Protocolo para Autenticação Quântica de Mensagens Clássicas

Rex Antonio da Costa Medeiros

Dissertação de Mestrado submetida à Coordenação dos Cursos de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Campina Grande como parte dos requisitos necessários para obtenção do grau de Mestre no domínio da Engenharia Elétrica.

Área de Concentração: Processamento da Informação - Comunicações

Prof. Francisco Marcos de Assis, Dr.
Orientador

Campina Grande, Paraíba, Brasil

M488p Medeiros,Rex Antonio da Costa
2004 Protocolo para autenticaçãoquântica de mensagens clássicas / Campina Grande: UFCG, 2004.
90 p.: il.
Dissertação (Mestrado em Eng. Elétrica) - UFCG/CCT
Inclui Bibliografia

1.AutenticaçãoQuântica 2.criptonafia Quântica 3.Funçôes Hash I.Titulo

CDU:003.26.530.145

PROTOCOLO PARA AUTENTICAÇÃO QUÂNTICA DE MENSAGENS CLÁSSICAS

REX ANTONIO DA COSTA MEDEIROS

Dissertação Aprovada em 22.06.2004

FRANCISCO MARCOS DE ASSIS, Dr., UFCG
Orientador

RUBENS VIANA RAMOS, Dr., UFC
Componente da Banca

AÉRCIO FERREIRA LIMA, Dr., UFCG
Componente da Banca

ANTONIO MARCUS NOGUEIRA LIMA, Dr., UFCG
Componente da Banca

JOSÉ EWERTON POMBO DE FARIAS, Dr., UFCG
Componente da Banca

CAMPINA GRANDE - PB
Junho - 2004

Dedicatória

Este trabalho é dedicado a minha família, em especial aos meus pais, Eugênio e Neusa, a minha irmã, Giselli e a minha noiva, Luciana, que sempre me apoiaram de forma incondicional.

Agradecimentos

- A Deus, por tudo;
- À minha família, por tudo que me ensinaram, pelo carinho, pelo apoio, enfim, pela minha existência;
- À minha noiva, Luciana, pela paciência "quase" que infinita;
- Aos meus futuros sogro e sogra, Geraldo e Tânia, e a vó Mary, que sempre confiaram em mim;
- Às tias queridas, Ana Rosalina, Inácia Terezinha e Marizete, pelas palavras de incentivo;
- Ao meu orientador, Prof. Francisco Marcos, por ter acreditado no meu potencial;
- A todos os professores do Departamento, pelos ensinamentos;
- Aos companheiros da QQ - Quinta Quântica, Prof. Francisco Marcos, Prof. Bernado Lula e Prof. Acrcio Lima, pelas discussões até mesmo filosóficas;
- Aos funcionários do DEE, em especial à Angela, sempre eficiente na resolução dos entraves burocráticos;
- Aos colegas do GEPOTI, que estiveram comigo no Mestrado e, possivelmente, estarão também no Doutorado;
- A todos os meus amigos, principalmente à Edmar e ao pessoal do cafezinho, Prof. Bruno Albert, Luiz Felipe, Luíz Gonzaga Jr., Alfranque Amaral e José Alves, pela companhia;
- Ao CNPQ, pelo apoio financeiro.

Resumo

Nos dias atuais, os sistemas de criptografia e autenticação desempenham um papel fundamental em aplicações que envolvem a manipulação de informações sigilosas, tais como movimentações financeiras, comércio eletrônico, aplicações militares e proteção de arquivos digitais.

A popularização do uso dos sistemas de criptografia e autenticação se deve, em grande parte, a descrição de esquema de criptografia por chave pública. A segurança de tais sistemas é baseada na intratabilidade computacional (clássica) de problemas da teoria dos números, como a fatoração em produtos de primos e o problema do logaritmo discreto. A partir da formulação da Mecânica Quântica, foram demonstrados algoritmos que, executados em um computador quântico e consumindo tempo e recursos polinomiais, são capazes de resolver tais problemas. A construção de um computador quântico inviabilizaria, portanto, o uso de sistemas de criptografia e autenticação por chave pública.

Nesta dissertação é discutido o problema da autenticação quântica de mensagens clássicas. É proposto um protocolo híbrido que alcança segurança incondicional, mesmo que um criptoanalista disponha de recursos computacionais infinitos, sejam eles clássicos ou quânticos. Através de uma prova matemática formal, é mostrado que o nível de segurança pode ser feito tão alto quanto desejado. Tal segurança é garantida pelos princípios fundamentais da mecânica quântica.

Abstract

Nowadays, cryptography and authentication play a central role in applications that manipulates confidential information, like financial transactions, e-commerce, military applications and digital data protection.

The explosive growth of cryptosystems is mostly due to the discovery of the so-called public-key cryptosystems. The security of such systems is based on the intractability of some problems from number theory, like factorization and the discrete logarithm problem. After the formulation of the quantum mechanics, several protocols were described in order to solve these problems in time and resources polynomials in their arguments. So, one can conclude that public-key cryptosystems are not secure in a scenario where an eavesdropper makes use of quantum computers.

In this work it is discussed the problem of quantum authenticating classical messages. It is proposed a non-interactive hybrid protocol reaching information-theoretical security, even when an eavesdropper possesses both infinite quantum and classical computer power. It is presented a mathematical proof that it is always possible to reach a desirable level of security. This security is due to the quantum mechanics proprieties of non-orthogonal quantum states.

Lista de Símbolos e Abreviaturas

$|\psi\rangle$ - Estado quântico ou vetor de estado com rótulo ψ.

$\langle\psi|$ - Vetor dual ou conjugado Hermitiano de $|\psi\rangle$.

$\langle v|w\rangle$ - Produto interno entre os vetores $|v\rangle$ e $|w\rangle$.

$|v\rangle\langle w|$ - Produto externo entre os vetores $|v\rangle$ e $|w\rangle$.

$|||\psi\rangle||$ - Norma do vetor $|\psi\rangle$.

ρ - Operador de densidade.

δ_{ij} - Função delta. Assume o valor 1 se e somente se $i = j$. Caso contrário, $\delta_{ij} = 0$.

$A^\dagger$ - Conjugado Hermitiano da matriz A.

U - Matriz unitária ou operador unitário.

I - Matriz identidade.

X, Y, Z - Matrizes de Pauli.

M_m - Operador de medição correspondente a saída m.

P_m - Projetor sobre o subespaço gerado pelos autovetores com autovalor m correspondente.

$p(m)$ - Probabilidade de se obter a saída m em uma medição.

$|\psi\rangle^{\otimes n}$ - n produtos tensoriais do estado $|\psi\rangle$.

POVM - *Positive Operator-Valued Measure.*

$|M|$ - Cardinalidade do conjunto M.

$\mathbb{Z}_p^*$ - Campo de Galois formado pelos inteiros $\{0, 1, \ldots, p-1\}$ sobre adição e multiplicação *módulo* p, em que p é um primo.

RQ_p - Conjunto dos resíduos quadráticos *módulo* p.

x_0, y_0 - Sementes para um gerador de seqüências pseudo-aleatórias.

$x_0(n)$ - n–ésima subseqüência gerada a partir da semente x_0.

$h \in H$ - função h pertencente a um conjunto H de funções *hash*.

Lista de Figuras

Conteúdo

Capítulo 1

Introdução

Nas últimas duas décadas houve um enorme crescimento do número de aplicações que utilizam sistemas de criptografia e de autenticação por chave pública [14, 26, 31]. Tais sistemas possibilitaram por exemplo, que transações financeiras e comerciais fossem feitas pela a Internet, o chamado comércio eletrônico.

O desenvolvimento de esquemas de criptografia e autenticação por chave pública começou na década de 1970. Alguns pesquisadores viram a possibilidade de basear a segurança de tais esquemas em problemas que acreditavam ser computacionalmente intratáveis. A teoria dos números é o principal arcabouço de onde é extraída a fundamentação teórica dos sistemas de criptografia e de autenticação por chave pública.

O nascimento da teoria da mecânica quântica no século passado impulsionou estudos em diversas áreas de conhecimento. Em particular, Wiesner publicou em 1983 um trabalho mostrando que as propriedades da mecânica quântica poderiam ser usadas em criptografia. Este foi o marco inicial da chamada criptografia quântica.

Paralelamente ao desenvolvimento da criptografia quântica, Shor descreveu em 1984 um algoritmo que, executado em um computador quântico, resolveria eficientemente uma classe importante de problemas da teoria dos números, o problema do subgrupo escondido. A construção de tal computador, portanto, tornaria vulneráveis todas as aplicações que utilizam sistemas por criptografia ou autenticação por chave pública.

Esta dissertação aborda o problema da autenticação quântica de mensagens clássicas. A principal contribuição deste trabalho é a descrição de um novo protocolo cuja segurança é assegurada pelas propriedades inerentes à mecânica quântica.

As principais características do esquema proposto são:

1. O protocolo apresenta segurança incondicional, ou seja, um criptoanalista não consegue forjar uma mensagem e esta ser aceita como autêntica, mesmo quando recursos computacionais infinitos estão disponíveis, tanto clássicos como quânticos. Uma prova matemática formal é apresentada;

2. Utilização de chaves secretas relativamente pequenas, especialmente quando se deseja enviar um grande número de mensagens autênticas;

3. Diferentemente de outros protocolos descritos na literatura, o protocolo proposto é passível de implementação, por exemplo, usando o estado da arte da tecnologia fotônica;

4. A fundamentação teórica em que a sua segurança se baseia é bastante conhecida.

1.1 Organização do trabalho

A contribuição desta dissertação encontra-se no Capítulo 5. Leitores familiarizados com os conceitos básicos de mecânica quântica (Capítulo 2) e com os códigos de autenticação de mensagens (Capítulo 4), em especial, com o código proposto por Wegman e Carter em 1981, podem ler diretamente tal capítulo.

Esta dissertação está organizada como segue.

No Capítulo 2 é feita uma introdução aos fundamentos da mecânica quântica. O capítulo se inicia com um histórico da teoria quântica. Em seguida, é introduzida a notação de Dirac e fundamentos de álgebra linear e espaços de Hilbert. O capítulo segue apresentando os postulados da mecânica quântica. Ao final, os postulados são reescritos em termos de operadores de densidade.

O Capítulo 3 faz uma revisão da literatura dos principais protocolos para autenticação quântica de mensagens. É mostrada uma parte da análise de segurança para o primeiro protocolo e são tecidos também comentários acerca da sua implementação.

No Capítulo 4 são apresentados dois códigos clássicos de autenticação de mensagens. O primeiro é o já conhecido esquema de Wegman e Carter, o qual apresenta segurança incondicional. O segundo é uma modificação do primeiro, que permite a redução do tamanho da chave secreta em troca da redução do nível de segurança. É feita também uma introdução à geração de seqüências pseudo-aleatórias.

Um protocolo novo para autenticação quântica de mensagens clássicas é apresentado no Capítulo 5. É feita também uma análise da sua segurança considerando criptoanalistas que possuem recursos computacionais clássicos e quânticos infinitos. Também é dado um exemplo da utilização do protocolo.

Por fim, o Capítulo 6 apresenta as conclusões, bem como sugestões para trabalhos futuros.

Capítulo 2

Fundamentos de Mecânica Quântica

Este capítulo apresenta um resumo dos principais conceitos da Mecânica Quântica, necessários ao entendimento da dissertação. Uma abordagem mais detalhada destes conteúdos pode ser encontrada em livros específicos [23].

2.1 Breve histórico da teoria quântica

No final do século XIX, os físicos possuíam uma imagem abrangente do modo pelo qual o mundo funcionava. Uns poucos grandes homens haviam solucionado os grandes problemas da física. A tarefa de seus sucessores seria preencher os detalhes, estender as medições até à casa decimal seguinte. Nada de glorioso. Tendo explicado tudo, a física clássica parecia ter encerrado as suas atividades. "A física acabou, meu jovem. É uma rua sem saída", disse o professor de Max Planck, e aconselhou-o a optar pela carreira de pianista.

Os físicos clássicos explicam os fenômenos do universo usando apenas duas entidades físicas: a matéria e os campos. Se uma maçã cai do alto de uma árvore é porque o campo gravitacional da terra age sobre ela. A física clássica reconhecia apenas dois campos – o eletromagnético e o gravitacional. A física moderna acrescentou somente dois: o campo forte, que mantém o núcleo atômico unido, e o campo fraco, que destrói o núcleo em certos tipos de desintegração radioativa. Essas entidades seriam tudo o que a natureza necessitava para produzir um universo como o nosso.

A fim de descrever o modo de funcionamento do mundo clássico, são necessários dois tipos de leis: as leis de movimento e as leis de campo. No século XVII, Isaac Newton descobriu as leis do movimento da matéria. Os campos de força deslocam a

matéria segundo trajetórias prescritas com exatidão pelas leis de Newton. O mesmo também descobrira a primeira lei da física relacionada a um campo: a razão inversa do quadrado das distâncias no comportamento do campo de gravidade.

Para completar a teoria clássica, James C. Maxwell uniu as leis da eletricidade e do magnetismo, verificando que os campos elétricos e magnéticos eram duas instâncias de um único campo eletromagnético. As leis de Maxwell permitem calcular a velocidade das ondas num campo eletromagnético. Ao calcular a velocidade de deslocamento dessas ondas, Maxwell achou um valor idêntico ao da velocidade luz, que supôs ser uma onda eletromagnética. A produção subseqüente de ondas eletromagnéticas de baixa freqüência por Heinrich Hertz confirmou as suposições de Maxwell.

Antes de apresentar o formalismo matemático da teoria quântica, serão descritos nas subseções seguintes alguns fenômenos físicos observados desde os primórdios da ciência até o início do século XX que, de alguma forma, não poderiam ser descritos pela teoria da física clássica.

2.1.1 O corpo-negro e a emissão de luz

Objetos coloridos possuem uma cor intrínseca; objetos negros, não. Contudo, quando um objeto negro é aquecido, ele adquire um certo fulgor. Por exemplo, o ferro aquecido por volta de mil e trezentos graus emite uma luz de cor vermelho-cereja. Para os físicos, o enigma do corpo-negro consiste em calcular a coloração desse fulgor para diferentes temperaturas.

Um corpo-negro é formado de pequenas porções de matéria. Quando essas porções se movem, elas geram um campo eletromagnético que os olhos interpretam como luz colorida. Quando um objeto torna-se mais quente, suas partículas componentes movimentam-se com maior velocidade. Na visão de um físico clássico, quanto mais rapidamente as partículas se movem, maior é a freqüência da luz emitida.

Durante um quarto de século após Maxwell ter anunciado a conexão luz-matéria, os físicos atacaram o enigma do corpo-negro, obtendo sempre a mesma resposta: os corpos-negro deveriam ter um fulgor azul brilhante em todas as temperaturas. Hoje, é sabido que a luz é produzida por elétrons em movimento. Contudo, os elétrons não obedecem às leis clássicas, inclusive à lei dos movimentos de Newton.

