ aplicações lineares com

$$\varphi_0(b_i) = 0 \text{ se } i \neq i_0, \quad \varphi_0(b_{i_0}) = 1$$

$$\varphi'_0(b'_i) = 0 \text{ se } i \neq j_0, \quad \varphi'_0(b'_{j_0}) = 1$$

onde (i_0, j_0) é um par de índices que aparecem em (4.3). A composição

$$V \otimes V' \xrightarrow{\varphi_0 \otimes \varphi'_0} R \otimes R \xrightarrow{\text{mult}} R$$

leva o lado esquerdo em (4.3) em $r_{i_0 j_0}$ que, portanto, deve ser zero.

□

Temos as seguintes relações:

1. $R \otimes V \cong V \cong V \otimes R$
2. $V_1 \otimes V_2 \cong V_2 \otimes V_1$
3. $V_1 \otimes (V_2 \otimes V_3) \cong (V_1 \otimes V_2) \otimes V_3$
4. $(V_1 \oplus V_2) \otimes V_3 \cong V_1 \otimes V_3 \oplus V_2 \otimes V_3$

Existe uma generalização natural da definição de produto tensorial para o caso de um número maior de espaços e um lema de propriedade universal correspondente:

Lema 4.2.3 *Para cada aplicação multilinear $f : V_1 \times \cdots \times V_k \rightarrow W$ existe uma única aplicação linear $\bar{f} : V_1 \otimes \cdots \otimes V_k \rightarrow W$ tal que $f = \bar{f} \circ \pi$.*

Outras propriedades do produto tensorial são as seguintes:

1. O produto tensorial é linear em cada entrada, no sentido de que para todo $\alpha, \beta \in \mathbb{C}$, $\psi, \phi \in \mathcal{H}_1$ e $\xi \in \mathcal{H}_2$,

$$(\alpha\psi + \beta\phi) \otimes \xi = (\alpha\psi) \otimes \xi + (\beta\phi) \otimes \xi$$

e se $\psi \in \mathcal{H}_1$ e $\phi, \xi \in \mathcal{H}_2$,

$$\psi \otimes (\alpha\phi + \beta\xi) = \psi \otimes (\alpha\phi) + \psi \otimes (\beta\xi)$$

e

$$\alpha(\psi \otimes \phi) = (\alpha\psi) \otimes \phi = \psi \otimes (\alpha\phi)$$

2. Existem vetores em $\mathcal{H}_1 \otimes \mathcal{H}_2$ que não podem ser escritos como um único produto $\psi \otimes \phi$, quaisquer que sejam $\psi \in \mathcal{H}_1$ e $\phi \in \mathcal{H}_2$. Um exemplo é

$$|\psi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}},$$

ou seja, não existem estados $|a\rangle$ e $|b\rangle$ tais que $|\psi\rangle = |a\rangle \otimes |b\rangle$. Quando tal fato ocorrer, diremos que ψ é um estado **emaranhado** (discutiremos emaranhamento (entanglement) na seção 4.5). Entretanto, todo vetor em $\mathcal{H}_1 \otimes \mathcal{H}_2$ pode ser escrito como uma soma de tais produtos.

Em particular, se $\{e_1, e_2, \dots, e_{N_1}\}$ e $\{f_1, f_2, \dots, f_{N_2}\}$ são bases para \mathcal{H}_1 e \mathcal{H}_2 , respectivamente, então uma base para $\mathcal{H}_1 \otimes \mathcal{H}_2$ é o conjunto de

vetores $e_i \otimes f_j$, $i = 1, 2, \dots, N_1$, $j = 1, 2, \dots, N_2$. Portanto, o vetor mais geral $\psi \in \mathcal{H}_1 \otimes \mathcal{H}_2$ tem a forma

$$\psi = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \psi_{ij} e_i \otimes f_j$$

onde $\psi_{ij} \in \mathbb{C}$. Em particular, isso mostra que a dimensão do espaço de Hilbert $\mathcal{H}_1 \otimes \mathcal{H}_2$ é o produto das dimensões de \mathcal{H}_1 e \mathcal{H}_2 .

3. O **produto interno** é definido em vetores produto por

$$\langle \psi_1 \otimes \psi_2, \phi_1 \otimes \phi_2 \rangle := \langle \psi_1, \phi_1 \rangle_{\mathcal{H}_1} \langle \psi_2, \phi_2 \rangle_{\mathcal{H}_2}$$

onde os produtos no lado direito são calculados nos espaços de Hilbert indicados. A expressão é estendida para somas de vetores se definimos

$$\begin{aligned} & \langle \psi_1 \otimes \psi_2, (\alpha \phi_1 \otimes \phi_2 + \beta \phi_3 \otimes \phi_4) \rangle \\ & := \alpha \langle \psi_1, \phi_1 \rangle_{\mathcal{H}_1} \langle \psi_2, \phi_2 \rangle_{\mathcal{H}_2} + \beta \langle \psi_1, \phi_3 \rangle_{\mathcal{H}_1} \langle \psi_2, \phi_4 \rangle_{\mathcal{H}_2} \end{aligned}$$

4. O **produto tensorial de operadores** pode ser definido da seguinte maneira. Sejam A_1 e A_2 operadores em \mathcal{H}_1 e \mathcal{H}_2 , respectivamente. O produto $A_1 \otimes A_2$ é definido primeiro em vetores produto

$$(A_1 \otimes A_2)\psi_1 \otimes \psi_2 := (A_1\psi_1) \otimes (A_2\psi_2)$$

e depois estendido para somas de produtos de maneira linear:

$$(A_1 \otimes A_2)\psi := \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \psi_{ij} (A_1 e_i) \otimes (A_2 f_j)$$

Da mesma forma que no caso de vetores, existem operadores em $\mathcal{H}_1 \otimes \mathcal{H}_2$ que não podem ser escritos na forma $A_1 \otimes A_2$. Entretanto, todos os operadores podem ser descritos por uma soma de tais operadores produto.

Enunciamos a definição do produto tensorial apenas para obter uma exposição completa. Não precisaremos lembrar da construção formal que fizemos anteriormente, apenas de suas propriedades operacionais. Em particular, será útil definir uma representação matricial, chamada **produto de Kronecker**. Sejam A uma matriz $m \times n$ e B uma matriz $p \times q$. Então temos a seguinte representação matricial:

$$A \otimes B := \begin{pmatrix} A_{11}B & A_{12}B & \cdots & A_{1n}B \\ A_{21}B & A_{22}B & \cdots & A_{2n}B \\ \vdots & \vdots & \vdots & \vdots \\ A_{m1}B & A_{m2}B & \cdots & A_{mn}B \end{pmatrix}$$

Ou seja, o termo $A_{ij}B$ é o elemento A_{ij} da matriz A multiplicado pela matriz B . Por exemplo, o produto tensorial dos vetores $(2, 3)$ e $(4, 5)$ é

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} \otimes \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 2 \times 4 \\ 2 \times 5 \\ 3 \times 4 \\ 3 \times 5 \end{bmatrix} = \begin{bmatrix} 8 \\ 10 \\ 12 \\ 15 \end{bmatrix}$$

4.3 Operador densidade

Suponha que um sistema quântico está em um estado $|\psi_i\rangle$ dentre vários estados possíveis, onde i é um índice, com respectivas probabilidades p_i . Dizemos que $\{p_i, |\psi_i\rangle\}$ é um **conjunto (ensemble) de estados**.

Definição O **operador densidade**, também chamado **matriz densidade** de um sistema é definido pela equação

$$\rho := \sum_i p_i |\psi_i\rangle \langle \psi_i|$$

O seguinte teorema fornece uma caracterização de operadores densidade:

Teorema 4.3.1 *Um operador ρ é o operador densidade associado a um ensemble $\{p_i, |\psi_i\rangle\}$ se e somente se satisfaz*

1. $tr(\rho) = 1$
2. ρ é um operador positivo

Prova Suponha que $\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i|$ é um operador densidade. Então

$$tr(\rho) = \sum_i p_i tr(|\psi_i\rangle \langle \psi_i|) = \sum_i p_i = 1$$

Suponha que $|\varphi\rangle$ é um vetor qualquer no espaço de estados. Então

$$\langle \varphi | \rho | \varphi \rangle = \sum_i p_i \langle \varphi | \psi_i \rangle \langle \psi_i | \varphi \rangle = \sum_i p_i |\langle \varphi | \psi_i \rangle|^2 \geq 0$$

Reciprocamente, suponha que ρ é um operador satisfazendo a condição do traço e de positividade dadas acima. Como ρ é positiva, então possui uma decomposição espectral

$$\rho = \sum_j \lambda_j |j\rangle \langle j|$$

onde os vetores $|j\rangle$ são ortogonais e os λ_j são autovalores reais não negativos de ρ . Pela condição do traço, vemos que $\sum_j \lambda_j = 1$. Portanto, um sistema no estado $|j\rangle$ com probabilidade λ_j terá ρ como o operador densidade correspondente. Ou seja, o ensemble $\{\lambda_j, |j\rangle\}$ é um ensemble de estados que induz o operador densidade ρ .

□

A utilidade do operador densidade é a de descrever subsistemas de um sistema quântico composto. Para isso, definimos o operador densidade reduzido. Suponha que temos dois sistemas físicos A e B cujos estados são descritos por um operador densidade ρ^{AB} . O **operador densidade reduzido** para o sistema A é definido por

$$\rho^A := tr_B(\rho^{AB}),$$

onde tr_B é o **traço parcial** sobre B , definido por

$$tr_B(|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2|) := |a_1\rangle\langle a_2| tr(|b_1\rangle\langle b_2|)$$

onde $|a_1\rangle$ e $|a_2\rangle$ são vetores no espaço de estados de A e $|b_1\rangle$ e $|b_2\rangle$ são vetores no espaço de estados de B . O operador traço aparecendo no lado direito é o operador traço usual para o sistema B , e logo, $tr(|b_1\rangle\langle b_2|) = \langle b_2|b_1\rangle$.

Definição Dizemos que sistema quântico encontra-se em um **estado puro** quando não há incerteza quanto ao conhecimento do estado do sistema. Ou seja, seu operador densidade é dado simplesmente por $\rho = |\psi\rangle\langle\psi|$. Costumamos dizer também que o operador ρ é um estado puro. Caso contrário, o estado é dito **misturado** (mixed).

O seguinte critério nos permite dizer se um estado é puro:

Lema 4.3.2 *Seja ρ um operador densidade. Então $tr(\rho^2) \leq 1$ e vale a igualdade se e somente se ρ é um estado puro.*

Prova Um cálculo simples mostra que $\rho^2 = \sum_{i=1}^n p_i^2 |i\rangle\langle i|$ se $\rho = \sum_{i=1}^n p_i |i\rangle\langle i|$. Daí $tr(\rho^2) = \sum p_i^2 \leq 1$, claramente. É igual a 1 se e somente se existe k tal que $p_k = 1$ (e portanto, $p_i = 0, i \neq k$), ou seja, se e somente se $\rho = |k\rangle\langle k|$

□

4.4 Postulados da mecânica quântica

Enunciamos a seguir os postulados da mecânica quântica. Mais detalhes podem ser vistos em [32]. Veremos que os postulados também possuem uma formulação interessante em termos de operadores densidade.

Postulado 1. Associado a cada sistema físico existe um espaço vetorial complexo com produto interno, chamado **espaço de estados** do sistema. O sistema é completamente descrito pelo seu **vetor de estado**, que é um vetor unitário no espaço de estados do sistema.

Em termos de operadores densidade, o postulado 1 pode ser escrito da seguinte maneira:

Postulado 1'. Associado a cada sistema físico isolado existe um espaço vetorial complexo com produto interno, dito espaço de estados do sistema. O sistema é completamente descrito pelo seu operador densidade. Se um sistema quântico está no estado ρ_i com probabilidade p_i , então o operador densidade para o sistema é $\sum_i p_i \rho_i$.

O postulado a seguir se refere à evolução temporal do sistema.

Postulado 2. A evolução de um sistema quântico fechado é dada por uma transformação unitária, ou seja, o estado $|\psi\rangle$ do sistema no tempo t_1 está relacionado com o estado $|\psi'\rangle$ do sistema no tempo t_2 por um operador unitário U que depende apenas dos tempos t_1 e t_2 :

$$|\psi'\rangle = U|\psi\rangle$$

Em termos de operadores densidade, temos:

Postulado 2'. A evolução de um sistema quântico fechado é dada por uma transformação unitária, ou seja, o estado ρ do sistema no tempo t_1 está relacionado com o estado ρ' no tempo t_2 por um operador unitário U que depende apenas dos tempos t_1 e t_2 :

$$\rho' = U\rho U^\dagger$$

Este postulado descreve como os estados de um sistema quântico fechado em dois tempos diferentes estão relacionados. Podemos postular, de maneira mais refinada, como ocorre a evolução do sistema em tempo contínuo.

Postulado 2''. A evolução no tempo de um estado de um sistema quântico fechado é descrita pela equação de Schrödinger

$$i\hbar \frac{d|\psi\rangle}{dt} = H|\psi\rangle$$

O fator \hbar é a constante de Planck. O termo H é um operador hermitiano fixado, que chamamos de **Hamiltoniano** do sistema fechado.

Vemos a relação entre os postulados 2 e 2'' quando escrevemos a solução da equação de Schrödinger

$$|\psi(t_2)\rangle = \exp\left[\frac{-iH(t_2 - t_1)}{\hbar}\right]|\psi(t_1)\rangle = U(t_1, t_2)|\psi(t_1)\rangle,$$

onde

$$U(t_1, t_2) := \exp\left[\frac{-iH(t_2 - t_1)}{\hbar}\right]$$

Vale que U definido desta forma é unitário e que todo operador unitário pode ser escrito na forma $U = \exp(iK)$ para algum operador hermitiano K .

Postulado 3. Medições quânticas são descritas por uma coleção $\{M_m\}$ de **operadores de medição**, que são operadores agindo no espaço de estados do sistema. O índice m se refere aos resultados das medições que podem ocorrer. Se o estado do sistema quântico é $|\psi\rangle$ antes que a medição seja realizada, então a probabilidade de que o resultado m ocorra é dado por

$$p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle$$

Em termos de operadores densidade, se o estado do sistema quântico é ρ , temos

$$p(m) = \text{tr}(M_m^\dagger M_m \rho)$$

O estado do sistema após a medição é

$$\frac{M_m|\psi\rangle}{\sqrt{\langle\psi|M_m^\dagger M_m|\psi\rangle}}$$

ou então

$$\frac{M_m \rho M_m^\dagger}{\text{tr}(M_m^\dagger M_m \rho)}$$

em termos de operadores densidade. Ainda, os operadores de medição satisfazem a equação de completude

$$\sum_m M_m^\dagger M_m = I.$$

Um sistema quântico é dito **composto** quando é formado por subsistemas quânticos.

Postulado 4. O espaço de estados de um sistema composto é dado pelo produto tensorial dos espaços de estados dos subsistemas componentes. Além disso, se temos sistemas numerados de 1 a n , e o sistema i encontra-se no estado $|\psi_i\rangle$, então o estado do sistema total é $|\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_n\rangle$ (em termos de operadores densidade, se o sistema i encontra-se no estado ρ_i , então o estado do sistema total é $\rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n$).

4.5 Estados emaranhados

Definição Considere um sistema composto e seja $|\psi\rangle$ um estado desse sistema. Sabemos da seção 4.2 que $|\psi\rangle$ pode ser escrito como uma soma de produtos tensoriais. Se $|\psi\rangle$ não pode ser escrito como um único produto dos seus estados componentes, então dizemos que $|\psi\rangle$ é um estado **emaranhado** (entangled).

Um exemplo de estado emaranhado é

$$|\psi\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}(|0\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle),$$

ou seja, não existem estados $|a\rangle$ e $|b\rangle$ tais que $|\psi\rangle = |a\rangle \otimes |b\rangle$.

Veremos agora uma decomposição que nos permite medir, em um certo sentido, a quantidade de emaranhamento entre dois sistemas.

Teorema 4.5.1 (*Decomposição de Schmidt*) *Suponha que $|\psi\rangle$ é um estado puro de um sistema composto AB . Então existem estados ortonormais $|i_A\rangle$ para o sistema A , e estados ortonormais $|i_B\rangle$ para o sistema B tais que*

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle |i_B\rangle,$$

onde os λ_i são números reais não negativos satisfazendo $\sum_i \lambda_i^2 = 1$, chamados coeficientes de Schmidt.

Observe a seguinte aplicação deste teorema. Seja $|\psi\rangle$ um estado puro de um sistema composto AB . Então pela decomposição de Schmidt, $\rho^A = \sum_i \lambda_i^2 |i_A\rangle \langle i_A|$ e $\rho^B = \sum_i \lambda_i^2 |i_B\rangle \langle i_B|$ e portanto os autovalores de ρ^A e ρ^B são iguais, ou seja, λ_i^2 para os dois operadores densidade. Muitas propriedades de sistemas quânticos são descritas completamente pelos autovalores do operador densidade reduzido e no caso de um estado puro de um sistema composto, tais propriedades continuarão valendo para os seus subsistemas. Por exemplo, considere o estado de dois q-bits

$$(|00\rangle + |01\rangle + |11\rangle)/\sqrt{3}$$

Este estado não possui nenhuma simetria evidente, mas no entanto, vale que $tr((\rho^A)^2) = tr((\rho^B)^2) = 7/9$. Isso é uma consequência simples da decomposição de Schmidt.

Para provar a decomposição de Schmidt, precisamos dos seguintes resultados. Lembramos que um operador linear é dito **normal** se $AA^* = A^*A$.

Teorema 4.5.2 (*Decomposição espectral*) *Todo operador normal M em um espaço vetorial V de dimensão finita é diagonal com respeito a uma base ortonormal para V . Reciprocamente, todo operador diagonalizável é normal.*

A decomposição espectral é um resultado conhecido de álgebra linear e sua prova pode ser vista, por exemplo, em [32].

Teorema 4.5.3 (*Decomposição Polar*) *Seja A um operador linear em um espaço vetorial V de dimensão finita. Então existe um operador unitário U e operadores positivos J e K tais que*

$$A = UJ = KU$$

onde os únicos operadores positivos J e K satisfazendo estas equações são $J := \sqrt{A^*A}$ e $K := \sqrt{AA^*}$. Além disso, se A é inversível então U é único. A decomposição $A = UJ$ é dita **decomposição polar à esquerda** de A e $A = KU$ é a **decomposição polar à direita** de A .

Prova Note que $J := \sqrt{A^*A}$ é um operador positivo, e então admite uma decomposição espectral $J = \sum_i \lambda_i |i\rangle\langle i|$, $\lambda_i \geq 0$. Defina $|\psi_i\rangle := A|i\rangle$. Pela definição, vemos que $\langle \psi_i | \psi_i \rangle = \lambda_i^2$. Considerando apenas os i tais que $\lambda_i \neq 0$, defina $|e_i\rangle := |\psi_i\rangle / \lambda_i$. Desta forma os $|e_i\rangle$ estão normalizados e além disso, tais vetores são ortogonais, pois se $i \neq j$, então $\langle e_i | e_j \rangle = \langle i | A^*A | j \rangle / \lambda_i \lambda_j = \langle i | J^2 | j \rangle / \lambda_i \lambda_j = 0$.

Acima consideramos os i tais que $\lambda_i \neq 0$. Agora use o procedimento de Gram-Schmidt para estender o conjunto ortonormal $|e_i\rangle$ para obter uma base ortonormal, que também chamaremos de $|e_i\rangle$. Defina um operador unitário $U := \sum_i |e_i\rangle\langle i|$. Quando $\lambda_i \neq 0$, temos $UJ|i\rangle = \lambda_i |e_i\rangle = |\psi_i\rangle = A|i\rangle$. Quando $\lambda_i = 0$ temos $UJ|i\rangle = 0 = |\psi_i\rangle$. Com isso provamos que a ação de A e de UJ é a mesma na base $|i\rangle$ e portanto, obtemos $A = UJ$.

O operador J é único, pois multiplicando $A = UJ$ à esquerda pela equação adjunta $A^* = JU^*$ fornece $J^2 = A^*A$, donde obtemos que $J = \sqrt{A^*A}$. Um cálculo simples mostra que se A é inversível então J também é e então U é unicamente determinado pela equação $U = AJ^{-1}$. A prova da decomposição polar à direita segue pois $A = UJ = UJU^*U = KU$, onde $K := UJU^*$ é um operador positivo. Como $AA^* = KUU^*K = K^2$, devemos ter que $K = \sqrt{AA^*}$, o que conclui a prova.

□

Corolário 4.5.4 (*Decomposição em valores singulares*) Seja A uma matriz quadrada. Então existem matrizes unitárias U e V e uma matriz diagonal D com entradas não negativas tais que

$$A = UDV$$

Os elementos na diagonal de D são ditos **valores singulares** de A .

Prova Pela decomposição polar, $A = SJ$ para S unitária e J positiva. Pelo teorema espectral, $J = TDT^*$, para T unitária e D diagonal com entradas não-negativas. Fazendo $U := ST$ e $V := T^*$, o resultado segue.

□

Prova do teorema 4.5.1 Faremos a prova no caso em que os sistemas A e B tem espaços de estado de mesma dimensão. O caso geral é análogo. Sejam $|j\rangle$ e $|k\rangle$ bases ortonormais para os sistemas A e B , respectivamente. Então $|\psi\rangle$ pode ser escrito como

$$|\psi\rangle = \sum_{jk} c_{jk} |j\rangle |k\rangle$$

para uma certa matriz complexa C com entradas c_{jk} . Pela decomposição em valores singulares, $C = UDV$, onde $D = (d_{ij})$ é uma matriz diagonal com entradas não negativas, e $U = (u_{ij})$ e $V = (v_{ij})$ são matrizes unitárias. Então

$$|\psi\rangle = \sum_{ijk} u_{ji} d_{ii} v_{ik} |j\rangle |k\rangle$$

Definindo $|i_A\rangle := \sum_j u_{ji} |j\rangle$, $|i_B\rangle := \sum_k v_{ik} |k\rangle$ e $\lambda_i := d_{ii}$, vemos que

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle |i_B\rangle$$

Vale que $|i_A\rangle$ é um conjunto ortonormal, pelo fato de que U é unitária e pela ortonormalidade de $|j\rangle$, e analogamente vale que $|i_B\rangle$ é um conjunto ortonormal.

□

As bases $|i_A\rangle$ e $|i_B\rangle$ são chamadas **bases de Schmidt** para A e B , respectivamente, e o número de coeficientes de Schmidt λ_i não nulos é dito **número de Schmidt** para o estado $|\psi\rangle$. O número de Schmidt é uma propriedade importante de um sistema quântico composto que, em um certo sentido, quantifica o emaranhamento entre os sistemas A e B . Para entender como, considere a seguinte propriedade: o número de Schmidt é preservado por transformações unitárias no sistema A ou no sistema B individualmente. Para ver porque, note que se $\sum_i \lambda_i |i_A\rangle |i_B\rangle$ é a decomposição de Schmidt para $|\psi\rangle$ então $\sum_i \lambda_i (U|i_A\rangle) |i_B\rangle$ é a decomposição de Schmidt para $U|\psi\rangle$, onde U é um operador unitário agindo apenas no sistema A . Com isso, podemos provar o seguinte:

Proposição 4.5.5 *Seja $|\psi\rangle$ um estado de um sistema composto AB . São equivalentes:*

1. $|\psi\rangle$ é um estado produto.
2. $|\psi\rangle$ possui número de Schmidt igual a 1.
3. ρ^A (e portanto ρ^B) são estados puros.

Prova A equivalência entre 1 e 2 é evidente pois o número de Schmidt de $|\psi\rangle$ ser igual a 1 equivale a dizer $|\psi\rangle = |k_A\rangle |k_B\rangle$ para algum $k \in \{1, \dots, n\}$. A equivalência entre 1 e 3, por exemplo, segue direto da definição de traço parcial, pois

$$|\psi\rangle = |i_A\rangle |i_B\rangle \Leftrightarrow \rho^{AB} = |i_A\rangle \langle i_A| \otimes |i_B\rangle \langle i_B| \Leftrightarrow \rho^A = |i_A\rangle \langle i_A|, \quad \rho^B = |i_B\rangle \langle i_B|$$

□

Outra técnica relacionada com o emaranhamento de estados é a seguinte. Seja ρ^A um estado de um sistema quântico A . É possível introduzir um outro sistema, denotado por R , e definir um estado puro $|AR\rangle$ para o sistema conjunto AR tal que $\rho^A = \text{tr}_R(|AR\rangle \langle AR|)$. Ou seja, o estado puro $|AR\rangle$ se reduz a ρ^A quando olhamos apenas para o sistema A . Este processo é chamado **purificação**, e nos permite associar estados puros a estados misturados de maneira natural. O sistema R é dito sistema de referência.

Dado um estado ρ^A qualquer, mostraremos como construir um sistema R e uma purificação $|AR\rangle$. Suponha que ρ^A possui uma decomposição ortonormal $\rho^A = \sum_i p_i |i^A\rangle \langle i^A|$. Para purificar ρ^A introduzimos um sistema R que possui o mesmo espaço de estados que o sistema A , com estados ortonormais $|i^R\rangle$ e definimos um estado puro para o sistema combinado

$$|AR\rangle := \sum_i \sqrt{p_i} |i^A\rangle |i^R\rangle.$$

Agora calculamos o operador densidade reduzido para o sistema A correspondente ao estado $|AR\rangle$:

$$\begin{aligned} \text{tr}_R(|AR\rangle\langle AR|) &= \sum_{ij} \sqrt{p_i p_j} |i^A\rangle\langle j^A| \text{tr}(|i^R\rangle\langle j^R|) = \sum_{ij} \sqrt{p_i p_j} |i^A\rangle\langle j^A| \delta_{ij} \\ &= \sum_i p_i |i^A\rangle\langle i^A| = \rho^A \end{aligned}$$

Portanto, $|AR\rangle$ é uma purificação de ρ^A .

Note a relação entre a decomposição de Schmidt e o método de purificação: o procedimento usado para se purificar um estado misturado do sistema A é definir um estado puro cuja base de Schmidt para o sistema A é simplesmente a base em que o estado misturado é diagonal, com os coeficientes de Schmidt sendo a raiz quadrada dos autovalores do operador densidade que está sendo purificado.

Capítulo 5

Entropia

5.1 Entropia de Shannon

Observação Neste texto, escrevemos $\log x$ para denotar o logaritmo de x na base 2. O logaritmo natural será denotado por $\ln x$.

A incerteza de uma coleção de estados possíveis a_i com uma distribuição de probabilidade $p(a_i)$ é dada pela sua entropia,

$$H(p) := - \sum_i p(a_i) \log p(a_i),$$

chamada **entropia de Shannon**. Estamos interessados em comparar duas distribuições de probabilidade distintas, e para este fim, introduzimos a noção de entropia relativa.

Definição Suponha que temos dois conjuntos de eventos discretos a_i e b_j com distribuições de probabilidade correspondentes $p(a_i)$ e $p(b_j)$. A **entropia relativa** (de Shannon) entre estas duas distribuições é dada por

$$H(p(a) \| p(b)) := \sum_i p(a_i) \log \frac{p(a_i)}{p(b_i)}$$

A entropia relativa é não negativa, $H(p(x) \| q(x)) \geq 0$, e vale a igualdade se e somente se $p(x) = q(x)$, para todo x , e se X é uma variável aleatória com d resultados possíveis, então $H(X) \leq d$, e vale a igualdade se e somente se X é uniformemente distribuída.

Temos que

$$H(p(x, y) \| p(x)p(y)) = H(p(x)) + H(p(y)) - H(p(x, y))$$

Um conceito importante derivado da entropia relativa está relacionado com a obtenção de informação. Quando um sistema aprende alguma informação a partir de outro, dizemos que seus estados estão correlacionados. A grandeza que mede a correlação entre esses estados é a informação mútua.

Definição A **informação mútua** (de Shannon) entre duas variáveis aleatórias A e B que possuem uma distribuição de probabilidade conjunta $p(a_i, b_j)$, e portanto distribuições de probabilidade marginais $p(a_i) = \sum_j p(a_i, b_j)$ e $p(b_j) = \sum_i p(a_i, b_j)$ é definida por

$$I_H(A : B) := H(p(a)) + H(p(b)) - H(p(a, b))$$

Podemos escrever I_H em termos da entropia relativa de Shannon. Neste sentido, ela representa uma distância entre a distribuição $p(a, b)$ e o produto das marginais $p(a) \times p(b)$. Vale que

$$I_H(A : B) = H(p(a, b) \| p(a) \times p(b))$$

Com respeito a entropia mútua de Shannon, vemos que ela descreve a correlação de dois observáveis, ou seja, tal grandeza é inerentemente clássica.

