Notas em Informação Quântica

Uma abordagem geométrica

Marcos Gabriel Alpino
mg.alpino@fisica.ufmg.br

Universidade Federal de Minas Gerais
Programa de pós-graduação em Física - Enlight

Belo Horizonte - 2024

Notas de aula feitas pro Marcos Gabriel Alpino - início 07/24.

Contents

1 Ruído e Operações em Mecânica Quântica **2**
- 1.1 Ruído Clássico . . . 2
- 1.2 Teoria de Probabilidade . . . 3
  - 1.2.1 Probabilidades Condicionais . . . 3
- 1.3 Entropia e informação . . . 6
  - 1.3.1 O que é informação? . . . 7
- 1.4 Teoria de Comunicação . . . 12

2 Formalismo Quântico **14**
- 2.1 O Espaço Projetivo Complexo . . . 15
  - 2.1.1 Estrutura de Kähler . . . 16
  - 2.1.2 Métrica de Fubini-Study . . . 16
  - 2.1.3 Motivação . . . 17
- 2.2 Esfera de Bloch . . . 17
- 2.3 Estatística e Distância de Fubini-Study . . . 20
- 2.4 Estrutura do Espaço de Hilbert . . . 20
  - 2.4.1 Evolução e Reversão Temporal . . . 23
- 2.5 Estados Clássicos e Quânticos . . . 23

3 Estados Coerentes e Espaço de Fase Quântico **26**
- 3.1 Estados Coerentes. . . . 26
- 3.2 Espaço de Fase Quântico . . . 30

4 Operações Quânticas **32**
- 4.1 Medidas e POVMs . . . 32
- 4.2 *Reshaping* e *Reshuffling* . . . 34
- 4.3 Mapas Positivos e Completamente Positivos . . . 35
- 4.4 Mapas em Qubits . . . 38

Chapter 1

Ruído e Operações em Mecânica Quântica

A menos do universo como um todo, sistemas físicos tendem a interagir com outras entidades físicas externas ao sistema de interesse. Quando tais interações externas existem, dizemos que o sistema físico é aberto. Neste cenário, surge a necessidade de suplementar tais interações à teoria quântica o tratamento de sistemas físicos abertos.

Para tratar de sistemas como os citados anteriormente, usaremos o formalismo de Operações Quânticas (OQ). Contudo, antes de desenvolvermos a contrapartida quântica de estados abertos, veremos de maneira explícita o tratamento clássico ao ruído.

1.1 Ruído Clássico

Imaginemos um certo aparato capaz de codificar certa informação em um bit (i.e. 0 ou 1). Se tal aparato for não-interagente, no sentido de que nada o motiva a mudar de estado, esperaríamos que a tempo infinito o bit registrado fosse sempre o mesmo. Entretanto, quando tal disco é exposto ao contato do ambiente[1] a informação contida no bit pode mudar com uma certa distribuição de probabilidade. Por exemplo, imaginemos que a informação está codificada num sistema de dois níveis da seguinte maneira:

- O momento magnético do sistema está alinhado para baixo em relação a um eixo fixo e conhecido (condição de saída 0);
- O momento magnético do sistema está para baixo em relação ao eixo bem definido (condição de saída 1).

supondo ainda que o material é paramagnético, segue que a inclusão de um campo na direção do eixo bem definido pode (ou não) mudar a informação contida no disco.

A abordagem anterior de modelar tanto o ambiente quanto a interação entre sistema-ambiente é uma maneira de lidar com o ruído tanto na teoria clássica quanto na

[1] Por ambiente entende-se qualquer outro ente físico não presente no sistema fechado anterior.

quântica. Contudo, tal formalismo possui aplicabilidade limitada, uma vez que é necessário conhecer por completo as variáveis do ambiente.

Uma discussão um pouco mais sofisticada e elucidativa ocorre quando abordamos o problema de um ponto de vista da dinâmica estoscástica. Consideremos um computador clássico no qual tentamos implementar uma certa tarefa computacional. A tarefa imbuída ao computador é a aplicação sucessivamente duas portas lógicas do tipo *NOT*. A primeira porta lógica, devido ao ruído do computador com seu externo, não funciona de maneira determinística e produz um estado intermediário (Y). Entretanto, é experimentado que a segunda porta funciona de maneira precisa, ou seja, retornando um certo valor Z. Como resultado desta computação, temos a sequência (dita de sequência de Markov) $X \to Y \to Z$. Fisicamente, tal procedimento pode ocorrer, pois, mesmo sob à ação do mesmo ambiente, as duas portas agem em locais distintos (em primeira aproximação, bem distintos), portanto, dizendo que o ambiente atua de maneira independente e local, temos a possibilidade de haver o fenômeno aqui descrito.

O processo Markoviano de multi-estágios descrito anteriormente, é amplamente usado em teoria estocástica e permite transitarmos de uma teoria de operador densidade para uma teoria probabilística. A ideia central desta formulação é manipular a distribuição de probabilidade através de uma matriz de transferência satisfazendo o mapeamento $\vec{q} = E\vec{p}$), onde $\vec{p}$ e $\vec{q}$, são distribuição de probabilidades para entrada e saída do processo estocástico.

1.2 Teoria de Probabilidade

Na seção anterior, motivamos o estudo do ruído em sistemas clássicos. Vimos que explorar tais sistemas está intrinsecamente ligado à computação de distribuições de probabilidade, portanto, nesta seção discutiremos mais afundo sobre a teoria de probabilidade e alguns aspectos da teoria de informação.

1.2.1 Probabilidades Condicionais

Consideremos um evento A, com possíveis saídas (resultados) descritas por um conjunto $\{a_i\}$. Se o conjunto $\{a_i\}$ contém todas as possibilidades de saída do evento A, então a probabilidade de $A = a_i$ é dada por $P(a_i)$ e ainda

$$\sum_i P(a_i) = 1.$$

Neste cenário dizemos que a_i é a descrição estatística completa de A. Adicionando um segundo evento ao nosso conjunto universo, digamos B, com suas possíveis saídas sendo descritas por uma distribuição $\{b_i\}$, então A e B são bem definidos pelas probabilidades $P(a_i)$ e $P(b_j)$. Contudo, tais probabilidades não descrevem toda informação sobre o conjunto de eventos.

A descrição mais completa, dos possíveis eventos contidos nesse universo, é dada pelo que definimos como conjunto de probabilidades conjuntas, isto é $\{P(a_i, b_j)\}$, onde a vírgula surge como um denotador lógico da sentença "e". Logo, $P(a_i, b_j)$ lê-se como a probabilidade de $A = a_i$ e $B = b_j$.

É intuitivo entender que para um conjunto universo, onde há eventos independentes, a probabilidade conjunta das saídas é o produto individual de cada probabilidade. Simbolicamente,

$$P(a_i, b_j) = P(a_i)P(b_j),$$

para $A = a_i$ e $B = b_j$ disjuntos.

Dada uma distribuição de probabilidade conjunta de dois eventos, é possível recuperar as distribuições de cada evento separadamente[2]:

$$P(a_i) = \sum_j P(a_i, b_j)$$
$$P(b_j) = \sum_i P(a_i, b_j).$$

Outra questão fundamental surge ao pensarmos: é possível inferir alguma informação a respeito de um evento, dado que temos informação sobre outro? A resposta é sim e formalizaremos a discussão a seguir.

Suponhamos que a saída de um certo evento A é bem conhecida, digamos a_0. Neste cenário, estamos interessados em saber como as probabilidades de um conjunto distinto B estão condicionadas ao conhecimento de A. Denotaremos esta quantidade como a probabilidade condicional, isto é, o conjunto: $\{P(b_j|a_0)\}$[3]. Evidentemente, este conjunto deve ter alguma relação com o conjunto das probabilidades conjuntas, uma vez que as probabilidades conjuntas carregam toda informação sobre o universo. Formalmente,

$$P(b_j|a_0) \propto P(a_0, b_j).$$

Num caso geral, conforme descrito no esquema abaixo, obtemos:

$$P(a_i, b_j) = P(b_j|a_i)P(a_i). \tag{1.1}$$

Contudo, note que em nenhum momento até aqui houve uma preferência pelo conjunto A ou B, desta forma, com um raciocínio análogo ao desenvolvido há pouco, devemos ter: $P(a_i, b_j) = P(a_i|b_j)P(b_j)$. Destas observações, enunciemos o teorema de Bayes:

Teorema 1.1: Teorema de Beyes

Dados dois eventos A e B cujas possíveis saídas são descritíveis pelas distribuições $\{a_i\}$ e $\{b_j\}$, temos a relação fundamental

$$P(a_i|b_j) = \frac{P(b_j|a_i)P(a_i)}{P(b_j)}, \tag{1.2}$$

onde $P(a_i|b_j)$ ($P(b_j|a_i)$) é a probabilidade de se obter a_i (b_j) dado que $B = b_j$ ($A = a_i$) e $P(a_i)$ ($P(b_j)$) é a probabilidade de $A = a_i$ ($B = b_j$).

[2] As distribuições de probabilidade de eventos individuais, dentro de um universo de eventos, são chamadas de distribuições marginais.

[3] $P(b_j|a_0)$ lê-se como: a probabilidade de obtermos b_j dado que $A = a_0$ ou a probabilidade b_j dado a_0.

Rotineiramente, temos que lidar com casos onde a aprendizagem de um certo aspecto, i.e. obtenção de informação, interfere diretamente na distribuição de saídas deste evento para experimentações futuras. Descrevemos formalmente este conceito através da probabilidade de Fisher (Fisher's likelihood), cuja definição é:

Definição 1.1: Fisher likelihood

Dados dois eventos A e B com possíveis saídas $\{a_i\}$ e $\{b_j\}$, respectivamente. A probabilidade de se obter $A = a_i$ dado que $B = b_j$ é dada por

$$P(a_i|b_j) = \ell(a_i|b_j)P(a_i), \tag{1.3}$$

onde $l(a_i, b_j)$ é a likelihood de a_i dado b_j.

A definição acima pode ser comparada com o enunciado do Teorema de Bayes, obtendo:

$$\ell(a_i|b_j) \propto P(b_j|a_i).$$

Além disso, fica claro que para casos ondes os eventos em questão são independentes que a likelihood deve ter valor unitário.

Para uma sequência de eventos a likelihood torna-se um produtório. Os efeitos do conhecimento de $B = b_j$ são expressos como $P(a_i, b_j) \propto \ell(a_i|b_j)P(a_i)$, ou ainda, adicionando um novo evento no universo de valor bem definido, por exemplo, $C = c_k$, temos que

$$P(a_i|b_j, c_k) \propto \ell(a_i|b_j)\ell(a_i|c_k)P(a_i).$$

Exemplo 1.1: Probabilidade de Fisher na genética

Suponhamos que exista uma certa espécie de rato que possua duas colorações diferentes, por exemplo, preto e marrom. Se a cor preta, representada aqui por B, é dominante em relação à marrom, denotada por b, sem informação ancestral sobre os pais, um ratinho filhote tem uma distribuição equiprovável de ser preto ou marrom; sabendo que um certo rato de teste é preto, seu genes segue a seguinte distribuição de probabilidade:

$$P(BB) = \frac{1}{2} \qquad P(Bb) = \frac{2}{3}.$$

Perceba que se a prole deste ratinho teste com um marrom for marrom, então o rato teste é Bb. Entretanto, se seus filhotes forem todos pretos há um aumento de probabilidade dele ter o genes BB. Formalmente, se x_i denota a cor do i-ésimo ratinho nascido do cruzamento do rato teste com um marrom, então a probabilidade de Fisher associada a um dos ratos genitores ser BB (Bb) é dada, respectivamente, por:

$$\begin{aligned} \ell(BB|x_i = B) &\propto P(x_i = B|BB) \\ &= 1; \\ \ell(Bb|x_i = B) &\propto P(x_i = B|Bb) \\ &= \frac{1}{2}. \end{aligned}$$

Concatenando o resultado acima com a definição da probabilidade de Fisher, temos

$$\begin{aligned} P(BB|x_i = B) &\propto \ell(BB|x_i = B)P(BB) \\ &= \frac{1}{3} \\ P(Bb|x_i = B) &\propto \ell(Bb|x_i = B)P(Bb) \\ &= \frac{1}{3}, \end{aligned}$$

ou ainda, normalizando:

$$P(BB|x_i = B) = P(Bb|x_i = B) = \frac{1}{2}.$$

Confirmando que o nascimento de um filhote preto aumenta a probabilidade do ratinho teste ter genes BB.