Em 1900, Max Planck, que se formara em física e não em música, atacou o enigma do corpo-negro. Como premissa simplificadora, ele decidiu não deixar que as partículas

materiais vibrassem da maneira que quisessem; ao contrário, restringiu artificialmente as freqüências das partículas, obrigando-as a seguirem uma regra simples:

$$E = n\hbar f, \tag{2.1}$$

em que E é a energia das partículas, n é qualquer número inteiro, f é a freqüência de vibração da partícula e $\hbar$ é uma constante a ser escolhida por Planck. Ele assim restringiu a energia das partículas a certos múltiplos da freqüência de vibração. A constante de Planck $\hbar$ seria mais tarde denominada "quantum de ação", porque sua dimensão é de "energia × tempo".

Planck verificou que, como todo mundo, obtinha o mesmo brilho azul quando $\hbar$ tendia para zero. Contudo, descobriu que atribuindo a $\hbar$ um valor particular, os cálculos correspondiam exatamente à experimentação. Os físicos clássicos ignoraram o trabalho de Planck, pois a restrição imposta à energia era estranha à física clássica. As leis de Newton permitem que as partículas possuam uma quantidade arbitrária de energia.

2.1.2 As teorias de Einstein

Em 1905, Albert Einstein, que trabalhava na Suíça como funcionário do serviço de patentes, publicou três artigos no jornal alemão *Annalen der Physik* que causaram grande impacto na comunidade dos físicos.

No primeiro artigo, Einstein explicava o efeito fotoelétrico (poder da luz para expulsar os elétrons de um metal) usando o novo quantum de ação definido por Planck. O segundo artigo de Einstein explicava o movimento browniano[1] das partículas microscópicas em suspensão num líquido. A descrição de tal movimento remonta à discussão sobre a existência dos átomos. No terceiro artigo, Einstein afirmava que as noções de comprimento e tempo são relativas, e dependem da velocidade do observador. A velocidade da luz é elevada à condição de velocidade limite do universo, a qual nenhum sinal pode exceder. A teoria especial da relatividade, como Einstein a denominou, teve profundas conseqüências para a física e para a filosofia, pois mostrou que algumas noções até então inquestionáveis estavam simplesmente erradas e deveriam ser substituídas por outras maneiras de pensar inteiramente novas.

Do modo em que são postas, as leis de Newton e de Maxwell são incompatíveis com a afirmação de que, a despeito do movimento da terra, a velocidade da luz permanecia

[1]Movimento frenético e aleatório. Em homenagem ao botânico escocês Robert Brown, que em 1828 descobriu tal movimento.

constante em qualquer direção. Para resolver isto, Einstein conservou as leis de Maxwell mas substituiu as leis de Newton por suas próprias leis relativistas do movimento. As leis de Maxwell, juntamente com as leis relativistas do movimento descrevem de modo completo e consistente todos os movimentos clássicos, mesmo aqueles que envolvem altas velocidades.

2.1.3 A dualidade onda-partícula

Os físicos clássicos imaginavam que quanto maior a amplitude de uma onda de luz que atingisse um metal maior era a energia do elétron ejetado da superfície do mesmo. Contudo, os resultados das experiências fotoelétricas não comprovavam essa hipótese. Os experimentalistas obtinham sempre elétrons com mesma energia para uma determinada freqüência da luz. Uma maior intensidade do raio luminoso apenas aumentava o número de elétrons emitidos. Para aumentar a energia de cada elétron, era necessário aumentar a freqüência da luz que incidia no metal.

Einstein esclareceu esse estranho fato supondo que a luz se comportava como uma chuva de partículas, denominadas de fótons, tendo cada uma a energia E dada pela expressão de Planck:

$$E = \hbar f, \tag{2.2}$$

em que $\hbar$ é a constante de ação de Planck e f é a freqüência da partícula.

Paralelamente, o físico Arthur Compton fez incidir raios-X (luz de alta freqüência) sobre um gás, onde os elétrons estão frouxamente ligados. Compton verificou novamente o comportamento da luz como partícula. Analisando não a absorção da luz mas a sua emissão, Compton detectou tanto a ejeção do elétron quando o ricochete do fóton, em que os resultados experimentais eram compatíveis com a teoria de Planck-Einstein desde que se atribuísse ao fóton uma quantidade de movimento

$$p = \hbar k, \tag{2.3}$$

em que k é a freqüência espacial da luz. Ou seja, a luz se comporta como uma partícula que possui quantidade de movimento e energia determinadas pelas leis quânticas.

Pouco depois, o francês Louis de Broglie apresentava sua tese de doutorado em que propunha que para cada partícula fundamental de matéria estaria associada uma onda cujas freqüências temporais e espaciais, f e k, seriam determinadas pela fórmula de Planck-Einstein $E = \hbar f$ e pela relação de Compton $p = \hbar k$, em que E e p seriam,

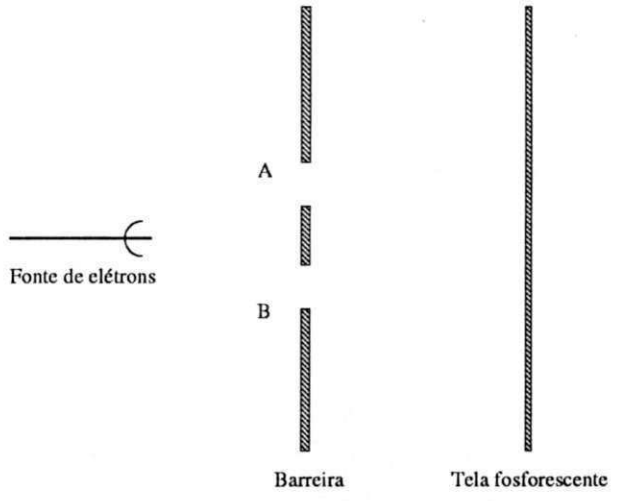


Figura 2.1: Aparato experimental para verificar o comportamento ondulatório das partículas.

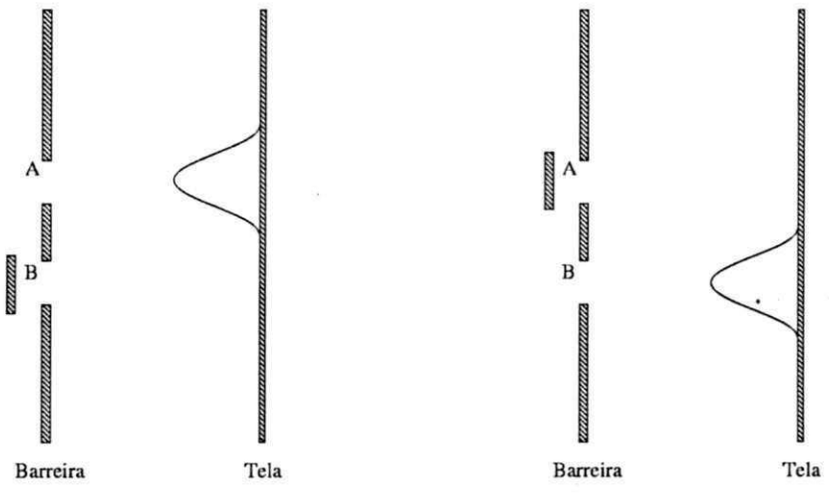
respectivamente, a energia e a quantidade de movimento da partícula. De Broglie argumentava que, assim como Einstein mostrara que as ondas de luz possuíam propriedades inerentes às partículas, estas também teriam propriedades ondulatórias. Dois anos depois, pesquisadores dos Laboratórios Bell mediram o "comprimento de onda de De Broglie" de um elétron.

2.1.4 Um exemplo

O comportamento onda-partícula pode ser verificado experimentalmente com montagens geralmente simples, que podem ser feitas em qualquer laboratório de física. O exemplo já "clássico" a seguir utiliza uma fonte de elétrons, uma tela fosforescente e, entre elas, uma barreira física que possui duas pequenas aberturas A e B. A fonte de elétrons deve apontar na direção perpendicular a barreira e estar localizada no ponto médio entre as aberturas. O esquema é ilustrado na Figura 2.1.

Quando os elétrons são emitidos um de cada vez, a maior parte deles deve colidir contra a barreira, mas alguns passam através das aberturas e colidem contra a tela. É justamente o padrão da distribuição formado na tela que sugere que a partícula, no caso o elétron, se comporta como uma onda enquanto viaja de encontro à tela sem ser observado, pelo menos até a colisão.

Primeiro considere o caso em que uma das aberturas é fechada, tendo o elétron



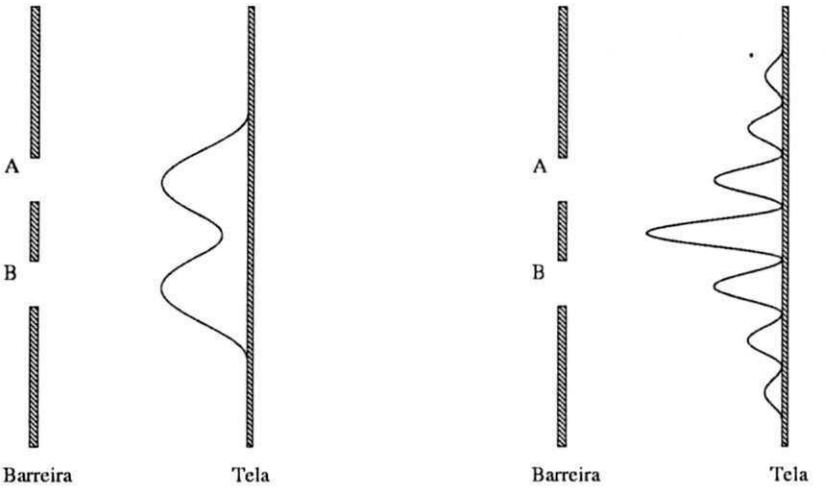
(a) A distribuição A (b) A distribuição B

Figura 2.2: A ou B aberto: A distribuição A. (b) A distribuição B.

a possibilidade de passar somente pela outra. Quando a abertura inferior é fechada, Figura 2.2(a), é obtida uma distribuição que é consistente com a hipótese de que cada elétron passa somente através de A. Chame essa distribuição de distribuição A. Se a passagem A é fechada e os elétrons são enviados contra a barreira, uma distribuição B análoga é obtida, coerente com a noção de que cada elétron que atinge a tela passa através de B (Figura 2.2(b)).

Um físico clássico deve imaginar que se as duas passagens estão abertas, e que os elétrons estão sendo enviados um a um, então a distribuição das partículas na tela é aproximadamente a soma da distribuição A e da distribuição B, chamada de distribuição A-ou-B (Figura 2.3(a)). Isto porque o elétron sendo uma partícula, hora ele passaria exclusivamente através de A hora exclusivamente através de B. Na verdade, a distribuição A-ou-B é obtida quando os elétrons são, um a um, obrigados a passar exclusivamente por uma das duas aberturas, bastando para isso fechar aleatoriamente uma das duas passagens no instante anterior à emissão de cada elétron pela fonte. Porém, quando as duas passagens estão abertas ao mesmo tempo, a distribuição obtida é semelhante àquela da Figura 2.3(b).

Uma conclusão imediata do fenômeno expresso na Figura 2.3(b) é que pelo menos alguns elétrons não passaram, exclusivamente, por nenhuma das aberturas, desde que



(a) A distribuição A-ou-B

(b) A distribuição interferência

Figura 2.3: (a) A distribuição A-ou-B. (b) A distribuição interferência.

se considere que a distribuição A-ou-B caracteriza a passagem de cada partícula, exclusivamente, através de uma das aberturas. A pergunta óbvia a fazer é por qual trajetória seguiram essas partículas?

De acordo com a formulação de von Neumann-Dirac da mecânica quântica, o elétron não passa, exclusivamente, através de uma ou de outra abertura; ao invés disso, a teoria afirma que cada partícula segue uma **superposição** de trajetórias, que neste caso é uma superposição do elétron passando pela abertura A e do elétron passando pela abertura B. O elétron neste estado possui características físicas **observáveis** que diferem fortemente de todas as possibilidades clássicas possíveis.

Enquanto um elétron está se movendo em direção à tela, ele é pensado como sendo uma onda, e da mesma forma que não há sentido em pensar na posição pontual de uma onda, a teoria afirma que a partícula antes de ser observada não possui uma posição determinada. É como se parte da onda passasse através de A e outra parte através de B. Esses dois "pacotes" de onda se encontrariam e causariam interferência entre si na região entre a barreira e a tela. Neste ponto da história, e os físicos não sabem dizer exatamente quando e como isto acontece[2], a partícula deixa de se comportar como uma onda. Para cada partícula neste estágio, o que se observa na tela não é o efeito

[2]Este problema é conhecido como o problema da medição.

global que se esperaria de uma onda colidindo com a tela, e sim o de uma partícula atingindo um único ponto determinado. Por outro lado, o comportamento de onda de cada partícula individual é visto na distribuição geral de partículas (Figura 2.3(b)), quando o experimento básico é realizado diversas vezes.

Na realidade, a teoria quântica afirma que a onda associada a cada partícula determina a **probabilidade** de a partícula ser encontrada na região da tela. Mais precisamente, a teoria prediz que se ψ representa uma onda complexa associada a uma determinada partícula, então a probabilidade de encontrar a partícula em uma região R é igual a integral de $|\psi|^2$ sobre R.

2.1.5 A formulação teórica adotada

A primeira descrição da teoria quântica foi dada por Werner Heisenberg em 1925. A teoria de Heisenberg é conhecida hoje como mecânica matricial, e muito dos detalhes da teoria foram sendo inseridos aos poucos por diversos físicos, entre eles Wolfgang Pauli, Max Born, E. P. Jordan e P. A. M. Dirac.

Ervin Schrödinger descreveu em 1926 uma formulação alternativa para a mecânica quântica. A sua teoria foi baseada na suposição de De Broglie de que a relação entre fótons e ondas eletromagnéticas deveria ser generalizada para incluir todas as partículas materiais, e que todas elas deveriam exibir características semelhantes às ondas.

Pouco depois, von Neumann apresentou um formalismo matemático rigoroso para representar estados de uma partícula quântica, bem como a sua dinâmica. Para tal formulação, são considerados dois cenários gerais. No primeiro cenário, os estados quânticos da formulação de Heisenberg são representados por vetores constantes de um espaço de Hilbert apropriado. O espaço escolhido para a representação dependeria das quantidades físicas que se desejasse representar.

No outro cenário, os estados de um sistema quântico da formulação de Schrödinger são representados por vetores unitários pertencentes a um espaço de Hilbert. Novamente, o espaço escolhido para a representação depende das quantidades físicas que se está interessado em representar.