A **entropia conjunta** (de Shannon) de X e Y é definida de maneira natural por

$$H(X, Y) := - \sum_{x, y} p(x, y) \log p(x, y)$$

Vale também uma propriedade subaditiva, ou seja,

$$H(X, Y) \leq H(X) + H(Y)$$

com igualdade se e somente se X e Y são variáveis independentes.

A **entropia condicional** (de Shannon) é definida por

$$H(X|Y) := H(X, Y) - H(Y)$$

e a informação mútua de X e Y é

$$H(X : Y) := H(X) + H(Y) - H(X, Y)$$

Nestes moldes, uma **cadeia de Markov** é uma seqüência $\{X_i\}$ de variáveis aleatórias tais que X_{n+1} independe de X_1, X_2, \dots, X_{n-1} , dado X_n . Formalmente,

$$p(X_{n+1} = x_{n+1} | X_n = x_n, \dots, X_1 = x_1) = p(X_{n+1} = x_{n+1} | X_n = x_n)$$

Proposição 5.1.1 *Valem as seguintes propriedades da entropia de Shannon:*

1. $H(X, Y) = H(Y, X)$, $H(X : Y) = H(Y : X)$
2. $H(Y|X) \geq 0$, e portanto $H(X : Y) \leq H(Y)$ com igualdade se e somente se Y é função de X (isto é, se Y ocorre sempre que X ocorre).
3. $H(X) \leq H(X, Y)$, com igualdade se e somente se Y é função de X .
4. $H(X, Y) \leq H(X) + H(Y)$ com igualdade se e somente se X e Y são variáveis aleatórias independentes (subaditividade).
5. $H(Y|X) \leq H(Y)$ e portanto $H(X : Y) \geq 0$ com igualdade se e somente se X e Y são variáveis independentes.
6. $H(X_1, X_2, X_3) + H(X_2) \leq H(X_1, X_2) + H(X_2, X_3)$, com igualdade se e somente se $\{X_3, X_2, X_1\}$ formam uma cadeia de Markov (subaditividade forte).
7. $H(X|Y, Z) \leq H(X|Y)$ (condicionamento diminui a entropia).

Prova

1. Evidente.
2. Como $p(x, y) = p(x)p(y|x)$, temos

$$\begin{aligned} H(X, Y) &= - \sum_{xy} p(x, y) \log p(x)p(y|x) \\ &= - \sum_x p(x) \log p(x) - \sum_{xy} p(x, y) \log p(y|x) \\ &= H(X) - \sum_{xy} p(x, y) \log p(y|x) \end{aligned}$$

Logo, $H(Y|X) = - \sum_{xy} p(x, y) \log p(y|x)$. Mas $-\log p(y|x) \geq 0$ e portanto $H(Y|X) \geq 0$ com igualdade se e somente se Y é uma função de X .

3. Segue do item anterior.
4. Usaremos o fato de que $\log(x) \ln 2 \leq x - 1$ para todo x positivo, com igualdade se e somente se $x = 1$. Assim,

$$\sum_{x,y} p(x, y) \log \frac{p(x)p(y)}{p(x, y)} \leq \frac{1}{\ln 2} \sum_{x,y} p(x, y) \left(\frac{p(x)p(y)}{p(x, y)} - 1 \right)$$

$$= \frac{1}{\ln 2} \sum_{xy} p(x)p(y) - p(x, y) = \frac{1-1}{\ln 2} = 0$$

A desigualdade segue. Note que a igualdade vale se e somente se $p(x, y) = p(x)p(y)$ para todo x e y .

5. Segue do item anterior.
6. Como na prova de subaditividade, usaremos o fato de que $\log(x) \ln 2 \leq x-1$ para todo x positivo, com igualdade se e somente se $x = 1$. Assim,

$$\begin{aligned} & \sum_{x_1, x_2, x_3} p(x_1, x_2, x_3) \log \frac{p(x_1, x_2)p(x_2, x_3)}{p(x_2)p(x_1, x_2, x_3)} \leq \\ & \leq \frac{1}{\ln 2} \sum_{x_1, x_2, x_3} p(x_1, x_2, x_3) \left(\frac{p(x_1, x_2)p(x_2, x_3)}{p(x_2)p(x_1, x_2, x_3)} - 1 \right) \\ & = \frac{1}{\ln 2} \sum_{x_1, x_2, x_3} \left(\frac{p(x_1, x_2)p(x_2, x_3)}{p(x_2)} - p(x_1, x_2, x_3) \right) \\ & = \frac{1}{\ln 2} \left(\sum_{x_1, x_2} p(x_1, x_2) - \sum_{x_1, x_2, x_3} p(x_1, x_2, x_3) \right) = 0 \end{aligned}$$

e a desigualdade segue. E note que a igualdade vale se e somente se

$$\frac{p(x_1, x_2)p(x_2, x_3)}{p(x_2)p(x_1, x_2, x_3)} = 1 \Leftrightarrow \frac{p(x_1, x_2, x_3)}{p(x_2, x_3)} = \frac{p(x_1, x_2)}{p(x_2)} \Leftrightarrow p(x_1|x_2, x_3) = p(x_1|x_2)$$

ou seja, se e somente se $\{X_3, X_2, X_1\}$ é uma cadeia de Markov.

7. Basta observar que

$$H(X|Y, Z) \leq H(X|Y) \Leftrightarrow H(X, Y, Z) - H(Y, Z) \leq H(X, Y) - H(Y)$$

que é a subaditividade forte.

□

O seguinte teorema nos diz como uma cadeia de Markov perde informação sobre os seus valores anteriores, à medida que o tempo cresce.

Teorema 5.1.2 *Suponha que $\{X_i\}$ é uma cadeia de Markov. Então*

$$H(X_i : X_{i+2}) \leq H(X_i : X_{i+1}) \leq H(X_i)$$

Prova A segunda desigualdade é o item 2 da proposição 5.1.1. Pelas definições, temos que $H(X_1 : X_3) \leq H(X_1 : X_2)$ é equivalente a $H(X_1|X_2) \leq H(X_1|X_3)$. Se $\{X_i\}$ é uma cadeia de Markov, então a seqüência inversa também é, ou seja, para $k \geq 1$,

$$p(X_{n+1} = x_{n+1}|X_n = x_n, \dots, X_1 = x_1) = p(X_{n+1} = x_{n+1}|X_n = x_n) \Rightarrow$$

$$\Rightarrow p(X_n = x_n|X_{n+1} = x_{n+1}, \dots, X_{n+k} = x_{n+k}) = p(X_n = x_n|X_{n+1} = x_{n+1})$$

e portanto $H(X_1|X_2) = H(X_1|X_2, X_3)$. Então o problema é reduzido a mostrar que

$$H(X_1, X_2, X_3) - H(X_2, X_3) = H(X_1|X_2, X_3) \leq H(X_1|X_3) = H(X_1, X_3) - H(X_3)$$

Ora, mas essa desigualdade é apenas a subaditividade forte (item 6, proposição 5.1.1).

□

Como um corolário do teorema acima, temos que se $\{X_i\}$ é cadeia de Markov, então

$$H(X_{i+2} : X_i) \leq H(X_{i+2} : X_{i+1})$$

Intuitivamente, isso significa que qualquer informação de X_{i+2} compartilhada com X_i deve ser uma informação que X_{i+2} compartilha com X_{i+1} .

5.2 Entropia de von Neumann

Dada uma função $f : \mathbb{C} \rightarrow \mathbb{C}$, é possível definir uma função matricial da seguinte maneira. Seja $A = \sum_a a|a\rangle\langle a|$ a decomposição espectral para um operador normal A . Defina

$$f(A) := \sum_a f(a)|a\rangle\langle a|$$

Desta forma, vale que $f(A)$ está unicamente determinado. Assim podemos definir, por exemplo, o logaritmo de operadores positivo-definidos e a exponencial de operadores normais. Usaremos este tipo de construção para definir a entropia de von Neumann.

Vamos considerar os análogos quânticos da seção anterior. A grandeza que está associada com as correlações de todo o sistema é a informação mútua de von Neumann, que definiremos a seguir. Devido a sua natureza global, não é difícil imaginar que esse valor depende da matriz densidade.

Começamos definindo a entropia quântica, mais conhecida como entropia de von Neumann.

Definição A **entropia de von Neumann** de um sistema quântico descrito por uma matriz densidade ρ é definida por

$$S(\rho) := -tr(\rho \log \rho)$$

Se λ_i são os autovalores de ρ então a entropia de von Neumann pode ser escrita como

$$S(\rho) = - \sum_i \lambda_i \log \lambda_i \quad (5.1)$$

Para provar esta afirmação, note que se

$$\rho = \sum \lambda_i |i\rangle\langle i|,$$

então

$$\log \rho = \sum \log \lambda_i |i\rangle\langle i|$$

e

$$\rho \log \rho(v) = \rho \left(\sum \log \lambda_i |i\rangle\langle i|v \right) = \sum \log \lambda_i \langle i|v \rangle \rho(|i\rangle) = \sum \lambda_i \log \lambda_i \langle i|v \rangle$$

Portanto,

$$-tr(\rho \log \rho) = -tr \left(\sum \lambda_i \log \lambda_i |i\rangle\langle i| \right) = - \sum \lambda_i \log \lambda_i tr(|i\rangle\langle i|) = - \sum \lambda_i \log \lambda_i$$

A entropia de von Neumann pode ser considerado o análogo quântico da entropia de Shannon [34]. Para provar algumas de suas propriedades, iremos introduzir primeiro a entropia relativa associada.

Definição A **entropia relativa** (de von Neumann) entre dois estados σ e ρ é dada por

$$S(\rho||\sigma) := tr(\rho \log \rho) - tr(\rho \log \sigma)$$

A propriedade básica é a seguinte desigualdade:

Lema 5.2.1 (*Desigualdade de Klein*). *A entropia relativa quântica é não negativa:*

$$S(\rho||\sigma) \geq 0,$$

e vale a igualdade se e somente se $\rho = \sigma$.

Prova Sejam $\rho = \sum_i p_i |i\rangle\langle i|$ e $\sigma = \sum_j q_j |j\rangle\langle j|$ decomposições ortonormais para ρ e σ . Pela definição de entropia relativa, temos

$$S(\rho\|\sigma) = \sum_i p_i \log p_i - \sum_i \langle i|\rho \log \sigma|i\rangle$$

Nesta última equação, usamos as equações $\langle i|\rho = p_i\langle i|$ e

$$\langle i|\log \sigma|i\rangle = \langle i|\left(\sum_j \log(q_j)|j\rangle\langle j|\right)|i\rangle = \sum_j \log(q_j)P_{ij},$$

onde $P_{ij} = \langle i|j\rangle\langle j|i\rangle \geq 0$, e daí obtemos

$$S(\rho\|\sigma) = \sum_i p_i \left(\log p_i - \sum_j P_{ij} \log(q_j) \right).$$

Note que P_{ij} satisfaz $P_{ij} \geq 0$, $\sum_i P_{ij} = 1$ e $\sum_j P_{ij} = 1$ (a matriz com entradas P_{ij} é duplamente estocástica). Como $\log(\cdot)$ é uma função estritamente concava, segue que $\sum_j P_{ij} \log q_j \leq \log(\sum_j P_{ij} q_j)$ com igualdade se e somente se existe um valor de j para o qual $P_{ij} = 1$. Portanto,

$$S(\rho\|\sigma) \geq \sum_i p_i \log \frac{p_i}{\sum_j P_{ij} q_j}$$

com igualdade se e somente se existe um valor de j para o qual $P_{ij} = 1$, ou seja, se e somente se P_{ij} é uma matriz de permutação. Esta expressão tem a mesma forma da entropia relativa clássica. Portanto deduzimos que

$$S(\rho\|\sigma) \geq 0$$

com igualdade se e somente se $p_i = \sum_j P_{ij} q_j$, para todo i e P_{ij} é uma matriz de permutação. Para simplificar essa condição de igualdade, note que trocando os nomes dos autoestados de σ se necessário, podemos supor que P_{ij} é a matriz identidade e então σ e ρ são diagonais na mesma base. A condição $p_i = \sum_j P_{ij} q_j$ nos diz que os autovalores correspondentes a ρ e σ são idênticos e portanto a condição de igualdade se reduz a afirmar que $\rho = \sigma$.

□

Proposição 5.2.2

$$S(\rho^A\|\sigma^A) \leq S(\rho^{AB}\|\sigma^{AB})$$

A prova usa o seguinte lema:

Lema 5.2.3 *Existe um conjunto de matrizes unitárias U_j e uma distribuição de probabilidade p_j tal que para qualquer matriz A ,*

$$\sum_i p_i U_i A U_i^\dagger = \text{tr}(A) \frac{I}{d},$$

onde d é a dimensão do espaço de Hilbert onde o operador A está definido.

Juntamente com a concavidade estrita da entropia, este lema pode ser usado para provar que o estado I/d em um espaço de dimensão d é o único estado de máxima entropia.

Prova da proposição 5.2.2 Pelo lema, existem transformações unitárias U_j no espaço B e probabilidades p_j tais que

$$\rho^A \otimes \frac{I}{d} = \sum_j p_j U_j \rho^{AB} U_j^\dagger$$

para todo ρ^{AB} . Pela convexidade de entropia relativa, obtemos

$$S\left(\rho^A \otimes \frac{I}{d} \parallel \sigma^A \otimes \frac{I}{d}\right) \leq \sum_j p_j S\left(U_j \rho^{AB} U_j^\dagger \parallel U_j \sigma^{AB} U_j^\dagger\right)$$

Mas a entropia relativa é invariante por conjugação unitária, então

$$S\left(\rho^A \otimes \frac{I}{d} \parallel \sigma^A \otimes \frac{I}{d}\right) \leq \sum_j p_j S\left(\rho^{AB} \parallel \sigma^{AB}\right) = S(\rho^{AB} \parallel \sigma^{AB})$$

Esta desigualdade juntamente com o fato de que

$$S\left(\rho^A \otimes \frac{I}{d} \parallel \sigma^A \otimes \frac{I}{d}\right) = S(\rho^A \parallel \sigma^A)$$

fornece a monotonicidade da entropia relativa. □

Prosseguimos na análise da entropia de von Neumann.

Definição A entropia conjunta (de von Neumann) $S(A, B)$ para um sistema composto com duas componentes A e B é definido de maneira natural por

$$S(A, B) := -\text{tr}(\rho^{AB} \log \rho^{AB})$$

onde ρ^{AB} é a matriz densidade do sistema AB .

Proposição 5.2.4 *Suponha que A e B são sistemas quânticos distintos que possuem um estado conjunto ρ^{AB} . Então a entropia conjunta para os dois sistemas satisfaz as desigualdades*

$$|S(A) - S(B)| \leq S(A, B) \leq S(A) + S(B)$$

A primeira desigualdade é chamada desigualdade triangular, ou **desigualdade de Araki-Lieb**. É o análogo quântico da desigualdade $H(X, Y) \geq H(X)$ para a entropia de Shannon. A segunda desigualdade é dita **desigualdade subaditiva**, e vale a igualdade se e somente se os sistemas A e B não estão correlacionados, ou seja, se $\rho^{AB} = \rho^A \otimes \rho^B$.

Prova da proposição 5.2.4 A prova da desigualdade subaditiva é uma aplicação simples da desigualdade de Klein, que podemos escrever como

$$S(\rho) \leq -\text{tr}(\rho \log \sigma)$$

Fazendo $\rho = \rho^{AB}$ e $\sigma = \rho^A \otimes \rho^B$, note que

$$\begin{aligned} -\text{tr}(\rho \log \sigma) &= -\text{tr}(\rho^{AB}(\log \rho^A + \log \rho^B)) \\ &= -\text{tr}(\rho^A \log \rho^A) - \text{tr}(\rho^B \log \rho^B) = S(A) + S(B) \end{aligned}$$

A desigualdade de Klein nos fornece $S(A, B) \leq S(A) + S(B)$, como queríamos. A condição de igualdade $\sigma = \rho$ para a desigualdade de Klein nos fornece condições de igualdade $\rho^{AB} = \rho^A \otimes \rho^B$ para a subaditividade.

Para provar a desigualdade de Araki-Lieb, introduza um sistema auxiliar R que purifica os sistemas A e B (vimos purificação na seção 4.5). Aplicando a subaditividade, obtemos

$$S(R) + S(A) \geq S(A, R)$$

Como ABR encontra-se em um estado puro, $S(A, R) = S(B)$ e $S(R) = S(A, B)$. A desigualdade anterior pode ser reescrita como

$$S(A, B) \geq S(B) - S(A)$$

Pela simetria entre os sistemas A e B , obtemos também $S(A, B) \geq S(A) - S(B)$.

□

Suponha que os ρ_i são estados de um sistema A . Introduza o **sistema auxiliar** B cujo espaço de estados possui uma base ortonormal $|i\rangle$ correspondente ao índice i dos operadores densidade ρ_i . Defina um estado conjunto de AB por

$$\rho^{AB} := \sum_i p_i \rho_i \otimes |i\rangle\langle i|$$

Para provar a concavidade de S , usaremos a propriedade subaditiva. Note que para a matriz densidade ρ^{AB} , temos

$$S(A) = S\left(\sum_i p_i \rho_i\right)$$

$$S(B) = S\left(\sum_i p_i |i\rangle\langle i|\right) = H(p_i)$$

$$S(A, B) = H(p_i) + \sum_i p_i S(\rho_i)$$

Aplicando a desigualdade subaditiva $S(A, B) \leq S(A) + S(B)$, obtemos

$$\sum_i p_i S(\rho_i) \leq S\left(\sum_i p_i \rho_i\right)$$

ou seja, S é concava. O método de se introduzir um sistema auxiliar, como foi usado aqui, e na prova da desigualdade de Araki-Lieb, é frequentemente aplicado em teoria da informação quântica.

Teorema 5.2.5 *Suponha que P_i é um conjunto completo de projetores ortogonais e ρ um operador densidade. Então a entropia do estado $\rho' := \sum_i P_i \rho P_i$ do sistema após a medição é tal que*

$$S(\rho') \geq S(\rho)$$

e vale a igualdade se e somente se $\rho = \rho'$. Em outras palavras, medições projetivas aumentam a entropia.

Prova Aplique a desigualdade de Klein para ρ e ρ' .

$$0 \leq S(\rho' || \rho) = -S(\rho) - \text{tr}(\rho \log \rho')$$

O resultado segue se provarmos que $-\text{tr}(\rho \log \rho') = S(\rho')$. Para isso, aplicamos as relações $\sum_i P_i = I$, $P_i^2 = P_i$ e a propriedade cíclica do traço para obter

$$-\text{tr}(\rho \log \rho') = -\text{tr}\left(\sum_i P_i \rho \log \rho'\right) = -\text{tr}\left(\sum_i P_i \rho \log \rho' P_i\right)$$

Note que $\rho'P_i = P_i\rho P_i = P_i\rho'$, ou seja, P_i comuta com ρ' e portanto com $\log \rho'$ e daí

$$-tr(\rho \log \rho') = -tr\left(\sum_i P_i \rho P_i \log \rho'\right) = -tr(\rho' \log \rho') = S(\rho')$$

□

Teorema 5.2.6 *Suponha que $\rho = \sum_i p_i \rho_i$, onde p_i é um conjunto de probabilidades e ρ_i são operadores densidade. Então*

$$S(\rho) \leq \sum_i p_i S(\rho_i) + H(p_i)$$

com igualdade se e somente se os estados ρ_i tem suporte em subespaços ortogonais.

Prova Primeiro suponha que temos um estado puro $\rho_i = |\psi_i\rangle\langle\psi_i|$. Suponha que os ρ_i são estados de um sistema A e introduza um sistema auxiliar B com uma base ortonormal $|i\rangle$ correspondendo ao índice i nas probabilidades p_i . Defina

$$|AB\rangle := \sum_i \sqrt{p_i} |\psi_i\rangle \otimes |i\rangle$$

Como $|AB\rangle$ é um estado puro, temos

$$S(B) = S(A) = S\left(\sum_i p_i |\psi_i\rangle\langle\psi_i|\right) = S(\rho)$$

Suponha que realizamos uma medição projetiva no sistema B na base $|i\rangle$. Após a medição, o estado do sistema B é

$$\rho^{B'} = \sum_i p_i |i\rangle\langle i|$$

Mas pelo teorema anterior, medições projetivas nunca diminuem a entropia e portanto, $S(\rho) = S(B) \leq S(B') = H(p_i)$. Observando que $S(\rho_i) = 0$ para o caso de estado puro, provamos que

$$S(\rho) \leq H(p_i) + \sum_i p_i S(\rho_i)$$

onde os estados ρ_i são estados puros. Além disso vale a igualdade se e somente se $B = B'$, que ocorre se e somente se os estados $|\psi_i\rangle$ são ortogonais.

Considere agora o caso de estados misturados. Se $\rho_i = \sum_j p_j^i |e_j^i\rangle\langle e_j^i|$ são as decomposições ortonormais para os estados ρ_i , temos que $\rho = \sum_{ij} p_i p_j^i |e_j^i\rangle\langle e_j^i|$. Aplicando o resultado para estados puros e a observação de que $\sum_j p_j^i = 1$ para cada i , temos

$$\begin{aligned} S(\rho) &\leq - \sum_{ij} p_i p_j^i \log(p_i p_j^i) = - \sum_i p_i \log p_i - \sum_i p_i \sum_j p_j^i \log p_j^i \\ &= H(p_i) + \sum_i p_i S(\rho_i) \end{aligned}$$

A condição de igualdade para o estado misturado segue direto da condição do caso de estados puros.

□

Corolário 5.2.7 *Suponha que p_i são probabilidades, $|i\rangle$ são estados ortogonais para um sistema A e ρ_i é um conjunto de operadores densidade para um outro sistema B . Então*

$$S\left(\sum_i p_i |i\rangle\langle i| \otimes \rho_i\right) = H(p_i) + \sum_i p_i S(\rho_i)$$

e portanto

$$S(\rho \otimes \sigma) = S(\rho) + S(\sigma)$$

para quaisquer operadores densidade ρ e σ .

Definição A **informação mútua** (de von Neumann) entre dois subsistemas ρ_U e ρ_V de um sistema conjunto ρ_{UV} é definido por

$$I_S(\rho_U : \rho_V; \rho_{UV}) := S(\rho_U) + S(\rho_V) - S(\rho_{UV})$$

Algumas propriedades da entropia de Shannon não valem para a entropia de von Neumann e isso acarreta algumas conseqüências. Por exemplo, para variáveis aleatórias X e Y , vale a desigualdade $H(X) \leq H(X, Y)$. Esta desigualdade é intuitiva, ou seja, é natural imaginar que há menos incerteza quanto ao estado de X do que incerteza quanto ao estado do sistema conjunto formado por X e Y .

No entanto, esta intuição falha para estados quânticos. Considere um sistema AB de dois q-bits no estado emaranhado $(|00\rangle + |11\rangle)/\sqrt{2}$. Este é um estado puro e portanto, $S(A, B) = 0$. Por outro lado, o sistema A possui operador densidade $I/2$ e portanto possui entropia igual a 1.

5.3 Subaditividade forte

Esta seção é baseada em [32]. Mostraremos a desigualdade chamada **subaditividade forte** para a entropia de von Neumann.

Teorema 5.3.1 *Sejam A, B e C estados quânticos. Então*

$$S(A, B, C) + S(B) \leq S(A, B) + S(B, C)$$

Para provar este resultado, precisamos saber alguns fatos sobre concavidade de funções.

Definição Sejam A e B matrizes e $f(A, B)$ uma função real. Dizemos que f é **conjuntamente concava** em A e B se para todo $0 \leq \lambda \leq 1$,

$$f(\lambda A_1 + (1 - \lambda)A_2, \lambda B_1 + (1 - \lambda)B_2) \geq \lambda f(A_1, B_1) + (1 - \lambda)f(A_2, B_2)$$

Toda função conjuntamente concava é concava em cada uma das variáveis, mas a recíproca não vale.

Teorema 5.3.2 (Lieb) *Seja X uma matriz, e $0 \leq t \leq 1$. Então a função*

$$f(A, B) := \text{tr}(X^\dagger A^t X B^{1-t})$$

é conjuntamente concava sobre matrizes positivas A e B .

Proposição 5.3.3 *A entropia relativa $S(\rho||\sigma)$ é conjuntamente convexa em seus argumentos.*

Prova Para matrizes A e X quaisquer no mesmo espaço, defina

$$I_t(A, X) := \text{tr}(X^\dagger A^t X A^{1-t}) - \text{tr}(X^\dagger X A)$$

O primeiro termo nesta expressão é concavo em A , pelo teorema de Lieb, e o segundo termo é linear em A . Portanto, $I_t(A, X)$ é concavo em A . Defina

$$I(A, X) := \left. \frac{d}{dt} \right|_{t=0} I_t(A, X) = \text{tr}(X^\dagger (\log A) X A) - \text{tr}(X^\dagger X (\log A) A)$$

Observando que $I_0(A, X) = 0$ e usando a concavidade de $I_t(A, X)$ em A , temos

$$I(\lambda A_1 + (1 - \lambda)A_2, X) = \lim_{\delta \rightarrow 0} \frac{I_\delta(\lambda A_1 + (1 - \lambda)A_2, X)}{\delta} \geq$$

$$\geq \lambda \lim_{\delta \rightarrow 0} \frac{I_\delta(A_1, X)}{\delta} + (1 - \lambda) \lim_{\delta \rightarrow 0} \frac{I_\delta(A_2, X)}{\delta} = \lambda I(A_1, X) + (1 - \lambda) I(A_2, X)$$

Ou seja, $I(A, X)$ é uma função concava de A . Definindo as matrizes em bloco

$$A := \begin{pmatrix} \rho & 0 \\ 0 & \sigma \end{pmatrix} \quad X := \begin{pmatrix} 0 & 0 \\ I & 0 \end{pmatrix}$$

vemos que $I(A, X) = -S(\rho \parallel \sigma) = -(tr(\rho \log \rho) - tr(\rho \log \sigma))$. A convexidade conjunta de $S(\rho \parallel \sigma)$ segue da concavidade de $I(A, X)$ em A .

□

Definição A entropia condicional (de von Neumann) é

$$S(A|B) := S(A, B) - S(B)$$

Corolário 5.3.4 *Seja AB um sistema composto com componentes A e B . Então a entropia condicional $S(A|B)$ é concava no estado ρ^{AB} de AB .*

Prova Seja d a dimensão do sistema A . Note que

$$\begin{aligned} S\left(\rho^{AB} \parallel \frac{I}{d} \otimes \rho^B\right) &= -S(A, B) - tr\left(\rho_{AB} \log\left(\frac{I}{d} \otimes \rho^B\right)\right) \\ &= -S(A, B) - tr(\rho^B \log \rho^B) + \log d = -S(A|B) + \log d \end{aligned}$$

Portanto, $S(A|B) = \log d - S(\rho^{AB} \parallel I/d \otimes \rho^B)$. A concavidade de $S(A|B)$ segue da convexidade conjunta da entropia relativa.

□

Prova do teorema (5.3.1) Mostraremos que para quaisquer sistemas quânticos,

$$S(A) + S(B) \leq S(A, C) + S(B, C)$$

$$S(A, B, C) + S(B) \leq S(A, B) + S(B, C)$$

Estas desigualdades são equivalentes. Usaremos a concavidade da entropia condicional para provar a primeira, e a seguir mostramos que isso implica a validade da segunda. Defina a seguinte função de operadores densidade no sistema ABC :

$$T(\rho^{ABC}) := S(A) + S(B) - S(A, C) - S(B, C) = -S(C|A) - S(C|B)$$

Pela concavidade da entropia condicional, vemos que $T(\rho^{ABC})$ é uma função convexa de ρ^{ABC} . Seja $\rho^{ABC} = \sum_i p_i |i\rangle\langle i|$ uma decomposição espectral de

ρ^{ABC} . Pela convexidade de T , $T(\rho^{ABC}) \leq \sum_i p_i T(|i\rangle\langle i|)$. Mas $T(|i\rangle\langle i|) = 0$ pois para estados puros vale que $S(A, C) = S(B)$ e $S(B, C) = S(A)$. Segue que $T(\rho^{ABC}) \leq 0$ e portanto

$$S(A) + S(B) - S(A, C) - S(B, C) \leq 0$$

que é a primeira desigualdade que queríamos provar.