1.3 Entropia e informação

Nesta seção motivaremos o porquê da entropia surgir como um bom quantificador de informação

1.3.1 O que é informação?

Suponhamos que A descreve um evento cujas saídas estão codificadas numa certa distribuição completa $\{a_i\}$. Veja que, se $A = a_0$ com toda certeza, i.e. $P(a_0) = 1$. Desta forma, descobrir (medir) a saída (resultado da medição) não nos traz nenhum tipo de informação sobre o sistema. Por exemplo, imagine uma superfície sem atrito na qual foi lançada uma bolinha de massa m com velocidade constante $\vec{v}$ à direta, sabemos pela segunda lei de Newton que o móvel estará a direta do seu ponto inicial de movimento. Logo, se o objetivo é saber se o corpo está à direta ou à esquerda do ponto inicial, medir sua posição não proverá mais informação do que já conhecida.

Uma segunda situação ocorre quando uma certa saída do evento tem uma probabilidade muito alta de acontecer, por exemplo, um camundongo foi envenenado, logo, após um grande período de tempo, a probabilidade do rato estar morto é alta. Contudo, não é certeza, ou seja, observar os sinais vitais do camundongo trará um pouco de informação.

A terceira situação ocorre quando duas bolas de cores diferentes, digamos uma azul e outra vermelha, são alocadas dentro de um saco e misturadas. Depois da mistura, Alice envia uma das bolas para Bob que está no Brasil e a outra para Charlie que vive no Japão. No primeiro momento, a remetente apenas pode dizer que cada um dos seus colegas tem 50% de chance de obter cada cor de bola. Todavia, caso um de seus colegas comunique a cor da bola que recebeu, Alice saberá com certeza qual a cor da bolinha do outro colega. Neste cenário, a informação adquirida pós "medição" é máxima.

Através da discussão anterior é intuitivo usar uma hipótese de concavidade sobre a função que quantificará a informação adquirida. Outra importante propriedade satisfeita por esse mensurador de informação deve ser a aditividade, uma vez que obter informação sobre eventos independentes não deve afetar o conhecimento do outro. Formalmente, se denotarmos por $h(a_i)$ o quantificador de informação sobre uma distribuição de probabilidade $\{a_i\}$, e sendo $\{b_j\}$ uma distribuição disjunta em relação a $\{a_i\}$, deve valer:

$$h(a_i, b_j) = h(a_i) + h(b_j).$$

Pelas propriedades requeridas, o funcional $h[\cdot]$ deve ser tal que

$$h(a_i) = -K \ln(a_i). \tag{1.4}$$

Definindo a entropia de informação. Além disso, em princípio, há diversas saídas possíveis a_i.Logo, um processo mais fidedigno à informação sobre o sistema seria uma média das entropias relacionadas a cada saída. Neste contexto, rotineiramente, definimos a entropia de informação como uma média nas probabilidades:

Definição 1.2: Entropia de Shannon

Dado um evento A descritível por uma distribuição $\{P(a_i)\}$ para cada saída a_i. Definimos a entropia de Shannon (ou entropia de informação) como

$$H(A) = -\sum_i P(a_i)\log(P(a_i)). \tag{1.5}$$

onde $\log(\cdot)$ é definido a menos de uma base. Em geral, encontramos na literatura duas bases mais comum: a base binária (i.e. $\log \equiv \log_2$) na onde a entropia tem dimensão de bits; e a base neperiana (i.e. $\log \equiv \ln$) cuja dimensão é dada por Nats.

Exemplo 1.2: Sistemas de dois níveis

Um evento é classificado com dois níveis quando o mesmo tem duas saídas possíveis, por exemplo, um elétron possui spin-$1/2$, pois seu spin na direção Z ou está para "cima" ou para "baixo". Nestas ocasiões, a entropia de Shannon se reduz ao que chamamos de entropia binária que, na base binária possui a seguinte forma:

$$H = -p\log_2 p - (1-p)\log_2(1-p). \tag{1.6}$$

Utilizando a fórmula do exemplo acima, obtemos os gráficos contidos na figura 1.1. Através dos plots podemos observar que a entropia é máxima quando $p = 1-p = 0.5$, ou seja, quando as duas saídas são equiprováveis. No caso do elétron de spin meio, quando eles está numa superposição equiprovável. Outro ponto importante retirado desta análise é que a entropia é nula quando a probabilidade de uma saída ocorrer é unitária (o que condiz com a heurística descritas no início da seção).

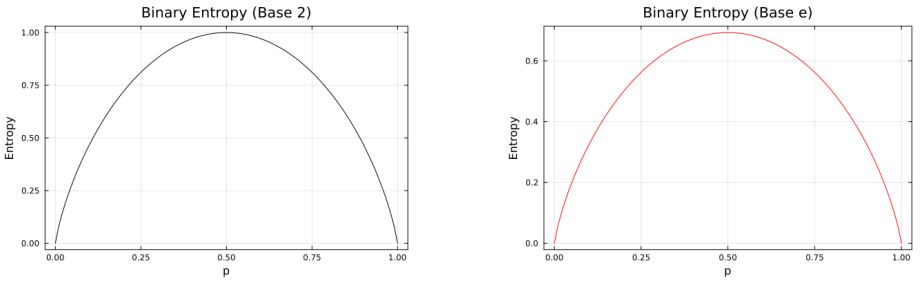

Figure 1.1: À esquerda, o plot da entropia binária na base 2; à direita o plot da entropia binária na base neperiana.

Na discussão acima, motivamos que a entropia deve ter um carácter côncavo, mas afinal, o que é isso? Por definição dizemos que uma função $f = f(x)$ é convexa dentro de

um intervalo se, e somente se, para um intervalo ordenado em, isto é
$X \coloneqq \{x_i | x_{i+1} \geqslant x_i, i \in \mathbb{N}\}$, vale

$$f\left(\sum_{i=1}^{n} p_i x_i\right) \leqslant \sum_{i=1}^{n} p_i f(x_i). \qquad \forall x_i \in X \tag{1.7}$$

Sendo $-f(x)$ convexa, então $f(x)$ é dita côncava.

Estudemos agora o que ocorre com a entropia quando o conjunto universo é composto, por exemplo, por dois eventos. Sejam A e B eventos probabilísticos descritos pelas respectivas distribuições $\{a_i\}$ e $\{b_j\}$. Então, podemos descrever a entropia do sistema total em termos da probabilidade conjunta dos eventos, formalmente

$$H(A,B) = -\sum_{ij} P(a_i, b_j) \log P(a_i, b_j). \tag{1.8}$$

As entropias de informação associadas a cada evento individual podem ser reescritas através das probabilidades marginais:

$$H(A) = -\sum_{ij} P(a_i, b_j) \log \left(\sum_{k} P(a_i, b_k)\right);$$
$$H(B) = -\sum_{ij} P(a_i, b_j) \log \left(\sum_{l} P(a_l, b_k)\right).$$

Neste contexto, computemos a diferença entre a soma das entropias individuais e a entropia conjunta do universo:

$$\begin{aligned} H(A) + H(B) - H(A,B) &= \sum_{ij} P(a_i, b_j) \left[\log P(a_i, b_j) - \log P(a_i) - \log P(b_j)\right] \\ &= \sum_{ij} P(a_i, b_j) \log \frac{P(a_i, b_j)}{P(a_i) P(b_j)}. \end{aligned} \tag{1.9}$$

Para a quantidade descrita em (**??**) damos o nome de informação mútua e a denotaremos por:

$$H(A:B) \coloneqq H(A) + H(B) - H(A,B) = H(\{P(a_i, b_j) || P(a_i) P(b_j)\}).$$

A medida acima mede a correlação entre as distribuições $\{P(a_i, b_j\}, \{P(a_i)P(b_j)\}$. Ademais, da própria definição segue que $H(A:B) = 0$ se, e somente se, A e B são eventos independentes. Uma outra maneira de escrever tal informação é em função da probabilidade de Fisher:

$$H(A:B) \coloneqq \sum_{ij} P(a_i, b_j) \log \ell(a_i | b_j), .$$

Lembrando que $\ell(a_i|b_j) = 1$, para a_i independente de b_j, recuperamos $H(A:B) = 0$.

Além disso, podemos mostrar que a informação mútua possui um limite relacionado com a entropia individual de cada evento:

$$\begin{aligned} H(A,B) - H(A) &= -\sum_{ij} P(a_i, b_j) \log P(a_i, b_j)/P(a_i) \\ &= \sum_i P(a_i) \left[\sum_j -P(b_j|a_i) \log P(b_j|a_i) \right] \end{aligned} \tag{1.10}$$

onde usamos a definição de probabilidade conjunta e o teorema de Bayes na última passagem[4]. Como os elementos da expressão acima são probabilidades é facilmente checado que (1.10) é uma medida positiva semi-definida. O fatos de que $H(A,B) - H(A)$ ($H(A,B) - H(B)$) possuir uma quota inferior nula e $H(A:B) \geqslant 0$, implicam:

$$0 \leqslant H(A:B) \leqslant \inf(H(A), H(B)). \tag{1.11}$$

Ainda sobre o significado físico da quantidade $H(A,B) - H(A)$, definindo

$$H(B|a_i) \coloneqq -\sum_j P(b_j|a_i) \log P(b_j|a_i),$$

como a quantidade de informação obtida sobre B dado que $A = a_i$. Concluímos que

$$\begin{aligned} H(B|A) &\coloneqq \sum_j P(a_i) H(B|a_i) \\ &H(A,B) - H(A), \end{aligned} \tag{1.12}$$

é a quantidade de informação de B condicionada a um conhecimento de A. Para qual nomeamos de entropia condicional de B em A[5].

E sobre a distribuição de probabilidade? Sabemos que, quando não há conhecimento sobre as possíveis saídas de um evento estocástico, o máximo que podemos dizer é que as saídas são equiprováveis entre si. Veja que esta forma de distribuir as probabilidades é equivalente a maximizar a entropia do evento [vide (1.4)]. Entretanto, como agir quando sabemos alguma informação sobre o sistema? Uma proposta é usar o método de máxima entropia originalmente proposto por Jayne. Para fixar e elucidar sobre essa ideia, façamos um exemplo

[4] De maneira análoga, é possível mostrar que:

$$H(A,B) - H(B) = \sum_j P(b_j) \left[-\sum_i P(a_i|b_j) \log P(a_i|b_j) \right]$$

.

[5] Utilizando este conceito, reescrevemos:

$$H(A,B) = H(A) + H(B|A).$$

Uma expressão análoga ao teorema de Bayes, mas, agora, para um conjunto de entropias.