Esta última formulação matemática será usada ao longo do trabalho, principalmente neste capítulo onde serão apresentados os fundamentos da teoria quântica na forma de postulados. Aliada a esta formulação está uma poderosa notação, conhecida como notação de Dirac, e que foi especialmente criada para ser usada na mecânica quântica.

2.2 Álgebra linear e espaços de Hilbert

Embora a Álgebra Linear seja um tópico bem conhecido em Engenharia, existem diferenças importantes de notação quando essa teoria matemática é empregada na mecânica quântica. A notação de Dirac, como é conhecida, além de elegante, mostra ser altamente eficiente quando usada para descrever sistemas quânticos e suas evoluções. Essa notação é amplamente conhecida pelos físicos e é padrão nos textos de Informação e Computação Quânticas [24]. Sua introdução será gradativa neste capítulo, conforme as definições forem sendo apresentadas.

Definição 1 (Espaço vetorial [23]) *Seja F um corpo algébrico. Um espaço vetorial V sobre F, com elementos (vetores) representados por $|v\rangle$, é uma estrutura composta por um conjunto e duas operações binárias, $+ : V \times V \longrightarrow V$ e $\cdot : F \times V \longrightarrow V$, tais que*

1. $(|v\rangle + |w\rangle) + |u\rangle = |v\rangle + (|w\rangle + |u\rangle)$ *para todo* $|v\rangle, |w\rangle, |u\rangle \in V$; .
2. $|v\rangle + |w\rangle = |w\rangle + |v\rangle$ *para todo* $|v\rangle, |w\rangle \in V$;
3. *Existe um elemento* $\mathbf{0} \in V$ *tal que* $|v\rangle + \mathbf{0} = |v\rangle$ *para todo* $|v\rangle \in V$;
4. *Para qualquer* $|v\rangle \in V$, *existe um elemento* $|w\rangle \in V$ *tal que* $|v\rangle + |w\rangle = \mathbf{0}$;
5. $k_1 \cdot (k_2 \cdot |v\rangle) = (k_1 k_2) \cdot |v\rangle$ *para todo* $k_1, k_2 \in F$ *e* $|v\rangle \in V$;
6. $1 \cdot |v\rangle = |v\rangle$ *para todo* $|v\rangle \in V$;
7. $k \cdot (|v\rangle + |w\rangle) = (k \cdot |v\rangle) + (k \cdot |w\rangle)$ *para todo* $k \in F$ *e* $|v\rangle, |w\rangle \in V$;
8. $(k_1 + k_2) \cdot |v\rangle = (k_1 \cdot |v\rangle) + (k_2 \cdot |v\rangle)$ *para todo* $k_1, k_2 \in F$ *e* $|v\rangle \in V$.

Os elementos de V são chamados de vetores e o elemento $\mathbf{0} \in V$ é chamado de vetor nulo de V.

Na Definição 1, o símbolo $|v\rangle$ foi usado para designar um vetor arbitrário em V, em que v é o rótulo do vetor $|v\rangle$. Na notação de Dirac, o elemento $|\cdot\rangle$ é chamado de *ket*. Observe que para o vetor nulo o *ket* não é usado. Esta terminologia não será usada neste trabalho.

Um subespaço vetorial de um espaço V é um subconjunto W de V, tal que W também é um espaço vetorial, ou seja, W deve satisfazer as oito condições da Definição 1.

Um conjunto de vetores $|v_1\rangle, \ldots, |v_n\rangle$ não nulos, pertencentes a um espaço vetorial V sobre um campo F, é dito ser linearmente independente se para quaisquer escalares $a_1, a_2, \ldots, a_n \in F$,

$$a_1|v_1\rangle + a_2|v_2\rangle + \ldots + a_n|v_n\rangle = 0$$

implica $a_1 = a_2 = \ldots = a_n = 0$. Caso contrário, o conjunto é dito ser linearmente dependente.

Um conjunto de vetores $\beta = \{|v_1\rangle, \ldots, |v_n\rangle\}$ gera o espaço vetorial V se qualquer vetor $|v\rangle \in V$ pode ser escrito como uma combinação linear $|v\rangle = \sum_i a_i|v_i\rangle$, em que $a_i \in F$. O conjunto β é denominado base de V se os vetores de β são linearmente independentes. A dimensão do espaço V, $\dim(V)$, é definida como sendo igual à cardinalidade da base β, $||\beta||$.

2.2.1 Produto interno

Seja V um espaço vetorial para o qual F é o campo dos números complexos C. Este espaço é de particular interesse no estudo da mecânica quântica. Para tal espaço, define-se produto interno como segue:

Definição 2 (Produto interno [23]) *Um produto interno em um espaço vetorial V sobre o campo C dos números complexos é uma função $(\ ,\ ) : V \times V \longrightarrow C$ tal que, para todo $k_1, k_2 \in C$ e $|v_1\rangle, |v_2\rangle, |v\rangle, |w\rangle \in V$, as propriedades abaixo são verificadas:*

1. $(|w\rangle, k_1|v_1\rangle + k_2|v_2\rangle) = k_1(|w\rangle, |v_1\rangle) + k_2(|w\rangle, |v_2\rangle)$[3]*;*
2. $(|v\rangle, |w\rangle) = (|w\rangle, |v\rangle)^*$*, em que* $(^*)$ *denota conjugação complexa;*
3. $(|v\rangle, |v\rangle) \geq 0$*, e* $(|v\rangle, |v\rangle) = 0$ *se e somente se* $|v\rangle = \mathbf{0}$.

A notação para produto interno na Definição acima **não** é a padrão em mecânica quântica. A notação padrão para o produto interno $(|v\rangle, |w\rangle)$ é $\langle v|w\rangle$, por razões que ficarão claras mais adiante. A notação $\langle v|$ se refere ao vetor dual de $|v\rangle$, a ser definido nas subseções seguintes.

Os vetores $|v\rangle$ e $|w\rangle$ são ortogonais se o produto interno $\langle v|w\rangle$ for igual a zero. A norma do vetor $|v\rangle$ é definida como sendo

$$|||v\rangle|| \equiv \sqrt{\langle v|v\rangle}. \tag{2.4}$$

[3]Alguns autores impõem a linearidade no primeiro argumento ao invés do segundo.

Um vetor unitário $|v\rangle$ é tal que $|||v\rangle|| = 1$. Define-se também um vetor $|v'\rangle = |v\rangle/|||v\rangle||$ como sendo a normalização do vetor $|v\rangle$. O conjunto de vetores $|i\rangle$, com índices i, é ortonormal se cada vetor é unitário, e vetores distintos no conjunto são ortogonais, i.e., $\langle i|j\rangle = \delta_{ij}$, em que i e j são escolhidos do conjunto de índices.

As definições que se seguem são necessárias à definição de espaços de Hilbert.

Definição 3 (Espaços métricos [23]) *Um espaço métrico, denotado por (X, d), é definido como sendo composto de duas partes: um conjunto X e uma métrica $d(x, y)$, $x, y, z \in X$, satisfazendo os axiomas seguintes:*

1. $d(x, y) \geq 0$ *e* $d(x, x) = 0$, $\forall x, y \in X$;
2. *Se* $d(x, y) = 0$, *então* $x = y$, $\forall x, y \in X$;
3. $d(y, x) = d(x, y)$ $\forall x, y \in X$;
4. $d(x, y) \leq d(x, z) + d(z, y)$, $\forall x, y, z \in X$.

Definição 4 (Seqüências de Cauchy [23]) *Uma seqüência $\{x_m\}$ em um espaço métrico (X, d) é dita ser uma seqüência de Cauchy se para cada $\epsilon > 0$ existe um N tal que $d(x_n, x_m) \leq \epsilon$ para quaisquer $n, m \geq N$.*

Como exemplo, considere o espaço métrico que consiste de todos os pontos no intervalo $(0, 1]$, $X = \{x \in R : 0 < x \leq 1\}$, e da métrica usual, $d(x, y) = |x - y|$. A seqüência $\{1/n\} = \{1, 1/2, 1/4, \ldots\}$ em X é uma seqüência de Cauchy. Seja $N \geq 2/\epsilon$. Se $n, m \geq N$, então $1/n \leq \epsilon/2$ e $1/m \leq \epsilon/2$. Conseqüentemente, $|1/n - 1/m| \leq 1/n + 1/m \leq \epsilon$ para todo $n, m \geq N$.

Definição 5 ([23]) *Um espaço métrico (X, d) é dito ser completo se cada seqüência de Cauchy em (X, d) é convergente.*

Por definição, todo espaço vetorial com produto interno possui uma métrica, o que implica que tais espaços vetoriais são também espaços métricos.

Definição 6 (Espaço de Hilbert [23]) *Um espaço de Hilbert é um espaço vetorial completo com produto interno.*

Como mencionado anteriormente, o espaço vetorial de maior interesse no estudo da mecânica quântica é o espaço vetorial das n-tuplas de números complexos, $(z_1, z_2, \ldots, z_n)$,

denotado por C^n. A notação de matriz coluna também será usada para referenciar esses vetores.

$$|z\rangle \equiv \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}. \tag{2.5}$$

O produto interno usual C^n é definido por

$$\langle y|z\rangle \equiv \begin{bmatrix} y_1^* & \dots & y_n^* \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}, \tag{2.6}$$

em que $(y_1, \dots, y_n)$ e $(z_1, \dots, z_n)$ são, respectivamente, as coordenadas dos vetores $|y\rangle$ e $|z\rangle$ com relação a uma mesma base ortonormal.

É fácil verificar que o espaço vetorial C^n, acompanhado do produto interno definido pela Equação (2.6), é um espaço de Hilbert de dimensão n [23]. Como será visto mais adiante, os estados quânticos pertencentes a um determinado sistema quântico são representados por vetores $|v\rangle$ pertencentes a um espaço de Hilbert de dimensão n. Com essa notação, considere que $|v\rangle = \sum_i v_i|i\rangle$ e $|w\rangle = \sum_j w_j|j\rangle$ são representações dos estados $|v\rangle$ e $|w\rangle$ com relação a uma mesma base ortonormal $|i\rangle$. Desde que $\langle i|j\rangle = \delta_{ij}$,

$$\begin{aligned} \langle v|w\rangle &= \left(\sum_i v_i|i\rangle, \sum_j w_j|j\rangle\right) = \sum_{ij} v_i^* w_j \langle i|j\rangle = \sum_i v_i^* w_i \\ &= \begin{bmatrix} v_1^* & \dots & v_n^* \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}. \end{aligned} \tag{2.7}$$

A equação acima mostra que o produto interno entre dois vetores é igual ao produto interno entre as representações matriciais desses vetores com relação a uma mesma base ortonormal. Outra observação interessante é a interpretação para o vetor dual $\langle v|$, como sendo um vetor linha cujas componentes são conjugadas complexas das componentes correspondentes do vetor $|v\rangle$.

De acordo com as definições precedentes, os vetores

$$|0\rangle \equiv \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{e} \quad |1\rangle \equiv \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{2.8}$$

formam uma base ortonormal para o espaço de Hilbert de dimensão 2. Ou seja, qualquer vetor

$$|v\rangle = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} \tag{2.9}$$

pode ser escrito como uma combinação linear $|v\rangle = a_0|0\rangle + a_1|1\rangle$ dos vetores $|0\rangle$ e $|1\rangle$.

Em mecânica quântica, a base $\{|0\rangle, |1\rangle\}$ é chamada de base computacional para o espaço de Hilbert de dimensão dois. Tal denominação é uma analogia dos estados quânticos desta base com os bits clássicos 0 e 1, respectivamente. Uma base computacional para um espaço de Hilbert de dimensão n é da forma $\{|0\rangle, |1\rangle, \ldots, |n-1\rangle\}$, em que $|k\rangle \equiv \begin{bmatrix} a_0 = 0 & a_1 = 0 & \ldots & a_k = 1 & \ldots & a_{n-1} = 0 \end{bmatrix}^T$.

2.2.2 Operadores lineares

Um operador linear entre os espaços vetoriais V e W é definido como sendo uma função $A : V \longrightarrow W$ linear com relação às suas entradas:

$$A\left(\sum_i a_i|v_i\rangle\right) = \sum_i a_i A|v_i\rangle. \tag{2.10}$$

É usual utilizar a notação $A|v\rangle$ ao invés de $A(|v\rangle)$. Dois operadores lineares importantes são o operador identidade I e o operador 0, que são tais que $I|v\rangle \equiv |v\rangle$ e $0|v\rangle \equiv \mathbf{0}$, respectivamente.

Os operadores lineares podem ser representados por matrizes complexas. Suponha que $A : V \longrightarrow W$ é um operador linear do espaço V para o espaço W. Seja $|v_i\rangle, \ldots, |v_m\rangle$ uma base para V e $|w_1\rangle, \ldots, |w_n\rangle$ uma base para W, em que m é a dimensão de V e n é a dimensão de W. Então, para cada j no intervalo $1, \ldots, m$, existem números complexos $A_{1j}, \ldots, A_{nj}$ tal que

$$A|v_j\rangle = \sum_i A_{ij}|w_i\rangle. \tag{2.11}$$

A matriz com entradas A_{ij} é dita ser a representação matricial do operador A. Note que a representação matricial de um operador linear está sujeita à especificação de bases para os espaços V e W. Quando não mencionado, subentende-se que são usadas as bases computacionais para os dois espaços vetoriais.

Outra maneira interessante de representar um operador linear, conhecida como representação por produto externo, é via produto interno. Suponha que $|v\rangle$ e $|w\rangle$

pertencem aos espaços vetoriais com produto interno V e W, respectivamente. Defina $|w\rangle\langle v|$ como sendo um operador linear de V para W cuja ação é definida por

$$(|w\rangle\langle v|)(|v'\rangle) \equiv |w\rangle\langle v|v'\rangle = \langle v|v'\rangle|w\rangle. \tag{2.12}$$

Esta poderosa notação sugere duas interpretações para a expressão acima. A primeira é a aplicação do operador $|w\rangle\langle v|$ ao vetor $|v'\rangle$ e a segunda é a multiplicação do número complexo $\langle v|v'\rangle$ pelo vetor $|w\rangle$.