Para obter a segunda desigualdade, introduza um sistema auxiliar R purificando o sistema ABC . Usando a desigualdade recém provada, temos

$$S(R) + S(B) \leq S(R, C) + S(B, C) \quad (5.2)$$

Como $ABCR$ é um estado puro, $S(R) = S(ABC)$ e $S(R, C) = S(A, B)$ e portanto (5.2) se torna

$$S(A, B, C) + S(B) \leq S(A, B) + S(B, C)$$

o que conclui a prova. □

5.4 Cadeias quânticas: entropia de Shannon

Relembrando: seja Λ uma σ -álgebra de subconjuntos de Ω e seja $A: \Lambda \rightarrow \mathbb{C}$ uma medida complexa com $A(\Omega) = 1$. O número complexo $A(\Delta)$ é dito **amplitude quântica** do evento $\Delta \in \Lambda$. Além disso, estaremos supondo que A é σ -aditiva. A probabilidade P de que um evento $\Delta \in \Lambda$ ocorra é definida por $P(\Delta) := |A(\Delta)|^2$. Lembre que de fato tal P não é uma probabilidade sobre Ω .

Iremos supor que $\Omega = \{s_0, s_1, \dots, s_n\}^{\mathbb{N}}$, ou seja, cada elemento $\omega \in \Omega$ é o resultado de uma seqüência de medições. Se $(X_t)_{t \in \mathbb{N}}$ é tal seqüência, escreveremos

$$\overline{a_0 a_1 \cdots a_t} := \{\omega \in \Omega | X_0(\omega) = a_0, X_1(\omega) = a_1, \dots, X_t(\omega) = a_t\} \quad (5.3)$$

Estamos interessados em calcular algum tipo de entropia para as cadeias quânticas, de maneira análoga a que é feita ao se calcular a entropia associada a uma cadeia de Markov usual.

Seja

$$\vec{X}_t = (A(X_t^{-1}(s_0)), \dots, A(X_t^{-1}(s_{n-1}))) \in \mathbb{C}^n$$

e

$$\mathcal{A}_{jk} = A(X_2^{-1}(s_j) | X_1^{-1}(s_k))$$

A amplitude de (5.3) é

$$A(\overline{s_{j(0)}s_{j(1)} \cdots s_{j(t)}}) = \mathcal{A}_{j(t)j(t-1)} \cdots \mathcal{A}_{j(2)j(1)} \mathcal{A}_{j(1)j(0)} (\overline{X_0})_{j(0)}$$

Vamos calcular a entropia da probabilidade associada a uma cadeia de Markov quântica. Como no caso usual de entropia para o shift, podemos definir $\sigma : \Omega \rightarrow \Omega$, $\sigma(\omega_1\omega_2 \cdots) = (\omega_2\omega_3 \cdots)$, a partição $\alpha = \{\overline{s_0}, \overline{s_1}, \dots, \overline{s_n}\}$ de Ω e a entropia (de Shannon) desta partição, que no caso de cadeias quânticas tem como expressão, fazendo $\mu = |A|^2$,

$$\begin{aligned} H\left(\bigvee_{k=0}^{t-1} \sigma^{-k} \alpha\right) &= - \sum_{s_0 s_1 \dots s_{t-1}} \mu(\overline{s_0 s_1 \cdots s_{t-1}}) \log \mu(\overline{s_0 s_1 \cdots s_{t-1}}) \\ &= - \sum_{s_0 s_1 \dots s_{t-1}} |A_{s_{t-1} s_{t-2}} \cdots A_{s_2 s_1} A_{s_1 s_0} (\overline{X_0})_{s_0}|^2 \\ &\quad \times \log \left| A_{s_{t-1} s_{t-2}} \cdots A_{s_2 s_1} A_{s_1 s_0} (\overline{X_0})_{s_0} \right|^2 \\ &= - \sum_{s_0 s_1 \dots s_{t-1}} |A_{s_{t-1} s_{t-2}} \cdots A_{s_2 s_1} A_{s_1 s_0} (\overline{X_0})_{s_0}|^2 \\ &\quad \times \left(\log |A_{s_{t-1} s_{t-2}}|^2 + \cdots + \log |A_{s_2 s_1}|^2 + \log |A_{s_1 s_0}|^2 + \log |(\overline{X_0})_{s_0}|^2 \right) \\ &= - \sum_{i=1}^{n-1} |(\overline{X_0})_{s_i}|^2 \log |(\overline{X_0})_{s_i}|^2 - (t-1) \sum_{k,l=1}^{n-1} |A_{s_k s_l} (\overline{X_0})_{s_l}|^2 \log |A_{s_k s_l}|^2 \end{aligned}$$

Portanto,

$$h_\mu(\sigma) = h_{|A|^2}(\sigma) = \lim_{t \rightarrow \infty} \frac{1}{t} H\left(\bigvee_{k=0}^{t-1} \sigma^{-k} \alpha\right) = - \sum_{k,l=1}^{n-1} |A_{s_k s_l} (\overline{X_0})_{s_l}|^2 \log |A_{s_k s_l}|^2$$

Observação: uma possível definição de entropia complexa. Podemos nos perguntar: por que não considerar uma “entropia complexa”? Uma construção desse tipo é possível, mas veremos que a dificuldade para a obtenção de uma expressão simples reside no fato de que o argumento de um número complexo é linear apenas sobre certas condições. Para entender isso melhor, analisemos a seguinte construção. Seja $U_0 = \mathbb{C} - \{x \in \mathbb{R} : x \leq 0\}$.

Podemos definir em U_0 um ramo do argumento da seguinte maneira: seja $a : (-\pi, \pi) \rightarrow S^1 - \{-1\}$, $a(t) = e^{it}$. Vale que tal aplicação é uma bijeção sobre $S^1 - \{-1\}$. Seja $b : S^1 - \{-1\} \rightarrow (-\pi, \pi)$ a inversa de a . Vale que b é contínua e limitada. Defina $Arg : U_0 \rightarrow (-\pi, \pi)$ por $Arg(z) = b(z/|z|)$, que é um ramo do argumento em U_0 , o ramo principal do argumento. O ramo do logaritmo associado a Arg é

$$Log(z) := \log |z| + iArg(z) \quad , \quad z \in U_0,$$

que é o ramo principal do logaritmo.

Fazendo o cálculo de entropia análogo ao feito acima, temos:

$$\begin{aligned} H_{\mathbb{C}}\left(\bigvee_{k=0}^{t-1} \sigma^{-k} \alpha\right) &= - \sum_{s_0 s_1 \dots s_{t-1}} A(\overline{s_0 s_1 \dots s_{t-1}}) Log A(\overline{s_0 s_1 \dots s_{t-1}}) \\ &= - \sum_{s_0 s_1 \dots s_{t-1}} A_{s_{t-1} s_{t-2}} \dots A_{s_2 s_1} A_{s_1 s_0} (\overrightarrow{X_0})_{s_0} Log A_{s_{t-1} s_{t-2}} \dots A_{s_2 s_1} A_{s_1 s_0} (\overrightarrow{X_0})_{s_0} \\ &= - \sum_{s_0 s_1 \dots s_{t-1}} A_{s_{t-1} s_{t-2}} \dots A_{s_2 s_1} A_{s_1 s_0} (\overrightarrow{X_0})_{s_0} \\ &\times \left(\log |(A_{s_{t-1} s_{t-2}} \dots A_{s_2 s_1} A_{s_1 s_0} (\overrightarrow{X_0})_{s_0})| + iArg(A_{s_{t-1} s_{t-2}} \dots A_{s_2 s_1} A_{s_1 s_0} (\overrightarrow{X_0})_{s_0}) \right) \\ &= - \sum_{s_0 s_1 \dots s_{t-1}} A_{s_{t-1} s_{t-2}} \dots A_{s_2 s_1} A_{s_1 s_0} (\overrightarrow{X_0})_{s_0} \\ &\times \left(\log |A_{s_{t-1} s_{t-2}}| + \dots + \log |A_{s_2 s_1}| + \log |A_{s_1 s_0}| + \log |(\overrightarrow{X_0})_{s_0}| \right. \\ &\quad \left. + iArg(A_{s_{t-1} s_{t-2}} \dots A_{s_2 s_1} A_{s_1 s_0} (\overrightarrow{X_0})_{s_0}) \right) \\ &= - \sum_{i=1}^{n-1} (\overrightarrow{X_0})_{s_i} \log |(\overrightarrow{X_0})_{s_i}| - (t-1) \sum_{k,l=1}^{n-1} A_{s_k s_l} (\overrightarrow{X_0})_{s_l} \log |A_{s_k s_l}| \\ &- i \sum_{s_0 s_1 \dots s_{t-1}} A_{s_{t-1} s_{t-2}} \dots A_{s_2 s_1} A_{s_1 s_0} (\overrightarrow{X_0})_{s_0} Arg(A_{s_{t-1} s_{t-2}} \dots A_{s_2 s_1} A_{s_1 s_0} (\overrightarrow{X_0})_{s_0}) \end{aligned}$$

Se z_1, \dots, z_n são números complexos tais que $Re(z_i) > 0$ e $Re(z_1 z_2 \cdots z_n) > 0$, $1 \leq i \leq n$, então o argumento é aditivo, ou seja, $Arg(z_1 \cdots z_n) = Arg(z_1) + \cdots + Arg(z_n)$. Em particular, $Log(z_1 \cdots z_n) = Log(z_1) + \cdots + Log(z_n)$. Em geral, matrizes de Dirichlet não satisfazem o critério dado acima. Por exemplo, se

$$M(j, k) = \frac{1}{\sqrt{n}} e^{ia\pi(j-k)^2/n},$$

fazendo $a = 1$ e $n = 3$, podemos escrever

$$M(3, 1) = \frac{1}{\sqrt{3}} (\cos(4\pi/3) + i \sin(4\pi/3)),$$

mas $Re(M(3, 1)) = (1/\sqrt{3}) \cos(4\pi/3) < 0$.

Supondo que as entradas de A satisfazem a hipótese observada acima, temos:

$$h_A(\sigma) = \lim_{t \rightarrow \infty} \frac{1}{t} H_{\mathbb{C}} \left(\bigvee_{k=0}^{t-1} \sigma^{-k} \alpha \right) = - \sum_{k,l=1}^{n-1} A_{s_k s_l} (\vec{X}_0)_{s_l} (\log |A_{s_k s_l}| + i Arg A_{s_k s_l})$$

5.5 Cadeias quânticas: entropia de von Neumann

Pelo que vimos nas seções 2.2 e 4.3, podemos facilmente obter a entropia de von Neumann de uma cadeia de Markov quântica. Vamos considerar um ensemble para cada tempo t , ou seja, se

$$S = \{s_0, s_1, \dots, s_{n-1}\} = \{|s_0\rangle, |s_1\rangle, \dots, |s_{n-1}\rangle\}$$

é o conjunto de estados possíveis, e $P_t(k) = P_t(s_k)$ denota a probabilidade de ocorrer o estado s_k no tempo t , consideramos os pares $\{P_t(k), |s_k\rangle\}$, $t \in \mathbb{N}$. O operador densidade associado à cadeia de Markov quântica no tempo t é

$$\begin{aligned} \rho_t &= \sum_{k=0}^{n-1} P_t(k) |s_k\rangle \langle s_k| = \sum_{k=0}^{n-1} |(A^t \vec{X}_0)_k|^2 |s_k\rangle \langle s_k| \\ &= \sum_{k=0}^{n-1} \left| \sum_j (\bar{\psi}_j)_0 \lambda_j^t (\psi_j)_k \right|^2 |s_k\rangle \langle s_k|, \end{aligned} \quad (5.4)$$

onde, lembramos, $\{X_t\}$ é uma N -cadeia quântica,

$$\vec{X}_t = (A[X_t = s_0], \dots, A[X_t = s_{n-1}]) \in \mathbb{C}^n,$$

$\lambda_0, \dots, \lambda_{n-1}$ são os autovalores da matriz de amplitude de transição $A_{jk} = A[X_2 = s_j | X_1 = s_k]$ e $\psi_0, \dots, \psi_{n-1}$ a base ortonormal de autovetores correspondente.

Calculando a entropia de von Neumann, temos

$$\begin{aligned} S(\rho_t) &= -tr(\rho_t \log \rho_t) = -\sum_{k=0}^{n-1} \left| \sum_j (\overline{\psi_j})_0 \lambda_j^t (\psi_j)_k \right|^2 \log \left| \sum_j (\overline{\psi_j})_0 \lambda_j^t (\psi_j)_k \right|^2 \\ &= -\sum_{k=0}^{n-1} \left| (\mathcal{A}^t \overrightarrow{X_0})_k \right|^2 \log \left| (\mathcal{A}^t \overrightarrow{X_0})_k \right|^2 \end{aligned}$$

Observação Note que a expressão para $(\mathcal{A}^t \overrightarrow{X_0})_k$ usada na equação (5.4) acima,

$$(\mathcal{A}^t \overrightarrow{X_0})_k = \sum_j (\overline{\psi_j})_0 \lambda_j^t (\psi_j)_k$$

é válida apenas fazendo a suposição de que $\overrightarrow{X_0} = (1, 0, \dots, 0)$, conforme obtida no final da seção 2.2. Para o caso geral de um $\overrightarrow{X_0}$ qualquer, procedemos da seguinte maneira. Fazemos os cálculos em dimensão 2, o caso geral é análogo, com uma notação mais complicada.

Sejam λ_0, λ_1 os autovalores de A e sejam ψ_0, ψ_1 a base ortonormal de autovetores correspondente. Então,

$$\begin{aligned} A^t \overrightarrow{X_0} &= A^t \sum_{j=0,1} \langle \overrightarrow{X_0}, \psi_j \rangle \psi_j = \sum_{j=0,1} \left[A(X_0 = \overline{0}) (\overline{\psi_j})_0 + A(X_0 = \overline{1}) (\overline{\psi_j})_1 \right] A^t \psi_j \\ &= \left[A(X_0 = \overline{0}) (\overline{\psi_0})_0 + A(X_0 = \overline{1}) (\overline{\psi_0})_1 \right] \lambda_0^t \psi_0 \\ &\quad + \left[A(X_0 = \overline{0}) (\overline{\psi_1})_0 + A(X_0 = \overline{1}) (\overline{\psi_1})_1 \right] \lambda_1^t \psi_1 \end{aligned}$$

No caso particular em que $\overrightarrow{X_0} = (1, 0)$, obtemos a expressão vista em [1], página 37, que em dimensão 2 é

$$A^t(1, 0) = \overline{(\psi_0)_0} \lambda_0^t \psi_0 + \overline{(\psi_1)_0} \lambda_1^t \psi_1$$

o que implica

$$P(X_t = k) = |\overline{(\psi_0)_0} \lambda_0^t (\psi_0)_k + \overline{(\psi_1)_0} \lambda_1^t (\psi_1)_k|^2, \quad k = 0, 1.$$

e em dimensão n ,

$$P(X_t = k) = \left| \sum_j (\overline{\psi_j})_0 \lambda_j^t (\psi_j)_k \right|^2, \quad k = 0, 1, \dots, n-1.$$

Mas no caso geral para \vec{X}_0 qualquer, em dimensão n , temos apenas que

$$A^t \vec{X}_0 = \sum_{j=0}^{n-1} \left[\sum_{k=0}^{n-1} A(X_0 = \bar{k}) \overline{(\psi_j)_k} \right] \lambda_j^t \psi_j.$$

Capítulo 6

Formalismo Termodinâmico

Neste capítulo analisamos um problema variacional para cadeias de Markov finitas. Se H denota a entropia, U é um potencial, e λ é um certo autovalor associado a U , temos um princípio variacional para energia livre:

$$H(\mu) - \int U d\mu \leq \log \lambda$$

onde vale a igualdade se e somente se μ for uma determinada medida especial $\nu = \nu^{(U)}$, que chamamos **medida de Markov** (faremos os detalhes a seguir). Dado este teorema para cadeias de Markov reais, podemos nos perguntar se existe algum análogo para cadeias quânticas.

6.1 Introdução

Apresentamos aqui um teorema de O. E. Lanford e D. Ruelle [26] numa versão de F. Spitzer [38] que caracteriza (em um caso simplificado) o estado de Gibbs (no reticulado unidimensional \mathbb{Z}) como sendo aquele que maximiza a energia livre. Essa energia é baseada em um potencial U que descreve de alguma forma a interação entre elementos vizinhos no reticulado \mathbb{Z} . Vamos assumir aqui que esta U tem uma expressão simples (um elemento no reticulado depende da interação de um número finito fixo de vizinhos) e assim ela determina uma matriz de transição M de uma cadeia de Markov [33] e esta define uma certa probabilidade estacionária sobre o espaço de Bernoulli. Esta probabilidade ν será o estado de Gibbs associado a U . Deste modo, será possível apresentar ao leitor uma versão matematicamente rigorosa de algumas idéias básicas e resultados fundamentais que aparecem em certos problemas simples de Mecânica Estatística. Esta exposição é baseada em [38].

6.2 Caracterização variacional

6.2.1 Introdução

A entropia é uma grandeza que mede a caoticidade, ou complexidade de um sistema: quanto maior for a entropia, mais caótico é o sistema. Esse conceito aparece na Física e está associado com o princípio de que a natureza tende a maximizar a entropia. Se em um tempo inicial t_0 consideramos partículas de gás concentradas em um dos cantos de uma caixa fechada, então após algum tempo (após o equilíbrio) as partículas tenderão a uma situação, onde elas estarão espalhadas na forma mais aleatória possível. Isso significa que decorrido algum tempo, o gás terá uma distribuição uniforme na caixa.

Um sistema de partículas é muito mais aleatório (tem mais entropia) se estiver uniformemente distribuído na caixa do que se estivesse concentrado em um dos cantos. Vemos assim que o equilíbrio é atingido em configurações de máxima entropia.

A entropia também está relacionada com a Teoria de Informação, a partir dos trabalhos de Shannon. Se quisermos transmitir uma mensagem através de um certo meio de comunicação usando um determinado alfabeto de n símbolos $\{1, \dots, n\}$, cada símbolo com uma certa probabilidade p_1, \dots, p_n de ocorrer, $\sum_{i=1}^n p_i = 1$ (suponha que a ocorrência dos símbolos seja independente), então a entropia deste sistema é a entropia do shift de Bernoulli $B(p_1, p_2, \dots, p_n)$, ou seja, $\sum_{i=1}^n -p_i \log p_i$.

A valor da entropia para medidas mais gerais tem uma expressão mais complexa. Nossa proposta aqui é a de usar a entropia de Shannon-Kolomogrov como uma ferramenta matemática para se estudar a Mecânica Estatística. O que apresentamos aqui é baseado na abordagem de Bowen-Ruelle-Sinai para entender reticulados em uma dimensão e veremos que essa proposta inclui estudar a pressão topológica do shift. Tal teoria é o que chamamos atualmente de Formalismo Termodinâmico. Uma dos objetos desta teoria é o operador de Ruelle-Perron-Frobenius, que é uma generalização natural (para o espaço de funções contínuas) de uma matriz com todas as entradas positivas. No caso de dimensão finita, o teorema de Perron-Frobenius agindo em \mathbb{R}^n basta para se obter o que se necessita. O operador de Ruelle-Perron-Frobenius possui diversas aplicações em outras áreas da matemática, como por exemplo Geometria e Teoria dos Números. Algumas vezes, na Física, este operador é denominado matriz de transferência.

6.2.2 Energia livre

Para motivar o problema que vamos analisar em breve vamos apresentar inicialmente o modelo mais simples possível. Considere um sistema físico com estados $\{1, \dots, n\}$, e sejam U_1, \dots, U_n as energias desses estados, respectivamente. Suponha que colocamos o sistema em contato com uma fonte de calor muito maior, que está a uma temperatura T . Sendo assim, a energia irá transitar entre o sistema original e a fonte de calor, e a temperatura T permanecerá constante, pois a fonte é muito maior que o nosso sistema. O problema físico que estamos considerando não é determinístico, e nós podemos apenas falar da probabilidade de um certo estado fixo, digamos j , ocorrer, onde $j \in \{1, 2, \dots, n\}$. Após esperar que o sistema se encontre em equilíbrio, se realizarmos uma sequência de observações, notaremos que o estado j irá ocorrer numa determinada proporção de vezes.

Por exemplo, se fizermos 1000 observações e em 112 delas aparece o estado 2, diremos que existe evidência de que 2 tem probabilidade $P_2 = \frac{112}{1000}$.

Então o que queremos saber, para cada j , é o valor dessa proporção quando o número de observações vai a infinito. É um fato conhecido da Mecânica Estatística (a partir de observações) que a probabilidade P_j de que o estado j ocorra é dado pela **distribuição de Gibbs**:

$$P_j = \frac{e^{-BU_j}}{\sum_{i=1}^n e^{-BU_i}}, \quad j \in \{1, \dots, n\},$$

onde $B = \frac{1}{kT}$ e k é uma constante, chamada **constante de Boltzmann**.

Uma formulação variacional do que foi dito acima pode ser feita da seguinte maneira. Seja

$$F(p_1, \dots, p_n) = - \sum_{i=1}^n p_i \log p_i - \sum_{i=1}^n p_i BU_i$$

definida no simplexo em \mathbb{R}^n dado por

$$\left\{ (p_1, \dots, p_n) : p_i \geq 0, i \in \{1, \dots, n\}, \sum_{i=1}^n p_i = 1 \right\}$$

Usando multiplicadores de Lagrange, podemos mostrar que o máximo de F no simplexo é obtido em

$$P_j = \frac{e^{-BU_j}}{\sum_{i=1}^n e^{-BU_i}}, \quad j \in \{1, \dots, n\},$$

ou seja, de acordo com o valor P_j dado acima.

A função

$$S(p_1, \dots, p_n) = - \sum_{i=1}^n p_i \log p_i$$

é a **entropia** da distribuição (p_1, \dots, p_n) . Defina $U(p_1, \dots, p_n) = - \sum_{i=1}^n p_i U_i$ como sendo a **energia média**. Então podemos dizer que a distribuição de Gibbs maximiza o valor

$$S(p_1, \dots, p_n) + BU(p_1, \dots, p_n)$$

Nesse contexto, a expressão $S + BU$ é o que chamaremos de **energia livre**. Logo, podemos dizer que a natureza minimiza a energia livre. Quando fazemos a temperatura T tender a $+\infty$, isto é, se B tender a 0, maximizamos a entropia.

Após a análise do sistema mais simples descrito acima, vamos considerar um caso um pouco mais complexo.

O modelo proposto por Ruelle é o seguinte (usaremos um modelo semelhante na próxima seção). Considere um reticulado unidimensional \mathbb{Z} . Cada inteiro está associado a um estado $1, \dots, n$ e uma configuração do sistema é uma sequência $\omega_i \in \{1, \dots, n\}$, $i \in \mathbb{Z}$. O nosso espaço de configurações é, então, $\Omega = \{1, \dots, n\}^{\mathbb{Z}}$. Seja $\sigma : \Omega \rightarrow \Omega$ a aplicação shift, $\tilde{\omega} = \sigma(\omega)$, dada por $\tilde{\omega}_k = (\sigma(\omega))_k = \omega_{k+1}$, $\omega = (\dots, \omega_{-1}, \omega_0, \omega_1, \omega_2, \dots)$, $k \in \mathbb{Z}$.

Em outras palavras, se $\omega = (\dots, \omega_{-2}, \omega_{-1}, \omega_0, \omega_1, \omega_2, \omega_3, \dots)$, onde na posição zero do reticulado \mathbb{Z} temos o valor ω_0 , então

$$\tilde{\omega} = \sigma(\omega) = (\dots, \tilde{\omega}_{-2}, \tilde{\omega}_{-1}, \tilde{\omega}_0, \tilde{\omega}_1, \tilde{\omega}_2, \tilde{\omega}_3, \dots),$$

onde na posição zero do reticulado \mathbb{Z} temos o valor $\tilde{\omega}_0$, e $\tilde{\omega}_k = \omega_{k+1}$.

Seja \mathcal{T} o espaço de probabilidades invariantes para o shift, i.e., $\mu \in \mathcal{T} \Leftrightarrow \mu(A) = \mu(\sigma^{-1}(A))$, A conjunto de Borel. Este é o modelo da Mecânica Estatística no reticulado \mathbb{Z} via o shift de Bernoulli. Um modelo mais apropriado seria sobre o reticulado tridimensional \mathbb{Z}^3 , mas aqui vamos evitar situações mais complexas.

Seja $U : \Omega \rightarrow \mathbb{R}$ uma função contínua, que contém a informação relacionada com alguma grandeza física (energia, temperatura, campo magnético, etc.). Queremos obter agora uma maneira de determinar a distribuição de Gibbs no reticulado unidimensional infinito de uma forma semelhante a usada no caso finito que vimos acima.

A distribuição de Gibbs associado a U será uma probabilidade ν sobre o espaço $\Omega = \{1, \dots, n\}^{\mathbb{Z}}$.

Por exemplo, considere uma determinada distribuição de spins $+$ ou $-$ de partículas no reticulado unidimensional \mathbb{Z} (poderíamos ter escrito 0 e 1 no lugar de $+$ e $-$, por exemplo). Devemos considerar o espaço de Bernoulli de dois símbolos $\Omega = \{+, -\}^{\mathbb{Z}}$, e probabilidades μ em Ω . fixado U , qual probabilidade ν é a de Gibbs asociada a U ? É apropriado considerar apenas probabilidades em \mathcal{T} porque não há uma razão natural para destacar um determinado ponto do reticulado como sendo o valor $i = 0$. Assim, a probabilidade de Gibbs deve ser invariante por translação.

Vamos denotar por $\overline{++} = \{\omega = (\dots, \omega_{-1}, \omega_0, \omega_1, \omega_2, \omega_3, \dots) \mid \text{tal que } \omega_1 = +, \omega_2 = +\}$. Da mesma forma, $\overline{+-} = \{\omega = (\dots, \omega_{-1}, \omega_0, \omega_1, \omega_2, \omega_3, \dots) \mid \text{tal que } \omega_1 = +, \omega_2 = -, \omega_3 = +\}$, e assim por diante...

O estado de Gibbs ν vai dizer, por exemplo, qual a probabilidade de ocorrer no reticulado o arranjo $\omega_1 = +, \omega_2 = -, \omega_3 = +$, através de $\nu(\overline{+-})$. Na verdade, vamos considerar o espaço $\Omega = \{+, -\}^{\mathbb{N}}$ e não $\{+, -\}^{\mathbb{Z}}$.

Exemplo. Seja $\Omega = \{+, -\}^{\mathbb{N}}$. Suponha que U é constante em cada um dos cilindros $\overline{++}$, $\overline{+-}$, $\overline{-+}$ e $\overline{--}$. Vamos identificar $+$ com 2 e $-$ com 1. Sejam $p_{11} \geq 0, p_{12} \geq 0, p_{21} \geq 0, p_{22} \geq 0, p_{11} + p_{12} = 1, p_{21} + p_{22} = 1$ e defina U da seguinte maneira:

$$U(\omega) = \begin{cases} -\log p_{22}, & \text{se } \omega \in \overline{++} \\ -\log p_{21}, & \text{se } \omega \in \overline{+-} \\ -\log p_{12}, & \text{se } \omega \in \overline{-+} \\ -\log p_{11}, & \text{se } \omega \in \overline{--} \end{cases}$$

Neste caso, assumimos que no reticulado \mathbb{Z} existe uma probabilidade p_{22} de obter um $+$ à direita de um $+$ e uma probabilidade p_{21} de obter um $-$ à direita de um $+$.

Qual seria o estado de Gibbs ν associado a tal U ?

Voltemos agora ao caso geral.

Dada uma função U contínua em Ω , vamos analisar o seguinte problema variacional. Seja

$$P(U) = \sup_{\mu \in \mathcal{T}} \left\{ S(\mu) + \int U d\mu \right\},$$

onde $S(\mu)$ é a entropia da probabilidade μ (ver definição na próxima seção). Diremos que $P(U)$ é a **pressão topológica** associada a U .