Exemplo 1.3: Método de Entropia Máxima

Um dado é uma figura cúbica de seis faces, onde cada face contém um número natural de um até seis. Para o caso onde o dado não seja viciado (sem informação inicial), temos que o valor médio para a saídas do dado é $\frac{1+2+3+4+5+6}{6} = 3.5$. Nesta situação, cada saída tem uma probabilidade de $1/6$. Agora, imaginemos que o valor médio das saídas é dado por 3.47 (talvez, um dado pouco viciado), como fica a distribuição de probabilidade associado às saídas? Para isso, iremos maximizar a entropia em relação aos vínculos: 1- a soma das probabilidades deve ser unitária; 2- o valor médio é dado por 3.47 :

$$H = -\sum_{n=1}^{6} p_n \ln p_n + \lambda \left(1 - \sum_{n=1}^{6} p_n\right) + \mu \left(3.47 - \sum_{n=1}^{6} n p_n\right),$$

ou ainda,

$$\delta H = -\sum_{n=1}^{6} \delta p_n \left(\ln p_n + 1 + \lambda + \mu n\right).$$

Desta forma, a solução estacionária (i.e. $\delta H = 0$):

$$p_n = \exp[-(1+\lambda)] \exp(-\mu n).$$

Devido ao vínculo $\sum_n p_n = 1$, obtemos:

$$p_n = \frac{\exp(-\mu n)}{\sum_{i=1}^{6} \exp(-\mu i)}.$$

Determinando λ. Para determinar μ, partimos do resultado para o valor médio:

$$3.47(e^{-\mu} + e^{-2\mu} + \cdots + e^{-6\mu}) = e^{-\mu} + 2e^{-2\mu} + \cdots + 6e^{-6\mu}.$$

Resolvendo via cálculo simbólico (Mathematica 12): $\mu \approx 0.1029$.

Imaginemos uma caixa retangular de volume total V_0:

1. Dentro da caixa é colocada uma molécula de gás;
2. Separamos a caixa em dois volumes $V_0/2$, os quais a molécula pode estar à direita ou à esquerda;
3. Um ser inteligente observa a caixa e infere sobre a localização da partícula;
4. Sabendo onde a partícula está localizada, acopla-se um sistema de polia-massa à repartição de tal forma que o movimento da molécula é capaz de executar trabalho e levantar o massa pendurada;

5. A quantidade de trabalho obtida da expansão isotérmica do gás é dada por: $W = K_B T \ln(2)$, onde K_B é a constante de Boltzman, T é a temperatura do reservatório onde há o banho térmico.

1.4 Teoria de Comunicação

A ideia central de um sistema de comunicação é conectar a informação gerada por um fonte (medição) e transmiti-lá ao receptor[6]. De maneira geral, este processo envolve ruído, de tal forma que Bob tem acesso somente a um distribuição de probabilidade sobre o evento que Alice transmitiu sua medida.

A informação deve sair da Alice e ser transportada via um canal de informação, por exemplo, celular, computador, etc[7]. Contudo, o processo de comunicação dentro de tais canais não é perfeito, portanto, o tratamento físico adequado é análogo a nossa discussão sobre sistemas abertos. Em Bob, o sinal transmitido deve ser decodificado e interpretado.

Imaginemos que A seja um determinado evento no laboratório da Alice, que ela deseja transmitir seu resultado para Bob. Seja o conjunto $\{a_i\}$ formado por todas as possibilidades de mensagens que Alice pode mandar para Bob. A probabilidade de Alice selecionar e mandar a mensagem a_i é dada por $P(a_i)$, quantidade esta que também é compartilhada com Bob. A recepção do sinal mensagem ocorre em Bob. Após a decodificação Bob obtém o valor b_j que contém o sinal obtido juntamente com as possíveis mensagens associadas a ele.

Neste cenário, a comunicação entre Bob e Alice, pode ser descrita especificando a probabilidade condicional de Bob obter b_j dado que Alice enviou a_i, isto é, o conjunto $\{P(b_j|a_i)\}$.[8]. O caso mais simples ocorre quando cada decodificação de Bob corresponde a única mensagem de Alice, ou seja,

$$P(a_i|b_j) = \delta_{ij}.$$

Entretanto, os processo ruidoso de transmissão corrobora com a introdução de erros. Portanto, é necessário suplementar à distribuição de probabilidade o carácter estocástico do canal.

Em 1948, Shannon introduz dois teoremas fundamentais para entender estes processos:

Teorema 1.2: Codificação não ruidosa de Shannon

Escrever

O teorema de codificação sem ruído de Shannon quantifica a quantidade de informação que pode ser comprimida sendo produzida por uma fonte clássica[9]. Além das fontes

[6] Rotineiramente, tal sistema é conhecido como Alice e Bob, onde a transmissão e seus processos ocorrem no laboratório da Alice e a recepção no do Bob.

[7] Em geral, um transmissor pode ser qualquer aparelho com capacidade de codificar um texto em sinais elétricos e transmiti-los numa rede.

[8] Ou no ponto de vista do Bob: a probabilidade dela ter enviado a_i dado que ele possui b_j, simbolicamente, $P(a_i|b_j)$.

[9] Por fonte clássica, entende-se um conjunto de variáveis aleatórias, por exemplo, $\{X_i\}_{i\in\mathbb{N}}$.

clássicas, assumimos que nossa fonte de informação gera variáveis independentes e igualmente distribuídas. No mundo real, o processo de informação ocorre um pouco diferente, em português, proparoxítonas devem receber acentuação (condicionando a distribuição do carácter acento agudo).

Vamos supor a existência de uma fonte clássica ideal (conforme descrita no parágrafo anterior) produzindo a sequencia de bits[10]: $X_1, X_2, X_3, \cdots$ cada um com certa probabilidade p de ser 0. A central do teorema de Shannon é que há dois tipos de sequências binárias possíveis:

1. Sequências típicas, uma sequencia de saídas $\{x_i\}_{i\in\mathbb{N}}$ que ocorre com alta probabilidade;

2. Sequências atípicas, que ocorrem esporadicamente.

Das definições acima, obtemos:

$$p(\{x_i\}) = p(x_1)p(x_2)\cdots p(x_n)$$

Em termos práticos, devido a cada variável ter uma probabilidade p de ser zero, quando $n \to \infty$ a probabilidade de uma certa parte da sequência ser composta por zeros é dada por np. Desta forma, reescrevemos a probabilidade equacionada acima como:

$$p(\{x_i\}) \approx p^{np}(1-p)^{(1-p)n}, \tag{1.13}$$

ou ainda,

$$-\log p(\{x_i\}) \approx nH(\{X_i\}). \tag{1.14}$$

Note que, nesta formulação a entropia é calculada sobre a distribuição da fonte gerar uma saída zero ou um para um dado sinal X_i. Especialmente, denominaremos a quantidade $H(\{X_i\})$ como o quociente de entropia.

Para uma sequência de N bits, há 2^N possibilidade de mensagens distintas.

[10] Devidos às variáveis serem bits, elas somente assumem valores 0 ou 1.

Chapter 2

Formalismo Quântico

No formalismo da Mecânica Quântica (MQ), tratamos o estado quântico como um operador positivo semidefinido, de traço unitário e Hermitiano. Rotineiramente, denotamos o estado quântico como ρ, cuja definição é a seguinte:

Definição 2.1: Operador Densidade

Existem duas classes de operadores densidade:

1. Seja $|\psi\rangle$ um estado normalizado bem definido na álgebra de Dirac sobre um certo espaço complexo normado $\mathcal{H}$. Formalmente, define-se o operador densidade puro como:

$$\psi \coloneqq |\psi\rangle\langle\psi| \tag{2.1}$$

2. Define-se a soma convexa de estados puros também como um estado[a], ou seja,

$$\rho \coloneqq \sum_i p_i \, |\psi_i\rangle\langle\psi_i| \,, \tag{2.2}$$

onde p_i são probabilidades e $|\psi_i\rangle\langle\psi_i|$ é um certo i-ésimo estado puro.

[a] Uma vez que todas as propriedades requeridas para um estado quântico são preservadas sob soma convexa

A dinâmica na MQ é imposta via evolução unitária, proveniente da equação de Schrödinger. A transposição deste conceito para o formalismo dos operadores densidade é realizada através da equação de evolução temporal de Liouville-Neumann:

Definição 2.2: Equação de Liouville-Neumann

Seja $\rho = \rho(t)$ um operador densidade em um certo instante de tempo. Então, determinamos sua evolução temporal, isto é, $\rho(t')$ ($t' \neq t$), através da solução da seguinte equação:

$$i\dot{\rho} = [H, \rho], \tag{2.3}$$

onde $i = \sqrt{-1}$, $\dot{\rho}$ é a derivada temporal de ρ, e $[H, \rho]$ é o comutador entre os operadores Hamiltoniano e densidade, respectivamente.

A conexão entre os componentes abstratos da álgebra, ρ, e a física de laboratório é feita através da seguinte definição:

Definição 2.3: Valor Médio

Quantidades físicas são representadas por operadores Hermitianos. Por exemplo, uma certa quantidade física $\mathcal{A}$ atua sobre as variáveis dinâmicas por meio da composição de um operador Hermitiano A. Os possíveis valores de medição de $\mathcal{A}$ são determinados pelos autovalores de A. Simbolicamente, a definição é expressa como:

$$\langle A \rangle \coloneqq \mathrm{Tr}(A\rho), \tag{2.4}$$

onde $\langle A \rangle$ é o valor esperado do operador A, $\mathrm{Tr}(\cdot)$ denota a operação de traço, e ρ descreve a preparação física na qual a medição será realizada.

2.1 O Espaço Projetivo Complexo

O **espaço projetivo complexo** $\mathbb{C}P^n$ é definido como o conjunto de todas as linhas complexas em $\mathbb{C}^{n+1}$[1]. Formalmente, $\mathbb{C}P^n$ é o quociente do espaço vetorial complexo $\mathbb{C}^{n+1}$ pelo grupo multiplicativo $\mathbb{C}^*$, ou seja:

$$\mathbb{C}P^n = \frac{\mathbb{C}^{n+1}\backslash\{0\}}{\mathbb{C}^*},$$

onde $\mathbb{C}^{n+1}\backslash\{0\}$ denota o conjunto dos vetores complexos não nulos em $\mathbb{C}^{n+1}$, e $\mathbb{C}^*$ é o grupo dos números complexos não nulos sob a multiplicação. A relação de equivalência é dada por:

$$\mathbf{z} \sim \mathbf{w} \quad \text{se, e somente se,} \quad \exists \lambda \in \mathbb{C}^* \text{ tal que } \mathbf{z} = \lambda \mathbf{w}.$$

[1] Dado $\vec{z} \in \mathbb{C}^{n+1}$, definimos a linha complexa gerada por $\vec{z}$ como o conjunto $\{\lambda\vec{z} \mid \lambda \in \mathbb{C}\backslash\{0\}\}$. Note que, devido à definição acima, vetores que diferem apenas por uma escala complexa geram linhas complexas equivalentes.

O espaço $\mathbb{C}P^n$ pode ser visto como uma variedade complexa, onde cada ponto é uma linha projetiva complexa, ou seja, uma classe de equivalência de vetores em $\mathbb{C}^{n+1}$ que diferem apenas por uma multiplicação por um número complexo diferente de zero.

2.1.1 Estrutura de Kähler

O espaço projetivo complexo $\mathbb{C}P^n$ é equipado com uma estrutura de Kähler. Uma **variedade de Kähler** é um objeto central na interseção entre a geometria complexa, a geometria simplética e a geometria Riemanniana. Ela representa uma classe especial de variedades que possuem uma estrutura rica e compatível, permitindo a aplicação de técnicas de diversas áreas da matemática.