Teorema 1 (Relação de completude para vetores ortonormais [24]) *Seja* $|i\rangle$ *uma base ortonormal qualquer para o espaço* V*, então*

$$\sum_i |i\rangle\langle i| = I. \tag{2.13}$$

Prova [24] Seja $|v\rangle$ um vetor arbitrário em V. Então $|v\rangle$ pode ser escrito como uma combinação linear $|v\rangle = \sum_i v_i|i\rangle$ dos vetores da base $|i\rangle$. Note que $v_i = \langle i|v\rangle$ e portanto

$$\left(\sum_i |i\rangle\langle i|\right)|v\rangle = \sum_i |i\rangle\langle i|v\rangle = \sum_i v_i|i\rangle = |v\rangle. \tag{2.14}$$

O resultado é obtido bastando observar que a última equação é válida para todo $|v\rangle \in V$. ■

Suponha então que $A : V \longrightarrow W$ é um operador linear e que $|v_i\rangle$ e $|w_j\rangle$ são bases ortonormais para os espaços V e W, respectivamente. Usando a relação de completude, tem-se que

$$\begin{aligned} A &= I_W A I_V \\ &= \sum_{ij} |w_j\rangle\langle w_j|A|v_i\rangle\langle v_i| \\ &= \sum_{ij} \langle w_j|A|v_i\rangle |w_j\rangle\langle v_i|, \end{aligned} \tag{2.15}$$

que é a representação por produto externo para A.

2.2.3 Matrizes de Pauli

Abaixo, encontram-se definidas quatro matrizes de ordem 2, conhecidas na literatura como matrizes ou operadores de Pauli [24]. Tais matrizes desempenham um papel fundamental no estudo de teoria da informação e computação quânticas.

$$\sigma_0 \equiv I \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \sigma_1 \equiv X \equiv \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{2.16}$$

$$\sigma_2 \equiv Y \equiv \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \text{ e} \qquad \sigma_3 \equiv Z \equiv \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{2.17}$$

2.2.4 Autovetores e autovalores

Um autovetor de um operador linear A em um espaço vetorial V é definido como sendo um vetor não nulo $|v\rangle$ tal que $A|v\rangle = \lambda|v\rangle$, em que λ é um número complexo denominado de autovalor associado ao autovetor $|v\rangle$.

Os autovalores de um operador A são encontrados resolvendo a equação característica $c(\lambda) \equiv \det|A - \lambda I|$. Pode-se mostrar que a equação característica depende somente do operador A, e não de uma representação matricial específica para A. Pelo teorema fundamental da álgebra, todo polinômio de grau n possui exatamente n raízes no corpo algébrico dos complexos. Assim, todo operador A possui pelo menos um autovalor e um autovetor correspondente. Define-se também o autoespaço associado ao autovalor λ como sendo o subespaço vetorial de V gerado a partir dos autovetores $|v\rangle$ associados ao autovalor λ.

Uma representação diagonal para um operador A em um espaço V é definida como sendo $A = \sum_i \lambda_i |i\rangle\langle i|$, em que $\{|i\rangle\}$ é um conjunto de autovetores ortonormais de A, com autovalores λ_i correspondentes. Um operador é dito ser diagonalizável se ele possui uma representação diagonal.

Como exemplo, considere o operador de Pauli X. Tal operador possui autovalores ± 1, com autovetores

$$|\psi_{+1}\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \tag{2.18}$$

$$|\psi_{-1}\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \tag{2.19}$$

Note que os autovetores devem estar normalizados. Dessa forma, o operador X pode ser escrito como

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = |\psi_{+1}\rangle\langle\psi_{+1}| - |\psi_{-1}\rangle\langle\psi_{-1}|, \tag{2.20}$$

em que a representação é feita com relação à base ortonormal $\{|0\rangle,|1\rangle\}$. A representação diagonal é também chamada de decomposição ortonormal.

2.2.5 Operadores hermitianos e unitários

Seja A um operador qualquer em um espaço de Hilbert V. É possível mostrar [24] que existe um único operador $A^\dagger$ em V tal que para quaisquer vetores $|v\rangle, |w\rangle \in V$,

$$(|v\rangle, A|w\rangle) = (A^\dagger|v\rangle, |w\rangle). \tag{2.21}$$

O operador $A^\dagger$ é chamado de adjunto ou conjugado hermitiano do operador A. Pela definição é fácil ver que $(AB)^\dagger = B^\dagger A^\dagger$. Por convenção, se $|v\rangle$ é um vetor, então o seu dual será $|v\rangle^\dagger \equiv \langle v|$. A partir dessa definição tem-se que $(A|v\rangle)^\dagger = \langle v|A^\dagger$. O conjugado hermitiano da representação matricial de um operador é uma matriz conjugada-transposta da matriz A, $A^\dagger \equiv (A^*)^T$, em que $(*)$ indica conjugação complexa e T indica transposição.

Um operador A é dito ser hermitiano ou auto-adjunto se a $A^\dagger = A$. Os projetores formam uma classe importante de operadores hermitianos. Suponha que W é um subespaço vetorial de dimensão k de um espaço V de dimensão d. Usando o procedimento de Gram-Schmidt [23] é possível construir uma base ortonormal $|1\rangle, \ldots, |d\rangle$ para V de tal forma que $|1\rangle, \ldots, |k\rangle$ formam uma base ortonormal para W. Por definição

$$P \equiv \sum_{i=1}^{k} |i\rangle\langle i| \tag{2.22}$$

é um projetor sobre o subespaço W de V. O complemento ortogonal de P é um operador hermitiano $Q = I - P$ sobre o subespaço ortogonal a W, chamado aqui de complemento ortogonal de W, gerado pelos vetores $|k+1\rangle, \ldots, |d\rangle$.

Um operador A é dito ser normal se $AA^\dagger = A^\dagger A$. Claramente, um operador hermitiano é também normal. É possível mostrar que um operador normal M qualquer em um espaço vetorial V é diagonal com relação a alguma base ortonormal de V e, equivalentemente, qualquer operador diagonalizável é normal [24, pp. 72]. Isso implica que todo operador normal M pode ser escrito como

$$M = \sum_{i=1}^{d} \lambda_i |i\rangle\langle i|, \tag{2.23}$$

em que $|i\rangle$ são autovetores de M com autovalores λ_i, d é a dimensão de V e os vetores $|i\rangle$ formam uma base ortonormal para V. A Equação (2.23) é chamada de decomposição espectral de M. Todo operador hermitiano é normal e possui, portanto, uma decomposição espectral.

Os operadores unitários, definidos a seguir, são de extrema importância no estudo da evolução dos sistemas quânticos.

Definição 7 (Operadores Unitários [24]) *Um operador U é unitário se $U^\dagger U = UU^\dagger = I$.*

Geometricamente, operadores unitários são importantes porque preservam o produto interno entre vetores, ou seja, se $|v\rangle$ e $|w\rangle$ são vetores quaisquer em V, então

$$(U|v\rangle, U|w\rangle) = \langle v|U^\dagger U|w\rangle = \langle v|I|w\rangle = \langle v|w\rangle. \tag{2.24}$$

A propriedade acima sugere uma representação elegante em termos de produto externo para um operador unitário U em V. Se $|v_i\rangle$ é uma base ortonormal para V, então $|w_i\rangle = U|v_i\rangle$ é também uma base ortonormal para V. Fazendo $|w_i\rangle\langle v_i| = U|v_i\rangle\langle v_i|$, então

$$\begin{aligned} U &= \sum_i U|v_i\rangle\langle v_i| \\ &= \sum_i |w_i\rangle\langle v_i|. \end{aligned} \tag{2.25}$$

Definição 8 (Operadores positivos [24]) *Um operador hermitiano A em um espaço vetorial V é positivo se, para qualquer vetor $|v\rangle \in V$, o número $\langle v|A|v\rangle$ é real e não negativo. Se $\langle v|A|v\rangle$ é real e maior que zero para todo $|v\rangle \neq 0$, o operador A é dito ser positivo definido.*

Os operadores positivos são importantes pois, além de serem hermitianos, possuem decomposição espectral $\sum_i \lambda_i |i\rangle\langle i|$ com autovalores λ_i não negativos.

2.2.6 Produto tensorial

O produto tensorial é usado para agrupar espaços vetoriais em um espaço vetorial de dimensão maior. Como será visto posteriormente, o estado de um sistema quântico composto é descrito por um vetor em um espaço de Hilbert que é o produto tensorial dos espaços de Hilbert de cada sistema separadamente.

Definição 9 (Produto Tensorial [24]) *Sejam V e W espaços de Hilbert de dimensão m e n, respectivamente. Então, $V \otimes W$ (lê-se V tensorial W) é um espaço de Hilbert de dimensão mn. Os elementos de $V \otimes W$ são combinações lineares de produtos tensoriais $|v_i\rangle \otimes |w_i\rangle$ de elementos $|v_i\rangle$ de V e $|w_i\rangle$ de W, satisfazendo as seguintes propriedades:*

P1. $z(|v\rangle \otimes |w\rangle) = (z|v\rangle) \otimes |w\rangle = |v\rangle \otimes (z|w\rangle)$, $z \in C$, $|v\rangle \in V$ *e* $|w\rangle \in W$;

P2. $(|v_1\rangle + |v_2\rangle) \otimes |w\rangle = |v_1\rangle \otimes |w\rangle + |v_2\rangle \otimes |w\rangle$, $|v_1\rangle, |v_2\rangle \in V$ *e* $|w\rangle \in W$;

P3. $|v\rangle \otimes (|w_1\rangle + |w_2\rangle) = |v\rangle \otimes |w_1\rangle + |v\rangle \otimes |w_2\rangle$, $|v\rangle \in V$, $|w_1\rangle, |w_2\rangle \in W$.

Se A e B são operadores lineares em V e W, respectivamente, então

$$(A \otimes B)(|v\rangle \otimes |w\rangle) \equiv A|v\rangle \otimes B|w\rangle, \tag{2.26}$$

em que $|v\rangle \in V$ e $|w\rangle \in W$. Naturalmente,

$$(A \otimes B)\left(\sum_i a_i|v_i\rangle \otimes |w_i\rangle\right) \equiv \sum_i a_i A|v_i\rangle \otimes B|w_i\rangle, \tag{2.27}$$

para $a_i \in C$, $|v_i\rangle \in V$ e $|w_i\rangle \in W$.

Dependendo do contexto, as notações para o produto tensorial de operadores e vetores podem variar. As notações a seguir serão utilizadas nesta dissertação. Se A e B são operadores lineares em V e W, respectivamente, as notações $A \otimes B$ e $A_V B_W$ são equivalentes, ou seja,

$$A \otimes B \equiv A_V B_W. \tag{2.28}$$

Observe que os índices nos operadores indicam em que espaço de Hilbert eles atuam. Se $|v\rangle \in V$ e $|w\rangle \in W$, então é usual escrever

$$|v\rangle \otimes |w\rangle \equiv |v\rangle|w\rangle \equiv |v, w\rangle \equiv |vw\rangle. \tag{2.29}$$

Dessa forma, se A atua em V e B atua em W, então as equações seguintes são equivalentes:

$$(A \otimes B)(|v\rangle \otimes |w\rangle) \equiv A_V B_W|v\rangle|w\rangle \equiv A_V B_W|vw\rangle. \tag{2.30}$$

O produto interno em $V \otimes W$ é definido de maneira natural, em termos de produtos internos em V e W separadamente:

$$\left(\sum_i a_i|v_i\rangle \otimes |w_i\rangle, \sum_j b_j|v'_j\rangle \otimes |w'_j\rangle\right) \equiv \sum_{ij} a_i^* b_j \langle v_i|v'_j\rangle\langle w_i|w'_j\rangle, \tag{2.31}$$

em que $a_i, b_j \in C$, $|v_i\rangle, |v'_j\rangle \in V$ e $|w_i\rangle, |w'_j\rangle \in W$. A partir da definição de produto interno acima, é fácil verificar que, se $|v_i\rangle$ e $|w_i\rangle$ são duas bases ortonormais para V e W, respectivamente, então o produto $|v_i\rangle \otimes |w_i\rangle$ é uma base ortonormal para $V \otimes W$.

Em termos de representação matricial, o produto tensorial entre matrizes de operadores A e B é equivalente ao produto de Kronecker entre essas matrizes. Assim, se A é uma matriz $m \times n$ e B é $p \times q$, então

$$A \otimes B \equiv \overbrace{\begin{bmatrix} A_{11}B & A_{12}B & \dots & A_{1n}B \\ A_{21}B & A_{22}B & \dots & A_{2n}B \\ \vdots & \vdots & \vdots & \vdots \\ A_{m1}B & A_{m2}B & \dots & A_{mn}B \end{bmatrix}}^{nq} \Bigg\} mp. \tag{2.32}$$

Como exemplos, considere o produto tensorial entre os operadores de Pauli Y e Z,

$$Y \otimes Z = \begin{bmatrix} 0 \cdot Z & -i \cdot Z \\ i \cdot Z & 0 \cdot Z \end{bmatrix} = \begin{bmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \\ i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{bmatrix}, \tag{2.33}$$

e o produto tensorial entre os vetores unitários $|\theta\rangle = \cos(\theta)|0\rangle + \sin(\theta)|1\rangle$ e $|w\rangle = \frac{1}{2}|0\rangle + \sqrt{\frac{3}{4}}|1\rangle$,

$$|\theta\rangle \otimes |w\rangle = \begin{bmatrix} \cos\theta \ \frac{1}{2} \\ \cos\theta \sqrt{\frac{3}{4}} \\ \sin\theta \ \frac{1}{2} \\ \sin\theta \sqrt{\frac{3}{4}} \end{bmatrix}. \tag{2.34}$$

É extremamente importante observar que as representações matriciais dos vetores $|\theta\rangle$ e $|w\rangle$ são feitas com relação à base ortonormal $\beta = \{|0\rangle, |1\rangle\}$ para o espaço de Hilbert de dimensão dois. Dessa forma, o resultado expresso pela matriz coluna na Equação (2.34) diz respeito às coordenadas do vetor $|\theta\rangle|w\rangle$ na base ortonormal $\beta \otimes \beta = \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, ou seja,

$$\begin{aligned} |\theta w\rangle &= \frac{\cos\theta}{2}\begin{bmatrix}1\\0\\0\\0\end{bmatrix} + \cos\theta\sqrt{\frac{3}{4}}\begin{bmatrix}0\\1\\0\\0\end{bmatrix} + \frac{\sin\theta}{2}\begin{bmatrix}0\\0\\1\\0\end{bmatrix} + \sin\theta\sqrt{\frac{3}{4}}\begin{bmatrix}0\\0\\0\\1\end{bmatrix} \\ &= \frac{\cos\theta}{2}|00\rangle + \cos\theta\sqrt{\frac{3}{4}}|01\rangle + \frac{\sin\theta}{2}|10\rangle + \sin\theta\sqrt{\frac{3}{4}}|11\rangle. \end{aligned} \tag{2.35}$$

Tal resultado pode ser obtido usando diretamente a propriedade distributiva do produto tensorial (Definição 9 - P2):

$$\begin{aligned} |\theta\rangle|w\rangle &= (\cos(\theta)|0\rangle + \sin(\theta)|1\rangle) \otimes \left(\frac{1}{2}|0\rangle + \sqrt{\frac{3}{4}}|1\rangle\right) \\ &= \frac{\cos\theta}{2}|00\rangle + \cos\theta\sqrt{\frac{3}{4}}|01\rangle + \frac{\sin\theta}{2}|10\rangle + \sin\theta\sqrt{\frac{3}{4}}|11\rangle. \end{aligned} \tag{2.36}$$

O leitor pode verificar que operações como transposição e conjugação complexa são distributivas com relação ao produto tensorial. Em seguida, verifica-se que o produto tensorial de duas matrizes unitárias é uma matriz unitária; de dois operadores hermitianos é um operador hermitiano; de dois operadores positivos é um operador positivo e de dois projetores é um projetor.