Gostaríamos de encontrar um probabilidade ν , definida em todo o espaço Ω , que assuma o supremo mencionado acima. Tal probabilidade será chamada

de **estado de equilíbrio**, ou **estado de Gibbs** associado ao potencial U . Em outras palavras,

$$P(U) = S(\nu) + \int U d\nu \geq S(\mu) + \int U d\mu, \quad \forall \mu \in \mathcal{T}.$$

O estado de equilíbrio ν será definido, portanto, por meio de um princípio de máximo, ou seja, ν maximiza algo. O potencial U do exemplo particular acima mencionado descreve uma certa interação entre spins no reticulado, mas o problema faz sentido para U qualquer, não necessariamente como no exemplo. Poderia, por exemplo, depender de mais coordenadas, não apenas duas. Vamos considerar na próxima seção o caso simples em que $U(\omega)$ depende apenas de duas, ou seja de ω_0 e ω_1 , onde $\omega = (\cdots \omega_{-1} \omega_0 \omega_1 \cdots)$, ou seja, cada spin depende apenas do vizinho à direita. No exemplo dado abaixo, a solução pode ser obtida através de Álgebra Linear, ou seja, pela teoria de Cadeias de Markov e pelo Teorema de Perron-Frobenius (descrito no Apêndice deste capítulo). Se o U depende de infinitas coordenadas, aí o procedimento via Álgebra Linear não resolve o problema; é necessário utilizar [35].

Ruelle mostrou que o que os físicos denominam de estado de Gibbs, no caso do reticulado \mathbb{Z} , pode ser obtido via o procedimento acima através de uma escolha correta de U . Diferentes problemas de Mecânica Estatística requerem diferentes U .

A análise de questões mais gerais em Mecânica Estatística pode ser encontrada em [10] [40].

6.2.3 Caracterização variacional

Para simplificar, no lugar de $\Omega = \{1, \dots, n\}^{\mathbb{Z}}$, vamos considerar

$$\Omega = \{1, \dots, n\}^{\mathbb{N}}.$$

O problema para um U geral definido em $\Omega = \{1, \dots, n\}^{\mathbb{Z}}$ pode ser reduzido a este $\{1, \dots, n\}^{\mathbb{N}}$ (ver Proposition 1.2 [35]).

Seja uma função potencial do tipo $U : \{1, \dots, n\}^{\mathbb{N}} \rightarrow \mathbb{R}$, com $U(\omega) = U(\omega_1, \omega_2)$, onde $\omega = (\omega_1 \omega_2 \omega_3 \cdots)$, isto é, U só depende de duas coordenadas. Observe que se consideramos um potencial U que depende de n coordenadas, com $2 < n < \infty$, podemos modificar o espaço Ω de modo a fazer com que U dependa de apenas duas coordenadas.

Por exemplo, suponha que $\Omega = \{1, 2\}^{\mathbb{N}}$ e que temos um potencial U que depende de 3 coordenadas, isto é, $U(\omega) = U(\omega_1, \omega_2, \omega_3)$. Considere então o

espaço $\hat{\Omega} = \{1^*, 2^*, 3^*, 4^*\}^{\mathbb{N}}$ e o relacionamos com Ω fazendo as identificações

$$11 \leftrightarrow 1^* \quad 12 \leftrightarrow 2^*$$

$$21 \leftrightarrow 3^* \quad 22 \leftrightarrow 4^*$$

Note agora que certas relações estão proibidas. O par 1^*3^* não é permitido. O par 1^*2^* é permitido.

Podemos pensar que U está definido em $\hat{\Omega}$ como uma certa V . De fato, V definida de maneira natural em $\hat{\Omega}$ depende apenas de duas coordenadas, pois, por exemplo $U(1, 1, 2, \dots) = U(1, 1, 2) = V(1^*, 2^*)$, ainda $U(1, 2, 2, \dots) = U(1, 2, 2) = V(2^*, 4^*)$ e assim por diante.

Desta maneira se pode fazer recair o caso em que U depende de finitas coordenadas no espaço de Bernoulli ao caso em que dependa de apenas duas, como será analisado a seguir.

Tal procedimento não pode ser feito se o potencial considerado depende de infinitas coordenadas.

Um cilindro de Ω de comprimento k é um conjunto A da seguinte forma: fixe $(a_1, \dots, a_k) \in \{1, \dots, n\}^k$.

$$A = \{\omega = (\omega_1, \dots, \omega_k, \omega_{k+1}, \dots, \omega_n, \dots) \in \Omega : (\omega_1, \dots, \omega_k) = (a_1, \dots, a_k)\}$$

Iremos denotar tal cilindro por

$$A = \overline{a_1 \cdots a_k}$$

Seja \mathcal{F} a σ -álgebra gerada pelos cilindros de Ω de todos os comprimentos possíveis.

Seja $\sigma : \Omega \rightarrow \Omega$ a aplicação shift dada por $(\sigma\omega)_k = \omega_{k+1}$, $k \in \mathbb{N}$.

Em outras palavras, se $\omega = (\omega_1, \omega_2, \omega_3, \dots)$,

$$\sigma(\omega) = (\omega_2, \omega_3, \omega_4, \dots).$$

Por exemplo, em $\Omega = \{1, 2, 3\}^{\mathbb{N}}$, temos que

$$\sigma(1, 2, 1, 1, 3, 2, \dots) = (2, 1, 1, 3, 2, \dots).$$

Iremos a seguir considerar medidas μ sobre a σ -álgebra \mathcal{F} . Diremos que $\mu \in \mathcal{T} \Leftrightarrow (\Omega, \mathcal{F}, \mu)$ é espaço de probabilidade, com μ invariante pelo shift (i.e., $\mu(\sigma^{-1}(A)) = \mu(A)$, $A \in \mathcal{F}$), e diremos que \mathcal{T} é o conjunto das medidas invariantes por translação. Sabe-se que \mathcal{T} é compacto se consideramos a convergência fraca de medidas [7] [6].

Seja k fixo e A_k o conjunto dos cilindros de comprimento k , $k \geq 1$, que é uma partição de Ω . Defina a **entropia da partição** por

$$S(A_k) = - \sum_{a \in A_k} \mu(a) \log \mu(a)$$

onde escrevemos $a = \overline{a_1 a_2, \dots, a_k} \in A_k$.

Definimos a **entropia da medida** (com relação ao shift) por

$$s(\mu) = \lim_{n \rightarrow +\infty} \frac{S(A_n)}{n}, \quad \mu \in \mathcal{T}, \quad (6.1)$$

Definindo a função potencial $U : \Omega \rightarrow \mathbb{R}$, diremos que $U(i, j)$, $1 \leq i, j \leq n$ é a energia de interação entre ω_k e ω_{k+1} se $\omega_k = i$, $\omega_{k+1} = j$. Defina a **energia média** da partição A_k como sendo

$$E(A_k) = \sum_{a \in A_k} \left(\mu(a) \sum_j U(a_j, a_{j+1}) \right)$$

A **energia específica** da partição com respeito a U é

$$e_U(\mu) = \lim_{n \rightarrow +\infty} \frac{E(A_n)}{n} \quad (6.2)$$

que é uma função contínua e afim. Afirmamos que

$$e_U(\mu) = \int U d\mu$$

De fato, pelo teorema ergódico de Birkhoff (proposição 6.4.3) e aplicando o teorema da convergência dominada para a sequência de funções $\frac{1}{N} \sum_{k=1}^{N-1} U \circ \sigma^k$, temos:

$$\begin{aligned} \int U d\mu &= \int \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{j=1}^{N-1} U(\sigma^j(\omega)) d\mu(\omega) = \lim_{N \rightarrow +\infty} \int \frac{1}{N} \sum_{j=1}^{N-1} U(\sigma^j(\omega)) d\mu(\omega) = \\ &= \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{j=1}^{N-1} \int U(\sigma^j(\omega)) d\mu(\omega) = \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_j \sum_{a \in A_k} \mu(a) U(\sigma^j(a)) = \\ &= \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{a \in A_k} \mu(a) \sum_j U(a_j, a_{j+1}) = e_U(\mu) \end{aligned}$$

Assim, se a **energia livre** em A_k é

$$F(A_k) = S(A_k) - E(A_k),$$

definimos a **energia livre específica** como sendo

$$f_U(\mu) = s(\mu) - e_U(\mu) = s(\mu) - \int U d\mu \quad (6.3)$$

Vamos estar interessados nas medidas μ que maximizam tal valor.

O conjunto $\mathcal{M} \subset \mathcal{T}$ das medidas de Markov será definido da seguinte maneira. Diremos que $\mu \in \mathcal{M} \Leftrightarrow \exists$ matriz $n \times n$ linha **estocástica** $M = M_{ij}$ (i.e., $\sum_j M_{ij} = 1$) estritamente positiva tal que

$$\mu(\overline{a_1, \dots, a_N}) = \varphi_{a_1} M(a_1, a_2) M(a_2, a_3) \cdots M(a_{N-1}, a_N), \quad a_1, \dots, a_N \in \{1, \dots, n\} \quad (6.4)$$

onde φ é o único vetor de probabilidade invariante à esquerda para M (ou seja, $\varphi_k > 0$, $1 \leq k \leq n$, $\sum \varphi_i = 1$ e $\varphi M = \varphi$). Definindo μ sobre os cilindros, fica determinada de maneira única, pelo teorema de Kolmogorov, uma medida de probabilidade sobre a σ -álgebra gerada pelos cilindros.

Voltando à função potencial U dada acima, seja a matriz $Q = Q_U$ definida por

$$Q(i, j) = e^{-U(i, j)}, \quad 1 \leq i, j \leq n \quad (6.5)$$

Seja $\lambda = \lambda(U)$ o maior autovalor (positivo) de Q e sejam l, r os autovetores à esquerda e à direita de Q correspondentes a λ , e normalizados de modo que $l_i > 0$, $r_i > 0$, $1 \leq i \leq n$ e $\langle l, r \rangle = \sum_i l_i r_i = 1$ (proposição 6.4.1).

Defina a seguinte matriz:

$$M(i, j) = \frac{1}{\lambda} Q(i, j) \frac{r_j}{r_i}, \quad 1 \leq i, j \leq n \quad (6.6)$$

A matriz $M = M(U)$ é positiva e é linha estocástica. De fato:

$$Qr = \lambda r \Rightarrow Q \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} = \lambda(r_1, \dots, r_n) \Rightarrow \sum_j Q(i, j) r_j = \lambda r_i$$

Então

$$\sum_j M(i, j) = \sum_j \frac{1}{\lambda} Q(i, j) \frac{r_j}{r_i} = \frac{1}{\lambda r_i} \sum_j Q(i, j) r_j = \frac{1}{\lambda r_i} \lambda r_i = 1$$

Observe que M tem um vetor de probabilidade invariante

$$\varphi_i = l_i r_i, \quad 1 \leq i \leq n \quad (6.7)$$

De fato,

$$\begin{aligned} \varphi M &= (l_1 r_1, \dots, l_n r_n) M \Rightarrow \\ \Rightarrow (\varphi M)_i &= \sum_k l_k r_k M(k, i) = \sum_k l_k r_k \frac{1}{\lambda} Q(k, i) \frac{r_i}{r_k} = \\ &= \frac{r_i}{\lambda} \sum_k l_k Q(k, i) = \frac{r_i}{\lambda} \lambda l_i = r_i l_i = \varphi_i \Rightarrow \varphi M = \varphi \end{aligned}$$

Note que l_i e r_i são definidos a menos de constante multiplicativa e acima escolhemos $\varphi = (\varphi_1, \dots, \varphi_n)$ como sendo $\varphi_i = l_i r_i$, $i = 1, \dots, n$ e $\sum \varphi_i = 1$, $\varphi_i > 0$.

Vamos formular a caracterização variacional de ν . Para U fixo, seja $\nu = \nu^{(U)}$ a medida de Markov definida pela matriz de transição $M = M(U)$ obtida acima, isto é,

$$\nu(a) = \varphi_{a_1} M(a_1, a_2) M(a_2, a_3) \cdots M(a_{k-1}, a_k), \quad a \in A_k$$

Dizemos que σ é **ergódica** para μ (invariante) se $\sigma^{-1}(A) = A$ implica $\mu(A) = 0$ ou 1 . Pode-se mostrar que a probabilidade ν acima é invariante e ergódica para o shift σ seção 11.6 [36].

Afirmamos que $\sup_{\mu \in \mathcal{T}} f_U(\mu) = f_U(\nu)$. Mais precisamente, temos o seguinte:

Teorema 6.2.1 *Seja $U : \Omega \rightarrow \mathbb{R}$ contínua. Então:*

$$f_U(\mu) \leq \log \lambda(U) = \log(\lambda), \quad \forall \mu \in \mathcal{T} \quad (6.8)$$

e vale a igualdade se e somente se μ for a medida de Markov $\nu = \nu^{(U)}$ definida acima.

Uma maneira de provar o teorema consiste em 3 passos:

1. Mostrar a desigualdade, e que vale a igualdade se $\mu = \nu^{(U)}$.
2. Mostrar a desigualdade estrita se $\mu \neq \nu^{(U)}$, μ ergódica.
3. Mostrar a desigualdade estrita se $\mu \neq \nu^{(U)}$, μ qualquer.

Na demonstração a seguir, escreveremos $\nu^{(U)} = \nu$.

Prova Mostremos o primeiro passo. Seja ν a medida de Markov definida por M , e $\mu \in \mathcal{T}$ qualquer. Sejam M e φ definidos por (5), (6), (7), $\nu(a) = \varphi_{a_1} M(a_1, a_2) M(a_2, a_3) \cdots M(a_{k-1}, a_k)$.

Suponha ainda que os cilindros considerados tem comprimento N , isto é, consideramos os cilindros de A_N . Então

$$E(A_N) = \sum_{a \in A_N} \left(\mu(a) \sum_{k=1}^{N-1} U(a_k, a_{k+1}) \right)$$

Escrevendo $\lambda(U) = \lambda$, temos:

$$\begin{aligned} \sum_{k=1}^{N-1} U(a_k, a_{k+1}) &= \sum_{k=1}^{N-1} -\log Q(a_k, a_{k+1}) = -\log \prod_{k=1}^{N-1} Q(a_k, a_{k+1}) = \\ &= -\log \left(\prod_{k=1}^{N-1} M(a_k, a_{k+1}) \lambda \frac{r_{a_k}}{r_{a_{k+1}}} \right) = \\ &= - \left(\log \prod_{k=1}^{N-1} M(a_k, a_{k+1}) + \log \prod_{k=1}^{N-1} \lambda + \log \prod_{k=1}^{N-1} r_{a_k} - \log \prod_{k=1}^{N-1} r_{a_{k+1}} \right) = \\ &= -(N-1) \log \lambda - \log \left(\prod_{k=1}^{N-1} M(a_k, a_{k+1}) \right) - \log r_{a_1} + \log r_{a_N} = \\ &= -(N-1) \log \lambda - \log \frac{\nu(a)}{\varphi_{a_1}} - \log r_{a_1} + \log r_{a_N} \\ &= -(N-1) \log \lambda - \log \nu(a) + \log l_{a_1} r_{a_1} - \log r_{a_1} + \log r_{a_N} \\ &= -(N-1) \log \lambda - \log \nu(a) + \log l_{a_1} + \log r_{a_N}. \end{aligned}$$

Calculando a energia média com respeito a μ obtemos

$$\begin{aligned} E(A_N) &= \sum_{a \in A_N} \mu(a) \sum_{k=1}^{N-1} U(a_k, a_{k+1}) = \\ &= \sum_{a \in A_N} \mu(a) (-(N-1) \log \lambda - \log \nu(a) + \log l_{a_1} + \log r_{a_N}) \end{aligned}$$

Como o número de $a_i \in \{1, 2, \dots, n\}$ é finito e os r_{a_i} e l_{a_i} são positivos, então existe c_1 e c_2 tal que para todo i vale $c_1 < \log r_{a_i}, \log l_{a_i} < c_2$.

Sendo assim,

$$\begin{aligned} & \sum_{a \in A_N} \mu(a) (-(N-1) \log \lambda - \log \nu(a) + 2c_1) \leq \\ \leq E_N(\mu) &= \sum_{a \in A_N} \mu(a) (-(N-1) \log \lambda - \log \nu(a) + \log l_{a_1} + \log r_{a_N}) \leq \\ & \leq \sum_{a \in A_N} \mu(a) (-(N-1) \log \lambda - \log \nu(a) + 2c_2). \end{aligned}$$

Observando que o somatório

$$\sum_{a \in A_N} \mu(a) \log \frac{\nu(a)}{\mu(a)} \quad (6.9)$$

é negativo (veja a proposição 6.4.2), temos

$$\frac{1}{N} F(A_N) = \frac{1}{N} (S(A_N) - E(A_N)) = \quad (6.10)$$

$$= \frac{1}{N} \sum_{a \in A_N} \mu(a) \log \frac{\nu(a)}{\mu(a)} + \log \lambda - \frac{\log \lambda}{N} - \frac{(\log l_{a_1} + \log r_{a_N})}{N} \leq \quad (6.11)$$

$$\leq \log \lambda - \frac{\log \lambda}{N} - 2 \frac{c_1}{N} \quad (6.12)$$

Fazendo $N \rightarrow +\infty$, obtemos a desigualdade procurada, isto é, $f_U(\mu) \leq \log \lambda$. Para mostrar a igualdade, observe que o somatório (6.9) é igual a zero se $\mu = \nu$. Segue daí que

$$\log \lambda - \frac{\log \lambda}{N} - 2 \frac{c_2}{N} \leq \frac{1}{N} F(A_N) \leq \log \lambda - \frac{\log \lambda}{N} - 2 \frac{c_1}{N} \quad (6.13)$$

Fazendo $N \rightarrow +\infty$, obtemos a igualdade, ou seja, $f_U(\mu) = \log \lambda$ se $\mu = \nu^{(U)}$ (isto é, se μ for a medida de Markov). Isso prova o primeiro passo.

Agora vamos supor provado o segundo passo (ou seja, suponha que vale a desigualdade estrita se $\mu \neq \nu^{(U)}$, μ ergódica). Este segundo passo é bastante técnico e referimos o leitor para uma prova geral em [35].

Mostremos o terceiro passo, ou seja, que vale a desigualdade estrita se $\mu \neq \nu^{(U)}$, μ qualquer.

Primeiro, observamos que \mathcal{T} é convexo [39]. De fato, se $\mu_1, \mu_2 \in \mathcal{T}$, então $\mu = (1 - \lambda)\mu_1 + \lambda\mu_2$ é tal que

$$\mu(\sigma^{-1}(A)) = (1-\lambda)\mu_1(\sigma^{-1}(A)) + \lambda\mu_2(\sigma^{-1}(A)) = (1-\lambda)\mu_1(A) + \lambda\mu_2(A) = \mu(A)$$

e então $\mu \in \mathcal{T}$, logo \mathcal{T} é convexo.

Vale que s é semicontínua superiormente (ver [42]), e como e_U é contínua, $f_U = s - e_U$ é uma função semicontínua superiormente. Seja

$$\mathcal{K} = \mathcal{K}_U = \{\mu \in \mathcal{T} : f_U(\mu) \text{ assume o valor máximo } M\}$$

(Sabemos que \mathcal{T} é compacto e f_U é semicontínua superiormente; isto implica que f_U assume um máximo em \mathcal{T} , conforme [27]). Vimos no primeiro passo que $\nu \in \mathcal{K}$. Afirmo que \mathcal{K} é fechado. De fato, seja $\mu = \lim \mu_n$, $\mu_n \in \mathcal{K}$. Então

$$M = \limsup_{n \rightarrow \infty} f_U(\mu_n) \leq f_U(\mu)$$

e logo, $f_U(\mu)$ assume o máximo $\Rightarrow \mu \in \mathcal{K}$. Logo, \mathcal{K} é fechado. Segue daí que \mathcal{K} é compacto, pela compacidade de \mathcal{T} .

Como f_U é uma aplicação afim, vale que \mathcal{K} é convexo [39]. Com efeito, sejam $\mu_1, \mu_2 \in \mathcal{K}$. Se $\mu = (1 - \lambda)\mu_1 + \lambda\mu_2$, temos

$$f_U(\mu) = f_U((1-\lambda)\mu_1 + \lambda\mu_2) = (1-\lambda)f_U(\mu_1) + \lambda f_U(\mu_2) = (1-\lambda)M + \lambda M = M$$

e logo, $(1 - \lambda)\mu_1 + \lambda\mu_2 \in \mathcal{K} \Rightarrow \mathcal{K}$ é convexo.

Para mostrar a desigualdade estrita, é suficiente mostrar que $\mathcal{K} = \{\nu\}$. Como \mathcal{K} é compacto e convexo, pelo teorema de Krein-Milman (ver o Apêndice deste capítulo), isso ocorrerá se e somente se \mathcal{K} tem um único ponto extremo. Afirmamos que todo ponto extremo de \mathcal{K} é um ponto extremo de \mathcal{T} . Com efeito, suponha por absurdo que η é ponto extremo de \mathcal{K} que não é ponto extremo de \mathcal{T} . Então η pode ser escrito como

$$\eta = \lambda\mu_1 + (1 - \lambda)\mu_2, \quad 0 < \lambda < 1, \quad \mu_1 \notin \mathcal{K} \text{ ou } \mu_2 \notin \mathcal{K}$$

Suponha que $\mu_1 \notin \mathcal{K}$. Então, $f_U(\mu_1) < M$ e assim,

$$\begin{aligned} M &= f_U(\eta) = f_U(\lambda\mu_1 + (1 - \lambda)\mu_2) = \lambda f_U(\mu_1) + (1 - \lambda)f_U(\mu_2) < \\ &< \lambda M + (1 - \lambda)f_U(\mu_2) \leq \lambda M + (1 - \lambda)M = M, \end{aligned}$$

absurdo. Logo, todo ponto extremo de \mathcal{K} é um ponto extremo de \mathcal{T} .

Um fato conhecido sobre o conjunto \mathcal{T} é que os seus pontos extremos são justamente as medidas ergódicas [42]. Então o problema é reduzido a mostrar que \mathcal{K} não contém outras medidas ergódicas além de ν . Ora, mas pelo segundo passo, se supomos que a medida $\mu \neq \nu$ considerada é ergódica, então vale a desigualdade estrita, ou seja, μ não é tal que f_U assume um máximo. Isso completa a prova.

□

6.3 Observações sobre processos quânticos

Sabemos de [15] que a matriz de amplitude $[A_{jk}]$ é uma matriz unitária, suas entradas são não-nulas e é coluna-estocástica, ou seja $\sum_j A_{jk} = 1$.

Seja A a matriz de amplitude de transição de uma N-cadeia. Defina a matriz de amplitude de transição com um potencial por $A^V : S \times S \rightarrow \mathbb{C}$,

$$A_{q'q}^V = A^V(q, q') := e^{-iV(q)} A(q, q') = e^{-iV(q)} A_{q'q}$$

Interpretamos que $A_{q'q}^V$ é, como no caso de cadeias de Markov, a amplitude de transição de q até q' . Supomos que o potencial depende apenas da posição onde a partícula se encontra, ou seja, $V = V(q)$.

Dizemos que A^V é a amplitude correspondente a uma partícula evoluindo sob a influência de um potencial. Note que se $V = 0$, A^V se reduz à função de amplitude livre A .

A seguir, escreveremos $S = \{s_0, \dots, s_{n-1}\}$ e $V(s_k) = V(k)$.

Lema 6.3.1 *Vale que A^V é uma matriz unitária com entradas não nulas. Além disso, A^V é coluna-estocástica apenas quando consideramos potenciais V do tipo*

$$V(s_k) = 2\pi c_k, \quad c_k \in \mathbb{Z}, \quad k = 0, \dots, n-1.$$

Prova De fato,

$$\begin{aligned} (A^{V*} A^V)_{jk} &= \sum_r (\bar{A}^V)_{jr}^T (A^V)_{rk} = \sum_r \bar{A}_{rj}^V A_{rk}^V \\ &= \sum_r e^{iV(j)} \bar{A}_{rj} e^{-iV(k)} A_{rk} = e^{i(V(j)-V(k))} \sum_r \bar{A}_{rj} A_{rk} \\ &= \begin{cases} 0 & , \text{ se } j \neq k \\ 1 & , \text{ se } j = k \end{cases} \end{aligned}$$

onde a última igualdade, no caso $j \neq k$, segue do fato que A é N-cadeia quântica (e portanto unitária) e, no caso em que $j = k$, do fato que

$$\sum_r |A_{rj}|^2 = \sum_r |A(X_2^{-1}(s_r) \cap X_1^{-1}(s_k))|^2 = 1$$

Como A tem entradas não nulas, é claro que A^V também.

Vale que A^V é coluna-estocástica apenas quando consideramos potenciais V do tipo

$$V(s_k) = 2\pi c_k, \quad c_k \in \mathbb{Z}, \quad k = 0, \dots, n-1,$$

pois

$$\sum_j A_{jk}^V = \sum_j e^{-iV(k)} A_{jk} = e^{-iV(k)} \sum_j A_{jk} = e^{-iV(k)}$$

□

A seguir, fazemos outras observações.

1) Considere as matrizes de Dirichlet

$$M(n, a) = \frac{1}{\sqrt{n}} e^{i\pi a(j-k)^2/n}, \quad j, k = 0, 1, \dots, n-1$$

vistas na seção 2.3. Ali, mostra-se uma maneira de transformar tal matriz em uma outra, $M'(n, a)$, que é estocástica desde que na seja par. O método visto ali não funciona para $M(3, 1)$ por exemplo. Para resolver esse caso, tentamos usar uma idéia baseada em [38], onde se usa o maior autovalor de uma dada matriz positiva A e o seu autovetor à direita para transformá-la em estocástica. Mas aqui temos apenas matrizes complexas, então necessitamos de algumas adaptações.

A matriz $M(3, 1)$ é

$$M(3, 1) = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & \frac{1+i\sqrt{3}}{2} & \frac{-1-i\sqrt{3}}{2} \\ \frac{1+i\sqrt{3}}{2} & 1 & \frac{1+i\sqrt{3}}{2} \\ \frac{-1-i\sqrt{3}}{2} & \frac{1+i\sqrt{3}}{2} & 1 \end{pmatrix}$$

Os autovalores e autovetores são, respectivamente,

$$-i, \quad (1, -1, 1);$$

$$\frac{i + \sqrt{3}}{2}, \quad (1, 1, 0), (-1, 0, 1).$$

Gostaríamos de escolher de alguma maneira natural um dos autovalores de $M(3, 1)$. Note que a norma dos dois autovalores é 1.

Defina $T : M(3, \mathbb{C}) \rightarrow M(3, \mathbb{R})$ a aplicação

$$A \xrightarrow{T} A_{\mathbb{R}},$$

onde $(A_{\mathbb{R}})_{ij} = \operatorname{Re}(A_{ij})$.

Defina a aplicação $t_{max} : M(3, \mathbb{R}) \rightarrow \mathbb{R}$ dada por

$$A \xrightarrow{t_{max}} \text{maior autovalor de } A.$$

A idéia é associar à matriz $M(3, 1)$, uma outra, $T(M(3, 1))$, que possui coeficientes reais e tomar o seu maior autovalor. Esse autovalor está associado de maneira natural com um dos autovalores da matriz original $M(3, 1)$ da seguinte maneira. Definindo

$$\lambda_0 := t_{max}(T(M(3, 1))),$$

tomamos o autovalor λ_1 de $M(3, 1)$, onde $\lambda_1 = \gamma(1)$, e $\gamma : \mathbb{C} \rightarrow \mathbb{C}$ é o único caminho que leva um autovalor de $M(3, 1)$ até $\lambda_0 = \gamma(0)$. Vamos explicar com mais precisão esta última idéia.

Defina a família de matrizes $M(3, 1; t)$, $t \in [0, 1]$ dada por

$$M(3, 1; t) = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & \frac{1+ti\sqrt{3}}{2} & \frac{-1-ti\sqrt{3}}{2} \\ \frac{1+ti\sqrt{3}}{2} & 1 & \frac{1+ti\sqrt{3}}{2} \\ \frac{-1-ti\sqrt{3}}{2} & \frac{1+ti\sqrt{3}}{2} & 1 \end{pmatrix}$$

É claro que $M(3, 1; 1) = M(3, 1)$ e que $M(3, 1; 0) = T(M(3, 1))$. O espectro da família $M(3, 1; t)$ é formado por três curvas, η , ζ e γ , onde uma delas (que chamamos de γ acima) tem como um de seus extremos o maior dos autovalores de $T(M(3, 1))$, que chamamos $\lambda_0 = \gamma(0)$. Então o que fazemos é escolher o autovalor no outro extremo desta curva, ou seja $\lambda_1 = \gamma(1)$.