Formalmente, uma variedade de Kähler é uma variedade diferenciável M que é dotada de três estruturas inter-relacionadas:

1. **Estrutura Complexa** J: Uma aplicação $J : TM \to TM$ (onde TM é o fibrado tangente de M) que satisfaz $J^2 = -\text{id}$. Esta condição dá a M a estrutura de uma variedade complexa, permitindo que M seja tratada como um espaço onde funções e tensores complexos podem ser definidos e manipulados.

2. **Métrica Riemanniana** g: Uma métrica Riemanniana g sobre M que é compatível com J. Esta compatibilidade é expressa pela condição:

   $$g(JX, JY) = g(X, Y),$$

   para quaisquer campos vetoriais X e Y em M. Isto significa que J é uma isometria, preservando o comprimento e o ângulo entre vetores, e que a métrica g é hermitiana em relação à estrutura complexa J.

3. **Forma Simplética** ω: Uma 2-forma fechada ω em M, conhecida como a *forma de Kähler*, definida por:

   $$\omega(X, Y) = g(JX, Y).$$

   A condição de que ω seja fechada ($d\omega = 0$) garante que ω define uma estrutura simplética em M, o que implica que M possui uma geometria que pode ser descrita tanto em termos de coordenadas locais quanto em termos de fluxos Hamiltonianos.

Essas três estruturas não são independentes, mas sim interligadas de maneira que M se torna uma variedade complexa, Riemanniana e simplética ao mesmo tempo. A geometria de Kähler é, portanto, uma fusão dessas três áreas, onde cada uma influencia e enriquece as outras.

Um exemplo clássico de uma variedade de Kähler é o espaço projetivo complexo $\mathbb{C}P^n$, que é equipado com a métrica de Fubini-Study. Esta métrica é uma métrica de Kähler natural, invariável sob o grupo de simetria $SU(n + 1)$.

2.1.2 Métrica de Fubini-Study

Sejam $\mathbf{z}, \mathbf{w} \in \mathbb{C}^{n+1}$ dois vetores complexos não nulos que representam pontos no espaço projetivo $\mathbb{C}P^n$. A métrica de Fubini-Study entre os pontos projetivos associados a esses vetores é definida por:

$$\cos^2\left(\frac{d_{FS}(\mathbf{z},\mathbf{w})}{2}\right) = \frac{|\langle \mathbf{z},\mathbf{w}\rangle|^2}{\langle \mathbf{z},\mathbf{z}\rangle\langle \mathbf{w},\mathbf{w}\rangle},$$

onde $d_{FS}(\mathbf{z},\mathbf{w})$ é a distância de Fubini-Study entre os pontos projetivos, e $\langle \mathbf{z},\mathbf{w}\rangle$ denota o produto interno Hermitiano usual em $\mathbb{C}^{n+1}$. Explicitamente, podemos escrever:

$$d_{FS}(\mathbf{z},\mathbf{w}) = \arccos\left(\frac{|\langle \mathbf{z},\mathbf{w}\rangle|}{\sqrt{\langle \mathbf{z},\mathbf{z}\rangle\langle \mathbf{w},\mathbf{w}\rangle}}\right).$$

Essa métrica mede o ângulo entre as linhas complexas geradas por $\mathbf{z}$ e $\mathbf{w}$ em $\mathbb{C}^{n+1}$. Como tal, a métrica de Fubini-Study pode ser interpretada como uma generalização da métrica esférica (ou métrica de corda) para o espaço projetivo complexo.

2.1.3 Motivação

A métrica de Fubini-Study surge naturalmente em diversas áreas da Física. Uma de suas principais motivações vem do estudo de estados puros em mecânica quântica. Nesse contexto, os estados de um sistema quântico são representados por vetores em um espaço de Hilbert, mas estados que diferem apenas por uma fase global são fisicamente indistinguíveis. Isso leva à identificação dos estados quânticos com os pontos do espaço projetivo complexo $\mathbb{C}P^n$.

A métrica de Fubini-Study fornece uma maneira natural de medir distâncias entre estados quânticos puros, o que é crucial para entender a evolução quântica, o emaranhamento, e outras propriedades físicas que dependem da geometria do espaço de estados. Além disso, esta métrica é invariante sob transformações unitárias, uma propriedade desejável em física, pois as transformações unitárias correspondem a mudanças de base que não alteram as propriedades físicas do sistema.

Em geometria, a métrica de Fubini-Study é um exemplo canônico de uma métrica de Kähler, que combina as estruturas complexa, simplética e Riemanniana de maneira compatível. Esta métrica é utilizada em vários contextos, incluindo a teoria dos espaços de moduli e a geometria algébrica, onde desempenha um papel fundamental na descrição de variedades complexas projetivas.

2.2 Esfera de Bloch

Um importante sistema Físico de estudo é o sistema de dois níveis, isto é, sistemas do tipo Qbit. O exemplo canônico de aplicação deste sistema é uma partícula de spin-$1/2$ passando por um Stern-Gerlach[2], por exemplo. Além disso, como veremos mais adiante nas notas, os Qbits possuem enorme importância na implementação da Teoria de Informação Quântica (TIQ) e Computação Quântica (CQ). Neste cenário, surge o interesse em como representar de maneira similar o conjunto dos estados Qbits (i.e. uma parametrização que contenha tanto os estados puros quanto mistos). A representação que cumpre o proposto é a esfera de Bloch:

[2] Como a partícula tem spin-$1/2$, então as possibilidades de medição de spin numa certa direção são fixas e iguais a dois.

Definição 2.4: Esfera de Bloch

A esfera de Bloch é uma esfera de raio unitária que parametriza o conjunto de estados quânticos via

$$\rho = \frac{\mathbb{1} + \vec{r} \cdot \vec{\sigma}}{2}, \tag{2.5}$$

onde $\mathbb{1}$ é a identidade do espaço de matrizes 2×2, $\vec{r}$ é um vetor 3-dimensional real de módulo menor ou igual à unidade, e $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ tal que $\{\sigma_j\}_{j=x,y,z}$ são as j-ésima matrizes de Pauli geradoras de $SU(2)$. Para nos elucidarmos sobre essa questão provemos algumas de suas propriedades:

1. Seja ρ um operador densidade qualquer. Então, (2.5) representa ρ.

   Proof. Se ρ pertence à classe de matrizes densidades, logo, $\mathrm{Tr}(\rho) = 1$, $\rho \geqslant 0$ e $\rho^\dagger = \rho$.

   Se ρ está parametrizado na representação de Bloch. Então ele é Hermitiano, uma vez que tanto a identidade quanto o produto interno são invariantes sob à aplicação da tranposta-conjugada;

   A verificação do traço ser unitário é direta, pois, vale a linearidade do traço, $\mathrm{Tr}(\sigma_j) = 0$ para todo j e $\mathrm{Tr}(\mathbb{1}/2) = 1$;

   Para verificar a semi-positividade, calculamos os autovalores de ρ:

   $$\rho \doteq \frac{1}{2}\begin{pmatrix} 1 + r_z & r_x - ir_y \\ r_x + ir_y & 1 - r_z \end{pmatrix}.$$

   Cujo os autovalores são:

   $$\lambda_\pm = \frac{1 \pm |\vec{r}|}{2}. \tag{2.6}$$

   Além disso, pela definição de $\vec{r}$, temos que $|r| \leqslant 1$ implicando que os autovalores pertencem ao intervalo $[0, 1]$.

   □

2. A esfera é centrada no estado de máxima mistura.

   Proof. O estado de máxima mistura é aquele descrito pela combinação equiprovável dos projetores ortogonais da base, ou seja

   $$\rho_m = \frac{1}{2}\left(|+\rangle\langle+| + |-\rangle\langle-|\right), \tag{2.7}$$

   onde $|+\rangle$ ($|-\rangle$) é o autoestado de spin-up (down) de σ_j. Comparando diretamente a equação acima com a parametrização de Bloch, obtemos: $\rho \to \rho_m$ se $|r| \to 0$. Mostrando que o estado maximamente misto encontra-se no centro da esfera.

   □

Além do spin-$1/2$, a polarização de um fóton (com momentum fixo) também pode ser vista como um sistema de dois níveis. Para nos certificarmos da correspondência entre os casos utilizamos a projeção estereográfica de pontos em uma esfera de Riemann, representando esses pontos com um número complexo:

$$z = x + iy = \frac{\sin\theta e^{i\phi}}{1 + \cos\theta} = \tan\frac{\theta}{2}e^{i\phi}, \tag{2.8}$$

onde x e y são coordenadas estereográficas, θ e ϕ são a latitude e longitude (contadas a partir do polo norte).

Proof. Vamos mostrar que a projeção estereográfica de um ponto na esfera unitária para o plano complexo é dada por $z = \tan\left(\frac{\theta}{2}\right) e^{i\phi}$.

A projeção estereográfica mapeia pontos da esfera unitária S^2 (com centro na origem) para o plano complexo. Considere um ponto P na esfera, com coordenadas esféricas (θ, ϕ), onde θ é a latitude (medida a partir do polo norte) e ϕ é a longitude.

A relação entre as coordenadas esféricas (θ, ϕ) e as coordenadas cartesianas (x, y, z) na esfera é dada por:

$$\begin{cases} x = \sin\theta\cos\phi, \\ y = \sin\theta\sin\phi, \\ z = \cos\theta. \end{cases}$$

A projeção estereográfica de um ponto $P(\theta, \phi)$ da esfera para o plano complexo $\mathbb{C}$ é feita a partir do polo norte $(0, 0, 1)$ até o ponto P. A projeção estereográfica é dada por:

$$z = \frac{x + iy}{1 - z}.$$

Substituindo x, y, e z nas coordenadas esféricas:

$$z = \frac{\sin\theta\cos\phi + i\sin\theta\sin\phi}{1 - \cos\theta}.$$

Fatorando $\sin\theta$ no numerador:

$$z = \frac{\sin\theta(\cos\phi + i\sin\phi)}{1 - \cos\theta}.$$

Usando a identidade trigonométrica $\cos\phi + i\sin\phi = e^{i\phi}$:

$$z = \frac{\sin\theta e^{i\phi}}{1 - \cos\theta}.$$

Porém, segue da identidade trigonométrica: $\tan\left(\frac{\theta}{2}\right) = \frac{\sin\theta}{1-\cos\theta}$:

$$z = \tan\left(\frac{\theta}{2}\right) e^{i\phi}.$$

□

Considerando os polos norte e sul como pontos correspondentes à polarização circular, temos uma nova esfera de Riemann com coordenadas $w = \sqrt{z}$. Para cada ponto z, exceto 0 e ∞, existem dois pontos w distintos, associados a elipses orientadas.

Dessa forma, obtemos uma correspondência um-a-um entre os estados de polarização do fóton e os pontos z na primeira esfera de Riemann. Isso nos leva ao espaço de estados $\mathbb{CP}^1$, conhecido como esfera de Poincaré. Pontos antípodas na equador da esfera de Poincaré correspondem a estados linearmente polarizados em direções perpendiculares, por exemplo.

2.3 Estatística e Distância de Fubini-Study

Nesta seção, discutiremos como distinguir dois estados quânticos. Em termos físicos, para compararmos estados, devemos fixar alguma observável para ser nossa métrica. Em MQ, os observáveis físicos são operadores que atuam no espaço de Hilbert. Neste cenário, deve-se realizar um conjunto finito de medições em ensembles distintos[3], determinando, por fim, a distância estatística entre os ensembles.