Por último, a notação $|\psi\rangle^{\otimes n}$ será usada para designar n vezes o produto tensorial por $|\psi\rangle$. Por exemplo, $|\psi\rangle^{\otimes 3} = |\psi\rangle \otimes |\psi\rangle \otimes |\psi\rangle$.

2.2.7 Funções de operadores

Existem situações onde se deseja calcular o valor de uma função em um operador qualquer, que na maioria das vezes é definido por uma matriz. Se uma função f é tal que $f : C \longrightarrow C$, é possível definir uma função correspondente para matrizes normais A, i.e, matrizes tais que $A^\dagger A = AA^\dagger$, seguindo a seguinte construção. Seja $A = \sum_a \lambda_a |v_a\rangle\langle v_a|$ a decomposição espectral para o operador normal A. A função $f(A)$ é definida como sendo

$$f(A) = \sum_a f(\lambda_a)|v_a\rangle\langle v_a|. \tag{2.37}$$

Para ilustrar a definição acima, será calculada a raiz quadrada $f(A) = \sqrt{A}$ da matriz normal abaixo:

$$A = \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}.$$

Resolvendo a equação $\det|A - \lambda I| = 0$ obtém-se $\lambda_1 = 1$ e $\lambda_2 = 7$. Lembre-se que os autovetores $|v_a\rangle$ devem estar normalizados:

$$|v_1\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{e} \quad |v_2\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Note que $|v_a\rangle$ forma uma base ortonormal para C^2 e que

$$\begin{aligned} A &= \sum_a \lambda_a |v_a\rangle\langle v_a| = 1 \times \frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}\begin{bmatrix} 1 & -1 \end{bmatrix} + 7 \times \frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}. \end{aligned}$$

Dessa forma,

$$\begin{aligned} \sqrt{A} &= \sum_a \sqrt{\lambda_a} |v_a\rangle\langle v_a| = \sqrt{1} \times \frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}\begin{bmatrix} 1 & -1 \end{bmatrix} + \sqrt{7} \times \frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1.8228 & 0.8228 \\ 0.8228 & 1.8228 \end{bmatrix}. \end{aligned} \tag{2.38}$$

Uma função de extrema importância em mecânica quântica é a função traço de uma matriz. O traço de uma matriz quadrada A é definido como sendo igual a soma dos elementos da diagonal principal,

$$\text{tr}\,(A) \equiv \sum_i A_{ii}. \tag{2.39}$$

Pela definição, é fácil verificar as seguintes propriedades para o traço, para matrizes quadradas A e B de mesma ordem e $z \in C$:

P1. $\text{tr}\,(A+B) = \text{tr}\,(A) + \text{tr}\,(B)$;

P2. $\text{tr}\,(zA) = z\text{tr}\,(A)$;

P3. $\text{tr}\,(AB) = \text{tr}\,(BA)$.

Uma última propriedade importante é que o traço é invariante sobre transformações unitárias de similaridade $A \rightarrow UAU^\dagger$, pois $\text{tr}\left(UAU^\dagger\right) = \text{tr}\left(U^\dagger UA\right) = \text{tr}\,(A)$.

Uma fórmula útil para o cálculo do traço de operadores na forma $A|\psi\rangle\langle\psi|$, em que $|\psi\rangle$ é um vetor unitário, pode ser obtida da seguinte forma. Use o procedimento de Gram-Schmidt para obter uma base ortonormal $|i\rangle$ que tenha $|\psi\rangle$ como primeiro vetor. Assim,

$$\begin{aligned} \text{tr}\,(A|\psi\rangle\langle\psi|) &= \sum_i \langle i|A|\psi\rangle\langle\psi|i\rangle \\ &= \langle\psi|A|\psi\rangle \end{aligned} \tag{2.40}$$

2.3 Postulados da mecânica quântica

A mecânica quântica é um arcabouço para o desenvolvimento de teorias físicas. Sendo assim, seus postulados não enunciam quais leis um sistema físico deve obedecer; ao invés disso, eles provêem um ambiente conceitual e matemático para o desenvolvimento de tais leis.

Nesta seção serão apresentados os quatro principais postulados da mecânica quântica e alguns resultados que são conseqüências diretas dos mesmos. Uma abordagem mais detalhada pode ser encontrada em Nielsen e Chuang (2000) [24].

2.3.1 Espaço de estados

O primeiro postulado estabelece o ambiente matemático onde os sistemas quânticos são definidos. Tal ambiente é o já conhecido espaço de Hilbert.

Postulado 1 *A todo sistema físico isolado está associado um espaço vetorial complexo com produto interno, isto é, um espaço de Hilbert, chamado de* espaço de estado *do sistema. O sistema é completamente descrito pelo seu* vetor de estado, *que é um vetor unitário no espaço de estado do sistema.*

O sistema quântico mais simples é o *qubit*, ou bit quântico. O qubit pertence a um espaço de estado de duas dimensões. Assim, um qubit arbitrário pode ser escrito como

$$|\psi\rangle = a|0\rangle + b|1\rangle, \tag{2.41}$$

em que a e b são números complexos e $|0\rangle$, $|1\rangle$ são definidos na Equação (2.8). O postulado impõe norma unitária para $|\psi\rangle$, $\langle\psi|\psi\rangle = 1$, o que implica dizer que $|a|^2+|b|^2 = 1$. O qubit será considerado, neste trabalho, o sistema quântico fundamental. Nada impede, porém, de trabalhar com outros sistemas quânticos fundamentais, como o *qutrit*, sistema quântico onde os vetores de estado pertencem a um espaço de estado de três dimensões. Existem diversos sistemas físicos que podem ser descritos em termos de qubits. A polarização de um fóton e o spin de um elétron são exemplos de tais sistemas. Por enquanto, é suficiente pensar em qubits em termos abstratos, não se importando com implementações ou realizações físicas.

A discussão sobre qubits será feita em torno de uma base ortonormal $\{|0\rangle, |1\rangle\}$. Em computação quântica, os estados $|0\rangle$ e $|1\rangle$ são, intuitivamente, análogos aos bits clássicos 0 e 1, respectivamente. A grande diferença é que os estados $|0\rangle$ e $|1\rangle$ podem coexistir

em um mesmo sistema $|\psi\rangle$, o que se denomina de superposição: $|\psi\rangle = a|0\rangle + b|1\rangle$. De maneira geral, uma combinação linear da forma $\sum_i \alpha_i|\psi_i\rangle$ é dita ser uma superposição dos estados $|\psi_i\rangle$, com amplitude α_i para o estado $|\psi_i\rangle$. Assim, por exemplo, o estado

$$\frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

é uma superposição dos estados $|0\rangle$ e $|1\rangle$, com amplitude $1/\sqrt{2}$ para o estado $|0\rangle$ e amplitude $-1/\sqrt{2}$ para o estado $|1\rangle$.

2.3.2 Evolução

O próximo postulado fornece uma descrição de como um estado $|\psi\rangle$ de um sistema quântico fechado evolui com o tempo.

Postulado 2 *A evolução de um sistema quântico isolado é descrita por transformações unitárias. Ou seja, o estado $|\psi_1\rangle$ do sistema no tempo t_1 está relacionado com o estado $|\psi_2\rangle$ do sistema no tempo t_2 por meio de um operador unitário U, que depende somente dos tempos t_1 e t_2,*

$$|\psi_2\rangle = U|\psi_1\rangle. \tag{2.42}$$

O Postulado 2 afirma que o sistema quântico deve estar isolado para que possa evoluir unitariamente. Naturalmente, qualquer sistema (exceto o universo como um todo) interage de alguma forma com outros sistemas. Entretanto, existem sistemas físicos que podem ser descritos com boa aproximação, durante um intervalo de tempo, como sendo fechados (Para mais informações consulte Nielsen e Chuang (2000) [24, Cap. 7]).

O Postulado 2 estabelece como estados quânticos de um sistema quântico fechado estão relacionados em dois instantes diferentes de tempo. O físico austríaco Ervin Schrödinger descreveu a evolução de um sistema quântico no tempo contínuo através de uma equação diferencial que ficaria conhecida como equação de Schrödinger. O Postulado 2 é então reescrito.

Postulado 2' *A evolução no tempo de um estado $|\psi\rangle$ num sistema quântico isolado é descrita pela equação de Schrödinger,*

$$i\hbar\frac{d|\psi\rangle}{dt} = H|\psi\rangle. \tag{2.43}$$

Nesta equação, $\hbar$ é a constante de Planck e H é um operador hermitiano fixo, conhecido como hamiltoniano do sistema quântico fechado.

Em princípio, se o hamiltoniano do sistema é conhecido, então a dinâmica do sistema quântico fica inteiramente determinada. Encontrar o hamiltoniano de um determinado sistema é geralmente uma tarefa difícil. Além do mais, não é objetivo deste trabalho descrevê-los.

Os enunciados para o Postulado 2 são equivalentes, ou seja, a partir de um é possível obter o outro. Para ver isso, tome a equação de Schrödinger como ponto de partida. Esta equação é uma equação diferencial e homogênea de primeira ordem. Reescrevendo-a como

$$\frac{d|\psi\rangle}{dt} + \frac{i}{\hbar}H|\psi\rangle = 0,$$

e tomando $|\psi(t_1)\rangle = |\psi_1\rangle$, tem-se que

$$|\psi_2\rangle = \exp\left[\frac{-iH(t_2 - t_1)}{\hbar}\right]|\psi_1\rangle. \tag{2.44}$$

Para mostrar que as Equações (2.42) e (2.44) são equivalente, basta mostrar que

$$U_H \equiv \exp\left[\frac{-iH_{2,1}}{\hbar}\right] \tag{2.45}$$

é uma matriz unitária. $H_{2,1}$ foi escrito em lugar de $H(t_2 - t_1)$ para não carregar a notação. Se $H_{2,1}$ é hermitiano,

$$H_{2,1} = H_{2,1}^{\dagger} = \sum_j \lambda_j |j\rangle\langle j| \tag{2.46}$$

para alguma base ortonormal $|j\rangle$ de autovetores com autovalores λ_j associados. Então,

$$U_H = \exp\left[-\frac{i}{\hbar}H_{2,1}\right] = \sum_j \exp\left[-\frac{i}{\hbar}\lambda_j\right]|j\rangle\langle j|. \tag{2.47}$$

Ainda,

$$\begin{aligned} U_H^{\dagger} &= \sum_j \exp\left[-\frac{i}{\hbar}\lambda_j\right]^* |j\rangle\langle j| \\ &= \sum_j \exp\left[\frac{i}{\hbar}\lambda_j\right]|j\rangle\langle j|. \end{aligned} \tag{2.48}$$

Por fim,

$$\begin{aligned} U_H^\dagger U_H &= U_H U_H^\dagger = \left[\sum_j \exp\left(-\frac{i}{\hbar}\lambda_j\right)|j\rangle\langle j|\right]\left[\sum_l \exp\left(\frac{i}{\hbar}\lambda_l\right)|l\rangle\langle l|\right] \\ &= \sum_j\sum_l \exp\left[-\frac{i}{\hbar}\lambda_j + \frac{i}{\hbar}\lambda_l\right]|j\rangle\langle j|l\rangle\langle l| \\ &= \sum_j \exp(0)|j\rangle\langle j| \\ &= I. \end{aligned} \tag{2.49}$$

Alguns operadores unitários são de particular interesse no estudo da teoria da computação e informação quânticas. O operador de Pauli X é também chamado de porta *NOT*, por analogia com a porta clássica *NOT*. Note que se $|\psi\rangle = a|0\rangle + b|1\rangle$, então $X|\psi\rangle = a|1\rangle + b|0\rangle$. Dependendo do contexto, as matrizes X e Z são chamadas de matrizes de troca de bit (*bit flip*) e troca de fase (*fase flip*). Outro operador importante é a porta de Hadamard, denotada por H e definida como sendo

$$H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{2.50}$$

Em computação quântica e informação quântica é comum usar a expressão "aplicar um operador unitário ao sistema quântico" para se referir à evolução do sistema segundo um determinado operador unitário.

2.3.3 Medições quânticas

A evolução de um sistema quântico que não interage com o mundo exterior é completamente descrita através de operadores unitários. Porém, quando o experimentador introduz um aparelho de medição para obter alguma informação sobre o estado atual do sistema, o mesmo se torna um sistema aberto e, assim, sujeito a interações não unitárias. O postulado seguinte descreve os efeitos das medições nos sistemas quânticos.

Postulado 3 *As medições em sistemas quânticos são descritas por um conjunto de operadores de medição $\{M_m\}$, que atuam no espaço de estado do sistema a ser medido. O índice m se refere ao resultado que pode ocorrer na medição. Assim, se o estado do sistema quântico imediatamente antes da medição é $|\psi\rangle$, então a probabilidade de m ocorrer é dada por*

$$p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle, \tag{2.51}$$

e o estado do sistema após a medição será

$$|\psi'\rangle = \frac{M_m|\psi\rangle}{\sqrt{\langle\psi|M_m^\dagger M_m|\psi\rangle}}. \tag{2.52}$$

Os operadores de medição devem satisfazer a equação de completude,

$$\sum_m M_m^\dagger M_m = I. \tag{2.53}$$

Note que a equação de completude é derivada da restrição imposta ao somatório das probabilidades:

$$1 = \sum_m p(m) = \sum_m \langle\psi|M_m^\dagger M_m|\psi\rangle = \langle\psi|\sum_m M_m^\dagger M_m|\psi\rangle. \tag{2.54}$$

Como exemplo, considere a medição de um qubit com relação à base computacional $\{|0\rangle, |1\rangle\}$. Esta medição pode gerar duas saídas possíveis e possui operadores de medição $M_0 = |0\rangle\langle 0|$, $M_1 = |1\rangle\langle 1|$. Observe que cada operador é hermitiano e, juntos, obedecem à equação de completude. Se o estado a ser medido é, inicialmente, $|\psi\rangle = a|0\rangle + b|1\rangle$, então a probabilidade de obter 0 na medição é

$$p(0) = \langle\psi|M_0^\dagger M_0|\psi\rangle = \langle\psi|M_0|\psi\rangle = |a|^2. \tag{2.55}$$

Neste caso, o estado do sistema logo após a medição será

$$|\psi\rangle_0 = \frac{M_0|\psi\rangle}{|a|} = \frac{a}{|a|}|0\rangle. \tag{2.56}$$

Da mesma forma, a probabilidade de obter 1 é $p(1) = |b|^2$, com o seguinte estado pós-medição:

$$|\psi\rangle_1 = \frac{M_1|\psi\rangle}{|a|} = \frac{b}{|b|}|1\rangle. \tag{2.57}$$

É comum usar o termo "colapsar" para se referir à transformação, às vezes não unitária, do estado $|\psi\rangle$ para o estado $|\psi'\rangle$, dada pela Equação (2.52). Como será visto na Seção 2.3.3, escalares tais como $a/|a|$, que possuem módulo igual a um, podem ser efetivamente ignorados, de tal forma que o estado pós-medição será $|0\rangle$ ou $|1\rangle$.