Continuando, obtemos

$$T(M(3, 1)) = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}$$

Daí, alguns cálculos simples mostram que $\lambda_0 = \sqrt{3}/2$ e que $\lambda_1 = (i + \sqrt{3})/2$ é o autovalor de $M(3, 1)$ encontrado de acordo com a regra postulada acima. Um autovetor r com entradas positivas, e tal que $\sum r_j^2 = 1$, é $r = (\sqrt{6}/6, \sqrt{6}/3, \sqrt{6}/6)$. Agora, procedendo de maneira análoga ao método em [38], definimos

$$M'(3, 1)(i, j) := \frac{1}{\lambda_1} M(3, 1) \frac{r_j}{r_i}$$

Escreva $\alpha = (1 + i\sqrt{3})/2$. Então

$$M'(3, 1)(1, 1) = M'(3, 1)(2, 2) = M'(3, 1)(3, 3) = 2/(3 + i\sqrt{3})$$

Mais algumas contas simples nos fornecem

$$M'(3, 1) = \frac{1}{3 + i\sqrt{3}} \begin{pmatrix} 2 & 4\alpha & -2\alpha \\ \alpha & 2 & \alpha \\ -2\alpha & 4\alpha & 2 \end{pmatrix}$$

que é linha estocástica, mas não é unitária.

2) Considere o seguinte lema [14]:

Lema 6.3.2 (Wielandt) *seja $A = (a_{ij})$ matriz irredutível não negativa de ordem n e $C = (c_{ij})$ uma matriz quadrada complexa de ordem n . Suponha que*

$$C^+ \leq A. \quad (6.14)$$

Então para todo autovalor γ de C vale que

$$|\gamma| \leq r, \quad (6.15)$$

onde r é o maior autovalor de A . Além disso, vale a igualdade se e somente se

$$C = e^{i\varphi} D A D^{-1}, \quad (6.16)$$

onde $e^{i\varphi} = \gamma/r$ e D é uma matriz diagonal cujos elementos não nulos tem módulo 1 (vale que $D^+ = I$).

A prova deste lema encontra-se no Apêndice deste capítulo. Um possível trabalho futuro baseado neste lema é o seguinte: dada uma matriz complexa, obter um método de normalização a partir da matriz módulo associada, o que pode fornecer um método mais geral do que o obtido no ítem 1).

3) Para que uma matriz estocástica com entradas não nulas seja a matriz de uma cadeia de Markov quântica, precisamos ainda que ela seja unitária. Para este fim, devemos determinar se é possível adaptar o método acima de modo a fornecer a propriedade unitária.

Podemos obter uma condição para que uma matriz normalizada na forma do ítem anterior seja unitária:

Lema 6.3.3 *Seja*

$$A(i, j) = \frac{1}{\lambda} M(i, j) \frac{r_j}{r_i}$$

onde M é uma matriz complexa, λ é um autovalor não nulo de M e r um autovetor associado a λ com entradas não nulas. Se vale que A é unitária, temos:

1. Para todo i ,

$$\sum_k |M(i, k)|^2 r_k^2 = |\lambda|^2 r_i^2$$

2. Para $i \neq j$,

$$\sum_k M(i, k) \overline{M}(j, k) r_k^2 = 0$$

Prova Temos:

$$\begin{aligned} \sum_{k=1}^n A_{ik} A_{kj}^* &= \sum_k A_{ik} \overline{A}_{jk} = \sum_k (1/\lambda) M(i, k) \frac{r_k}{r_i} (1/\overline{\lambda}) \overline{M}(j, k) \frac{r_k}{r_j} \\ &= \frac{1}{|\lambda|^2 r_i r_j} \sum_k M(i, k) \overline{M}(j, k) r_k^2 \end{aligned}$$

Se A é unitária então para $i = j$, vale a seguinte condição:

$$\sum_k |M(i, k)|^2 r_k^2 = |\lambda|^2 r_i^2$$

E se $i \neq j$, vale

$$\sum_k M(i, k) \overline{M}(j, k) r_k^2 = 0$$

□

6.4 Apêndice

Proposição 6.4.1 (Perron-Frobenius) *Seja $A = \| a_{ij} \|$ uma matriz com entradas estritamente positivas, $1 \leq i, j \leq n$. Então existem $\lambda > 0$ e vetores $u = (u_1, \dots, u_n)$ e $v = (v_1, \dots, v_n)$ tais que*

$$1. u_i > 0, v_i > 0, \quad 1 \leq i \leq n$$

$$2. \sum_{j=1}^n a_{ij} u_j = \lambda u_i, \quad 1 \leq i \leq n \quad \text{e} \quad \sum_{i=1}^n v_i a_{ij} = \lambda v_j, \quad 1 \leq j \leq n$$

(i.e., u é autovetor à direita de A e v é autovetor à esquerda de A).

Prova Esta prova segue [37]. Seja A matriz com entradas positivas. Mostremos que existe pelo menos um vetor u com coordenadas positivas, e $\lambda \geq 0$, tais que

$$\sum_{j=1}^n a_{ij}u_j = \lambda u_i, \quad 1 \leq i \leq n.$$

Considere o conjunto convexo \mathcal{H} de vetores $h = (h_1, \dots, h_n)$ tais que $h_i \geq 0$, $1 \leq i \leq n$ e $\sum_{i=1}^n h_i = 1$. A matriz A determina uma transformação contínua $A : \mathcal{H} \rightarrow \mathcal{H}$, dada por $Ah = h'$, onde

$$h'_i = \frac{\sum_{j=1}^n a_{ij}h_j}{\sum_{i=1}^n \sum_{j=1}^n a_{ij}h_j}$$

O teorema do ponto fixo de Brouwer [27] nos diz que tal aplicação possui pelo menos um ponto fixo. Se u é tal ponto fixo então $Au = u$, ou seja,

$$u_i = \frac{\sum_{j=1}^n a_{ij}u_j}{\sum_{i=1}^n \sum_{j=1}^n a_{ij}u_j}$$

Fazendo $\lambda = \frac{\sum_{i=1}^n \sum_{j=1}^n a_{ij}u_j}{\sum_{i=1}^n \sum_{j=1}^n a_{ij}u_j}$, obtemos o que queríamos.

Considere a matriz $A^t = \| a_{ij}^t \|$, $a_{ij}^t = a_{ji}$. Pela primeira parte do lema, podemos obter λ^* e v tais que $A^t v = \lambda^* v$, ou seja,

$$\sum_{j=1}^n a_{ji}v_j = \lambda^* v_i$$

e $v_i > 0$. Além disso,

$$\lambda \langle u, v \rangle = \langle Au, v \rangle = \langle u, A^t v \rangle = \lambda^* \langle u, v \rangle$$

o que implica $\lambda = \lambda^*$.

□

Proposição 6.4.2 *Se p_1, \dots, p_k e q_1, \dots, q_k são distribuições de probabilidade, com $p_i > 0$, $i = 1, \dots, k$, então*

$$\sum_{i=1}^k q_i \log \frac{q_i}{p_i} \geq 0$$

com igualdade valendo se e somente se $p_i = q_i$, $i = 1, \dots, k$ (por convenção, escrevemos $0 \log 0 = 0$).

Prova a desigualdade é evidente no caso em que $p_i = q_i$ porque $\log 1 = 0$. No caso geral, observe que a função $\varphi(x) = -x \log x$ é estritamente côncava, pois $\varphi''(x) < 0$. Como

$$\varphi\left(\sum_{i=1}^k p_i \frac{q_i}{p_i}\right) = \varphi\left(\sum_{i=1}^k q_i\right) = \varphi(1) = 0,$$

então

$$\varphi\left(\sum_{i=1}^k q_i\right) > \sum_{i=1}^k p_i \varphi\left(\frac{q_i}{p_i}\right) = \sum_{i=1}^k p_i \frac{-q_i}{p_i} \log \frac{q_i}{p_i} \Rightarrow \sum_{i=1}^k q_i \log \frac{q_i}{p_i} > 0.$$

□

Uma função f é dita **T -invariante** se $f(T(\omega)) = f(\omega)$. Diremos que uma propriedade vale em **quase toda parte** (q.t.p.), ou com **probabilidade 1**, se o conjunto dos elementos onde não vale a propriedade tem medida nula.

Proposição 6.4.3 (Teorema Ergódico de Birkhoff) [42] *Seja T uma transformação que preserva medida em $(\Omega, \mathcal{F}, \mu)$. Seja f função mensurável e integrável. Então existe K , $\mu(K) = 1$ tal que para todo $\omega \in K$,*

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^n f(T^{k-1}\omega) = \hat{f}(\omega)$$

(i.e., o limite existe q.t.p.), onde \hat{f} é T -invariante e vale que

$$\int \hat{f} d\mu = \int f d\mu$$

Se T é ergódica para μ então $\hat{f} = \int f d\mu$ q.t.p.

Um espaço vetorial X com uma topologia \mathcal{T} é um **espaço vetorial topológico** se a soma é uma função contínua de $X \times X$ em X e se a multiplicação por escalar é uma função contínua de $\mathbb{R} \times X$ em X . Um espaço vetorial topológico é dito **localmente convexo** se podemos obter uma base para a topologia formada por conjuntos convexos.

Proposição 6.4.4 (Krein-Milman)[39][6][10] *Seja K um conjunto convexo compacto em um espaço vetorial topológico localmente convexo. Então a intersecção de todos os conjuntos convexos fechados contendo os pontos extremos de K é o próprio K .*

6.4.1 Prova do lema de Wielandt

Aqui provamos o lema de Wielandt. Esta exposição é baseada em [14].

Sejam A e B matrizes reais retangulares de mesma dimensão $m \times n$, $A = (a_{ij})$, $B = (b_{ij})$. Escrevemos $A \leq B$ (ou $B \geq A$) se

$$a_{ij} \leq b_{ij} \quad , \quad i = 1, \dots, m \quad , \quad j = 1 \dots n.$$

Em particular, quando $A \geq 0$, diremos que A é não-negativa. Se o sinal de igualdade puder ser omitido em todas as desigualdades acima, escreveremos $A < B$ (analogamente, quando $A > 0$, diremos que A é positiva).

Denotamos por C^+ a matriz módulo C , obtida a partir de C quando todos os seus elementos são trocados pelos seus respectivos módulos.

Diremos que A é **reduzível** se pudermos escrever o conjunto de índices $\{1, \dots, n\}$ como sendo uma união de conjuntos complementares $\{i_1, \dots, i_\mu\}$ e $\{j_1, \dots, j_\nu\}$ (com $\mu + \nu = n$) tais que

$$a_{i_\alpha j_\beta} = 0 \quad , \quad \alpha = 1, \dots, \mu \quad , \quad \beta = 1, \dots, \nu.$$

Caso contrário, diremos que A é **irreduzível**.

Vamos obter outra caracterização para matrizes irreduzíveis. Uma permutação de uma matriz quadrada A significa uma permutação das linhas de A juntamente com a mesma permutação de colunas. Podemos então definir matriz reduzível da seguinte forma: a matriz A será reduzível se existir uma permutação que transforma A em um operador da forma

$$A = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$$

onde B e D são matrizes quadradas. Caso contrário, diremos que A é irreduzível. Um subespaço coordenado ν -dimensional de \mathbb{R}^n é um subespaço de \mathbb{R}^n que possui uma base $\{e_{k_1}, \dots, e_{k_\nu}\}$ ($1 \leq k_1 < k_2 < \dots < k_\nu \leq n$). Existem $\binom{n}{\nu}$ subespaços coordenados ν -dimensionais associados a uma base $\{e_1, \dots, e_n\}$ dada. Assim, podemos definir matriz reduzível de outra forma: uma matriz A é reduzível se e somente se possui um subespaço coordenado invariante ν -dimensional com $\nu < n$.

De agora em diante, a menos que seja especificado, iremos considerar matrizes A não negativas (i.e., $A \geq 0$).

Lema 6.4.5 *Se $A \geq 0$ é uma matriz irredutível de ordem n então*

$$(I + A)^{n-1} > 0$$

Prova É suficiente mostrar que para todo vetor $y \geq 0, y \neq 0$, vale

$$(I + A)^{n-1}y > 0.$$

Provaremos a desigualdade acima se mostrarmos que se $y \geq 0, y \neq 0$, o vetor $z = (I + A)y$ possui menos coordenadas nulas que y . Vamos supor que isso não vale. Como $z = y + Ay$ e $Ay \geq 0$, vale que as coordenadas positivas de y correspondem a coordenadas positivas de z . E segue da expressão acima que z não pode ter mais coordenadas nulas que y . Então y e z tem as mesmas coordenadas nulas. Sem perda de generalidade, suponha que

$$y = (u, 0), \quad z = (v, 0) \quad u > 0, v > 0.$$

(com u e v são colunas de mesma dimensão). Escrevendo

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

temos

$$y + Ay = z \Leftrightarrow \begin{pmatrix} u \\ 0 \end{pmatrix} + \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} u \\ 0 \end{pmatrix} = \begin{pmatrix} v \\ 0 \end{pmatrix}$$

e logo, $A_{21}u = 0$. Como $u > 0$, segue que $A_{21} = 0$, o que contradiz a irredutibilidade de A . Isso prova o lema.

□

Um corolário conhecido é a seguinte definição equivalente para matrizes irredutíveis.

Corolário 6.4.6 *Se $A \geq 0$ é irredutível então para todo i, j existe $p \in \mathbb{N}$ tal que $(a_{ij}^p) > 0$. Além disso, p pode ser escolhido de maneira que $p \leq m - 1$ se $i \neq j$ ou então de maneira que $p \leq m$ se $i = j$, onde m é o grau do polinômio minimal de A .*

Vamos enunciar o teorema de Frobenius.

Teorema 6.4.7 (Frobenius): *Toda matriz irredutível não negativa A de ordem n possui um autovalor positivo r que é uma raiz simples da equação característica, e a ele corresponde um autovetor com entradas positivas. O módulo dos demais autovalores é menor ou igual a r .*

Além disso, se A tem h autovalores $\lambda_1 = r, \lambda_2, \dots, \lambda_h$ de módulo r então esses valores são todos distintos, são raízes da equação

$$\lambda^h - r^h = 0,$$

e o espectro de A é invariante por rotação de ângulo $2\pi/h$.

Este é um importante teorema cuja demonstração pode ser encontrada em [14]. Estamos interessados em uma caracterização de matrizes complexas (limitadas em um certo sentido por uma matriz irredutível) em termos dos seus autovalores e que será dada de forma mais precisa a seguir (lema 6.4.10). Tal formulação está relacionada com o teorema de Frobenius.

Dado um vetor real $x = (x_1, \dots, x_n)$, $x \geq 0$ fixado, definimos

$$r_x = \min_{i \in \{1, \dots, n\}} \frac{(Ax)_i}{x_i}$$

onde

$$(Ax)_i = \sum_{j=1}^n a_{ij}x_j, \quad i = 1, \dots, n.$$

Nesta definição de mínimo, excluimos os valores de i onde $x_i = 0$. Segue diretamente da definição que $r_x \geq 0$ e é o maior número η tal que

$$\eta x \leq Ax.$$

Lema 6.4.8 *Existe $z \geq 0$ tal que o valor máximo r da função r_x é atingido, ou seja:*

$$r = r_z = \max_{x \geq 0} r_x = \max_{x \geq 0} \min_{i \in \{1, \dots, n\}} \frac{(Ax)_i}{x_i}. \quad (6.17)$$

Prova Pela definição de r_x , segue que ao multiplicarmos um vetor $x \geq 0$, $x \neq 0$, por um número λ , o valor de r_x não se altera. Então, para calcularmos o máximo de r_x , podemos nos restringir ao conjunto fechado

$$M = \{x : x \geq 0, \sum_{i=1}^n x_i^2 = 1\}.$$

Se a função r_x fosse contínua em M , poderíamos obter um máximo. Entretanto, esta função pode ser descontínua nos pontos de fronteira de M onde uma das coordenadas do vetor se anula. Sendo assim, vamos considerar o conjunto

$$N = \{y : y = (I + A)^{n-1}x, x \in M\}.$$

Este conjunto é fechado e limitado, e pelo lema 6.4.5, consiste apenas de vetores positivos.

Multiplicando a desigualdade

$$r_x x \leq Ax$$

por $(I + A)^{n-1}$, obtemos

$$r_x y \leq Ay,$$

onde $y = (I + A)^{n-1}x$. Então, pela definição de r_y , obtemos

$$r_x \leq r_y.$$

Assim, ao calcularmos o máximo de r_x , podemos nos restringir ao conjunto N que consiste apenas de vetores positivos. E em N , que é fechado e limitado, a função r_x é contínua e portanto assume um valor máximo para algum vetor $z \geq 0$.

□

Diremos que um vetor $z \geq 0$ tal que $r_z = r$ é um vetor extremal.

Lema 6.4.9 *O valor r definido no lema anterior é positivo, e é um autovalor de A (é o valor r mencionado no teorema de Frobenius). Todo vetor extremal z é positivo e é um autovetor de A para o autovalor r . Isto é,*

$$r > 0, \quad z > 0, \quad Az = rz.$$

Prova Seja $u = (1, 1, \dots, 1)$. Então

$$r_u = \min_{i \in \{1, \dots, n\}} \sum_{k=1}^n a_{ik}.$$

Daí, $r_u > 0$, porque nenhuma coluna de uma matriz irredutível pode ser formada por zeros apenas. Logo, como $r \geq r_u$, obtemos $r > 0$.

Agora, considere

$$x = (I + A)^{n-1}z,$$

onde z é um vetor extremal. Pelo lema 6.4.5, $x > 0$. Suponha por absurdo que $Az - rz \neq 0$. Então obtemos que

$$Az - rz > 0 \Rightarrow (I + A)^{n-1}(Az - rz) > 0 \Rightarrow Ax - rx > 0.$$

A última desigualdade contradiz a definição de r porque ela implicaria que $Ax - (r + \varepsilon)x > 0$ para $\varepsilon > 0$ suficientemente pequeno, ou seja, obteríamos $r_x \geq r + \varepsilon > r$. Logo, $Az = rz$. Então

$$0 < x = (I + A)^{n-1}z = (1 + r)^{n-1}z,$$

e portanto, $z > 0$.

□

Lema 6.4.10 seja $A = (a_{ij})$ matriz irredutível não negativa de ordem n e $C = (c_{ij})$ uma matriz quadrada complexa de ordem n . Suponha que

$$C^+ \leq A. \quad (6.18)$$

Então para todo autovalor γ de C vale que

$$|\gamma| \leq r, \quad (6.19)$$

onde r é o maior autovalor de A . Além disso, vale a igualdade se e somente se

$$C = e^{i\varphi} D A D^{-1}, \quad (6.20)$$

onde $e^{i\varphi} = \gamma/r$ e D é uma matriz diagonal cujos elementos não nulos tem módulo 1 (vale que $D^+ = I$).

Prova Considere y um autovetor de C correspondente ao autovalor γ :

$$C y = \gamma y, \quad \gamma \neq 0. \quad (6.21)$$

Como $C^+ \leq A$, vale que $C^+ y^+ \leq A y^+$. Além disso,

$$\begin{aligned} C y &= \left(\sum_i c_{1i} y_i, \dots, \sum_i c_{ni} y_i \right), \quad \gamma y = (\gamma y_1, \dots, \gamma y_n) \\ \Rightarrow C^+ y^+ &= \left(\sum_i |c_{1i} y_i|, \dots, \sum_i |c_{ni} y_i| \right), \quad |\gamma| y^+ = (|\gamma y_1|, \dots, |\gamma y_n|) \end{aligned}$$

De $C y = \gamma y$, obtemos:

$$|\gamma y_j| = \left| \sum_{i=1}^n c_{ji} y_i \right| \leq \sum_{i=1}^n |c_{ji} y_i|, \quad j = 1, \dots, n$$

o que implica

$$|\gamma| y^+ \leq C^+ y^+. \quad (6.22)$$

Logo,

$$|\gamma| y^+ \leq C^+ y^+ \leq A y^+ \quad (6.23)$$

pelo que vimos antes. Como por definição r_{y^+} é o maior número tal que

$$r_{y^+} y^+ \leq A y^+,$$

concluimos de (6.23) que

$$|\gamma| \leq r_{y^+} \leq r.$$

(a segunda desigualdade acima segue de (6.17)). Isso prova a desigualdade (6.19) que enunciamos.

Analisemos agora o caso em que $|\gamma| = r$. A última desigualdade nos fornece $r_{y^+} = r = \max_{x \geq 0} r_x$. Segue daí que y^+ é um vetor extremal para A , $y^+ > 0$ e que y^+ é autovetor de A para o autovalor r . Então $ry^+ = Ay^+$ e a desigualdade (6.23) se transforma em

$$|\gamma|y^+ = ry^+ = Ay^+ = C^+y^+. \quad (6.24)$$

Então como $C^+ \leq A$ e $y^+ > 0$, segue de $Ay^+ = C^+y^+ \Leftrightarrow (A - C^+)y^+ = 0$ que

$$C^+ = A. \quad (6.25)$$

Seja $y = (y_1, \dots, y_n)$, onde $y_j = |y_j|e^{i\varphi_j}$, $j = 1, \dots, n$. Defina a seguinte matriz diagonal

$$D = \{e^{i\varphi_1}, \dots, e^{i\varphi_n}\}.$$

Então, vale que

$$y = Dy^+.$$

Escrevendo $\gamma = re^{i\varphi}$ e substituindo a expressão acima em (6.21), obtemos

$$Fy^+ = ry^+, \quad (6.26)$$

onde

$$F = e^{-i\varphi}D^{-1}CD. \quad (6.27)$$

Comparando (6.24) com (6.26), obtemos

$$Fy^+ = C^+y^+ = Ay^+. \quad (6.28)$$

Mas por (6.27) e (6.25),

$$F^+ = C^+ = A.$$

Logo, obtemos a partir de (6.28) que

$$Fy^+ = F^+y^+.$$

Como $y^+ > 0$, afirmamos que isso implica que

$$F = F^+$$

Provamos esta última afirmação no caso de dimensão 2. Seja

$$F = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Rightarrow F^+ = \begin{pmatrix} |a| & |b| \\ |c| & |d| \end{pmatrix}$$

Então

$$Fy^+ = F^+y^+ \Rightarrow (F - F^+)y^+ = 0 \Rightarrow \begin{pmatrix} a - |a| & b - |b| \\ c - |c| & d - |d| \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Daí escrevendo $a = x + iy$, $b = z + iw$,

$$(a - |a|)y_1 + (b - |b|)y_2 = 0 \Leftrightarrow (x + iy - |a|)y_1 + (z + iw - |b|)y_2 = 0$$

o que implica, em particular, que

$$(x - |a|)y_1 + (z - |b|)y_2 = 0 \Leftrightarrow (x - \sqrt{x^2 + y^2})y_1 + (z - \sqrt{z^2 + w^2})y_2 = 0$$

Os termos entre parenteses na última expressão são menores ou iguais a zero, claramente. Mas não podem ser estritamente menores que zero (pois o fato de que $y_1, y_2 > 0$ implicaria que a expressão acima é negativa). Logo,

$$x - \sqrt{x^2 + y^2} = 0 \Rightarrow x = |a| \text{ e } y = 0 \Rightarrow a = |a|$$

$$z - \sqrt{z^2 + w^2} = 0 \Rightarrow z = |b| \text{ e } w = 0 \Rightarrow b = |b|$$

Analogamente para c e d . Logo, $F = F^+$.

Mas

$$F = F^+ \Leftrightarrow e^{-i\varphi} D^{-1} C D = A.$$

Logo,

$$C = e^{i\varphi} D A D^{-1}$$

□

Capítulo 7

C^* -álgebras e cadeias de Markov quânticas

Faremos aqui uma descrição breve de cadeias de Markov quânticas no contexto de álgebra de operadores. Essa formulação é baseada na construção original feita por L. Accardi em [1] e [2], por exemplo.

A vantagem de se estudar cadeias de Markov quânticas no contexto de álgebras reside no fato de que podemos fazer uso do instrumental construído nesse ambiente para se tentar descobrir estados de equilíbrio, um problema importante em mecânica estatística quântica. Discutiremos brevemente a relação entre os estados KMS e os estados de Gibbs, mencionados no capítulo anterior.

7.1 C^* -álgebras

Definição 7.1.1 *Uma álgebra A sobre \mathbb{C} é um espaço vetorial complexo equipado com uma operação bilinear e associativa $\bullet : A \times A \rightarrow A$, dita multiplicação. Para $a, b \in A$, denotaremos $\bullet(a, b)$ simplesmente por ab .*

Definição 7.1.2 *Uma álgebra normada é uma álgebra A sobre \mathbb{C} equipada com uma função norma $a \in A \mapsto \|a\| \in \mathbb{R}$, que torna A um espaço normado, ou seja, para $a, b \in A$ e $\lambda \in \mathbb{C}$, temos*

1. $\|a\| \geq 0$, e $\|a\| = 0 \Rightarrow a = 0$
2. $\|\lambda a\| = |\lambda| \|a\|$, onde $|\lambda|$ denota o módulo do número complexo λ
3. $\|a + b\| \leq \|a\| + \|b\|$,

e que além disso obedece a seguinte propriedade:

$$4. \|ab\| \leq \|a\|\|b\|$$

Naturalmente, podemos nos referir à distância entre dois elementos de uma álgebra normada, bastando para isso considerar a métrica induzida pela norma.

Definição 7.1.3 *Uma álgebra de Banach é uma álgebra normada completa.*

Definição 7.1.4 *Seja A uma álgebra de Banach. Uma involução em A é uma função $*$: $A \rightarrow A$ tal que para todo $a, b \in A$, $\lambda \in \mathbb{C}$, e denotando $c^* := *(c)$, $\forall c \in A$, temos*

1. $(a + b)^* = a^* + b^*$
2. $(\lambda a)^* = \bar{\lambda}a^*$
3. $(ab)^* = b^*a^*$
4. $(a^*)^* = a$
5. $\|a^*\| = \|a\|$

Definição 7.1.5 *Uma C^* -álgebra é uma álgebra de Banach equipada com uma involução para o qual vale*

$$\|a^*a\| = \|a\|^2, \quad \forall a \in A.$$

Por exemplo, a álgebra M_n das matrizes de ordem n sobre \mathbb{C} é uma C^* -álgebra se considerarmos as matrizes como sendo operadores no espaço euclidiano \mathbb{C}^n e se tomarmos a norma de operadores $\|\cdot\|$ sobre matrizes. A involução é dada pela matriz transposta conjugada.

7.2 Cadeias de Markov quânticas

A descrição de cadeias de Markov quânticas dada aqui segue [1] e [28], que é uma construção conhecida em álgebra de operadores. Nesse contexto, um **estado em \mathcal{B}** é simplesmente um funcional $\phi : \mathcal{B} \rightarrow \mathbb{C}$ tal que $\phi(1) = 1$.

Seja $\mathcal{B} = \mathcal{B}(H)$ uma C^* -álgebra para um certo espaço de Hilbert fixado H e seja

$$\mathcal{A} := \mathcal{B} \otimes \mathcal{B} \otimes \mathcal{B} \otimes \dots$$

onde \otimes é o C^* -produto tensorial induzido pelo produto tensorial usual de espaços de Hilbert.

Definição 7.2.1 Uma aplicação bilinear $\mathcal{E} : \mathcal{B} \otimes \mathcal{B} \rightarrow \mathcal{B}$ é dita **esperança de transição** se for completamente positiva e preservar a identidade.

Definição 7.2.2 (Cadeia de Markov quântica associada a um estado) Um estado ϕ em \mathcal{A} é uma **cadeia de Markov quântica** se existir um estado ϕ_0 em \mathcal{B} e uma esperança de transição \mathcal{E} tais que

$$\begin{aligned} & \phi(a_0 \otimes a_1 \otimes \cdots \otimes a_n \otimes 1 \otimes \cdots) \\ &= \phi_0(\mathcal{E}(a_0 \otimes \mathcal{E}(a_1 \otimes \cdots \otimes \mathcal{E}(a_{n-1} \otimes \mathcal{E}(a_n \otimes 1)) \cdots))), \end{aligned} \quad (7.1)$$

para quaisquer $a_i \in \mathcal{B}$, $i = 0, 1, \dots, n$.

Prova-se em [2] que toda esperança de transição tem a forma

$$\mathcal{E}(x) = \text{tr}_2\left(\sum_j K_j^* x K_j\right) \quad (7.2)$$

onde $x, K_j \in \mathcal{B} \otimes \mathcal{B}$, $j = 1, 2, \dots$ e tr_2 é o traço parcial com respeito ao segundo fator. O índice j percorrerá apenas um subconjunto finito de \mathbb{N} se H for de dimensão finita, e esse será o caso considerado aqui.