Seja A um operador de dimensão $n+1$ completo. Então, a base que expande o espaço amostral de tal observável é $\{e_i\}_{i=1}^{n}$. Como A é completo, dois estados distintos $|\psi\rangle$ e $|\phi\rangle$ podem ser escritos como

$$|\psi\rangle = \sum_{i=0}^{n} \sqrt{p_i} e^{i\mu_i} |e_i\rangle, \qquad |\phi\rangle = \sum_{i=0}^{n} \sqrt{q_i} e^{i\nu_i} |e_i\rangle,$$

onde $\sqrt{p_i}$ e $\sqrt{q_i}$ são probabilidades que somam para a unidade.

Geometricamente, uma medida de distância entre distribuições de probabilidade é a fidelidade. A fidelidade possui a conotação geométrica de computar o produto interno entre vetores com componentes $\{(p_i)^{1/2}\}$ e $\{(q_i)^{1/2}\}$. Formalmente:

$$F(p_i, q_i) \coloneqq \sum_{i=0}^{n} \sqrt{p_i}\sqrt{q_i}. \tag{2.9}$$

2.4 Estrutura do Espaço de Hilbert

Diferentemente da estrutura da Mecânica Clássica (MC), as quantidades físicas mensuráveis no laboratório são descritas através de operadores que atuam em um espaço complexo normado. Os operadores em MQ podem tanto transformar vetores (estados) em outros estados quanto em números complexos. Por exemplo, seja X^I um vetor e A um operador: $X^I \rightarrow A^I_J X^J$, ou ainda, na forma quadrática, $X^I \rightarrow X^I A_{IJ} X^J$, um número complexo. Devido a essa estrutura, precisamos suplementar a teoria com algum ente capaz de transformar os observáveis (subir e descer os índices). Inicialmente, consideremos um espaço de Hilbert de dimensão dois, cujos vetores são determinados por uma matriz coluna com seu α-ésimo elemento dado por Z^α, tais que $\{Z^\alpha \in \mathcal{C} \mid \sum_\alpha Z^\alpha = 1\}$. Reescrevendo em termos de números reais, isto é, $Z^\alpha = x^\alpha + iy^\alpha$, obtemos que todo elemento $X^I \in \mathscr{H} \longrightarrow X^I \doteq [x^\alpha\ y^\alpha]^\intercal$.

[3] Numa visão frequentista, com tal conjunto de medições é possível construir uma distribuição de probabilidade e, a partir dela, determinar a distância entre os estados.

Neste cenário, computemos o produto interno entre dois vetores quaisquer:

$$\begin{aligned}\langle X_1|X_2\rangle &= (Z_1^1)^* Z_2^1 + (Z_1^2)^* Z_2^2 \\ &= (x_1^1 - iy_1^1)(x_2^1 + iy_2^1) + (x_1^2 - iy_1^2)(x_2^2 + iy_2^2) \\ &= x_1^1 x_2^1 + y_1^1 y_2^1 + x_1^2 x_2^2 + y_1^2 y_2^2 \\ &+ i\left(x_1^1 y_2^1 + x_2^2 y_2^2 - y_1^1 x_2^1 - y_1^2 x_2^2\right) \\ &= [x_1^1\, x_1^2\, y_1^1\, y_1^2][x_2^1\, x_2^2\, y_2^1\, y_2^2]^\intercal \\ &+ i[x_1^1\, x_1^2\, y_1^1\, y_1^2][y_2^1\, y_2^2\, (-x_2^1)\, (-x_2^2)]^\intercal \\ &= X_1^I g_{IJ} X_2^J + i X_1^I \Omega_{IJ} X_2^J, \end{aligned} \tag{2.10}$$

onde $X_i^\alpha = [x_i^1\, x_i^2\, y_i^1\, y_i^2]$, I, J são índices variando sobre o espaço complexo do par X_1, X_2, e g_{IJ} e Ω_{IJ} são os tensores definidos como:

$$g_{IJ} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{2.11}$$

$$\Omega_{IJ} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}. \tag{2.12}$$

A generalização destes tensores para um espaço de Hilbert N-dimensional complexo é dada por:

$$g_{IJ} = \begin{pmatrix} \mathbb{1}_d & 0 \\ 0 & \mathbb{1}_d \end{pmatrix}, \tag{2.13}$$

$$\Omega_{IJ} = \begin{pmatrix} 0 & \mathbb{1}_d \\ -\mathbb{1}_d & 0 \end{pmatrix}, \tag{2.14}$$

onde $\mathbb{1}_d$ é o operador identidade d-dimensional. Veja que o tensor g é simétrico e caracteriza um tensor métrico; já o tensor Ω satisfaz as condições simpléticas

$$\Omega^{IK}\Omega_{KJ} = \delta^I_J.$$

Ademais, seja

$$J^I_J = \begin{pmatrix} 0 & -\mathbb{1}_d \\ \mathbb{1}_d & 0 \end{pmatrix}. \tag{2.15}$$

Verificamos que J satisfaz

$$J^I_J = \Omega^{IK} g_{KI}; \quad g_{IJ} = \Omega_{IK} J^K_J, \quad J^2 = -\mathbb{1}_d.$$

A seguir, construiremos os observáveis em MQ. Seja O um tensor simétrico que satisfaça a relação quadrática, isto é,

$$\langle O \rangle \coloneqq X^I O_{IJ} X^J. \tag{2.16}$$

Então, O é capaz de mapear os estados X^I no conjunto dos números reais. Entretanto, O também deve transformar estados em estados. Para isso, devemos postular um método para subir um dos índices do tensor O. Utilizando o tensor simplético:

$$\tilde{O}^I_J = \Omega^{IK} O_{KJ}. \tag{2.17}$$

Utilizando as definições anteriores, provamos que o conjunto dos operadores O forma uma álgebra sobre os números reais:

Proof. Mostraremos que há um isomorfismo entre os colchetes de Poisson e a relação de comutação envolvendo os operadores O_i ($i = 1, 2$). Por definição, o colchete de Poisson entre F e G é dado por:

$$\{F, G\} \coloneqq \partial_i F \Omega^{ij} \partial_j G. \tag{2.18}$$

Tomando $F = X^I$ e $G = \langle O \rangle$ em (2.18), obtemos:

$$\begin{aligned}
\{X^I, \langle O \rangle\} &= \Omega^{IJ} \partial_J \langle O \rangle \\
&= \partial_I X^K \Omega^{IJ} \partial_J (X^M O_{MN} X^N) \\
&= \delta^K_I \Omega^{IJ} (\partial_J X^M O_{MN} X^N + X^M O_{MN} \partial_J X^N) \\
&= \delta^K_I \Omega^{IJ} \delta^M_J O_{MN} X^N + \delta^K_I \Omega^{IJ} X^M O_{MN} \delta^N_J \\
&= 2\Omega^{IK} O_{KJ} X^J \\
&= 2\tilde{O}^I_J X^J,
\end{aligned}$$

onde usamos a simetria do tensor O da quarta para quinta igualdade.

Invertendo a equação (2.17), temos que $O_{IJ} = \Omega_{IK} \tilde{O}^K_J$. Logo,

$$\begin{aligned}
\partial_I \langle O_\alpha \rangle &= (O_\alpha)_{IN} X^N + X^N (O_\alpha)_{NI} \\
&= \Omega_{IK} (\tilde{O}_\alpha)^K_N X^N + X^N \Omega_{NK} (\tilde{O}_\alpha)^K_I.
\end{aligned}$$

Desta forma:

$$\begin{aligned}
\{\langle O_1 \rangle, \langle O_2 \rangle\} &= \partial_I \langle O_1 \rangle 2(\tilde{O}_2)^I_J X^J \\
&= 2 \left[\Omega_{IK} (\tilde{O}_1)^K_N X^N (\tilde{O}_2)^I_J X^J + X^N \Omega_{NK} (\tilde{O}_1)^K_I (\tilde{O}_2)^I_J X^J \right] \\
&= 2X^I \left[\Omega_{IM} \tilde{O}_1{}^M_N \tilde{O}_2{}^N_J - \Omega_{IM} \tilde{O}_2{}^M_N \tilde{O}_1{}^N_J \right] X^J \\
&= 2 \left\langle \Omega[\tilde{O}_1, \tilde{O}_2] \right\rangle.
\end{aligned} \tag{2.19}$$

□

Acima, mostramos que os operadores definidos pelo par de equações (2.16) - (2.17) formam uma estrutura algébrica bem definida e atendem aos requisitos propostos pelo formalismo da MQ. Contudo, tais observáveis O são um tanto quanto gerais, ao contrário do que discutimos no início das notas, onde mencionamos que os operadores que descrevem observáveis físicos devem ser Hermitianos.

A Hermiticidade dos operadores O surge quando vinculamos à teoria desenvolvida acima a condição de comutatividade com o operador complexo J[4]. Como resultado, os operadores O resultantes são Hermitianos.

2.4.1 Evolução e Reversão Temporal

A MQ no espaço $2N$-dimensional real é invariante sob a ação de fases globais. De maneira geral, a evolução temporal é determinada via equação de von Neumann e as operações permitidas por unitárias que agem sobre o EH. O conjunto de operações unitárias é aquele que mantém invariante a métrica (produto interno), ou seja, não alteram o módulo dos vetores de estado. Na subseção anterior motivamos que uma boa medida em quântica é formada pela estrutura métrica de g e o tensor simplético Ω. Dizemos que uma operação U é unitária se, e somente se:

$$UgU^{\intercal} = g; \quad U\Omega U^{\intercal} = \Omega. \tag{2.20}$$

Em contrapartida, dizemos que Θ é uma operação anti-unitária se:

$$\Theta g\Theta^{\intercal} = g; \quad \Theta\Omega\Theta^{\intercal} = -\Omega. \tag{2.21}$$

Formalmente,

$$\begin{aligned} \langle \Theta X|\Theta Y\rangle &= X^I(\Theta^{\intercal}g\Theta)_{IJ}Y^J + iX^I(\Theta^{\intercal}\Omega\Theta)_{IJ}Y^J \\ &= X^I g_{IJ}Y^J - iX^I\Omega_{IJ}Y^J \\ &= \langle Y|X\rangle . \end{aligned} \tag{2.22}$$

Um operador anti-unitário contém consigo a ambiguidade $\Theta^2 = \pm\mathbb{1}$. Fisicamente, para sistema do tipo cadeias de *spin*, a determinação do sinal é feita por meio do carácter bosônico ou fermiônico dos constituintes do sistema. Por exemplo, seja um sistema de spin fracionário, digamos $1/2$, o espaço de Hilbert associado é equivalente ao discutido na seção anterior, ou seja, é isomorfo à $\mathbb{R}P^2$, determinando que $\Theta^2 = 1$. Em geral, para todo n par, onde $\mathscr{H} = \mathbb{C}^{n+1}$, tomamos o sinal positivo. Para casos onde n é ímpar temos que a composição dupla da mesma anti-unitária não transforma o estado nele mesmo. Neste cenário, um estado $|\psi\rangle \longrightarrow \Theta\,|\psi\rangle$ forma um par distinto em $\mathbb{C}P^{2k+1}$.

2.5 Estados Clássicos e Quânticos

Nesta seção abordaremos as similaridades e diferenças entre o conceito de estado em MC e MQ. No laboratório, as quantidades Físicas são expressas através de medições cujo

[4]Veja que os operadores descritos até aqui atuam em um espaço real $2N$-dimensional. Devemos retornar à estrutura do EH N-dimensional, impondo, portanto, a comutatividade com o operador J.

resultados são expressos no corpos dos reais. Vimos em seções anteriores que ambos formalismo são compatíveis com tal descrição.