O Postulado 3 descreve as medições em sistemas quânticos da forma mais geral possível. Existem dois casos particulares deste postulado que são de interesse para o trabalho desenvolvido nesta dissertação. São eles: as medições projetivas e as medições POVM.

Medições projetivas

Medições projetivas formam uma classe especial de medições em sistemas quânticos.

Medições projetivas. *Uma medição projetiva é descrita por um* observável M, *que é um operador hermitiano no espaço de estado do sistema a ser medido. O observável possui uma decomposição espectral,*

$$M = \sum_m m P_m, \tag{2.58}$$

em que P_m é um projetor sobre o autoespaço de M com autovalor m. Os resultados possíveis da medição correspondem aos autovalores do observável. Quando o estado $|\psi\rangle$ é medido, a probabilidade de obter m é dada por

$$p(m) = \langle \psi | P_m | \psi \rangle. \tag{2.59}$$

Dado que a saída m ocorreu, o estado do sistema quântico imediatamente após a medição será

$$|\psi'\rangle = \frac{P_m |\psi\rangle}{\sqrt{p(m)}}. \tag{2.60}$$

As medições projetivas podem ser entendidas como um caso particular do Postulado 3. Suponha que os operadores de medição no Postulado 3, além de satisfazerem a equação de completude, satisfaçam também a condição de ortogonalidade, ou seja, M_m são hermitianos e $M_m M_{m'} = \delta_{m,m'} M_m$. Então, as medições descritas no Postulado 3 se reduzem às medições projetivas definidas acima.

Uma das propriedades interessantes das medições projetivas é o valor esperado das medições. Suponha que o estado $|\psi\rangle$ é continuamente preparado e sujeito a medições definidas por um observável $M = \sum_m m P_m$. Como visto, o resultado de cada medição será m com probabilidade $p(m) = \langle \psi | P_m | \psi \rangle$. Desta forma, o valor esperado quando um número grande de medições é feito é definido como sendo

$$\begin{aligned} \mathbf{E}(M) &= \sum_m m p(m) \\ &= \langle \psi | \left(\sum_m m P_m \right) | \psi \rangle \\ &= \langle \psi | M | \psi \rangle. \end{aligned} \tag{2.61}$$

É comum escrever $\langle M \rangle \equiv \langle \psi | M | \psi \rangle$. O desvio padrão é facilmente calculado:

$$\begin{aligned} \Delta(M) &= \langle (M - \langle M \rangle)^2 \rangle \\ &= \langle M^2 \rangle - \langle M \rangle^2. \end{aligned} \tag{2.62}$$

A definição de medições em termos de projetores e os conceitos de valor médio e desvio padrão de um observável formam a base principal para um dos resultados mais interessantes da mecânica quântica: o princípio da incerteza de Heisenberg (ver Nielsen and Chuang [24, pp 89]).

Como exemplo de medições projetivas, considere a medição do observável X, que possui autovalores $+1$ e -1 com autovetores $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ e $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, respectivamente. Observe que

$$P_{+1} = |+\rangle\langle +| = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{e} \quad P_{-1} = |-\rangle\langle -| = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}. \tag{2.63}$$

Medindo X sobre o estado $|\psi\rangle = |0\rangle$, obtém-se $+1$ e -1 com as respectivas probabilidades

$$p(+1) = \langle \psi | + \rangle\langle + | \psi \rangle = 1/2 \quad \text{e} \quad p(-1) = \langle \psi | - \rangle\langle - | \psi \rangle = 1/2. \tag{2.64}$$

Medições POVM

As medições descritas pelo Postulado 3 e pelas medições projetivas explicitam, além da probabilidade de ocorrer determinada saída, o estado do sistema logo após ser medido, chamado de estado pós-medição. Entretanto, existem situações onde somente as probabilidades de saídas são de interesse, não importando o estado posterior à medição. Esse tipo de medição recebe o nome de POVM (do inglês, *Positive Operator-Valued Measure*). Medições POVM serão bastante utilizadas ao longo desta dissertação.

Medições POVM *Considere uma medição quântica descrita no Postulado 3, com elementos de medição M_m. Defina*

$$E_m \equiv M_m^\dagger M_m. \tag{2.65}$$

As medições POVM são descritas por um conjunto de operadores POVM $\{E_m\}$, em que E_m é um operador positivo e tal que $\sum_m E_m = I$. Dessa forma, a probabilidade de se obter uma saída m quando o estado $|\psi\rangle$ é medido é $p(m) = \langle \psi | E_m | \psi \rangle$. O conjunto $\{E_m\}$ é comumente chamado de POVM.

É interessante verificar que, qualquer conjunto de operadores positivos $\{E_m\}$ satisfazendo a equação de completude, $\sum_m E_m = I$, define uma medição na forma do Postulado 3. Para ver isso, defina $M_m = \sqrt{E_m}$. É fácil ver que $\sum_m M_m^\dagger M_m^{\cdot} = \sum_m E_m = I$, de tal maneira que $\{M_m\}$ forma um conjunto de operadores de medição com POVM $\{E_m\}$ associado. A seguir são ilustradas duas aplicações de medições POVM.

A primeira aplicação é um resultado bastante conhecido em mecânica quântica, o de que "nenhum esquema de medição é capaz de distinguir perfeitamente entre dois estados quânticos não ortogonais".

A prova é mostrada para o caso em que os estados são qubits, i.e., pertencem ao espaço de Hilbert de dimensão dois, mas pode ser facilmente estendida para estados não ortogonais em um espaço de dimensão n. Considere $|\psi_1\rangle$ e $|\psi_2\rangle$ não ortogonais. Se for possível distinguir perfeitamente entre $|\psi_1\rangle$ e $|\psi_2\rangle$, significa que existe um POVM $\{E_1, E_2\}$ tal que

$$\langle\psi_1|E_1|\psi_1\rangle = 1 \quad \text{e} \quad \langle\psi_2|E_2|\psi_2\rangle = 1. \tag{2.66}$$

A prova é feita por contradição. Desde que $\sum E_m = I$ e $\sum_m \langle\psi_1|E_m|\psi_1\rangle = 1$, então $\langle\psi_1|E_2|\psi_1\rangle = 0$ e, conseqüentemente, $\sqrt{E_2}|\psi_1\rangle = 0$. É sempre possível fazer a decomposição $|\psi_2\rangle = \alpha|\psi_1\rangle + \beta|\varphi\rangle$, em que $|\varphi\rangle$ é ortogonal a $|\psi_1\rangle$ e $|\beta| < 1$. Note que $\sqrt{E_2}|\psi_2\rangle = \alpha\sqrt{E_2}|\psi_1\rangle + \beta\sqrt{E_2}|\varphi\rangle = \beta\sqrt{E_2}|\varphi\rangle$. Isso contradiz a Equação (2.66), pois

$$\begin{aligned} \langle\psi_2|E_2|\psi_2\rangle &= |\beta|^2\langle\varphi|E_2|\varphi\rangle \\ &\leq |\beta|^2 \\ &< 1. \end{aligned} \tag{2.67}$$

A penúltima desigualdade é porque $\langle\varphi|E_2|\varphi\rangle \leq \sum_m \langle\varphi|E_m|\varphi\rangle = \langle\varphi|\varphi\rangle = 1$.

A segunda aplicação de medições POVM concerne exatamente a distinguibilidade de estados quânticos não ortogonais. Imagine um jogo em que Alice prepara um dos estados $|\psi_a\rangle = |0\rangle$ ou $|\psi_b\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ e entrega a Bob que tem a missão de inferir, através de uma medição, qual foi o estado que Alice preparou. Claramente, trata-se de dois estados não ortogonais, pois $\langle\psi_a|\psi_b\rangle = 1/\sqrt{2}$. Embora não seja possível distingui-los sempre, é possível montar um esquema de medição POVM tal que Bob algumas vezes acerte qual o estado preparado por Alice, mas que ele nunca erra na sua escolha.

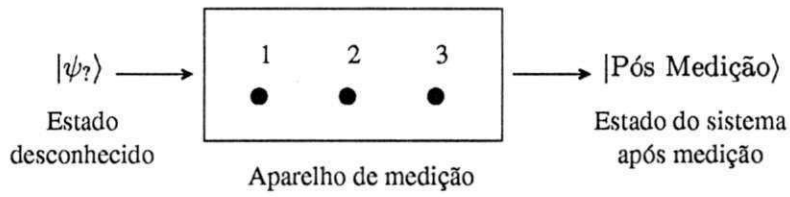


Figura 2.4: Aparelho de medição para distinguir os estados não ortogonais $|\psi_a\rangle$ e $|\psi_b\rangle$.

Para tanto, considere um POVM contendo os seguintes elementos:

$$E_1 \equiv \frac{\sqrt{2}}{1+\sqrt{2}}|1\rangle\langle 1|. \tag{2.68}$$

$$E_2 \equiv \frac{\sqrt{2}}{1+\sqrt{2}}\frac{(|0\rangle - |1\rangle)(\langle 0| - \langle 1|)}{2}. \tag{2.69}$$

$$E_3 \equiv I - E_1 - E_2. \tag{2.70}$$

O leitor pode verificar facilmente que tais operadores são positivos e satisfazem a equação de completude.

Suponha que Alice prepara o estado $|\psi_b\rangle$ e que uma medição definida pelo POVM $\{E_1, E_2, E_3\}$ é realizada. Note que, nesse caso, $p(2) = \langle\psi_b|E_2|\psi_b\rangle = 0$, i.e.,

$$\begin{aligned} p(2) &= \left[\frac{1}{\sqrt{2}}(\langle 0| + \langle 1|)\right]\frac{\sqrt{2}}{1+\sqrt{2}}\frac{(|0\rangle - |1\rangle)(\langle 0| - \langle 1|)}{2}\left[\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\right] \\ &= \frac{1}{4}\frac{\sqrt{2}}{1+\sqrt{2}}\left[\langle 0|0\rangle - \langle 0|1\rangle + \langle 1|0\rangle - \langle 1|1\rangle\right]\left[\langle 0|0\rangle - \langle 1|0\rangle + \langle 0|1\rangle - \langle 1|1\rangle\right] \\ &= 0. \end{aligned} \tag{2.71}$$

Note ainda que $p(1) = \langle\psi_b|E_1|\psi_b\rangle = 1 - \sqrt{2}/2$ e $p(3) = \langle\psi_b|E_3|\psi_b\rangle = \sqrt{2}/2$. Em resumo, se Bob efetua uma medição e obtém 2 como resultado, Alice com certeza não lhe entregou o estado $|\psi_b\rangle$, pois $\langle\psi_b|E_2|\psi_b\rangle = 0$. Logo ele pode afirmar que o estado que Alice preparou foi $|\psi_a\rangle = |0\rangle$. Da mesma forma, é fácil verificar que a probabilidade de Bob obter 1 na medição dado que $|\psi_a\rangle$ lhe foi entregue é igual a zero, $p(1) = \langle\psi_a|E_1|\psi_a\rangle = 0$. Bob conclui que Alice lhe entregou o estado $|\psi_b\rangle$ quando ele obtém 1 como resultado da medição. Algumas vezes, entretanto, Bob poderá obter 3 como resultado da medição. Neste caso, ele nada pode inferir sobre o estado que Alice lhe entregou. O mais importante é que Bob nunca erra na sua escolha.

O exemplo acima pode ser mais facilmente entendido observando o esquema na Figura 2.4. Aqui, o medidor é uma caixa preta que possui três lâmpadas, cada uma indicando um dos possíveis resultados da medição, i.e., 1, 2 ou 3, correspondentes

aos elementos do POVM, E_1, E_2 e E_3, respectivamente. Quando o estado quântico desconhecido $|\psi_?\rangle$ é apresentado ao aparelho, ele efetua a medição e acende a luz correspondente à saída, de acordo com a distribuição de probabilidade já mencionada.

Os elementos POVM foram definidos de tal forma que, se o estado apresentado ao medidor for $|\psi_2\rangle$, a lâmpada 1 nunca irá acender. Logicamente, quando tal lâmpada acender, indicará que o estado desconhecido é $|\psi_1\rangle$. Já quando o estado $|\psi_1\rangle$ é submetido à medição, a lâmpada 2 nunca acenderá. Assim, quando a lâmpada 2 acender, indicará que o estado na entrada do medidor é $|\psi_2\rangle$. Observe que o estado do sistema após a medição não é de interesse, o que caracteriza as medições POVM.

Fase global e fase relativa

A fase é um termo usado em mecânica quântica que possui diversos significados, dependendo do contexto em que é utilizada. As chamadas fase global e fase relativa estão intimamente relacionadas com medições de estados quânticos, e serão descritas aqui.

Considere o estado $e^{j\theta}|\psi\rangle$, em que $|\psi\rangle$ é um vetor de estado e θ é um número real. O número complexo $e^{j\theta}$ é chamado de fator de fase global. A fase global se caracteriza por não afetar as estatísticas de medição de um estado qualquer $|\psi\rangle$. Para ver isso, basta notar que se M_m é um operador de medição na forma do Postulado 3, então as probabilidades de se obter m na medição são iguais para $e^{j\theta}|\psi\rangle$ e $|\psi\rangle$, i.e., $\langle\psi|e^{-j\theta}M_m^\dagger M_m e^{j\theta}|\psi\rangle = \langle\psi|M_m^\dagger M_m|\psi\rangle$. Por essa razão, o estado $e^{j\theta}|\psi\rangle$ é dito ser igual a $|\psi\rangle$ a menos de uma fase global θ.