Seja $\dim(H) = d < \infty$ e seja $\{e_i\}_{n=1}^d$ uma base ortonormal para H . Neste caso, $\mathcal{B}(H) \cong M_d$, onde M_d é a álgebra das matrizes $d \times d$. Para simplificar, iremos supor $\mathcal{B}(H) = M_d$.

Para cada $K \in \mathcal{B} \otimes \mathcal{B}$, temos a expressão

$$K = \sum_{n,n'} |e_n\rangle\langle e_{n'}| \otimes K_{n,n'}, \quad (7.3)$$

onde $|e_n\rangle\langle e_{n'}|$ é a matriz cuja posição (n, n') é igual a 1, e as outras posições são nulas. Denotaremos por \mathcal{D} a sub-álgebra diagonal de M_d correspondente à base $\{e_i\}_{n=1}^d$.

Mostraremos agora como uma cadeia de Markov clássica $\{X_n\}_{n=1}^\infty$, com valores no conjunto $\{1, \dots, d\}$, em um espaço de probabilidade (Ω, \mathbf{P}) com distribuição inicial p e matriz de transição $P = (p_{ij})$ pode ser vista como sendo uma cadeia de Markov quântica. Procedemos da seguinte maneira:

1. Defina

$$K_{1,n,n} := \begin{pmatrix} \sqrt{p_{n,1}} & 0 & \cdots & 0 \\ 0 & \sqrt{p_{n,2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{p_{n,d}} \end{pmatrix} \quad (7.4)$$

$$K_{1,n,n'} := 0 \text{ , se } n \neq n' \text{ ,} \quad (7.5)$$

$$K_{m,n,n'} := 0 \text{ , } \forall m \geq 2. \quad (7.6)$$

2. Defina ϕ_0 em $\mathcal{D} \subset \mathcal{B}$ por

$$\phi_0(\cdot) := \text{tr} \left[\begin{pmatrix} p_1 & 0 & \cdots & 0 \\ 0 & p_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_d \end{pmatrix} \cdot \right] \quad (7.7)$$

É fácil verificar que

$$\mathcal{E}_1(x) := \text{tr}_2(K_1^* x K_1), \quad (7.8)$$

onde $K_1 := \sum_n |e_n\rangle\langle e_n| \otimes K_{1,n,n'}$, é uma esperança de transição, e é tal que

$$\mathcal{E}_1(f \otimes g) = f \mathcal{E}_1(1 \otimes g) \text{ , } \forall f, g \in \mathcal{D} \quad (7.9)$$

e vale que ϕ_0 é um estado em M_d . Portanto (ϕ_0, \mathcal{E}_1) é uma cadeia de Markov quântica e além disso tal cadeia quântica restrita à sub-álgebra diagonal \mathcal{D} é simplesmente a cadeia de Markov clássica dada inicialmente.

Observe que a escolha de (ϕ_0, \mathcal{E}) não é única. De fato, podemos trocar a matriz

$$\begin{pmatrix} p_1 & 0 & \cdots & 0 \\ 0 & p_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_d \end{pmatrix}$$

por qualquer outra matriz densidade ω_0 que possua na diagonal os elementos $\{p_1, \dots, p_d\}$. Além disso, podemos trocar $K_{1,n,n}$ por qualquer $K_{n,n}$ que possua a propriedade

$$K_{n,n} K_{n,n}^*(j, j) = p_{n,j}, \quad \forall j = 1, \dots, d.$$

Denote por

$$K_0 := \sum_n |e_n\rangle\langle e_n| \otimes K_{n,n}$$

e defina

$$\mathcal{E}_0(\cdot) := \text{tr}_2(K_0^* \cdot K_0).$$

Então, $(\text{tr}(\omega_0 \cdot), \mathcal{E}_0(\cdot))$ também é uma cadeia de Markov quântica obtida a partir da mesma cadeia clássica $(p, P = (p_{ij}))$ e além disso, restringindo as

duas cadeias quânticas $(tr(\omega_0 \cdot), \mathcal{E}_0(\cdot))$ e (ϕ_0, \mathcal{E}_1) ao conjunto \mathcal{D} , obtemos a mesma cadeia de Markov clássica $(p, P = (p_{ij}))$.

Observação Note que se

$$\mathcal{E}(\mathcal{D} \otimes \mathcal{D}) \subset \mathcal{D}, \quad (7.10)$$

então podemos considerar a matriz densidade inicial ω_0 como sendo diagonal, já que os elementos fora da diagonal não são usados neste caso.

Proposição 7.2.3 [28] *Uma esperança de transição \mathcal{E} leva $\mathcal{D} \otimes \mathcal{D}$ em \mathcal{D} se e somente se para cada $r, r', n \in \{1, \dots, d\}$, com $r \neq r'$, a matriz $\sum_j K_{j,n,r} K_{j,n,r'}^*$ é tal que os elementos de sua diagonal são iguais a zero.*

Definição 7.2.4 (Cadeia de Markov quântica associada a um processo estocástico clássico) *Dizemos que um processo estocástico clássico $\{X_n\}_{n=0}^\infty$, que toma valores em $\{1, \dots, d\}$ em um espaço de probabilidade (Ω, \mathbf{P}) é uma cadeia de Markov quântica (ϕ_0, \mathcal{E}) se \mathcal{E} satisfaz (7.10) e se as distribuições conjuntas são as mesmas, ou seja, para cada $n \in \mathbb{N}$, $i_k \in \{1, \dots, d\}$, $k = 1, \dots, n$, temos*

$$\begin{aligned} & P(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) \\ &= \phi(|e_{i_0}\rangle\langle e_{i_0}| \otimes |e_{i_1}\rangle\langle e_{i_1}| \otimes \dots \otimes |e_{i_n}\rangle\langle e_{i_n}|) \\ &= \phi_0(\mathcal{E}(|e_{i_0}\rangle\langle e_{i_0}| \otimes \mathcal{E}(|e_{i_1}\rangle\langle e_{i_1}| \otimes \dots \otimes \mathcal{E}(|e_{i_{n-1}}\rangle\langle e_{i_{n-1}}| \otimes \mathcal{E}(|e_{i_n}\rangle\langle e_{i_n}| \otimes 1)) \dots))) \end{aligned} \quad (7.11)$$

Definição 7.2.5 (Cadeia de Markov clássica associada a uma cadeia quântica) *Uma cadeia de Markov quântica (ϕ_0, \mathcal{E}) é uma cadeia de Markov clássica na sub-álgebra diagonal \mathcal{D} se (7.10) vale e se*

$$\mathcal{E}(a \otimes b) = a \mathcal{E}(1 \otimes b), \quad \forall a, b \in \mathcal{D}. \quad (7.12)$$

Teorema 7.2.6 *Seja $\{X_n\}_{n=1}^\infty$ um processo estocástico clássico $\{X_n\}_{n=0}^\infty$, que toma valores em $\{1, \dots, n\}$ em um espaço de probabilidade (Ω, \mathbf{P}) . Então $\{X_n\}_{n=1}^\infty$ é uma cadeia de Markov quântica se e somente se existe uma medida de probabilidade p_0 em $\{1, \dots, d\}$ e uma matriz cúbica (de 3 índices) $T = (T_{j,i,k})$ com as seguintes propriedades:*

$$T_{j,i,k} \geq 0, \quad \forall j, i, k \in \{1, \dots, d\} \quad (7.13)$$

$$\sum_{i,k} T_{j,i,k} = 1, \quad \forall j \in \{1, \dots, d\} \quad (7.14)$$

$$\mathbf{P}(X_0 = j) = \sum_{i,k} p_0(i) T_{j,i,k}, \quad \forall j \in \{1, \dots, d\}, \quad (7.15)$$

e tal que a distribuição conjunta é dada por

$$\begin{aligned} & \mathbf{P}(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) \\ &= \sum_{j, j_0, j_1, \dots, j_n} p_0(j) T_{j, i_0, j_0} T_{j_0, i_1, j_1} \cdots T_{j_{n-2}, i_{n-1}, j_{n-1}} T_{j_{n-1}, i_n, j_n}. \end{aligned} \quad (7.16)$$

Em [28], temos ainda a prova do seguinte:

Teorema 7.2.7 *Seja $\{X_n\}_{n=1}^\infty$ um processo estocástico clássico $\{X_n\}_{n=0}^\infty$, que toma valores em $\{1, \dots, n\}$ em um espaço de probabilidade (Ω, \mathbf{P}) . Então temos que $\{X_n\}_{n=1}^\infty$ é uma cadeia de Markov quântica se e somente se existir uma cadeia de Markov clássica $\{(Z_n, Y_n)\}_{n=1}^\infty$ tomando valores em $\{1, \dots, d\}^2$ em um espaço de probabilidade (\mathbf{P}', Ω') com as seguintes propriedades:*

1. *A probabilidade de transição*

$$p_{(j,i),(j',i')} := \mathbf{P}'((Z_n, Y_n) = (j', i') | (Z_{n-1}, Y_{n-1}) = (j, i)) \quad (7.17)$$

é independente de i .

2. *Existe uma distribuição de probabilidade p_0 em $\{1, \dots, d\}$ e uma distribuição inicial para $\{(Z_n, Y_n)\}_{n=0}^\infty$ dada por*

$$\mathbf{P}'((Z_0, Y_0) = (j, i)) = \sum_k p_0(k) T_{k,i,j} \quad (7.18)$$

(onde $T_{k,i,j}$ foi determinado no teorema 7.2.6), tais que

$$\mathbf{P}(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = \mathbf{P}'(Y_0 = i_0, Y_1 = i_1, \dots, Y_n = i_n). \quad (7.19)$$

Corolário 7.2.8 *Se a cadeia de Markov quântica (ϕ_0, \mathcal{E}) em $\bigotimes_{\mathbb{N}} M_d$ tem a propriedade de que \mathcal{E} leva $\mathcal{D} \otimes \mathcal{D}$ em \mathcal{D} então o processo estocástico clássico obtido pela restrição de (ϕ_0, \mathcal{E}) à sub-álgebra diagonal \mathcal{D} é estocasticamente equivalente à segunda componente de uma cadeia de Markov clássica*

$$\{(Z_n, Y_n)\}_{n=0}^\infty,$$

que toma valores em $\{1, \dots, d\}$, e que possui as propriedades 1 e 2 do teorema 7.2.7.

Teorema 7.2.9 *A cadeia de Markov quântica (ϕ_0, \mathcal{E}) é uma cadeia de Markov clássica sobre a sub-álgebra diagonal se e somente se os operadores K_j , dados por*

$$K_j = \sum_{n,n'} |e_n\rangle\langle e_{n'}| \otimes K_{j,n,n'},$$

tem a forma

$$K_j = \sum_n |e_n\rangle\langle e_n| \otimes K_{j,n,n}, \quad \forall j = 1, \dots, d. \quad (7.20)$$

Observação A fórmula (7.20) significa que cada um dos K_j é diagonal em blocos de ordem d^2 , ou seja,

$$\begin{pmatrix} K_{j,1,1} & 0 & \cdots & 0 \\ 0 & K_{j,2,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & K_{j,d,d} \end{pmatrix},$$

onde cada $K_{j,l,l}$, $l = 1, \dots, d$, é uma matriz de ordem d .

7.3 CAR C^* -álgebras e estados KMS

Fazemos aqui uma breve introdução às CAR C^* -álgebras e aos estados KMS. Uma referência básica para o assunto é [8].

Definição 7.3.1 *Uma $*$ -álgebra \mathcal{A} é uma álgebra associativa sobre os complexos, munida de uma involução que é um antiautomorfismo antilinear, ou seja,*

1. $(u^*)^* = u$
2. $(u + v)^* = u^* + v^*$
3. $(cv)^* = \bar{c}v^*$
4. $(uv)^* = v^*u^*$,

para todo $u, v \in \mathcal{A}$, $c \in \mathbb{C}$.

Um homomorfismo entre álgebras $f : \mathcal{A} \rightarrow \mathcal{B}$ é um **$*$ -homomorfismo** se for compatível com as involuções de \mathcal{A} e \mathcal{B} , ou seja, se

$$f(a^*) = f(a)^*, \quad \forall a \in \mathcal{A}.$$

E f será um **$*$ -isomorfismo** se for uma aplicação bijetora.

Definição 7.3.2 Uma C^* -álgebra \mathcal{A} tal que seus elementos satisfazem as relações de anticomutatividade canônicas

$$\{a_k, a_j\} = 0$$

$$\{a_k^*, a_j\} = \delta_{ij},$$

onde a_k , etc. são elementos de \mathcal{A} , e $\{\cdot, \cdot\}$ é o anticomutador, é dita CAR C^* -álgebra.

O próximo teorema, cuja prova pode ser vista em [8], é a base de estudo de tais álgebras.

Teorema 7.3.3 Seja \mathcal{H} um espaço de Hilbert e sejam \mathcal{H}_i , $i = 1, 2$ duas C^* -álgebras geradas pela identidade 1 e elementos $a_i(v)$, $v \in \mathcal{H}$ satisfazendo:

1. $v \mapsto a_i(v)$ é linear,
2. $\{a_i(v), a_i(w)\} = 0$,
3. $\{a_i(v), a_i(w)^*\} = (v, w)1$,

para todo $v, w \in \mathcal{H}$, $i = 1, 2$. Então existe um único $*$ -isomorfismo $\alpha : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ tal que

$$\alpha(a_1(v)) = a_2(v),$$

para todo $v \in \mathcal{H}$. Portanto, existe uma única C^* -álgebra $\mathcal{A} = \mathcal{A}(\mathcal{H}) = \mathcal{A}(\overline{\mathcal{H}})$, a menos de $*$ -isomorfismo, satisfazendo as relações de anticomutatividade canônicas sobre \mathcal{H} . Além disso, se \mathcal{H} tem dimensão n , temos que $\mathcal{A}(\mathcal{H})$ é isomorfa à C^* -álgebra das matrizes complexas $2^n \times 2^n$.

Seja \mathcal{U} a CAR-álgebra com geradores a_i e respectivos adjuntos a_i^* , $i \in I$, onde o conjunto de índices I é discreto, enumerável e totalmente ordenado, contendo possivelmente um menor elemento j_- e/ou um maior elemento j_+ . Se I não contém nem j_+ e nem j_- então podemos fazer a identificação $I \sim \mathbb{Z}$. Se contém apenas j_+ então $I \sim \mathbb{Z}_-$; se contém apenas j_- então $I \sim \mathbb{Z}_+$.

Os geradores $\{a_j, a_j^+\}_{j \in I}$ satisfazem as relações

$$\{a_j^+, a_k\} = \delta_{jk}, \quad \{a_j, a_k\} = \{a_j^+, a_k^+\} = 0, \quad j, k \in I$$

O automorfismo de paridade de \mathcal{U} é denotado por Θ . Para qualquer subconjunto $\Lambda \subset I$, a C^* -sub-álgebra de \mathcal{U} gerada por a_j, a_j^+ , para $j \in \Lambda$ é denotada por \mathcal{U}_Λ . Sabemos que \mathcal{U} é uma álgebra \mathbb{Z}_2 -graduada, com Θ sendo o automorfismo da graduação. Sabemos também que a CAR-álgebra é isomorfa ao produto C^* -tensorial infinito $\bigotimes_I M_2(\mathbb{C})$.

A fim de obter estados periódicos ou invariantes por translação, consideramos apenas esperanças quase-condicionais (ver definição a seguir) Θ -invariantes, a menos que especificado em contrário.

Uma notação comum em álgebra de operadores é a seguinte. Se $\mathcal{A}_n, \mathcal{B}$ são álgebras, $n \in \mathbb{N}$, escrevemos

$$\mathcal{A}_n \otimes \mathcal{B} \otimes \mathcal{A}_{(n)} := \mathcal{A}_1 \otimes \mathcal{A}_2 \otimes \cdots \otimes \mathcal{A}_{n-1} \otimes \mathcal{B} \otimes \mathcal{A}_{n+1} \otimes \mathcal{A}_{n+2} \otimes \cdots$$

Vale uma notação análoga para operadores.

Definição 7.3.4 *Sejam $\mathcal{W} \subset \mathcal{V} \subset \mathcal{U}$ C^* -álgebras com unidade. Uma **esperança quase-condicional** com respeito à essa tripla é uma aplicação $E : \mathcal{U} \rightarrow \mathcal{V}$ tal que é completamente positiva, preserva a identidade e*

$$E(wu) = wE(u), \quad u \in \mathcal{U}, w \in \mathcal{W}$$

Definição 7.3.5 *Um estado ϕ em \mathcal{U} é dito **estado de Markov** se para cada $j_- \leq j \leq j_+$ existe uma esperança quase-condicional E_n com respeito à tripla $\mathcal{U}_{(n-1)} \subset \mathcal{U}_n \subset \mathcal{U}_{(n+1)}$ satisfazendo*

$$\phi_{(n+1)} \circ E_n = \phi_n$$

$$E_n(\mathcal{U}_{[n, n+1]}) = \mathcal{U}_{\{n\}}.$$

Definição 7.3.6 *Uma **esperança condicional** (de Umegaki) $E : \mathcal{A} \rightarrow \mathcal{B} \subset \mathcal{A}$ é uma projeção de norma 1 da C^* -álgebra \mathcal{A} em uma C^* -sub-álgebra \mathcal{B} (com a mesma identidade I).*

Quando \mathcal{A} é uma álgebra de matrizes, a estrutura de tal esperança condicional é bem conhecida [3]. A prova da seguinte proposição pode ser vista em [3]:

Proposição 7.3.7 *Seja ϕ um estado da CAR-álgebra. São equivalentes:*

1. ϕ é um estado de Markov.
2. as propriedades listadas na definição 7.3.5 são satisfeitas se trocarmos as esperanças quase-condicionais E_n por esperanças condicionais de Umegaki \mathcal{E}_n .
3. Para cada $j < j_+$, existe uma esperança condicional $\mathcal{E}_n : \mathcal{U} \rightarrow \mathcal{R}(\mathcal{E}_n) \subset \mathcal{U}_n$ satisfazendo

$$\begin{aligned} \phi \circ \mathcal{E}_n &= \phi \\ \mathcal{E}_n(\mathcal{U}_{(n)}) &\subset \mathcal{U}_{\{n\}} \end{aligned} \tag{7.21}$$

4. As propriedades listadas no item 3 são satisfeitas se trocarmos as esperanças condicionais \mathcal{E}_n por esperanças quase-condicionais E_n .

Na prova desta proposição, consideramos a restrição $e_n := E_n|_{\mathcal{U}_{[n, n+1]}}$, que é uma aplicação completamente positiva, que preserva a identidade. Fazendo uso de um limite

$$\varepsilon_n := \lim_k \frac{1}{k} \sum_{t=0}^{k-1} (e_n)^t,$$

obtemos uma esperança condicional que deixa invariante o estado ϕ . Com isso, chegamos ao seguinte lema [3]:

Lema 7.3.8 *Seja ϕ um estado de Markov na CAR-álgebra e $\{\varepsilon_j\}_{j-\leq j \leq j+}$ a seqüência associada de esperanças condicionais. Então*

$$\phi(x_k \cdots x_l) = \phi((\varepsilon_k(x_k \varepsilon_{k+1}(x_{k+1} \cdots \varepsilon_{l-1}(x_{l-1} x_l) \cdots)))$$

para todo $k, l \in I$ com $k < l$ e $x_k x_{k+1} \cdots x_{l-1} x_l$ qualquer gerador linear de $\mathcal{U}_{[k, l]}$.

Definição 7.3.9 *Seja \mathcal{A} uma C^* -álgebra, $\beta \in \mathbb{R}$ e seja $\alpha : \mathbb{R} \rightarrow \text{Aut}(\mathcal{A})$ uma ação fortemente contínua. Um estado ϕ satisfaz a **condição KMS** para α à temperatura inversa β se*

$$\phi(v \alpha_{i\beta}(u)) = \phi(uv)$$

para todo $u, v \in \mathcal{A}$, com u inteira para α .

Em teoria quântica de campos descreve-se um sistema físico a partir de uma C^* -álgebra \mathcal{A} com unidade. Os elementos autoadjuntos de \mathcal{A} (tais que $u^* = u$) são ditos **observáveis** do sistema. Como vimos antes, um **estado** do sistema é definido com sendo um funcional \mathbb{C} -linear $\phi : \mathcal{A} \rightarrow \mathbb{C}$, com $\phi(uu^*) > 0$, tal que $\phi(1) = 1$. Determinar a existência e a unicidade de estados KMS é importante para se estudar a mecânica estatística de uma teoria quântica de campos. Por exemplo, considerando o espaço dos operadores compactos em um espaço de Hilbert dado, pode-se mostrar que o único estado sobre esse espaço que satisfaz a condição KMS acima com respeito ao valor β é o estado de equilíbrio canônico de Gibbs

$$\omega_{\beta\mu}(A) = \frac{\text{tr}(e^{-\beta K} A)}{\text{tr}(e^{-\beta K})},$$

onde $K = H - \mu N$, H é o hamiltoniano autoadjunto e N é o operador número, respectivamente. Para mais detalhes, veja [8].

7.4 Moeda quântica revisitada

Consideremos aqui a moeda quântica (ver seção 8.2). Tal sistema é bastante simples em dimensão 2, dependendo de um único parâmetro real $x \in (0, 1)$.

Como antes, sejam $u, v \in \mathbb{C}$ tais que $u + v = 1$, $|u|^2 + |v|^2 = 1$. Vimos que as probabilidades de transição independem do tempo, e mais especificamente que $P_t(0) = |u|^2$, $P_t(1) = |v|^2$. Vamos mostrar que a moeda quântica, vista como uma cadeia de Markov quântica algébrica (i.e., via álgebra de operadores) nos permite obter, como é desejável, estas mesmas probabilidades de transição.

Seguimos a construção feita no início da seção 7.2. Se p_{ij} é a probabilidade de transição do estado i para o estado j , temos que $p_{00} = p_{10} = |u|^2$, $p_{01} = p_{11} = 1 - |u|^2 = |v|^2$ e portanto,

$$K_{0,0,0} = K_{0,1,1} = \begin{pmatrix} \sqrt{p_{00}} & 0 \\ 0 & \sqrt{p_{01}} \end{pmatrix} = \begin{pmatrix} \sqrt{p_{10}} & 0 \\ 0 & \sqrt{p_{11}} \end{pmatrix} = \begin{pmatrix} |u| & 0 \\ 0 & |v| \end{pmatrix}$$

Para ilustrar os cálculos envolvidos mostremos, por exemplo, que

$$P(X_0 = 0) = \phi(|e_0\rangle\langle e_0|) = |u|^2,$$

e

$$P(X_0 = 0, X_1 = 1) = \phi(|e_0\rangle\langle e_0| \otimes |e_1\rangle\langle e_1|) = |u|^2|v|^2.$$

Temos:

$$\phi(|e_0\rangle\langle e_0|) = \phi_0(\mathcal{E}(|e_0\rangle\langle e_0| \otimes 1))$$

Mas

$$\begin{aligned} \mathcal{E}(|e_0\rangle\langle e_0| \otimes 1) &= \text{tr}_2(K_1^*(|e_0\rangle\langle e_0| \otimes 1)K_1) \\ &= \text{tr}_2((|e_0\rangle\langle e_0| \otimes K_{0,0,0})^*(|e_0\rangle\langle e_0| \otimes 1)(|e_0\rangle\langle e_0| \otimes K_{0,0,0})) \\ &= \text{tr}_2((|e_0\rangle\langle e_0| \otimes K_{0,0,0})(|e_0\rangle\langle e_0| \otimes K_{0,0,0})) \\ &= \text{tr}_2(|e_0\rangle\langle e_0| \otimes K_{0,0,0}^2) \\ &= |e_0\rangle\langle e_0| \text{tr} \begin{pmatrix} |u|^2 & 0 \\ 0 & |v|^2 \end{pmatrix} = |e_0\rangle\langle e_0|. \end{aligned}$$

Logo,

$$\begin{aligned} \phi(|e_0\rangle\langle e_0|) &= \phi_0(|e_0\rangle\langle e_0|) \\ &= \text{tr} \left[\begin{pmatrix} |u|^2 & 0 \\ 0 & |v|^2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right] = \text{tr} \begin{pmatrix} |u|^2 & 0 \\ 0 & 0 \end{pmatrix} = |u|^2, \end{aligned}$$

como era de se esperar.

Analogamente,

$$\begin{aligned} P(X_0 = 0, X_1 = 1) &= \phi(|e_0\rangle\langle e_0| \otimes |e_1\rangle\langle e_1|) \\ &= \phi_0(\mathcal{E}(|e_0\rangle\langle e_0| \otimes \mathcal{E}(|e_1\rangle\langle e_1| \otimes 1))) \end{aligned}$$

Mas

$$\begin{aligned} &\mathcal{E}(|e_1\rangle\langle e_1| \otimes 1) \\ = tr_2((|e_0\rangle\langle e_0| \otimes K_{0,0,0} + |e_1\rangle\langle e_1| \otimes K_{0,1,1})^*(|e_1\rangle\langle e_1| \otimes 1)(|e_0\rangle\langle e_0| \otimes K_{0,0,0} + \\ &\quad + |e_1\rangle\langle e_1| \otimes K_{0,1,1})) \\ &= tr_2((|e_0\rangle\langle e_0| \otimes K_{0,0,0} + |e_1\rangle\langle e_1| \otimes K_{0,1,1})^*(|e_1\rangle\langle e_1| \otimes K_{0,1,1})) \\ &= tr_2(|e_1\rangle\langle e_1| \otimes K_{0,1,1}^2) = |e_1\rangle\langle e_1|. \end{aligned}$$

Então

$$P(X_0 = 0, X_1 = 1) = \phi_0(\mathcal{E}(|e_0\rangle\langle e_0| \otimes |e_1\rangle\langle e_1|))$$

Mas

$$\begin{aligned} &\mathcal{E}(|e_0\rangle\langle e_0| \otimes |e_1\rangle\langle e_1|) \\ = tr_2((|e_0\rangle\langle e_0| \otimes K_{0,0,0} + |e_1\rangle\langle e_1| \otimes K_{0,1,1})^*(|e_0\rangle\langle e_0| \otimes |e_1\rangle\langle e_1|)(|e_0\rangle\langle e_0| \otimes K_{0,0,0} + \\ &\quad + |e_1\rangle\langle e_1| \otimes K_{0,1,1})) \\ &= tr_2((|e_0\rangle\langle e_0| \otimes K_{0,0,0} + |e_1\rangle\langle e_1| \otimes K_{0,1,1})^*(|e_0\rangle\langle e_0| \otimes |e_1\rangle\langle e_1| K_{0,0,0})) \\ &= tr_2(|e_0\rangle\langle e_0| \otimes K_{0,0,0}^2 |e_1\rangle\langle e_1|) = |e_0\rangle\langle e_0| tr(K_{0,0,0}^2 |e_1\rangle\langle e_1|) \\ &= |v|^2 |e_0\rangle\langle e_0|. \end{aligned}$$

Logo,

$$\begin{aligned} P(X_0 = 0, X_1 = 1) &= \phi_0(\mathcal{E}(|e_0\rangle\langle e_0| \otimes |e_1\rangle\langle e_1|)) \\ &= \phi_0(|v|^2 |e_0\rangle\langle e_0|) = tr \left[\begin{pmatrix} |u|^2 & 0 \\ 0 & |v|^2 \end{pmatrix} \begin{pmatrix} |v|^2 & 0 \\ 0 & 0 \end{pmatrix} \right] \\ &= tr \begin{pmatrix} |u|^2 |v|^2 & 0 \\ 0 & 0 \end{pmatrix} = |u|^2 |v|^2. \end{aligned}$$

Capítulo 8

Apêndice: Processos de Markov quânticos

Neste apêndice fazemos uma descrição de processos de Markov quânticos. Tal construção é mais elaborada do que a usada para cadeias de Markov quânticas e permite o estudo de sistemas mais complicados. Mostramos ainda como a moeda quântica é descrita nesse contexto. Esta exposição segue [16].

8.1 Notações e definições

Um **espaço de medida pontual** é um espaço de medida em que conjuntos pontuais (isto é, com um elemento apenas) são mensuráveis. Seja Ω um conjunto não vazio, chamado espaço amostral.

Definição Uma aplicação mensurável $X : \Omega \rightarrow S$ com imagem $R(X) \subset S$ é uma **medição** se:

1. $R(X)$ é o espaço base de um espaço de medida pontual $(R(X), \Sigma_X, \mu_X)$, onde Σ_X é uma σ -álgebra de $R(X)$ e μ_X é uma medida real sobre Σ_X .
2. para cada $x \in R(X)$, $X^{-1}(x)$ é o espaço base de um espaço de medida $(X^{-1}(x), \Sigma_X^x, \mu_X^x)$, onde Σ_X^x é uma σ -álgebra de $X^{-1}(x)$ e μ_X^x é uma medida real sobre Σ_X^x .