Tanto a teoria clássica quanto a quântica, baseiam-se na ideia de que os observáveis atuando em estados. Tal operação pode ser vista como um mapeamento dos estados para os reais, por exemplo. Contudo, nossa interpretação do que seria o "constituinte elementar" destas teorias é estritamente ligada a como axiomatizamos a teoria.

Independente do conjunto de axiomas, os estados em ambas teorias são descritíveis por distribuições de probabilidades. Em MC, os estados são descritos por vetores de probabilidades majorados no espaço de fase clássico Q, P. Em contrapartida, os estados em MQ são operadores lineares atuando sobre um espaço complexo normado.

Denotando V^+ como o cone convexo num certo espaço vetorial V tal que todo elemento $x \in V^+$ é dito positivo semi-definido: seja $e : V \longrightarrow \mathbb{R}$ um funcional linear. Definimos o espaço de estados como o conjunto de todos os elementos x que satisfazem à condição de normalização $e(x) = 1$.

Vejamos que a definição dada ao final do parágrafo anterior acomoda muito bem os estados clássicos e quânticos. Por exemplo, seja x equivalente à $\vec{p}$, um vetor de probabilidade membro do espaço de fase $\mathbb{R}$, tomando $e(x) = \sum_i p_i$ é fácil mostrar que $\vec{p}$ são estados. De maneira mais sofisticada, x poderia ser visto como um subconjunto dos reais com medida unitária, isto é $x \longrightarrow Y \subset \mathbb{R}$) tal que $\mu(Y) = 1$ para uma certa região do espaço de fase. Já em termos da Física Quântica, $x \longrightarrow \rho$ juntamente com $e(\rho) = \mathrm{Tr}\,\rho$. Caso ρ seja proporcional à identidade, a distribuição de probabilidade clássica pode ser recuperada na MQ.

Os axiomas de Segal definem o que são observáveis num contexto independente de teoria. Em nossa notação, $\mathscr{A}$ é um subconjunto distinguível do conjunto de matrizes densidade $\mathscr{M}$ cujo elementos são determinados por ρ e denominados de estados. Neste cenário, definimos a forma dos observáveis através dos axiomas:

1. O conjunto dos observáveis forma um espaço linear real;
2. Os observáveis formam uma álgebra comutativa;
3. Existe um mapa bilinear de $\mathscr{M} \times \mathscr{A} \longrightarrow \mathbb{R}$;
4. Os observáveis formam uma álgebra diferenciável (álgebra de Lie).

Um resultado direto desta abordagem é que os estados surgem como mapas do conjunto dos observáveis para os reais.

Como a teoria clássica é baseada numa distribuição de probabilidade real que é uma função do espaço de fase, mostrando que as condições 2 e 4 são satisfeitas. Além disso, pela própria definição de medida, as demais condições também são confirmadas.

Em termos do formalismo quântico, tomamos os observáveis como matrizes Hermitianas agindo sobre um espaço de Hilbert finito formando uma álgebra comutativa (axioma 2). A relação de comutação é dada pelo parêntese de Lie com unidade complexa (axioma 4). O mapa que permite $\mathscr{M} \times \mathscr{A} \longrightarrow \mathbb{R}$ é o traço: $\langle A \rangle = \mathrm{Tr}(\rho A)$, ou ainda, $\rho(A) = \mathrm{Tr}(\rho A)$.

Como conclusão, a abordagem axiomática de Segal estabelece uma definição (e distinção) formal para os conjuntos de observáveis e estados. No formalismo de C^*-álgebra,

o conjunto dos observáveis assume um papel crucial. Os estados são definidos como funcionais sobre $\mathscr{A}$. No caso da MC, $\vec{p}$ mapeia um vetor $\vec{y} \in \mathbb{R}^N$ nos reais por meio do produto interno, isto é $\vec{p}\cdot\vec{y} \in \mathbb{R}$, ou ainda, para uma função integrável qualquer, temos $\mu : f(X) \longrightarrow \int_X \mathrm{d}\mu\,(x) f(x)$. Em termos da Quântica, a distribuição de probabilidade é obtida pelo mapa $\langle A \rangle = \mathrm{Tr}(\rho A)$, estabelecendo um princípio de incerteza intrínseco sem correspondência clássica.

Em um ponto de vista dinâmico, ambas teorias repeitam uma evolução temporal similar, seja para elementos no espaço de fase ou para operadores abstratos. Além disso, ambos os conjuntos são convexos. Entretanto, mesmo satisfazendo propriedades similares, os conjuntos de estados Quânticos e Clássicos se distinguem substancialmente quando olhamos para o subconjunto de estados puros.

Fisicamente, podemos descrever um estado puro Quântico como uma combinação convexa de vetores. Notemos também que os estados puros são invariantes por rotações, formando assim um conjunto contínuo. Em contrapartida, os estados em MC podem ser fundamentalmente discretos, cara e coroa, por exemplo.

Chapter 3

Estados Coerentes e Espaço de Fase Quântico

3.1 Estados Coerentes.

Dizemos que dois operadores, $\hat{q}$ e $\hat{p}$, pois formam uma álgebra de Heisenberg, ou seja:

$$[\hat{q}, \hat{p}] = i\hbar\mathbb{1}. \tag{3.1}$$

Como os comutadores $[[\hat{q}, \hat{p}], \hat{q}] = [[\hat{q}, \hat{p}], \hat{p}]$ são identicamente nulos, é válida a expansão de Baker-Hausdorff:

$$e^{\hat{q}}e^{\hat{p}} = e^{[\hat{q},\hat{p}]/2}e^{\hat{q}+\hat{p}} = e^{[\hat{q},\hat{p}]}e^{\hat{p}}e^{\hat{q}}$$

.

Dos operadores de momentum e posição, definimos o operador de aniquilação e criação, respectivamente, como

$$a = (\hat{q} + i\hat{p})/2; \tag{3.2}$$

$$a^\dagger = (\hat{q} - i\hat{p})/2. \tag{3.3}$$

Por meio dos operadores definidos acima, podemos adicionar partículas ao estado de vácuo denotado pelo vetor $|0\rangle$[1]. Note que a adição de partículas torna o estado de vácuo um estado descritível por sua posição e momentum. Simbolicamente,

$$|q, p\rangle = \hat{U}(q, p)\,|0\rangle\,,$$

onde $\hat{U}(q, p) = \exp[i(p\hat{q} + q\hat{p})]$ é representação unitária dos elementos do grupo de Heisenberg formado por $\hat{q}e\hat{p}$.

Os estados $|q, p\rangle$ são ditos estados coerentes e são caracterizados com por meio de suas coordenados no "espaço de fase", ou seja, q e p. Tais coordenadas podem ser parametrizadas pela curva:

$$z \coloneqq (q + ip)/\sqrt{2}.$$

[1] Por definição, o estado de vácuo é aquele onde a ação do operador de aniquilação é nula, isto é, $a\,|0\rangle = 0$

Além disso,

$$q = \frac{z + z^*}{\sqrt{2}i}; \qquad p = \frac{z - z^*}{\sqrt{2}}.$$

Neste cenário, reescrevemos a unitária $\hat{U}(q,p)$ em termos de z, obtendo

$$\begin{aligned} |z\rangle &= \exp\left(za^\dagger - z^* a\right)|0\rangle \\ &= \exp\left(-|z|^2/2\right) \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{n!}} |n\rangle \end{aligned} \tag{3.4}$$

Expressão esta que define os n-estados de Fock.

Os estados definidos em (3.4) são auto-estados do operador de destruição:

$$\begin{aligned} a|z\rangle &= \exp\left(-|z|^2/2\right) \sum_{n=1}^{\infty} \frac{z^n}{\sqrt{n!}} n |n-1\rangle \\ &= \exp\left(-|z|^2/2\right) \sum_{m=0}^{\infty} \frac{z^n}{\sqrt{m!}} |m\rangle, \end{aligned}$$

onde usamos ————————- . Veja que tal conceito é diretamente ligado ao laboratório, uma vez que em óptica medimos os fótons absorbidos. Factualmente, lasers ópticos produzem estados coerentes. Os estados $|n\rangle$ podem ser determinados no espaço de posição por meio da solução da equação de Schrödinger para o oscilador harmônico quântico. As soluções no espaço de posição são geradas via os polinômios de Hermite.

A função de onda $\psi_\alpha(x) = \langle x|\alpha\rangle$ no espaço de coordenadas é dada por:

$$\psi_\alpha(x) = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} \psi_n(x),$$

onde a função de onda para o estado de número n é:

$$\psi_n(x) = \left(\frac{1}{2^n n!}\right)^{1/2} \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega x^2}{2\hbar}} H_n\left(\sqrt{\frac{m\omega}{\hbar}} x\right), \tag{3.5}$$

tal que $H_n(x)$ denota os n-ésimo polinômio de Hermite. Ademais, dado que $\alpha = \frac{q+ip}{\sqrt{2}}$, temos:

$$\mathrm{Re}(\alpha) = \frac{q}{\sqrt{2}}, \quad \mathrm{Im}(\alpha) = \frac{p}{\sqrt{2}}. \tag{3.6}$$

Substituindo na expressão da função de onda:

$$\psi_\alpha(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left(-\frac{m\omega}{2\hbar}\left(x - \sqrt{\frac{2\hbar}{m\omega}}\frac{q}{\sqrt{2}}\right)^2\right) \exp\left(i\sqrt{\frac{2m\omega}{\hbar}}\frac{p}{\sqrt{2}}x\right) \tag{3.7}$$

$$= \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left(-\frac{m\omega}{2\hbar}\left(x - \frac{q}{\sqrt{m\omega}}\right)^2\right) \exp\left(i\frac{p}{\hbar}x\right), \tag{3.8}$$

ou ainda, em unidades naturais:

$$\psi_\alpha(x) = \pi^{-1/4} \exp\big[ipx - (x-q)^2/2\big]. \tag{3.9}$$

Verificando que a densidade de probabilidade é uma Gaussiana centrada em q.

Outro aspecto importante relacionado aos estados coerentes é que eles formam uma relação de completeza:

$$\frac{1}{2\pi}\int \mathrm{d}q\,\mathrm{d}p\,|q,p\rangle\langle q,p| = \tag{3.10}$$

Além disso, podemos mostrar que o produto da incerteza dos operadores $\hat{q}$ e $\hat{p}$ é o mínimo permitido pelo princípio de incerteza de Heisenberg. Neste sentido, tais estados são os mais "clássicos" permitidos pela teoria quântica.

Proof. O operador de posição $\hat{q}$ e o operador de momento $\hat{p}$ podem ser expressos em termos dos operadores de criação $\hat{a}^\dagger$ e aniquilação $\hat{a}$ do oscilador harmônico da seguinte forma:

$$\hat{q} = \frac{1}{\sqrt{2}}(\hat{a} + \hat{a}^\dagger), \tag{3.11}$$

$$\hat{p} = \frac{i}{\sqrt{2}}(\hat{a}^\dagger - \hat{a}). \tag{3.12}$$

Os estados coerentes $|\alpha\rangle$ são definidos como autovetores do operador de aniquilação $\hat{a}$:

$$\hat{a}\,|\alpha\rangle = \alpha\,|\alpha\rangle\,, \tag{3.13}$$

onde $\alpha = \frac{q+ip}{\sqrt{2}}$ é um número complexo.