A fase relativa possui um significado um pouco diferente da fase global. Considere os estados

$$\frac{|0\rangle + |1\rangle}{\sqrt{2}} \qquad \text{e} \qquad \frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

No primeiro estado, a amplitude de $|1\rangle$ é $1/\sqrt{2}$. Para o segundo estado, esta a amplitude é $-1/\sqrt{2}$. Nos dois casos, a magnitude da amplitude é a mesma, mas diferem no sinal (fase). Note que as amplitudes acima são dependentes das bases utilizadas na representação dos estados. Assim, duas amplitudes a e b diferem por uma fase relativa se existe um número real θ tal que $a = e^{j\theta}b$. De forma mais geral, dois estados

$$|\psi_1\rangle = \sum_i a_i|i\rangle \qquad \text{e} \qquad |\psi_2\rangle = \sum_i b_i|i\rangle$$

diferem por uma fase relativa com relação à base $|i\rangle$, se $a_i = e^{j\theta_i}$. O leitor pode verificar que a fase relativa altera completamente as estatísticas de medição para um

determinado conjunto de operadores de medição $\{M_m\}$.

2.3.4 Composição de sistemas quânticos

Sistemas quânticos individuais podem interagir para formarem sistemas quânticos compostos. O postulado seguinte descreve como o espaço de estado do sistema composto é construído a partir dos espaços de estado dos sistemas individuais.

Postulado 4 *O espaço de estado de um sistema quântico composto é o produto tensorial dos espaços de estado dos sistemas individuais. Se o sistema composto é formado por n sistemas, e cada sistema individual é preparado no estado $|\psi_i\rangle$, então o estado do sistema composto é $|\psi_1\rangle \otimes |\psi_2\rangle \otimes \ldots \otimes |\psi_n\rangle$.*

A notação de índice é bastante usada na literatura para indicar estados e operadores nos diferentes sistemas. Por exemplo, em um sistema composto de três qubits, a notação X_2 se refere ao operador de Pauli X sendo aplicado ao segundo qubit.

O Postulado 4 implica na existência de um dos fenômenos mais intrigantes da mecânica quântica, que não possui análogo em toda a teoria clássica: o emaranhamento. Por definição, um estado composto é dito ser emaranhado se ele não puder ser decomposto em produtos tensoriais de estados dos sistemas individuais. Para entender melhor, considere o estado composto de dois qubits

$$|\psi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}. \tag{2.72}$$

Esse estado é emaranhado, pois ele não pode ser escrito como o produto tensorial de dois qubits quaisquer $|a\rangle$ e $|b\rangle$. Para provar isto, considere

$$|a\rangle = \alpha_a|0\rangle + \beta_a|0\rangle \qquad \text{e} \qquad |b\rangle = \alpha_b|0\rangle + \beta_b|0\rangle,$$

lembrando que α_x e β_x são números complexos tais que $|\alpha_x|^2 + |\beta_x|^2 = 1$. Fazendo

$$\begin{aligned} |a\rangle|b\rangle &= (\alpha_a|0\rangle + \beta_a|0\rangle) \otimes (\alpha_b|0\rangle + \beta_b|0\rangle) \\ &= \alpha_a\alpha_b|00\rangle + \beta_a\beta_b|11\rangle + \alpha_a\beta_b|01\rangle + \alpha_b\beta_a|10\rangle, \end{aligned} \tag{2.73}$$

é fácil ver que $|\psi\rangle \neq |a\rangle|b\rangle$, ou seja, não existem números complexos α_a, α_b, β_a e β_b tais que $|\psi\rangle = |a\rangle \otimes |b\rangle$.

No início do século XX, Einstein, Podolsky, Rosen e Bell descobriram propriedades interessantes de quatro estados emaranhados, que eram estranhas àquela época. Tais

estados formam uma base para o espaço de Hilbert de dimensão quatro. São eles:

$$|\beta_{00}\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \tag{2.74}$$

$$|\beta_{01}\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}}, \tag{2.75}$$

$$|\beta_{10}\rangle = \frac{|00\rangle - |01\rangle}{\sqrt{2}} \tag{2.76}$$

e

$$|\beta_{11}\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}. \tag{2.77}$$

Na literatura, esses estados são chamados de vários nomes: estados de Bell, estados EPR, pares EPR ou, finalmente, de base de Bell.

Apesar de estar emaranhado, o estado $|\beta_{00}\rangle$ ainda é um estado composto de dois qubits, como ilustrado na Figura 2.5. Isto que dizer que, após a sua criação, ele pode ser fisicamente separado em duas partes. Porém, o estado propriamente dito de quaisquer uma das partes é indeterminado, ou seja, somente o estado do sistema composto é determinado. Não é possível dizer, portanto, que o primeiro qubit está num estado $|a\rangle$ e o segundo num estado $|b\rangle$.

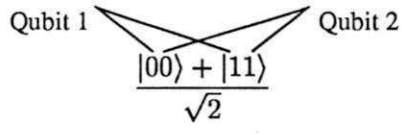


Figura 2.5: O estado de Bell $|\beta_{00}\rangle$.

Imagine que são criados dois sistemas físicos idênticos, cada um no estado $|0\rangle$, e que esses estados são colocados juntos para formar um sistema composto, cujo estado é claramente $|00\rangle$. Evidentemente, $|00\rangle$ não é um estado emaranhado, desde que $|00\rangle = |0\rangle \otimes |0\rangle$. Suponha agora que este sistema seja submetido a uma dinâmica descrita pelo operador unitário

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{bmatrix}. \tag{2.78}$$

O operador U é unitário, pois $UU^\dagger = U^\dagger U = I$. De acordo com o Postulado 2, o estado do sistema após a interação será

$$\begin{aligned} |\psi\rangle &= U|00\rangle \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 1 \end{bmatrix}^T \\ &= |\beta_{00}\rangle. \end{aligned} \tag{2.79}$$

O emaranhamento é mais do que uma singularidade da teoria quântica, ele é tido como sendo um recurso físico, assim como a capacidade de canal é um recurso físico em Teoria da Informação [11]. Isso quer dizer que a presença do emaranhamento é essencial em determinadas aplicações da teoria, assim como uma capacidade de canal maior do que zero possibilita a troca de informações entre duas partes. O teletransporte de estados quânticos é uma das aplicações mais interessantes do emaranhamento [24, pp. 26].

2.4 O operador de densidade

A teoria da mecânica quântica apresentada até aqui foi desenvolvida em termos de vetores que representam estados de sistemas quânticos em um espaço de Hilbert apropriado. Esses estados são chamados de **estados puros**, e expressam situações de ignorância mínima, em que não há nada mais a ser determinado sobre tais estados. Uma outra noção, a de **estado misto**, foi introduzida na teoria quântica para lidar com situações de maior ignorância, em particular

- com famílias $\mathcal{F}$ (do inglês, *ensembles*) em que o sistema em questão pode estar em qualquer um dois estados puros $|\psi_1\rangle, |\psi_2\rangle, \ldots$, com probabilidades $p_1, p_2, \ldots$;
- em situações em que o sistema em questão (chamado de A) é parte de um sistema maior (chamado de AB) que, por si só, está num estado puro Ψ emaranhado.

O formalismo matemático empregado nesta representação é chamado de operador de densidade. Esse formalismo fornece uma ferramenta adequada para descrever um sistema quântico cujo estado não é completamente conhecido. Mais precisamente,

Definição 10 (Operador de densidade [24].) *Suponha que um sistema quântico esteja em algum estado $|\psi_i\rangle$ com probabilidade p_i. O operador de densidade para esse sistema é definido como sendo*

$$\rho \equiv \sum_i p_i |\psi_i\rangle\langle\psi_i|. \tag{2.80}$$

O operador de densidade é também chamado de matriz densidade do sistema. Como será visto mais adiante, é possível enunciar os quatro postulados da Seção 2.3 em termos de operadores de densidade.

2.4.1 Propriedades gerais dos operadores de densidade

Os resultados descritos nesta seção ajudam na caracterização dos operadores de densidade e exibem propriedades importantes, como o grau de liberdade na representação de uma família em termos de operadores de densidade.

Teorema 2 (Caracterização de operadores de densidade [24].) *Um operador ρ é um operador de densidade associado a uma família $\{p_i, |\psi_i\rangle\}$ se e somente se as condições abaixo são satisfeitas:*

1. **(Condição do traço)** *ρ tem traço igual a 1;*

2. **(Condição de positividade)** *ρ é um operador positivo.*

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

$$\mathrm{tr}\,(\rho) = \sum_i p_i \mathrm{tr}\,(|\psi_i\rangle\langle\psi_i|) = \sum_i p_i = 1, \tag{2.81}$$

de forma que a condição do traço é satisfeita. Suponha que $|\varphi\rangle$ é um vetor qualquer do mesmo espaço de estado de ρ, então

$$\begin{aligned} \langle\varphi|\rho|\varphi\rangle &= \sum_i p_i \langle\varphi|\psi_i\rangle\langle\psi_i|\varphi\rangle \\ &= \sum_i p_i |\langle\varphi|\psi_i\rangle|^2 \\ &\geq 0, \end{aligned} \tag{2.82}$$

de forma que a condição de positividade é também satisfeita.

Equivalentemente, suponha que ρ é um operador de densidade satisfazendo as condições do traço e da positividade. Desde que ρ é positivo, ele deve ter decomposição espectral

$$\rho = \sum_j \lambda_j |j\rangle\langle j|, \tag{2.83}$$

em que os vetores $|j\rangle$ são ortonormais e λ_j são autovalores reais e não negativos de ρ. Pela condição do traço tem-se que $\sum_j \lambda_j = 1$. Portanto, um sistema que se encontra num estado $|j\rangle$ com probabilidade λ_j irá ter operador de densidade ρ. Isto é, a família $\{\lambda_j, |j\rangle\}$ possui operador de densidade ρ. ■

A caracterização de estados puros e mistos fica mais clara quando o conceito de operador de densidade é usado. Claramente, se o sistema se encontra em algum estado puro $|\psi\rangle$, seu operador de densidade é $\rho = |\psi\rangle\langle\psi|$. No exercício seguinte, mostra-se como identificar se um operador de densidade representa um sistema puro ou misto.

Exercício 11 (Sistemas puros e mistos.) *Seja ρ um operador de densidade. Mostre que $tr(\rho^2) \leq 1$, com igualdade se e somente se ρ é puro.*

Resolução. Se ρ é operador de densidade, então $\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$ e

$$\rho^2 = \sum_i \lambda_i^2 |\psi_i\rangle\langle\psi_i|.$$

Pela condição do traço, $\sum_i \lambda_i = 1$. Como $0 \leq \lambda_i \leq 0$ e $\lambda_i^2 \leq \lambda_i$, tem-se que

$$\begin{aligned} \mathrm{tr}\left(\rho^2\right) &= \mathrm{tr}\left(\sum_i \lambda_i^2 |\psi_i\rangle\langle\psi_i|\right) \\ &= \sum_i \lambda_i^2 \mathrm{tr}\left(|\psi_i\rangle\langle\psi_i|\right) \\ &= \sum_i \lambda_i^2 \\ &\leq \sum_i \lambda_i \\ &= 1. \end{aligned} \tag{2.84}$$

Se ρ é um estado puro, então $\rho = |\psi\rangle\langle\psi|$ e

$$\begin{aligned} \mathrm{tr}\left(\rho^2\right) &= \mathrm{tr}\left(|\psi\rangle\langle\psi|\right) \\ &= \langle\psi|\psi\rangle \\ &= 1. \end{aligned} \tag{2.85}$$

Equivalentemente, se $\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$ e tal que tr $(\rho^2) = 1$, então

$$\sum_i \lambda_i^2 = 1.$$

Tal condição só se verifica se e somente se $\lambda_k = 1$ e $\lambda_{i\neq k} = 0$, de formà que

$$\rho = |\psi_k\rangle\langle\psi_k|$$

é um estado puro. ■

Suponha que um sistema quântico possa estar no estado $|0\rangle$ com probabilidade 3/4 e no estado $|1\rangle$ com probabilidade 1/4. A matriz densidade de tal sistema é dada por

$$\rho = \frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1|. \tag{2.86}$$

Agora suponha que outro sistema quântico esteja em um dos estados

$$|a\rangle \equiv \frac{\sqrt{3}}{2}|0\rangle + \frac{1}{2}|1\rangle \text{ou} \tag{2.87}$$

$$|b\rangle \equiv \frac{\sqrt{3}}{2}|0\rangle - \frac{1}{2}|1\rangle \tag{2.88}$$

com probabilidade 1/2. Note que este último possui uma matriz de densidade

$$\rho = \frac{1}{2}|a\rangle\langle a| + \frac{1}{2}|b\rangle\langle b| = \frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1|. \tag{2.89}$$

Isto é, duas famílias diferentes dão origem a uma mesma matriz de densidade! O próximo teorema expõe quais as classes de famílias que possuem a mesma matriz de densidade.

Teorema 3 ([24]) *Os conjuntos* $\{|\tilde{\psi}_i\rangle\}$ *e* $\{|\tilde{\varphi}_i\rangle\}$ *geram a mesma matriz de densidade se e somente se*

$$|\tilde{\psi}_i\rangle = \sum_{ij} u_{ij}|\tilde{\varphi}_i\rangle, \tag{2.90}$$

em que u_{ij} *é uma matriz unitária de números complexos com índices* i *e* j. *O menor dos conjuntos* $|\tilde{\psi}_i\rangle$ *e* $|\tilde{\varphi}_i\rangle$ *é completado com vetores nulos* $\mathbf{0}$, *de forma que os dois conjuntos possuam o mesmo número de elementos.*

Finalmente, imagine que um sistema quântico é preparado num estado ρ_i com probabilidade p_i. Será provado a seguir que tal sistema pode ser descrito por um operador de densidade $\rho = \sum_i p_i\rho_i$.

Para tanto, considere que ρ_i representa a família de estados $\{p_{ij}, |\psi_{ij}\rangle\}$, ou seja, $\rho_i = \sum_j p_{ij}|\psi_{ij}\rangle$ (note que i é fixo). Assim, a probabilidade do sistema se encontrar no estado $|\psi_{ij}\rangle$ é $p_i p_{ij}$. A matriz de densidade para o sistema é, portanto,

$$\begin{aligned} \rho &= \sum_{ij} p_i p_{ij} |\psi_{ij}\rangle\langle\psi_{ij}| \\ &= \sum_i p_i \rho_i. \end{aligned} \tag{2.91}$$

O estado ρ é dito ser uma mistura de estados ρ_i com probabilidades p_i.

Com isso, é possível enunciar os postulados da mecânica quântica em termos do formalismo de operadores de densidade.

2.4.2 Postulados quânticos e operadores de densidade

Os quatro Postulados enunciados na Seção 2.3 são reescritos em termos de operadores de densidade:

Postulado 1: Associado a um sistema quântico qualquer está um espaço vetorial complexo com produto interno (isto é, um espaço de Hilbert), chamado de espaço de estado do sistema. O sistema é completamente descrito pelo seu operador de densidade, que é um operador positivo ρ com traço unitário, atuando no espaço de estado do sistema. Se o sistema quântico está no estado ρ_i com probabilidade p_i, então o operador de densidade para o sistema é $\rho = \sum_i p_i \rho_i$.