Observação Em processos estocásticos e em aplicações de mecânica quântica baseadas em amplitudes de transição, estamos interessados nos elementos de $R(X)$ (os resultados das medições), e não no conjunto Ω ou nas fibras $X^{-1}(x)$, $x \in R(X)$. Entretanto, seguiremos a descrição dada em [16] porque ela é útil para se descrever a interferência entre medições.

No caso de N-cadeias quânticas, supomos que $S = \{s_0, \dots, s_{n-1}\}$ é um conjunto finito, $\Omega = S^n$, e que naquele caso uma medição $X : \Omega \rightarrow S$ é uma aplicação tal que $X^{-1}(s_j) \in \Lambda$ (onde Λ é uma σ -álgebra fixada de Ω), $\sum_j P(X^{-1}(s_j)) = 1$. A menos que seja especificado, iremos supor que $\Omega = S^n$ para algum n . Ainda, iremos supor em geral que $S = R(X)$ é finito ou enumerável.

Alguns exemplos:

1. Moeda quântica. Neste caso, $\Omega = \{0, 1\}^n$, $S = \{0, 1\}$, $X_k : \{0, 1\}^n \rightarrow \{0, 1\}$ é $X_k(x_1, \dots, x_n) = x_k$, $R(X) = \{0, 1\}$ e μ_X, μ_X^x são medidas da contagem.
2. Mecânica quântica discreta. Neste caso, $S = \{s_0, \dots, s_k\}$ onde s_i é um estado que o sistema pode assumir, $\Omega = S^{n+1}$, $X_j : \Omega \rightarrow S$, $X_j(s_0, \dots, s_n) = s_j$, $R(X_j) = \{s_0, \dots, s_k\}$ e μ_j, μ_j^s são medidas da contagem em S e em $X_j^{-1}(s)$, respectivamente. Vemos que este exemplo possui uma estrutura semelhante ao da moeda quântica mencionada acima.
3. Mecânica quântica discreta com amplitudes de Feynman. Este é um caso particular do anterior. Sejam $a, m \in \mathbb{Z}$. Seja $\alpha = 2\pi/m$ e para $i = 0, \dots, m-1$, defina $k_i \in \mathbb{R}^2$, $\|k_i\| = 1$, $\angle(k_{i+1}, k_i) = \alpha$. Seja

$$V = \{v \in \mathbb{R}^2 : v = \sum_{j=1}^n e_j, e_j \in \{k_0, \dots, k_{m-1}\}, n \in \mathbb{N}\}$$

Pensamos em V como sendo um espaço de configuração discreto, e que está associado a um espaço de fase discreto

$$S = \{(v, k_j) : v \in V, j = 0, \dots, m-1\}$$

Então, $\Omega = S^{n+1}$, e $X_j : \Omega \rightarrow S$, $X_j((v_0, k_{i(0)}), \dots, (v_n, k_{i(n)})) = (v_j, k_{i(j)})$.

Definição Se X é mensurável, o conjunto de eventos de X é definido por

$$\mathcal{E}(X) := \{X^{-1}(B) : B \in \Sigma_X\},$$

o qual é uma σ -álgebra de subconjuntos de Ω .

Definição Uma função $a : \Omega \rightarrow \mathbb{C}$ é uma **densidade de amplitude** (para a medição X) se

$$A_X(x) := \int_{X^{-1}(x)} a d\mu_X^x \in L^2(R(X), \Sigma_X, \mu_X) \quad (8.1)$$

onde, é claro, supomos que $a|X^{-1}(x) \in L^1(X^{-1}(x), \Sigma_X^x, \mu_X^x)$, para cada $x \in R(X)$ e ainda que

$$\|A_X\|^2 := \int |A_X|^2 d\mu_X = 1 \quad (8.2)$$

Veremos no exemplo da moeda quântica que se $u, v \in \mathbb{C}$, com $u + v = 1$, $|u|^2 + |v|^2 = 1$, então definindo $a : \Omega \rightarrow \mathbb{C}$, $a(\omega) = u^k v^{n-k}$, onde n é o comprimento de ω e k é o número de “caras” (ou zeros) na seqüência, vale que a é uma densidade de amplitude para $X_j(x_1, x_2, \dots, x_n) = x_j$.

Definição O espaço de probabilidade quântica, denotado por $\mathcal{A}(\Omega, a)$, é o conjunto das medições para os quais $a : \Omega \rightarrow \mathbb{C}$ é uma densidade de amplitude.

Definição Seja $X \in \mathcal{A}(\Omega, a)$ uma medição. Um (X, a) -evento é um conjunto $C \subset \Omega$ tal que

$$A_X(C)(x) := \int_{C \cap X^{-1}(x)} a d\mu_X^x \in L^2(R(X), \Sigma_X, \mu_X) \quad (8.3)$$

onde supomos que $C \cap X^{-1}(x) \in \Sigma_X^x$, para todo $x \in R(X)$. Dizemos que $A_X(C)$ é a **densidade de amplitude de C determinada por X** . Note que $A_X(\Omega) = A_X$. Denotaremos o conjunto dos (X, a) -eventos por $\mathcal{E}(X, a)$.

Vale que $\mathcal{E}(X) \subset \mathcal{E}(X, a)$. Com efeito, seja $C \in \mathcal{E}(X)$, então $C = X^{-1}(B)$, $B \in \Sigma_X$. Daí,

$$C \cap X^{-1}(x) = X^{-1}(B) \cap X^{-1}(x) = \begin{cases} X^{-1}(x) & \text{se } x \in B \\ \emptyset & \text{se } x \notin B \end{cases}$$

Em qualquer caso, $C \cap X^{-1}(x) \in \Sigma_X^x$. É claro que a aplicação $A_X(C) \in L^2(R(X), \Sigma_X, \mu_X)$. Logo, $\mathcal{E}(X) \subset \mathcal{E}(X, a)$.

Definição Seja $C \in \mathcal{E}(X, a)$. Definimos a (X, a) -probabilidade de C como sendo

$$P_{X,a}(C) := \int |A_X(C)|^2 d\mu_X = \|A_X(C)\|^2$$

Lema 8.1.1 *Seja $C \in \mathcal{E}(X, a)$, $B \in \Sigma_X$. Então $X^{-1}(B) \cap C \in \mathcal{E}(X, a)$ e $A_X(X^{-1}(B) \cap C) = 1_B A_X(C)$.*

Prova É claro que $X^{-1}(B) \cap C \in \Sigma_X^x$, para todo $x \in R(X)$, e além disso,

$$X^{-1}(B) \cap C \cap X^{-1}(x) = \begin{cases} C \cap X^{-1}(x) & \text{se } x \in B \\ \emptyset & \text{se } x \notin B \end{cases}$$

Portanto,

$$A_X(X^{-1}(B) \cap C) = \int_{X^{-1}(B) \cap C \cap X^{-1}(x)} ad\mu_X^x = 1_B A_X(C)$$

Como $A_X(C) \in L^2(R(X), \Sigma_X, \mu_X)$, temos que $1_B A_X(C) \in L^2(R(X), \Sigma_X, \mu_X)$ e portanto (8.3) vale.

□

Aplicando o lema 8.1.1, obtemos

$$A_X(X^{-1}(B)) = A_X(X^{-1}(B) \cap \Omega) = 1_B A_X(\Omega) = 1_B A_X$$

$$P_{X,a}(X^{-1}(B)) = \int_{X^{-1}(B)} |A_X|^2 d\mu_X = \int_{X^{-1}(B)} \left| \int_{X^{-1}(x)} ad\mu_X^x \right|^2 d\mu_X(x) \quad (8.4)$$

Concluimos de (8.2) e (8.4) que $P_{X,a}$ é uma medida de probabilidade em Σ_X que chamamos de **distribuição** de X .

Definição Seja $X, Y \in \mathcal{A}(\Omega, a)$. Dizemos que X **não interfere** em Y se $\mathcal{E}(Y) \subset \mathcal{E}(X, a)$, e para cada $B \in \Sigma_Y$,

$$P_{X,a}(Y^{-1}(B)) = P_{Y,a}(Y^{-1}(B))$$

Nesse caso, a distribuição de Y é determinada quando realizamos a medição X . É fácil mostrar que não é uma relação simétrica em geral. Para ver como esta definição estende a de cadeias de Markov quânticas, considere X e Y medições. Pela primeira definição, temos que se X não interfere em Y , então

$$P[Y = s_j] = \sum_{k=0}^{n-1} P[Y = s_j, X = r_k]$$

O lado esquerdo da igualdade pode ser escrito como

$$P[Y = s_j] = \int_{Y^{-1}(s_j)} |A_Y|^2 d\mu_Y = P_{Y,a}(Y^{-1}(s_j))$$

E o lado direito como

$$\sum_x \int_{Y^{-1}(s_j) \cap X^{-1}(x)} |A_X|^2 d\mu_X^x = \int_{Y^{-1}(s_j)} |A_X|^2 d\mu_X = P_{X,a}(Y^{-1}(s_j))$$

Definição Dizemos que X é **independente** de Y se $\mathcal{E}(Y) \subset \mathcal{E}(X, a)$ e para cada $B \in \Sigma_X$ e $C \in \Sigma_Y$,

$$P_{X,a}(X^{-1}(B), Y^{-1}(C)) = P_{X,a}(X^{-1}(B))P_{X,a}(Y^{-1}(C))$$

Se X for independente de Y então $P_{X,a}(B|Y^{-1}(C)) = P_{X,a}(X^{-1}(B))$ para todo $B \in \Sigma_Y$ e $C \in \Sigma_Y$ com $P_{X,a}(Y^{-1}(C)) \neq 0$.

Seja T um subconjunto não vazio de \mathbb{R} e suponha que existe uma medição $X_t \in \mathcal{A}(\Omega, a)$ para cada $t \in T$.

Definição Dizemos que $(X_t)_{t \in T}$ é um **processo estocástico quântico** (QSP) se para cada $t, s_1, \dots, s_n \in T$, com $s_j \leq t, j = 1, \dots, n$ e para cada $B_j \in \Sigma_{s(j)}$ (onde $s(j) = s_j$), temos

$$\bigcap X_{s(j)}^{-1}(B_j) \in \mathcal{E}(X_t, a) \quad (8.5)$$

A equação (8.5) afirma que uma medição no presente pode ser usada para se obter uma informação sobre o passado. Isso é mais fraco do que a afirmação

$$\bigcap X_{s(j)}^{-1}(B_j) \in \mathcal{E}(X_t),$$

ou seja, que a informação do passado está contida no presente.

Seja $(X_t)_{t \in T}$ um QSP em $\mathcal{A}(\Omega, a)$. Para $t_1, \dots, t_n \in T$ com $t_1 < \dots < t_n$ e $x_j \in R_j, j = 1, \dots, n$, definimos (onde $t(n) = t_n$):

$$\begin{aligned} & A_{X_{t(n)}}[x_n | X_{t(n-1)} = x_{n-1}, \dots, X_{t(1)} = x_1] \\ & := \frac{A_{X_{t(n)}}[X_{t(n-1)} = x_{n-1}, \dots, X_{t(1)} = x_1](x_n)}{A_{X_{t(n-1)}}[X_{t(n-2)} = x_{n-2}, \dots, X_{t(1)} = x_1](x_{n-1})} \end{aligned}$$

quando o denominador não se anula; caso contrário, definimos o lado esquerdo como sendo igual a zero.

Definição Dizemos que $(X_t)_{t \in T}$ é um **processo de Markov quântico** (QMP) **quase-discreto** se

1.

$$\begin{aligned} & A_{X_{t(n)}}[x_n | X_{t(n-1)} = x_{n-1}, \dots, X_{t(1)} = x_1] \\ & = A_{X_{t(n)}}[x_n | X_{t(n-1)} = x_{n-1}] \end{aligned} \quad (8.6)$$

2.

$$A_{X_t}[X_s = x] \in L^1(R_t, \Sigma_t, \mu_t), \forall s, t \in T, s \leq t \quad (8.7)$$

3.

$$\begin{aligned} x \mapsto A_{X_t}[C \cap X_s^{-1}(x)](y) &\in L^{1,2}(R_s, \Sigma_s, \mu_s), \forall s \leq t, \quad y \in R_t, \\ \forall C \in \mathcal{E}(X_u), u \leq t, \text{ e } \int A_{X_t}[C \cap X_s^{-1}(x)](y) d\mu_s(x) &= A_{X_t}(C)(y) \end{aligned} \quad (8.8)$$

A seguir analisamos exemplos de QSP's quase-discretos.

8.2 Exemplos

Para analisar os exemplos desta seção, resumimos as expressões para amplitude e probabilidade obtidas na seção anterior.

A (X, a) -probabilidade de um conjunto $C \subset \Omega$ é

$$P_{X,a}(C) = \int \left| \int_{C \cap X^{-1}(x)} ad\mu_X^x \right|^2 d\mu_X(x) \quad (8.9)$$

e definimos também as amplitudes

$$A_X(x) = \int_{X^{-1}(x)} ad\mu_X^x \quad (8.10)$$

e

$$A_X(C)(x) = \int_{C \cap X^{-1}(x)} ad\mu_X^x \quad (8.11)$$

Iremos dar atenção maior para o exemplo da moeda quântica descrito a seguir. Primeiro, faremos a construção de acordo com [16], adotando a notação usada ali (isto é, com as fórmulas de amplitude e de probabilidade enunciadas acima). Depois, enunciaremos a construção em uma forma que é mais usual em teoria da medida. Existem outras construções na literatura, também ditas moedas quânticas, que são usadas com mais frequência [23].

Exemplo 1 Moeda quântica. Sejam $u, v \in \mathbb{C}$ tais que $u + v = 1$ e $|u|^2 + |v|^2 = 1$. Uma condição necessária e suficiente para que essas propriedades sejam satisfeitas é que $u = 1 - v$, $0 \leq \text{Re}(u) \leq 1$ e $\text{Im}(u) = \pm(\text{Re}(u))^{1/2}(1 - \text{Re}(u))^{1/2}$. Então, vemos que o número real $0 \leq \text{Re}(u) \leq 1$ determina u e v a menos de conjugação. Por exemplo, se $\text{Re}(u) = 1/2$, temos $u = (1 \pm i)/2$, $v = (1 \mp i)/2$.

Seja $n \in \mathbb{N}$ e seja $\Omega = \{0, 1\}^n$. Para $\omega = (x_1, x_2, \dots, x_n) \in \Omega$, defina $X_j(\omega) = x_j$, $j = 1, \dots, n$. Considerando a medida da contagem na imagem e nas fibras de X_j , vemos que X_j é uma medição, $j = 1, \dots, n$ (para considerações sobre cilindros em $\Omega = \{0, 1\}^n$, veja observação abaixo). Então

Ω representa n jogadas de moeda e X_j mede o resultado da j -ésima jogada. Defina a densidade de amplitude $a : \Omega \rightarrow \mathbb{C}$ por $a(\omega) = u^k v^{n-k}$, onde k é o número de zeros na seqüência ω . Pela definição (8.10), as funções A_{X_j} , $j = 1, \dots, n$ assumem os valores

$$A_{X_j}(0) = \int_{X_j^{-1}(0)} ad\mu_{X_j}^0 = u \sum_{k=0}^{n-1} \binom{n-1}{k} u^k v^{n-k-1} = u(u+v)^{n-1} = u$$

$$A_{X_j}(1) = v(u+v)^{n-1} = v$$

Como

$$\|A_{X_j}\|^2 = \int |A_{X_j}|^2 d\mu_{X_j} = |A_{X_j}(0)|^2 + |A_{X_j}(1)|^2 = |u|^2 + |v|^2 = 1,$$

segue que $a : \{0, 1\}^n \rightarrow \mathbb{C}$ é uma densidade de amplitude para X_j , e $X_j \in \mathcal{A}(\Omega, a)$, $j = 1, \dots, n$. Vale que $(X_j)_{j=1}^n$ é um QSP.

A distribuição de X_j é dada por

$$P_{X_j}(X_j^{-1}(0)) = \int \left| \int_{X_j^{-1}(0) \cap X_j^{-1}(x)} ad\mu_{X_j}^x \right|^2 d\mu_{X_j}(x)$$

$$= \left| \int_{X_j^{-1}(0) \cap X_j^{-1}(0)} ad\mu_{X_j}^x \right|^2 = |A_{X_j}(0)|^2 = |u|^2$$

$$P_{X_j}(X_j^{-1}(1)) = |A_{X_j}(1)|^2 = |v|^2$$

Logo, os X_j são identicamente distribuídos.

Vale que os X_j não interferem entre si. Por exemplo,

$$P_{X_j}(X_k^{-1}(0)) = \int \left| \int_{X_k^{-1}(0) \cap X_j^{-1}(x)} ad\mu_{X_j}^x \right| d\mu_{X_j}$$

$$= |A_{X_j}(X_k^{-1}(0))(0)|^2 + |A_j(X_k^{-1}(0))(1)|^2$$

$$= |u|^2 + |uv|^2 = |u|^2 = P_{X_k}(X_k^{-1}(0))$$

e

$$P_{X_j}(X_k^{-1}(1)) = |A_{X_j}(X_k^{-1}(1))(0)|^2 + |A_{X_j}(X_k^{-1}(1))(1)|^2$$

$$= |uv|^2 + |v|^2 = |v|^2 = P_{X_k}(X_k^{-1}(1))$$

Vale também que os X_j são independentes no sentido que definimos acima, pois por exemplo,

$$P_{X_j}(X_j = 0, X_k = 1) = \left| uv \sum_{i=0}^{n-1} \binom{n-2}{i} u^i v^{n-i-2} \right|^2 = |u|^2 |v|^2$$

Finalmente, vale que (X_j) é um processo de Markov. De fato, as equações (8.7) e (8.8) claramente valem. Para verificar (8.6), temos

$$A_{X_{t(m)}}(X_{t(m-1)} = x_{m-1}, \dots, X_{t(1)} = x_1)(x_m) = u^k v^{m-k}$$

onde k é o número de zeros na seqüência (x_1, \dots, x_m) . Portanto,

$$A_{X_{t(m)}}(x_m | X_{t(m-1)} = x_{m-1}, \dots, X_{t(1)} = x_1) = \begin{cases} u & \text{se } x_m = 0 \\ v & \text{se } x_m = 1 \end{cases}$$

E é claro que $A_{X_{t(m)}}(x_m | X_{t(m-1)} = x_{m-1})$ tem o mesmo valor.

Observação 1 Em teoria da medida, o exemplo acima admite uma construção muito mais simples. Seja \mathcal{B} a álgebra gerada pelos cilindros em $\{0, 1\}^{\mathbb{N}}$. Fixe $n \in \mathbb{N}$ e defina

$$B_{k,n} = \{\omega = (x_1, x_2, \dots) \in \{0, 1\}^{\mathbb{N}} : \text{o número de 0's em } (x_1, \dots, x_n) \text{ é } k\}$$

Vale que $B_{k,n}$ é uma união de cilindros, e portanto $B_{k,n} \in \mathcal{B}$. Então a densidade de amplitude de $B_{k,n}$,

$$a(B_{k,n}) = u^k v^{n-k}$$

pode ser vista como sendo uma medida complexa $a : \mathcal{B} \rightarrow \mathbb{C}$, e então obtemos uma única extensão para a σ -álgebra gerada pelos cilindros, pelo teorema da extensão de Kolmogorov. A probabilidade é obtida tomando o módulo ao quadrado desta medida, o que está de acordo com os cálculos de amplitudes feitos acima. Ainda, não precisamos nos preocupar em definir uma medida e uma σ -álgebra na imagem e nas fibras das medições.

Exemplo 2 Moeda quântica de 3 lados. Sejam $u, v, w \in \mathbb{C}$ tais que $u + v + w = 1$ e $|u|^2 + |v|^2 + |w|^2 = 1$. Seja $n \in \mathbb{N}$ e defina $\Omega = \{0, 1, 2\}^n$. Para $\omega = (x_1, \dots, x_n) \in \Omega$, defina $X_j(\omega) = x_j$, $j = 1, \dots, n$. Considerando a medida da contagem na imagem e nas fibras de X_j , temos que X_j é uma medição, $j = 1, \dots, n$. Defina a densidade de amplitude $a : \Omega \rightarrow \mathbb{C}$, $a(\omega) = u^{j_0} v^{j_1} w^{j_2}$, onde j_k é o número de k 's na seqüência ω . As funções A_{X_j} , $j = 1, \dots, n$ tem os valores

$$\begin{aligned} A_{X_j}(0) &= \int_{X_j^{-1}(0)} ad\mu_X^x = u \sum \left\{ \binom{n-1}{j_1 j_2 j_3} u^{j_0} v^{j_1} w^{j_2} : j_1 + j_2 + j_3 = n-1 \right\} \\ &= u(u + v + w)^{n-1} = u \end{aligned}$$

$$A_{X_j}(1) = v(u + v + w)^{n-1} = v$$

$$A_{X_j}(2) = w(u + v + w)^{n-1} = w$$

Como

$$|A_{X_j}(0)|^2 + |A_{X_j}(1)|^2 + |A_{X_j}(2)|^2 = |u|^2 + |v|^2 + |w|^2 = 1,$$

temos que $X_j \in \mathcal{A}(\Omega, a)$, $j = 1, \dots, n$, e é claro que $(X_j)_{j=1}^n$ é um QSP. A distribuição de X_j é dada por

$$\begin{aligned} P_{X_j}(X_j^{-1}(0)) &= \int \left| \int_{X_j^{-1}(0) \cap X^{-1}(x)} ad\mu_X^x \right|^2 d\mu_X(x) \\ &= \left| \int_{X_j^{-1}(0) \cap X_j^{-1}(0)} ad\mu_X^0 \right|^2 + \left| \int_{X_j^{-1}(0) \cap X_j^{-1}(1)} ad\mu_X^1 \right|^2 + \left| \int_{X_j^{-1}(0) \cap X_j^{-1}(2)} ad\mu_X^2 \right|^2 \\ &= |A_{X_j}(X_j^{-1}(0))(0)|^2 + 0 + 0 = |A_{X_j}(0)|^2 = |u|^2 \end{aligned}$$

Analogamente,

$$P_{X_j}(X_j^{-1}(1)) = |A_{X_j}(1)|^2 = |v|^2$$

$$P_{X_j}(X_j^{-1}(2)) = |A_{X_j}(2)|^2 = |w|^2$$

Logo, os X_j são identicamente distribuídos, e como no exemplo 1, eles são mutuamente independentes e formam um QMP. Entretanto, ao contrário do exemplo 1, os X_j interferem entre si. Para mostrar isso, seja $j, k \in \{1, \dots, n\}$, com $j \neq k$. Então

$$A_{X_j}(X_k \in \{0, 1\})(0) = \int_{(X_k^{-1}(0) \cup X_k^{-1}(1)) \cap X_j^{-1}(0)} ad\mu_{X_j}^0 = u(u + v)$$

$$A_{X_j}(X_k \in \{0, 1\})(1) = \int_{(X_k^{-1}(0) \cup X_k^{-1}(1)) \cap X_j^{-1}(1)} ad\mu_{X_j}^1 = v(u + v)$$

$$A_{X_j}(X_k \in \{0, 1\})(2) = \int_{(X_k^{-1}(0) \cup X_k^{-1}(1)) \cap X_j^{-1}(2)} ad\mu_{X_j}^2 = w(u + v)$$

Logo,

$$\begin{aligned} P_{X_j}(X_k \in \{0, 1\}) &= \int \left| \int_{X_j^{-1}(x) \cap (X_k^{-1}(0) \cup X_k^{-1}(1))} ad\mu_{X_j}^x \right|^2 d\mu_{X_j}(x) \\ &= \left| \int_{X_j^{-1}(0) \cap (X_k^{-1}(0) \cup X_k^{-1}(1))} ad\mu_{X_j}^0 \right|^2 + \left| \int_{X_j^{-1}(1) \cap (X_k^{-1}(0) \cup X_k^{-1}(1))} ad\mu_{X_j}^1 \right|^2 + \end{aligned}$$

$$\begin{aligned}
& + \left| \int_{X_j^{-1}(2) \cap (X_k^{-1}(0) \cup X_k^{-1}(1))} ad\mu_{X_j}^2 \right|^2 \\
& = |u(u+v)|^2 + |v(u+v)|^2 + |w(u+v)|^2 = |u+v|^2(|u|^2 + |v|^2 + |w|^2) \\
& = |u+v|^2
\end{aligned}$$

É claro que em geral, isso não é igual a

$$\begin{aligned}
P_{X_k}(X_k \in \{0, 1\}) & = \int \left| \int_{X_k^{-1}(x) \cap (X_k^{-1}(0) \cup X_k^{-1}(1))} ad\mu_X^x \right|^2 d\mu_{X_j}(x) \\
& = \left| \int_{X_k^{-1}(0) \cap (X_k^{-1}(0) \cup X_k^{-1}(1))} ad\mu_{X_k}^0 \right|^2 + \left| \int_{X_k^{-1}(1) \cap (X_k^{-1}(0) \cup X_k^{-1}(1))} ad\mu_{X_k}^1 \right|^2 + \\
& \quad + \left| \int_{X_k^{-1}(2) \cap (X_k^{-1}(0) \cup X_k^{-1}(1))} ad\mu_{X_k}^2 \right|^2 \\
& = |A_{X_k}(0)|^2 + |A_{X_k}(1)|^2 + 0 = |u|^2 + |v|^2
\end{aligned}$$

Logo, como $X_k^{-1}(\{0, 1\}) \in \Sigma_{X_k}$ segue pela definição de não interferência que X_j interfere em X_k .

Exemplo 3 Mecânica quântica discreta. Seja S um conjunto não vazio de estados que uma partícula pode assumir.

Definição Uma função $A_1 : S \times S \rightarrow \mathbb{C}$ é uma **amplitude de transição estocástica em um passo** se para cada $s_1, s_2 \in S$, temos

$$\sum_s A_1(s_1, s) \bar{A}_1(s_2, s) = \sum_s A_1(s, s_1) \bar{A}_1(s, s_2) = \delta_{s_1 s_2}$$

e

$$\sum_s A_1(s_1, s) = 1$$

onde a soma converge absolutamente.

Uma matriz de amplitude de transição de uma cadeia de Markov quântica induz uma função de amplitude de transição estocástica (a função induzida é simplesmente $A : S \times S \rightarrow \mathbb{C}$, $A(s_j, s_k) = A_{kj}$). Denote o conjunto das funções de amplitude de transição estocástica por $T(S)$.

Fixemos $\Omega = S^{n+1} = \{(s_0, s_1, \dots, s_n) : s_i \in S\}$. Seja $a_0 \in l^2(S)$ um vetor unitário representando a distribuição inicial de uma partícula quântica. Fixe $A_1 \in T(S)$. Para $\omega = (s_0, s_1, \dots, s_n) \in \Omega$, defina a densidade de amplitude

$$a(\omega) := a_0(s_0) A_1(s_0, s_1) \cdots A_1(s_{n-1}, s_n)$$

Para $j = 0, 1, \dots, n$, defina $X_j : \Omega \rightarrow S$ por

$$X_j(s_0, s_1, \dots, s_n) = s_j$$

Seja μ_j e μ_j^s a medida da contagem em S e em $X_j^{-1}(s)$, respectivamente. Equipado com essa estrutura, X_j é uma medição, $j = 0, 1, \dots, n$ e vale que $X_j \in \mathcal{A}(\Omega, a), j = 0, \dots, n$. Se $j < k$, vale que X_k não interfere em X_j (ver [16]). Mas exemplos simples mostram que X_j pode interferir em X_k . Além disso, X_j e X_k não são independentes em geral. Finalmente, vale que $(X_j)_{j=0}^n$ é um QMP quase-discreto (ver [16]).

Vamos analisar um modelo concreto para a mecânica quântica discreta em 2 dimensões. Sejam a e m inteiros positivos relativamente primos, m par. Seja $\alpha = 2\pi/m$ e sejam k_0, k_1, \dots, k_{m-1} vetores unitários em \mathbb{R}^2 tais que cada um forma um ângulo α com o anterior. Seja

$$V = \{v \in \mathbb{R}^2 : v = \sum_{j=1}^n e_j, e_j \in \{k_0, \dots, k_{m-1}\}, n \in \mathbb{N}\}$$

Pensamos em V como sendo um espaço de configuração discreto, e que está associado a um espaço de fase discreto

$$S = \{(v, k_j) : v \in V, j = 0, \dots, m-1\}$$

Definição A amplitude de transição de Feynman discreta em um passo é a aplicação $A_1 : S \times S \rightarrow \mathbb{C}$ dada por

$$A_1((v, k_r), (v + k_r, k_t)) = n^{-1/2} \exp\left[\frac{im\pi(t-r)^2}{n}\right]$$

e A_1 é zero, caso contrário.