A incerteza de q no estado $|\alpha\rangle$ é dado por:

$$\delta q = \sqrt{\langle\alpha|\hat{q}^2|\alpha\rangle - \langle\alpha|\hat{q}|\alpha\rangle^2}. \tag{3.14}$$

Calculando o valor esperado $\langle\alpha|\hat{q}|\alpha\rangle$:

$$\langle\alpha|\hat{q}|\alpha\rangle = \frac{1}{\sqrt{2}}\langle\alpha|(\hat{a} + \hat{a}^\dagger)|\alpha\rangle \tag{3.15}$$

$$= \frac{1}{\sqrt{2}}(\alpha + \alpha^*) = \frac{q}{\sqrt{2}}. \tag{3.16}$$

Agora, o valor esperado de $\hat{q}^2$:

$$\langle\alpha|\hat{q}^2|\alpha\rangle = \frac{1}{2}\langle\alpha|(\hat{a} + \hat{a}^\dagger)^2|\alpha\rangle \tag{3.17}$$

$$= \frac{1}{2}\left[\langle\alpha|\hat{a}^2|\alpha\rangle + \langle\alpha|(\hat{a}^\dagger)^2|\alpha\rangle + 2\langle\alpha|\hat{a}^\dagger\hat{a}|\alpha\rangle + 1\right]. \tag{3.18}$$

Como $\langle\alpha|\hat{a}|\alpha\rangle = \alpha$ e $\langle\alpha|\hat{a}^\dagger|\alpha\rangle = \alpha^*$, temos:

$$\langle\alpha|\hat{q}^2|\alpha\rangle = \frac{1}{2}\left[|\alpha|^2 + 1\right] = \frac{q^2 + p^2 + 1}{2}. \tag{3.19}$$

A incerteza de q:

$$\delta q = \sqrt{\frac{1}{2}} = \frac{1}{\sqrt{2}}. \tag{3.20}$$

De maneira análoga, o desvio padrão de p é dado por:

$$\delta p = \sqrt{\langle\alpha|\hat{p}^2|\alpha\rangle - \langle\alpha|\hat{p}|\alpha\rangle^2}. \tag{3.21}$$

Calculando o valor esperado de $\hat{p}$:

$$\langle\alpha|\hat{p}|\alpha\rangle = \frac{i}{\sqrt{2}}\langle\alpha|(\hat{a}^\dagger - \hat{a})|\alpha\rangle \tag{3.22}$$

$$= \frac{i}{\sqrt{2}}(\alpha^* - \alpha) = \frac{p}{\sqrt{2}}. \tag{3.23}$$

Agora, para $\hat{p}^2$:

$$\langle\alpha|\hat{p}^2|\alpha\rangle = \frac{1}{2}\left[|\alpha|^2 + 1\right] = \frac{q^2 + p^2 + 1}{2}. \tag{3.24}$$

Logo, a incerteza sobre p é:

$$\delta p = \frac{1}{\sqrt{2}}. \tag{3.25}$$

A relação de incerteza de Heisenberg estabelece que:

$$\delta q\, \delta p \geqslant \frac{1}{2}. \tag{3.26}$$

Como vimos, nos estados coerentes $\delta q = \delta p = \frac{1}{\sqrt{2}}$, de forma que:

$$\delta q\, \delta p = \frac{1}{\sqrt{2}} \times \frac{1}{\sqrt{2}} = \frac{1}{2}. \tag{3.27}$$

Portanto, o produto dos desvios atinge o limite mínimo permitido pela relação de incerteza de Heisenberg.

□

3.2 Espaço de Fase Quântico

Em geral, elementos numa teoria clássica podem ser obtidos por meio de integrações no espaço de fase. A transcrição deste conceito para teoria quântica pode ser obtido quando requeremos a existência de uma função quasi-probabilistisca tal que:

$$\langle q|\rho|q\rangle = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}p\, W(q,p), \tag{3.28}$$

onde ρ denota uma matriz densidade qualquer. Em termos de probabilidade:

$$P(q_1 \leqslant q \leqslant q_2) = \frac{1}{2\pi}\int_{q\in[q_1,q_2]} \mathrm{d}q \int_{-\infty}^{\infty} \mathrm{d}p\, W(q,p).$$

O espaço de fase definido pelos operadores $\hat{q}$ e $\hat{p}$ pode ser parametrizado em termos do parâmetro angular θ de tal forma que[2]

$$\hat{q_\theta} \coloneqq \hat{q}\cos\theta + \hat{p}\sin\theta; \qquad \hat{p_\theta} \coloneqq -\hat{q}\sin\theta + \hat{p}\cos\theta.$$

Teorema 3.1: Bertrand e Bertrand - Unicidade de $W(q,p)$

A função $W_\theta = W(q(\hat{q_\theta},\hat{p_\theta}), p(\hat{q_\theta},\hat{p_\theta}))$ é única e determinada quando satisfaz:

$$\langle \hat{q_\theta}|\rho|\hat{q_\theta}\rangle = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\hat{p_\theta}\, W_\theta. \qquad \forall\theta \tag{3.29}$$

Definição 3.1: Função de Wigner

A função $W = W(q,p)$ é definida como

$$W(q,p) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}u \int_{-\infty}^{\infty} \mathrm{d}v\, \tilde{W}(u,v)\exp(iuq+ivp), \tag{3.30}$$

onde $\tilde{W}(u,v) = \mathrm{Tr}[\rho\exp(-iu\hat{q}-iv\hat{p})]$ define a transformada de Fourier Quântica de ρ.

Antes de prosseguirmos com o desenvolvimento da função W, vejamos como determinar a transformada de Fourier Quântica no espaço de posição. Primeiramente, é claro que a expansão BH vale para o termo exponencial, tal que:

$$\exp(-iu\hat{q}-iv\hat{p}) = \exp(iuv/2)\exp(-iu\hat{q})\exp(-iv\hat{p})$$

[2] Esta parametrização define os operadores de posição e momentum em termos de uma parâmetro angular que pode ser muito bem definido no aparato experimental de, por exemplo, óptica quântica.

$$= \exp(iuv/2) \int \mathrm{d}q \exp(-iu\hat{q}) \left|q\right\rangle\!\left\langle q\right| \exp(-iv\hat{p}) \left|q\right\rangle\!\left\langle q\right|$$
$$= \exp(iuv/2) \int \mathrm{d}q \exp(-iuq) \left|q\right\rangle\!\left\langle q - v\right| .$$

Entretanto, tomando $q = q + v/2$, obtemos:

$$\exp(-iu\hat{q} - iv\hat{p}) = \int \mathrm{d}q \left|q + v/2\right\rangle\!\left\langle q - v/2\right| . \tag{3.31}$$

Portanto

$$\tilde{W}(u, v) = \int \mathrm{d}q \left\langle q - v/2|\rho|q + v/2\right\rangle \exp(-iuq) \tag{3.32}$$

Chapter 4

Operações Quânticas

Neste capítulo, trataremos os sistemas Quânticos com um grau de liberdade a mais, isto é, o ambiente.

4.1 Medidas e POVMs

Seja $\mathscr{A}$ um sistema Físico descrito por um espaço de Hilbert $\mathscr{H}_A$ de dimensão N. Ademais, as operações Quânticas podem ser separadas em classes, sendo elas divididas por $1-$ transformações unitárias; $2-$ transformações não unitárias. No caso de sistemas quânticos abertos[1] podemos descrever uma extensão (para o sistema $\mathscr{A}$) definindo um espaço de Hilbert estendido definido por $\mathscr{H} := \mathscr{H}_A \otimes \mathscr{H}_K$. Em termos do estado, sendo $\rho \in \mathscr{H}_A$ (o estado do sistema de interesse) e $\sigma \in \mathscr{H}_K$ de dimensão K (o estado que descreve o ambiente). Formalmente,

$$\rho' \longrightarrow \rho \otimes \sigma. \tag{4.1}$$

Neste cenário, uma terceira operação permitida : $3-$ redução, de tal forma que o estado do sistema Físico pode ser obtido pelo traço parcial sobre o ambiente. Por fim, $4-$ medições seletivas, as quais têm como resultado da medição um estado específico. [2]

Uma suposição possível é que o sistema de interesse é aberto (há contato com o ambiente), contudo, o sistema estendido é, de fato, isolado. Desta forma, operações sobre o espaço estendido são do tipo unitárias. Simbolicamente, obtemos

$$\begin{aligned} \rho' &= \mathrm{Tr}_K[U(\rho \otimes |\nu\rangle\langle\nu|)U^\dagger] \\ &= \sum_{\mu=1}^{K} \langle\mu|U|\nu\rangle \, \rho \, \langle\nu|U^\dagger|\mu\rangle \,, \end{aligned} \tag{4.2}$$

[1] Ou, simplesmente, sistemas abertos.

[2] O conjunto $(1-3)$ definem operações determinísticas. Já as operações permitidas em $4-$ formam o conjunto de medições seletivas.

onde $\{|\mu\rangle\}_{i\in I_k}$[3] denota uma base completa para o ambiente. Definindo um operador A_μ como

$$A_\mu = \langle\mu|U|\nu\rangle. \tag{4.3}$$

Veja que

$$\sum_{\mu=1}^{K}(A_\mu)^\dagger A_\mu = \sum_{\mu}\langle\nu|U^\dagger|\mu\rangle\langle\mu|U|\nu\rangle = \mathbb{1}.$$

Definição 4.1: Processo de medida

Seja o espaço de possíveis medidas descrito por um conjunto de k elementos. A cada elemento, associamos um operador A_i que satisfaz uma relação de completeza.
A medição de A_i sobre um estado inicial ρ produz um estado ρ_i com uma dada probabilidade p_i definidos da seguinte forma:

$$\rho \longrightarrow \rho_i = \frac{A_i\rho A_i^\dagger}{\mathrm{Tr}\Big(A_i\rho A_i^\dagger\Big)}; p_i(\rho=\rho_i) = \mathrm{Tr}\Big(A_i\rho A_i^\dagger\Big). \tag{4.4}$$

Definindo uma medição seletiva. Em contrapartida, caso não haja seleção, temos uma soma convexa sob as possíveis saídas pós a ação de A_i. Simbolicamente,

$$\rho \longrightarrow \rho' = \sum_{i=1}^{k} A_i\rho A_i^\dagger. \tag{4.5}$$

Um exemplo particular do formalismo descrito acima ocorre quando o conjunto $\{A_i\}_{i\in I_k}$ é formado por projetores ortogonais entre si. Neste caso, dizemos que a ação dos A_i's define um conjunto de medições projetivas.

Definição 4.2: *Positive Operator values Measure* (POVM)

Em geral, quando relaxamos a condição de ortogonalidade entre os $\{A_i\}$'s definimos a classe de *Positive Operator values Measure* (POVM). Formalmente, um POVM de k-elementos, sobre $\mathscr{H}_N$, é descrito pelo conjunto de operadores $\{E_i\}_{i\in I_k}$, hermitianos, positivos e que somam a uma partição da unidade, isto é:

$$\sum_{i=1}^{k} E_i = \mathbb{1};\ \ E_i = E_i^\dagger;\ \ E \geqslant 0. \tag{4.6}$$

[3] Aqui denotaremos I_j como um conjunto indexador de dimensão j.

Num processo de tomografia quântica qualquer é possível reconstruir um estado quântico desconhecido através de um conjunto de medições. Neste contexto, dizemos que um POVM é informacionalmente completo se, e somente se, é possível reconstruir a matriz densidade somente com o conjunto estatístico proveniente da implementação do POVM, isto é, através das probabilidades $p_i = \mathrm{Tr}(E_i\rho)$. Outra definição importante neste conceito de POVM é sobre a pureza do conjunto, dizemos que um POVM é puro se cada elemento E_i do POVM é um operador de rank unitário.