Pode-se mostrar que um múltiplo constante de A_1 , de módulo 1, está contido em $T(S)$. Como tal múltiplo não afeta as probabilidades, iremos assumir que $A_1 \in T(S)$.

Seja $n \in \mathbb{N}$ e seja $\Omega = S^{n+1} = \{(s_0, s_1, \dots, s_n) : s_i \in S\}$. Defina a densidade de amplitude $a : \Omega \rightarrow \mathbb{C}$ e as medições $X_j : \Omega \rightarrow S$ como antes. Neste caso, para

$$\omega = ((v_0, k_{i(0)}), \dots, (v_n, k_{i(n)}))$$

temos que $X_j(\omega) = (v_j, k_{i(j)})$, $j = 0, \dots, n$. Pelo que fizemos acima, concluimos que $(X_j)_{j=0}^n$ é um QMP em $\mathcal{A}(\Omega, a)$.

Vamos mostrar que se $n \rightarrow \infty$, A_1 se aproxima, em um certo sentido, da amplitude de Feynman usual para uma partícula livre. Seja

$$p = \{(x_0, k_{j_0}), \dots, (x_r, k_{j_r})\},$$

tal que $x_s + k_{j_s} = x_{s+1}$, $s = 0, \dots, r-1$. Seja $\beta_s = 2\pi(j_s - j_{s-1})/n$, $s = 1, \dots, r$ e suponha que os β_s são pequenos (ou próximos de 2π). Então a distância entre x_{s+1} e x_{s-1} é

$$\begin{aligned} \|x_{s+1} - x_{s-1}\|^2 &= \|x_s + k_{j_{s-1}} + k_{j_s} - x_s\|^2 \\ &= \|k_{j_{s-1}} + k_{j_s}\|^2 = 2 + 2k_{j_s} \cdot k_{j_{s-1}} = 2 + 2\cos\beta_s \approx 2 + 2\left(1 - \frac{\beta_s^2}{2}\right) = 4 - \beta_s^2. \end{aligned}$$

Portanto, $\beta_s^2 \approx 4 - \|x_{s+1} - x_{s-1}\|^2$. Se v_s denota a “velocidade” da partícula no tempo s , temos $v_s^2 \approx \|x_{s+1} - x_{s-1}\|^2/4$. Logo, $\beta_s^2 \approx 4(1 - v_s^2)$ e a amplitude do caminho p é

$$\begin{aligned} A(p) &= n^{-r/2} \exp\left[im\pi n^{-1} \sum_{s=1}^r (j_s - j_{s-1})^2\right] \\ &= n^{-r/2} \exp\left[imn(4\pi)^{-1} \sum_{s=1}^r \beta_s^2\right] \approx \\ &\approx n^{-r/2} \exp\left[imn\pi^{-1} \sum_{s=1}^r (1 - v_s^2)\right] \\ &= n^{-r/2} e^{imnr/\pi} \exp\left[-i2n\pi^{-1} \sum_{s=1}^r \frac{mv_s^2}{2}\right]. \end{aligned}$$

Se fizermos com que o número m corresponda à massa da partícula, então o somatório corresponde à integral da energia cinética sobre o caminho. Desta forma, $A_1(p)$ se aproxima da amplitude de Feynman usual (contínua) para uma partícula livre.

Exemplo 4 Ainda, de maneira análoga a feita para cadeias quânticas, podemos também calcular a entropia de processos de Markov. Usando a expressão para a probabilidade de um evento $B \in \Sigma_X$, que repetimos aqui,

$$\begin{aligned} P_{X,a}(B) &= \int_B |A|^2 d\mu_X = \int |1_B A|^2 d\mu_X = \int |A(X^{-1}(B))|^2 d\mu_X \\ &= \int \left| \int_{X^{-1}(B) \cap X^{-1}(x)} a d\mu_X^x \right|^2 d\mu_X(x), \end{aligned}$$

podemos calcular a entropia de Shannon e a de von Neumann dos exemplos que consideramos nas seções 8.1 e 8.2. No caso da moeda quântica, os dois casos que consideramos (número finito ou arbitrário de medições) possuíam a mesma distribuição estacionária, a saber, $P_j(0) = |A_j(0)|^2 = |u|^2$, e $P_j(1) = |A_j(1)|^2 = |v|^2$. O operador densidade associado é $\rho = |u|^2|0\rangle\langle 0| + |v|^2|1\rangle\langle 1|$ e portanto, por (5.1) (após definição de entropia de von Neumann),

$$S(\rho) = -\text{tr}(\rho \log \rho) = -|u|^2 \log |u|^2 - |v|^2 \log |v|^2$$

8.3 Processos de Markov quânticos quase-discretos

Nesta seção, $X_t, t \in T \subset \mathbb{R}$ será um QMP quase-discreto sobre $\mathcal{A}(\Omega, a)$. Para $s, t \in T$, defina $F_{s,t} : R_s \times R_t \rightarrow \mathbb{C}$ por

$$F_{s,t}(x, y) = A_t(y|X_s = x)$$

Se $A_s(x) \neq 0$ então

$$F_{s,t}(x, y) = \frac{A_t(X_s = x)(y)}{A_s(x)} = A_s(x)^{-1} \int_{X_s^{-1}(x) \cap X_t^{-1}(y)} a d\mu_{X_t}^y$$

e $F_{s,t}(x, y) = 0$, caso contrário. Podemos aplicar o item 3 da definição de QMP quase-discreto para calcular A_t em termos de A_s , para $s \leq t$:

$$A_t(y) = \int A_t(X_s = x)(y) d\mu_s(x) = \int A_s(x) F_{s,t}(x, y) d\mu_s(x) \quad (8.12)$$

Definição O kernel de amplitude de transição $K_{s,t} : R_s \times \Sigma_t \rightarrow \mathbb{C}$ para $s, t \in T, s \leq t$ é dado por

$$K_{s,t}(x, B) = \int_B F_{s,t}(x, y) d\mu_t(y)$$

Segue do item 2 da definição de QMP quase-discreto que $K_{s,t}$ existe e é finito, e pelo item 3 vale que $F_{s,t}$ é mensurável em ambas as variáveis. Fazendo uma analogia com um núcleo de Markov, vemos que $K_{s,t}(x, \cdot)$ é uma medida complexa limitada em Σ_t e que $K_{s,t}(\cdot, B)$ é mensurável em R_s . Além disso, $K_{s,t}(x, \cdot) \ll \mu_t$ e

$$\frac{dK_{s,t}(x, \cdot)}{d\mu_t}(y) = F_{s,t}(x, y)$$

Agora provamos uma versão do teorema de Chapman-Kolmogorov neste contexto.

Teorema 8.3.1 Para $s, u, t \in T$, $s \leq u \leq t$ e $x \in R_s$, $z \in R_t$, $B \in \Sigma_t$, temos

$$F_{s,t}(x, z) = \int F_{u,t}(y, z)F_{s,u}(x, y)d\mu_u(y) = \int F_{u,t}(y, z)K_{s,u}(x, dy)$$

$$K_{s,t}(x, B) = \int K_{u,t}(y, B)K_{s,u}(x, dy)$$

Prova Se $A_s(x) = 0$, a desigualdade vale claramente, então assumimos que $A_s(x) \neq 0$. Aplicando (8.6) e (8.8), temos

$$\begin{aligned} \int F_{u,t}(y, z)K_{s,u}(x, dy) &= \int F_{u,t}(y, z)F_{s,u}(x, y)d\mu_u(y) \\ &= \int A_t(z|X_u = y)A_u(y|X_s = x)d\mu_u(y) \\ &= \int A_t(z|X_u = y, X_s = x)A_u(y|X_s = x)d\mu(y) \\ &= \int \frac{A_t(X_u = y, X_s = x)(z)}{A_u(X_s = x)(y)} \frac{A_u(X_s = x)(y)}{A_s(x)} d\mu_u(y) \\ &= \frac{1}{A_s(x)} \int A_t(X_u = y, X_s = x)(z)d\mu_u(y) \\ &= \frac{1}{A_s(x)} A_t(X_s = x)(z) = F_{s,t}(x, z) \end{aligned}$$

Integrando a primeira igualdade, obtemos a segunda. □

Agora assumimos que os contradomínios coincidem, ou seja,

$$(R_t, \Sigma_t, \mu_t) = (R, \Sigma, \mu), \quad t \in T$$

Seja $H = L^2(R, \Sigma, \mu)$. Diremos que $(X_t)_{t \in T}$ é **estacionário** (homogêneo no tempo) se $F_{s+u, t+u} = F_{s,t}$ sempre que $s, t, s+u, t+u \in T$, $s \leq t$. Assuma que $(X_t)_{t \in T}$ é estacionário e que $T = [0, a]$, $0 < a < \infty$ ou $T = [0, \infty)$. Então definimos $F_t : R \times R \rightarrow \mathbb{C}$ por $F_t = F_{0,t}$, $t \in T$. Daí, temos $F_{s,t} = F_{t-s}$. Analogamente, definimos $K_t : R \times \Sigma \rightarrow \mathbb{C}$ por $K_t = K_{0,t}$ e daí $K_{s,t} = K_{t-s}$. Então

$$K_t(x, B) = \int_B F_t(x, y)d\mu(y)$$

Assim, podemos reescrever (8.12) na forma

$$F_t(y) = \int F_0(x)F_t(x, y)d\mu(x) \quad (8.13)$$

Além disso, fazendo $s = 0$ e trocando t por $s + t$ e u por s , a equação de Chapman-Kolmogorov se torna

$$F_{s+t}(x, z) = \int F_s(x, y)F_t(y, z)d\mu(y) \quad (8.14)$$

$$K_{s+t}(x, B) = \int K_t(y, B)K_s(x, dy) \quad (8.15)$$

Para $t \in T$ defina a aplicação $K_t : \Sigma \times R \rightarrow \mathbb{C}$ por

$$K_t(B, y) = \int_B F_t(x, y)d\mu(x)$$

Segue dos itens 2 e 3 da definição de QMP que $K_t(B, x)$ é mensurável na segunda variável e é uma medida complexa limitada na primeira variável. Além disso, $K_t(\cdot, y) \ll \mu$ e

$$\frac{dK_t(\cdot, y)}{d\mu}(x) = F_t(x, y)$$

Diremos que $(X_t)_{t \in T}$ é **unitário** se para cada $t \in T$ se para cada $t \in T$, $y \in R$, $B \in \Sigma$ com $\mu(B) < \infty$, temos

$$\int \overline{K_t}(x, B)F_t(x, y)d\mu(x) = \int \overline{K_t}(B, x)F_t(y, x)d\mu(x) = 1_B(y), \quad \mu - \text{q.t.p.}$$

Defina o operador linear $U_t : H \rightarrow H$ por

$$U_t g(y) = \int g(x)F_1(x, y)d\mu(x) = \int g(x)K_t(dx, y)$$

Aplicando (8.13) temos $A_t = U_t A_0$, $t \in T$. Agora mostraremos que neste caso $t \mapsto U_t$ é um semigrupo unitário a um parâmetro. Mostraremos também que $(X_t)_{t \in T}$ é unitário se e somente se U_t é unitário para todo $t \in T$.

Teorema 8.3.2 1. Se $(X_t)_{t \in T}$ é unitário então U_t é um operador unitário, $t \in T$ e $U_{s+t} = U_s U_t$ para todo $s, t \in T$ com $s + t \in T$.

2. Se U_t é unitário, $t \in T$ então $(X_t)_{t \in T}$ é unitário.

Prova 1. Suponha que $(X_t)_{t \in T}$ é unitário. Primeiro mostraremos que U_t é limitado, $t \in T$. Seja $g \in H$ uma função simples. Então existe $B_i \in \Sigma$, $i = 1, \dots, n$ com $B_i \cap B_j = \emptyset$ se $i \neq j$, $\mu(B_i) < \infty$ e $c_i \in \mathbb{C}$, $i = 1, \dots, n$ tais que $g = \sum c_i 1_{B(i)}$. Daí,

$$\begin{aligned} \|U_t g\|^2 &= \int |U_t g(y)|^2 d\mu(y) = \int \left[\int g(x) K_t(dx, y) \right] \left[\int \bar{g}(z) \bar{K}_t(dz, y) \right] d\mu(y) \\ &= \int \left[\int g(x) F_t(x, y) d\mu(x) \right] \left[\sum \bar{c}_i \bar{K}_t(B_i, y) \right] d\mu(y) \\ &= \sum \bar{c}_i \int g(x) \left[\int \bar{K}_t(B_i, y) F_t(x, y) d\mu(y) \right] d\mu(x) \\ &= \sum \bar{c}_i \int g(x) 1_{B(i)}(x) d\mu(x) = \sum |c_i|^2 \mu(B_i) = \|g\|^2 \end{aligned}$$

Portanto, U_t restrita ao subespaço denso S de funções simples possui norma 1. Logo, esta restrição possui uma única extensão linear e limitada \hat{U}_t para H de norma 1. Agora seja $g \in H$ qualquer. Então existe uma seqüência $g_i \in S$ tal que $|g_i(x)| \leq |g(x)|$ para todo $x \in R$ e $g_i \rightarrow g$ na convergência da norma. Segue que existe uma subseqüência, que também denotaremos por g_i tal que $g_i \rightarrow g$, μ -q.t.p.. Como $g, F_t(\cdot, y) \in H$, temos $g F_t(\cdot, y) \in L^1(R, \Sigma, \mu)$ e além disso,

$$|g_i(x) F_t(x, y)| \leq |g(x) F_t(x, y)|$$

para todo $x \in R$. Aplicando o teorema da convergência dominada, temos

$$\begin{aligned} \hat{U}_t g(y) &= \lim \hat{U}_t g_i(y) = \lim \int g_i(x) F_y(x, y) d\mu(x) \\ &= \int g(x) F_t(x, y) d\mu(x) = U_t g(y) \end{aligned}$$

Logo, $U_t = \hat{U}_t$ e portanto U_t é limitada.

Mostremos que U_t é unitária. Note que a adjunta de U_t é dada por

$$U_t^* g(y) = \int g(x) \bar{F}_t(y, x) d\mu(x) = \int g(x) \bar{K}_t(y, dx)$$

Novamente, se $g = \sum c_i 1_{B(i)}$ é uma função simples, temos

$$\begin{aligned} U_t U_t^* g(y) &= \int U_t^* g(x) F_t(x, y) d\mu(x) \\ &= \int \left[\int g(z) \bar{K}_t(x, dz) \right] F_t(x, y) d\mu(x) \end{aligned}$$

$$= \sum c_i \int \overline{K}_t(x, B_i) F_t(x, y) d\mu(x) = \sum c_i 1_{B(i)}(y) = g(y)$$

Analogamente,

$$\begin{aligned} U_t^* U_t g(y) &= \int U_t g(x) \overline{F}_t(y, x) d\mu(x) \\ &= \int \left[\int g(z) K_t(dz, x) \right] \overline{F}_t(y, x) d\mu(x) \\ &= \sum c_i \int K_t(B_i, x) \overline{F}_t(y, x) d\mu(x) = \sum c_i 1_{B(i)}(y) = g(y) \end{aligned}$$

Logo, $U_t U_t^* = U_t^* U_t = 1$ em S e portanto U_t é unitária. Finalmente, se $s, t, s+t \in T$, temos por (8.14) que

$$\begin{aligned} U_{s+t} g(y) &= \int g(x) F_{s+t}(x, y) d\mu(x) \\ &= \int g(x) \left[\int F_s(x, z) F_t(z, y) d\mu(z) \right] d\mu(x) \\ &= \int \left[\int g(x) F_s(x, z) d\mu(x) \right] F_t(z, y) d\mu(z) = U_t U_s g(y) \end{aligned}$$

2. Suponha que $U_t, t \in T$ é unitário. Se $B \in \Sigma$ com $\mu(B) < \infty$ então $1_B \in H$. Logo,

$$\begin{aligned} 1_B(x) &= U_t U_t^* 1_B(x) = \int \left[\int 1_B(z) \overline{K}_t(x, dz) \right] F_t(x, y) d\mu(x) \\ &= \int \overline{K}_t(x, B) F_t(x, y) d\mu(x) \end{aligned}$$

e

$$\begin{aligned} 1_B(x) &= U_t^* U_t 1_B(x) = \int \left[\int 1_B(z) K_t(dz, x) \right] \overline{F}_t(y, x) d\mu(x) \\ &= \int K_t(B, x) \overline{F}_t(y, x) d\mu(x) \end{aligned}$$

□

Seja $(X_t)_{t \in T}$ unitário com $T = [0, \infty)$. Definindo $U_{-t} = U_t^*$, temos que $t \mapsto U_t$ é um grupo unitário a um parâmetro em \mathbb{R} . Note que $F_t(x, \cdot) - F_0(x, \cdot) \in H$ e $\|F_t(x, \cdot) - F_0(x, \cdot)\|$ é mensurável. Dizemos que $(X_t)_{t \in T}$ é **contínuo** se para todo $\epsilon > 0$ existe $\delta > 0$ tal que $|t| < \delta$ implica que

$$\int \|F_t(x, \cdot) - F_0(x, \cdot)\|^2 d\mu(x) < \epsilon$$

Teorema 8.3.3 *Se $(X_t)_{t \in T}$ é contínuo, unitário e estacionário, então $t \mapsto U_t$ é fortemente contínuo.*

Prova Mostraremos que $t \rightarrow U_t$ é fracamente contínuo em 0, donde o resultado segue. Para $g, h \in H$ temos, aplicando a desigualdade de Schwarz, que

$$\begin{aligned} |\langle (U_t - I)g, h \rangle| &= \left| \int \int g(x)[F_t(x, y) - F_0(x, y)]d\mu(x)\bar{h}(y)d\mu(y) \right| \\ &\leq \int |g(x)| \int |F_t(x, y) - F_0(x, y)||h(y)|d\mu(y)d\mu(x) \\ &\leq \|h\| \int |g(x)| \|F_t(x, \cdot) - F_0(x, \cdot)\|d\mu(x) \\ &\leq \|h\| \|g\| \left[\int \|F_t(x, \cdot) - F_0(x, \cdot)\|^2 d\mu(x) \right]^{1/2} \end{aligned}$$

Logo, dado $\epsilon > 0$, existe $\delta > 0$ tal que $|t| < \delta$ implica

$$|\langle (U_t - I)g, h \rangle| \leq \|h\| \|g\| \epsilon$$

□

Sob as condições do teorema anterior, temos que $t \mapsto U_t$ é um grupo unitário a um parâmetro contínuo. Portanto, pelo teorema de Stone, $U_t = e^{itH}$ para um único operador auto-adjunto H . Chamamos tal processo de **processo Hamiltoniano**.

Teorema 8.3.4 *Se $(X_t)_{t \in T}$ é estacionário e unitário então X_t não interfere em X_s para $s \leq t$.*

Prova Para $B \in \Sigma$, temos por (8.8) e pela unitariedade de U_{t-s} que

$$\begin{aligned} P_t(X_s \in B) &= \int |A_t(X_s \in B)(y)|^2 d\mu(y) \\ &= \int \left| \int A_t[X_s^{-1}(B) \cap X_s^{-1}(x)](y) d\mu(x) \right|^2 d\mu(y) \\ &= \int \left| \int 1_B(x) A_t(X_s = x)(y) d\mu(x) \right|^2 d\mu(y) \\ &= \int \left| \int 1_B(x) A_s(x) F_{t-s}(x, y) d\mu(x) \right|^2 d\mu(y) \\ &= \int |U_{t-s} 1_B A_s(y)|^2 d\mu(y) = \|U_{t-s} 1_B A_s\|^2 \\ &= \|1_B A_s\|^2 = \int_B |A_s|^2 d\mu = P_s(B) \end{aligned}$$

□

Referências Bibliográficas

- [1] Accardi, L. A. Nonrelativistic quantum mechanics as a non-commutative Markov process. *Adv. Math*, Vol. 20, 1976.
- [2] Accardi, L. Topics in quantum probability, *Phys. Reports*, 77, No. 3 (1981), 169-192.
- [3] Accardi, L., Fidaleo, F., Mukhamedov, F. Markov states on the CAR algebra. *Math-ph/0411035*.
- [4] Arveson, W. The canonical anticommutation relations. *Lecture Notes for Mathematics* 208. November 1998.
- [5] Atiyah, M.F., Macdonald, I.G. *Introduction to Commutative Algebra*. Addison-Wesley, 1969.
- [6] Balakrishnan, A. V. *Applied Functional Analysis*. Springer-Verlag, 1976.
- [7] Billingsley, P. *Probability and Measure*. John Wiley & Sons, Inc., 1995.
- [8] Bratteli, O., Robinson, D. *Operator algebras and quantum statistical mechanics*, vols. 1 e 2. Second edition. Springer-Verlag, 1997.
- [9] Brin, M. Stuck, G. *Introduction to Dynamical Systems*. Cambridge University Press, 2002.
- [10] Ellis, R. S. *Entropy, Large Deviations, and Statistical Mechanics*. Springer-Verlag, 1985.
- [11] Fernandez, P. *Medida e Integração. Projeto Euclides*, 1996.
- [12] Feynman, R., Hibbs, A. *Quantum mechanics and path integrals*. McGraw-Hill, New York, N.Y., 1965.
- [13] Feynman, R. *QED: The Strange Theory of Light and Matter*. Princeton University Press, Princeton, New Jersey, 1985.

- [14] Gantmacher, F.R. The Theory of Matrices, vol. 2. Chelsea Publishing Company, New York, N.Y., 1959.
- [15] Gudder, S., Marbeau, J. Analysis of a Quantum Markov Chain. Annales de l'Institut Henri Poincaré - Physique Théorique, Vol 52, no. 1, 1990, p. 31-50.
- [16] Gudder, S., Schindler, C. Quasi-discrete Quantum Markov Processes. Annales de l'Institut Henri Poincaré - Physique Théorique, Vol 56, no. 2, 1992, p. 123-142.
- [17] Gudder, S., Marbeau, J. A quantum random walk. Annales de la Fondation Louis de Broglie, Vol. 14, no. 4, 1989.
- [18] Gudder, S. Quantum Probability. Academic Press, Inc., Boston, 1988.
- [19] Gudder, S. Discrete Quantum Mechanics. J. Math. Physics, 27 (7), 1986, p. 1782-1790.
- [20] Gustafson, S., Sigal, I. Mathematical Concepts of Quantum Mechanics. Springer-Verlag, 2003.
- [21] Isham, C. J. Lectures on Quantum Theory. Imperial College Press, 1995.
- [22] Karlin, S., Taylor, H. A first course in stochastic processes. Academic Press, Inc. 1975.
- [23] Kempe, J. Quantum random walks: an introductory overview. Contemporary Physics, vol. 44, number 4, july-august 2003.
- [24] Kitaev, A. Yu., Shen, A. H., Vyalyi, M. N. Classical and Quantum Computation. Graduate Studies in Mathematics, vol. 47. American Mathematical Society, 2002.
- [25] Kumjian, A., Renault, J. KMS states on C^* -algebras associated to expansive maps. Math.OA/0305044.
- [26] Lanford, O. E., Ruelle, D. Observables at infinity and states with short range correlation in statistical mechanics. Comm. Math. Phys. Vol. 13, 194-215, 1969.
- [27] Lima, E. L. Espaços Métricos. Projeto Euclides, 1996.
- [28] Lu, Y-G. Quantum Markov Chains and classical random sequences. Nagoya Math. Journal. Vol. 139 (1995), pp. 173-183.

- [29] Madsen, I., Tornehave, J. From calculus to cohomology: De Rham cohomology and characteristic classes. Cambridge University Press, 1997.
- [30] Nelson, E. Construction of quantum fields from Markov fields. Journal of Functional Analysis, Vol. 12, 1973.
- [31] Neshveyev, S., Stormer, E. The variational principle for a class of asymptotically abelian C^* -algebras.
http://www.math.uio.no/eprint/pure_math/2000/11-00.ps.
- [32] Nielsen, M., Chuang, I. Quantum computation and quantum information. Cambridge University Press, 2000.
- [33] Norris, J. R. Markov Chains. Cambridge University Press, 1998.
- [34] Ohya, M., Petz, D. Quantum entropy and its use, Texts and monographs in physics. Springer-Verlag, Berlin. 1993.
- [35] Parry, W., Pollicott, M. Zeta Functions and the Periodic Orbit Structure of Hyperbolic Dynamics. Société Mathématique de France. 187-188, Astérisque, 1990.
- [36] Pollicott, M. Yuri, M. Dynamical Systems and Ergodic Theory. Cambridge University Press, 1998.
- [37] Sinai, Ya. Probability Theory. Springer-Verlag, 1992.
- [38] Spitzer, F. A variational characterization of finite Markov chains. Annals of Math. Statistics, Vol 43, No. 1, 1972, p. 303-307.
- [39] Urruty, J. B. H. Fundamentals of Convex Analysis. Springer-Verlag, 2001.
- [40] van Enter, A. C. C., Fernandez R. and Sokal, A. D. Regularity properties and pathologies of position-space renormalization-group transformations: Scope and limitations of Gibbsian theory, *Journ. of Stat. Phys.*, Vol 72, N 5/6, 1993 pp 879-1187.
- [41] Vedral, V. The role of relative entropy in quantum information theory. Reviews of modern physics, Vol. 74, January 2002.
- [42] Walters, P. An Introduction to Ergodic Theory. Springer-Verlag, New York, 1982.

Índice Remissivo

- (X, a) -evento, 132
- (X, a) -probabilidade, 132, 135
- *-álgebra, 124
- C^* -álgebra, 4, 119
- Álgebra, 118
- Álgebra de Banach, 119
- Álgebra normada, 118

- Amplitude de transição, 8

- Bit, 54

- Cadeia de Markov, 12
- Cadeia de Markov quântica, 13, 120, 122
- Conjunto de eventos, 131

- Decomposição espectral, 67
- Decomposição polar, 67
- Decomposição de Schmidt, 66
- Decomposição em valores singulares, 68
- Densidade de amplitude, 131, 132
- Desigualdade de Araki-Lieb, 79
- Desigualdade de Klein, 76
- Distribuição de Gibbs, 93

- Energia específica, 98
- Energia livre, 99
- Energia média, 98
- Ensemble, 62
- Entropia condicional de Shannon, 72
- Entropia condicional de von Neumann, 84
- Entropia de Shannon, 71

- Entropia de uma medida, 98
- Entropia de uma partição, 98
- Entropia de von Neumann, 76
- Entropia quântica, 76
- Entropia relativa de Shannon, 71
- Entropia relativa de von Neumann, 76

- Equação de Schrödinger, 5, 64
- Espaço de probabilidade quântica, 132
- Esperança condicional (de Umegaki), 126
- Esperança de transição, 120
- Esperança quase-condicional, 126
- Estado (funcional), 119
- Estado (vetor-estado), 54
- Estado de Markov, 126
- Estado emaranhado, 60, 66
- Estado KMS, 127
- Estado misturado, 63
- Estado puro, 63
- Estados correlacionados, 72, 79

- Informação mútua de Shannon, 72
- Informação mútua de von Neumann, 82

- Involução, 119

- Lema de Wielandt, 107

- Matriz densidade, 62
- Matriz irredutível, 111
- Matriz redutível, 111
- Mecânica estatística quântica, 118
- Mecânica quântica discreta, 139

Medição, 10, 130
Medições independentes, 134
Medida ergódica, 100
Moeda quântica, 127, 135
Moeda quântica de 3 lados, 137

N-cadeia, 11
N-cadeia quântica, 13
N-cadeia unitária, 13
Número de Schmidt, 69
Não-interferência, 12, 133

Operador densidade, 62
Operador densidade reduzido, 63

Pressão topológica, 95
Processo de Markov quântico, 134
Processo estacionário, 143
Processo estocástico quântico, 134
Processo unitário, 144
Produto de Kronecker, 61
Purificação de um estado, 69

Q-bit, 54

Relações de anticomutatividade canônicas
(CAR), 125

Teorema de Birkhoff, 110
Teorema de Krein-Milman, 110
Teorema de Lieb, 83
Teorema de Perron-Frobenius, 108
Teoria quântica de campos, 127
Traço parcial, 63