Anteriormente, argumentamos que a estrutura do POVM generaliza as medições projetivas. Entretanto, é possível mostrar que, num espaço maior, o POVM é equivalente a uma medição projetiva. Em geral, dado um POVM puro de k elementos e um estado ρ sobre um espaço de Hilbert de dimensão N, é possível construir $\rho \otimes \rho_0 \in \mathscr{H}_N \otimes \mathscr{H}_K$ tal que, a estatística proveniente do POVM em $\mathscr{H}_N$ é equivalente a de uma medição projetiva no espaço estendido $\mathscr{H}_N \otimes \mathscr{H}_K$.

Teorema 4.1: Dilatação de Naimark

Qualquer POVM $\{E_i\}$ agindo sobre um espaço de Hilbert $\mathscr{H}$ pode ser *dilatado* como operador projetivo com resolução fracionária à unidade num espaço maior.

4.2 *Reshaping* e *Reshuffling*

O *reshape* de uma matriz retangular qualquer, consiste em reescrever a matriz num ordenamento *lexicographical*, isto é, linha após linha:

$$\vec{a}_k = A_{ij};\ \ k = (i-1)N + j;\ \ i \in I_M;\ \ j \in I_N. \tag{4.7}$$

Em geral, os elementos $\vec{a}_k$ podem se transformar linearmente sob a ação de um operador C de dimensão $MN \times MN$ (cujo os elementos denotaremos por $C_{kk'}$). Contudo, k são índices que correm sobre o ordenamento lexiocographical, de maneira geral, um operador neste espaço assume a seguinte representação tensorial:

$$C^{m\mu}_{n\nu};\ \ m, n = 1, \cdots, M;\ \ \mu, \nu = 1, \cdots, N.$$

Nesta notação, o traço parcial sobre o sistema B, numa bipartição AB, se escreve como

$$C^{A}_{mn} = C^{m,\mu}_{n,\mu}.$$

Dos conceitos definidos no parágrafo anterior somos capazes de demonstrar que o produto tri-matricial, isto é o produto de ABC onde os entes são matrizes, como

$$\Phi(\vec{B}) = ABC, \tag{4.8}$$

onde $\Phi \coloneqq A \otimes C^{\intercal}$, e $\vec{B}$ é a vetorização lexicographical da matriz B.

Consideremos $\tilde{U}$ uma matriz unitária de dimensão N^2. Executando um reshape em suas colunas, podemos obter uma base ortonormal $\{A_k\}$. Veja que para cada k há dois

índices (um para cada elemento da base do espaço estendido). Formalmente, um vetor X de tamanho N^2, no espaço estendido, assume a seguinte decomposição:

$$X = \sum_{i=1}^{N^2} \sum_{j=1}^{N^2} C_{ij} \left|A^i\right\rangle \left|A^j\right\rangle .$$

Da definição acima, segue que a matriz dos coeficientes

$$C_{n\nu}^{m\mu} \coloneqq \mathrm{Tr}\big[(A^{m\mu} \otimes A^{n\nu})^\dagger X\big],$$

onde usamos a mudança para índices duplos na forma: $i = N(m-1) + \mu$.

Por fim, podemos considerar que a matriz unitária é, de fato, a identidade no espaço reduzido. Desta forma, a matriz C, para um dado vetor X, é $C_{n\nu}^{m\mu} = X_{\mu\nu}^{mn}$, a qual definimos como o *reshuffle* de X. Na próxima seção discutiremos mais sobre os aspectos Físicos desta transformação.

Outra operação que virá a ser útil no contexto de Física, em especial, na mensurabilidade de Emaranhamento, é a transposição parcial. Seja um vetor $X \in \mathscr{H}_{HS} \otimes \mathscr{H}_{HS}$, definimos a transposição parcial por meio da ação do operador $T_A \coloneqq T \otimes \mathbb{1}$ (análogo para transposição parcial em B). Simbolicamente,

$$(X_{n\nu}^{m\mu})^{T_A} = X_{m\nu}^{n\mu} \qquad \& \qquad (X_{n\nu}^{m\mu})^{T_B} = X_{n\mu}^{m\nu}.$$

Note que vale a expansão

$$\rho = \sum_k p_k A^k \otimes B^k.$$

Desta forma, $((A^\intercal)^\dagger)_{ij} = A_{ij}^*$. Contudo, se ρ é hermitiano implica em A^k ser hermitiano, ou seja, $A_{ij}^* = A_{ji} = (A^\intercal)_{ij}$ mostrando que a transposição parcial preserva a hermiticidade. Em contrapartida, o espectro associado a ρ é afetado, uma vez que os elementos de matrizes não são iguais ao do operador original.

A última operação que definiremos nessa seção é o Swap. O Swap consiste em trocar os graus de liberdades no espaço estendido, por exemplo $S\left|01\right\rangle = \left|10\right\rangle$ onde S denota a operação de Swap. Formalmente,

$$S \coloneqq \sum_{i,j=1}^{N} \left|i,j\right\rangle\!\left\langle i,j\right| . \tag{4.9}$$

4.3 Mapas Positivos e Completamente Positivos

Seja $\rho \in \mathscr{M}^N$ um operador densidade agindo sobre um espaço de Hilbert N-dimensional. Nesta seção discutiremos os requisitos Físicos e Matemáticos para um mapa $\Phi : \mathscr{M}^N \longrightarrow \mathscr{M}^N$ deve satisfazer para possuir significado Físico.

O primeiro requisito é que Φ deve agir linearmente sobre os vetores de estado, isto é

$$\rho' = \Phi_{n\nu}^{m\mu} \rho_{n\nu}. \tag{4.10}$$

Notemos que ação do mapa sob uma não-homegenidade adicionando um termo proporcional à identidade tal que

$$(\Phi_{n\nu}^{m\mu})' = \Phi_{n\nu}^{m\mu}\rho_{n\nu} + \sigma_{n\nu}\delta_{m\mu}^{n\nu}. \tag{4.11}$$

As condições Físicas impostas sobre $\rho' \in \mathscr{M}^N$ implicam condições sobre os mapas. Formalmente,

1. A hermiticidade das matrizes densidade é reescrita como a hermiticidade do mapa:

$$\rho' = (\rho')^\dagger \iff \Phi_{n\nu}^{m\mu} = (\Phi^S)_{\nu n}^{\mu m};$$

2. Normalização da matriz densidade

$$\mathrm{Tr}\,\rho' \iff \Phi_{n\nu}^{mm} = \delta_{n\nu};$$

3. Critério de positividade

$$\rho' \geqslant 0 \iff \Phi_{n\nu}^{m\mu}\rho_{n\nu} \geqslant 0.$$

As condições acima são verdadeiras, contudo, pouco intuitivas ou de difícil computação. Desta forma definimos a matriz de transferência

Definição 4.3: Matriz Dinâmica

Definimos a matriz dinâmica como o reshuflle de um Mapa. Simbolicamente,

$$D_\Phi \coloneqq \Phi^R \implies D_{n\nu}^{m\mu} = (\Phi_{n\nu}^{m\mu})^R = \Phi_{\mu\nu}^{mn}. \tag{4.12}$$

Definida a matriz dinâmica, reescrevemos as condições físicas sobre a matriz densidade em termos da matriz D_Φ. A hermiticidade do estado quântico se traduz como a hermiticidade da matriz dinâmica, isto é, $D_{\mu\nu}^{mn} = (D_{\mu\nu}^{mn})^\dagger$; a unidade do traço de um operador densidade se reduz a $D_{m\nu}^{mn} = \delta_{n\nu}$; já a positividade do mapa implica na bloco positividade da matriz dinâmica (vide o teorema de Jamiolkowski)

Teorema 4.2: Jamiolkowski

Um mapa linear Φ é positivo se, e somente se, sua matriz dinâmica correspondente, D_Φ é bloco positiva.

Em mecânica Quântica, sempre é possível acoplar o sistema a um sistema auxiliar, i.e. $\rho \longrightarrow \rho \otimes \sigma$. Em geral, pode ocorrer de um certo mapa Φ agindo sobre um estado ρ ser

positivo, contudo, sua ação sobre ρ no espaço estendido não ser. Com isso em mente, definimos a classe de mapas *completamente positivos* como a classe composta por mapas positivos Φ tais que $\Phi \otimes \mathbb{1}$ é positivo. Em especial, dizemos que para um mapa ter significado físico (realizável no laboratório) ele deve ser complemente positivo.

Para testar a positividade completa de um mapa, utilizamos o Teorema de Choi:

Teorema 4.3: Choi

Um mapa linear Φ é completamente positivo se, e somente se, a sua matriz dinâmica D_Φ é positiva.

Dos conceitos descritos anteriormente, fica claro que os mapas completamente positivos são isomorfos ao conjunto das matrizes dinâmicas positivas. Além disso, dizemos que um mapa é traço preservante se sua matriz dinâmica satisfaz $\mathrm{Tr}(D_\Phi) = N$, onde D_Φ é uma matriz de tamanho N.

Uma matriz dinâmica positiva assuma uma decomposição na forma espectral:[4]

$$D_\Phi = \sum_i \left|A^i\right\rangle\!\left\langle A^i\right| \implies D^{mn}_{\mu\nu} = \sum_i A^i_{mn} \bar{A}^i_{\mu\nu} \tag{4.13}$$

Neste cenário, deve ser possível reescrever o resultado do teorema de Choi em termos de tais vetores (matrizes):

Teorema 4.4: Representação de Krauss

Um mapa linear Φ é completamente positivo se, e somente se, ele possui a seguinte forma

$$\rho \longrightarrow \rho' = \sum_i A_i \rho A_i^\dagger \tag{4.14}$$

Em especial, um mapa é traço preservante se, e somente se

$$\sum_i A_i^\dagger A_i = \mathbb{1}_N. \tag{4.15}$$

Vejamos também que a unicidade da decomposição não está garantida pelo teorema anterior. Para isso, precisamos de mais um estrutura formal:

[4] Como $D_\Phi \in \mathscr{H}_N \otimes \mathscr{H}_N$, temos que $\left|A^i\right\rangle$ é um vetor N^2-dimensional. Logo, via reshape, A_{ij} denota uma matriz $N \times N$.

Teorema 4.5: Forma Canônica de Krauss

Um mapa completamente positivo $\Phi : \mathscr{M}^N \longrightarrow \mathscr{M}^N$, cuja matriz dinâmica é dada por

$$D_\Phi = \sum_i d_i \, |\chi_i\rangle\langle\chi_i|$$

, pode ser representado como

$$\begin{aligned} \rho \longrightarrow \rho' &= \sum_{i=1}^{r} A_i \rho A_i^\dagger \\ &= \sum_{i=1}^{r \leqslant N^2} d_i \chi_i \rho \chi_i^\dagger, \end{aligned} \tag{4.16}$$

onde r denota o rank de Krauss e

$$\mathrm{Tr}\, A_i^\dagger A_j \coloneqq \sqrt{d_i d_j}\, \langle \chi_i | \chi_j \rangle = d_i \delta_{ij}.$$

Por meio dos teoremas anteriores, justificamos que, para Φ traço preservante temos $D_{m\nu}^{mn} = N$. Em especial, a representação em soma pode ser escrita como

$$\Phi = \sum_{i=1}^{N^2} A_i \otimes \bar{A}_i = \sum_{i=1}^{N^2} d_i \chi_i \otimes \bar{\chi}_i^\dagger \tag{4.17}$$

4.4 Mapas em Qubits

Bibliography