Postulado 2: A evolução de um sistema quântico fechado é descrita por transformações unitárias. Isto é, o estado ρ_1 do sistema no tempo t_1 está relacionado ao estado ρ_2 do sistema no tempo t_2 por intermédio de um operador unitário U, que depende somente dos tempos t_1 e t_2,

$$\rho_2 = U \rho_1 U^\dagger. \tag{2.92}$$

Postulado 3: As medições em sistemas quânticos são definidas por um conjunto $\{M_m\}$ de operadores de medição. Os operadores M_m atuam no espaço de estado do sistema sujeito à medição. O índice m se refere aos possíveis resultados da medição. Se o estado do sistema imediatamente antes da medição é ρ, então a probabilidade de se obter m é dada por

$$p(m) = \mathrm{tr}\left(M_m^\dagger M_m \rho\right), \tag{2.93}$$

e o estado do sistema imediatamente após a medição será

$$\rho' = \frac{M_m \rho M_m^\dagger}{\text{tr}\left(M_m^\dagger M_m \rho\right)}. \tag{2.94}$$

Os operadores de medição satisfazem a equação de completude,

$$\sum_m M_m^\dagger M_m = I. \tag{2.95}$$

Postulado 4: O espaço de estado de um sistema quântico composto é o produto tensorial dos espaços de estado dos sistemas quânticos componentes. Ainda, se cada um dos n sistemas componentes é preparado no estado ρ_i, então o estado conjunto do sistema composto é $\rho_i \otimes \rho_2 \otimes \ldots \otimes \rho_n$.

A formulação dos postulados em termos de operadores de densidade é, evidentemente, matematicamente equivalente à descrição em termos de vetores de estado. Por exemplo, suponha que a evolução de um sistema quântico fechado é descrita pelo operador unitário U. Se o sistema i está inicialmente no estado $|\psi_i\rangle$ com probabilidade p_i, então, após a evolução ter ocorrido, o sistema irá estar no estado $U|\psi_i\rangle$ com probabilidade p_i. Assim, a evolução do operador de densidade é descrita pela equação

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \xrightarrow{U} \sum_i p_i U|\psi_i\rangle\langle\psi_i|U^\dagger = U\rho U^\dagger. \tag{2.96}$$

O Postulado 3 para vetores de estado, que estabelece as regras para medições em sistemas quânticos, possui uma redução matemática interessante à linguagem dos operadores de densidade. Suponha que o conjunto $\{M_m\}$ define, de forma mais geral, uma medição em algum sistema quântico. Se o estado inicial do sistema for $|\psi_i\rangle$, então a probabilidade de se obter m é

$$p(m|i) = \langle\psi_i|M_m^\dagger M_m|\psi_i\rangle = \text{tr}\left(M_m^\dagger M_m |\psi_i\rangle\langle\psi_i|\right), \tag{2.97}$$

em que a Equação (2.40) foi usada para calcular o traço do operador. Pela lei da probabilidade total, a probabilidade de obter o resultado m é

$$\begin{aligned} p(m) &= \sum_i p_i p(m|i) \\ &= \sum_i p_i \text{tr}\left(M_m^\dagger M_m |\psi_i\rangle\langle\psi_i|\right) \\ &= \text{tr}\left(M_m^\dagger M_m \rho\right). \end{aligned} \tag{2.98}$$

Para calcular o operador de densidade do sistema após se obter m na medição, observe que, se o estado inicial era $|\psi_i\rangle$, então o estado posterior à medição será

$$|\psi_i^m\rangle = \frac{M_m|\psi_i\rangle}{\sqrt{\langle\psi_i|M_m^\dagger M_m|\psi_i\rangle}}.$$

Note que, após a medição que resultou em m, o sistema se encontra em algum dos estados $|\psi_i^m\rangle$ com probabilidades $p(i|m)$. O operador de densidade correspondente é

$$\rho_m = \sum_i p(i|m)|\psi_i^m\rangle\langle\psi_i^m| = \sum_i p(i|m)\frac{M_m|\psi_i\rangle\langle\psi_i|M_m^\dagger}{\langle\psi_i|M_m^\dagger M_m|\psi_i\rangle}. \tag{2.99}$$

Pela regra de Bayes, $p(i|m) = p(m,i)/p(m) = p(m|i)p(i)/p(m)$. Substituindo as Equações (2.97) e (2.98) na Equação (2.99),

$$\begin{aligned}\rho_m &= \sum_i p_i \frac{M_m|\psi_i\rangle\langle\psi_i|M_m^\dagger}{\mathrm{tr}\left(M_m^\dagger M_m \rho\right)} \\ &= \frac{M_m \rho M_m^\dagger}{\mathrm{tr}\left(M_m^\dagger M_m \rho\right)}.\end{aligned} \tag{2.100}$$

Dado o operador densidade de um sistema composto AB, como obter o operador de densidade do sistema A? A resposta é o operador de densidade reduzido, discutido na próxima seção.

2.4.3 O operador de densidade reduzido

Suponha que um sistema quântico composto AB seja descrito por um operador de densidade ρ^{AB}. O operador de densidade para o sistema A é definido por

$$\rho^A \equiv \mathrm{tr}_B\left(\rho^{AB}\right), \tag{2.101}$$

em que $\mathrm{tr}_B(\cdot)$ é um mapeamento entre operadores chamado de traço parcial sobre o sistema B. O traço parcial é definido por

$$\mathrm{tr}_B\left(|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2|\right) \equiv |a_1\rangle\langle a_2|\mathrm{tr}\left(|b_1\rangle\langle b_2|\right), \tag{2.102}$$

em que $|a_1\rangle$ e $|a_2\rangle$ são dois vetores quaisquer no espaço de estado do sistema A, e $|b_1\rangle$ e $|b_2\rangle$ são dois vetores quaisquer no espaço de estado do sistema B. A operação de traço no lado direito da Equação (2.102) é o traço usual no sistema B, de forma

que $\mathrm{tr}\left(|b_1\rangle\langle b_2|\right) = \langle b_2|b_1\rangle$. O traço parcial dado pela Equação (2.102) foi definido para uma classe especial de operadores em AB. Para um operador arbitrário em AB, basta considerar que o traço parcial é linear com relação às suas entradas.

Não é óbvio que o operador de densidade reduzido do sistema A possa descrever completamente tal sistema. A justificação física para isto é que operador de densidade reduzido é o único dentre os mapeamentos que fornece estatísticas de medição corretas para medições feitas em A [24, pp. 107].

Como exemplos, considere primeiramente o caso em que o sistema composto é um estado produto $\rho^{AB} = \rho \otimes \sigma$, em que ρ é um operador de densidade para um sistema A e σ é um operador de densidade para um sistema B. Então,

$$\rho^A = \mathrm{tr}_B\left(\rho \otimes \sigma\right) = \rho \mathrm{tr}\left(\sigma\right) = \rho, \tag{2.103}$$

que é o resultado esperado. Considere agora o estado de Bell $\beta_{01} = (|01\rangle + |10\rangle)/\sqrt{2}$. O operador de densidade para este estado é

$$\begin{aligned} \rho &= \left(\frac{|01\rangle + |10\rangle}{\sqrt{2}}\right)\left(\frac{\langle 01| + \langle 10|}{\sqrt{2}}\right) \\ &= \frac{|01\rangle\langle 01| + |01\rangle\langle 10| + |10\rangle\langle 01| + |10\rangle\langle 10|}{2}. \end{aligned} \tag{2.104}$$

Aplicando o traço parcial no segundo qubit, tem-se o operador de densidade reduzido para o primeiro qubit,

$$\begin{aligned} \rho^1 &= \mathrm{tr}_2\left(\rho\right) \\ &= \frac{\mathrm{tr}_2\left(|01\rangle\langle 01|\right) + \mathrm{tr}_2\left(|01\rangle\langle 10|\right) + \mathrm{tr}_2\left(|10\rangle\langle 01|\right) + \mathrm{tr}_2\left(|10\rangle\langle 10|\right)}{2} \\ &= \frac{|0\rangle\langle 0|\langle 1|1\rangle + |0\rangle\langle 1|\langle 1|0\rangle + |1\rangle\langle 0|\langle 0|1\rangle + |1\rangle\langle 1|\langle 0|0\rangle}{2} \\ &= \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} \\ &= \frac{1}{2} I. \end{aligned} \tag{2.105}$$

Este resultado é interessante, pois o estado composto ρ é um estado puro, enquanto que o estado do primeiro qubit é um estado misto, pois $\mathrm{tr}\left((I/2)^2\right) = 1/2 < 1$. Isto é, o estado do primeiro qubit é tal que não se pode precisar com certeza em que estado ele se encontra. Note que ρ é um estado puro mas emaranhado, ou seja $\rho \neq |a\rangle \otimes |b\rangle$. Assim, a interpretação do operador de densidade reduzido é compatível com o conceito de emaranhamento, já que não se pode dizer que o estado do primeiro qubit é $|a\rangle$, mas que ele é um estado misto $\rho^1 = I/2$.

2.5 Conclusões

Neste capítulo foi apresentada uma introdução à mecânica quântica. O capítulo se iniciou mostrando alguns experimentos realizados principalmente em meados do século XX cujos resultados não poderiam ser explicados pela teoria clássica. Foi então que Planck deu o primeiro passo para explicar a radiação do corpo-negro, introduzindo sua famosa constante, chamada de quantum de ação. Foram citados os trabalhos de Einstein, Compton e de Broglie, que juntos mostraram o comportamento dual onda-partícula.

Na segunda parte do capítulo, a formalização da teoria quântica, na forma de seus postulados, foi apresentada. Antes porém foram apresentados alguns conceitos de álgebra linear e espaços de Hilbert, fundamentais para o entendimento da teoria quântica.

No próximo capítulo será feita uma revisão bibliográfica dos principais protocolos para autenticação quântica de mensagens, especialmente quando são usados para autenticar mensagens clássicas, que é o tema deste trabalho de dissertação.

Capítulo 3

Autenticação Quântica de Mensagens

3.1 Introdução

A autenticação é um procedimento para verificar que uma mensagem recebida foi enviada por uma determinada entidade, sem alteração no seu conteúdo. A criptografia clássica apresenta diversas técnicas para implementar autenticação. Os códigos de autenticação de mensagens (MAC), por exemplo, pressupõem a existência de uma chave secreta compartilhada entre as duas partes, A (Alice) e B (Bob). O algoritmo de codificação gera uma etiqueta (*tag*), chamada de bloco de autenticação, que é função da mensagem e da chave secreta. A etiqueta é então enviada junto com a mensagem. O algoritmo de decodificação, no receptor, gera outra etiqueta, função novamente da mensagem e da chave secreta. A etiqueta gerada é comparada com a etiqueta recebida. Ao final, o algoritmo retorna um bit indicando se a mensagem é autêntica ou não [31].

A descoberta e a formalização da mecânica quântica no século passado impulsionaram estudos nos campos da ciência da computação e da teoria da informação [24, 30, 5]. Efeitos como o emaranhamento e a descoberta dos pares EPR possibilitaram o teletransporte de estados quânticos [4]. Alguns problemas computacionalmente intratáveis no universo clássico, como a fatoração, são resolvidos por algoritmos de ordem polinomial em um computador quântico. O desenvolvimento de tal tecnologia inviabilizaria, por exemplo, os sistemas de criptografia por chave pública, cuja segurança é baseada na ineficiência dos algoritmos clássicos de fatoração de números grandes em produtos de primos [31]. Uma das aplicações mais interessantes da teoria da informação

quântica é a criptografia quântica. Em 1970, Wiesner mostrou que as propriedades da mecânica quântica poderiam ser usadas para tal fim, mas seu trabalho só foi publicado em 1983 [33]. No ano seguinte, Bennett *et al.* descreveram um protocolo para distribuição de chave secreta usando um canal quântico, que ficou conhecido como BB84 [3]. Existem várias provas de que o BB84 é incondicionalmente seguro[21, 20, 29], mesmo quando sujeito a ataques coletivos [6].

Até o final da década de 90, a expressão "criptografia quântica" se referia basicamente aos protocolos para distribuição de chave secreta (QKD) usando um canal quântico. Recentemente, várias pesquisas vêm sendo feitas no sentido de explorar as propriedades da mecânica quântica na resolução de outros problemas ligados à segurança de dados. Os primeiros trabalhos diz respeito à verificação de chave secreta [35] e à autenticação de usuários [15, 34, 19]. A verificação de chave consiste em garantir a legitimidade das duas partes que compõem um sistema de distribuição de chave, e que a mesma é autêntica. A autenticação de usuário, conhecida também como identificação de usuário, permite que um sistema determine a identidade do usuário que deseja usá-lo.

Somente em 2001, foi descrito o primeiro protocolo para autenticação quântica de mensagens [12]. Tratava-se de um protocolo que permitia o envio, de forma autêntica, de mensagens clássicas binárias de comprimento unitário (bit), e cuja segurança era garantida pelas leis da mecânica quântica. Em seguida, os mesmos autores propuseram um protocolo para autenticação de qubits [13], que foi uma extensão natural do primeiro. Por último, Barnum *et al.* [1] descreveram um protocolo que permitia a autenticação de mensagens quânticas de comprimento arbitrário.

No estudo da criptografia quântica, os dois participantes do sistema são chamados de Alice (A) e Bob (B). A terceira parte, que representa os criptoanalistas, é também chamada de Eva (do inglês *Eve*, que lembra *eavesdropper* - espião).

Neste capítulo, será dada uma visão geral dos protocolos de autenticação descritos na literatura, com ênfase para o protocolo de autenticação quântica de mensagens clássicas. Este capítulo está organizado como segue. Na Seção 3.2, o protocolo de Curty e Santos é detalhado, sendo feita uma análise de sua segurança e alguns comentários serão tecidos com relação aos recursos quânticos utilizados pelo mesmo. Em seguida, são descritos de forma sucinta os outros protocolos de autenticação de mensagens quânticas. Por último, são apresentadas as conclusões.

3.2 O protocolo de Curty e Santos para autenticação quântica de mensagens clássicas

O primeiro protocolo encontrado na literatura para a autenticação quântica de mensagens clássicas foi descrito por Curty e Santos [12]. Mais especificamente, os autores propuseram um protocolo capaz de autenticar mensagens clássicas de comprimento unitário, ou seja, um bit. Esta seção visa descrever e analisar, em detalhes, o funcionamento de tal protocolo.

Suponha que Alice deseja enviar para Bob, por meio de um canal quântico, uma mensagem clássica certificada. Supondo que a mensagem é binária de comprimento unitário, existirão somente duas mensagens possíveis, "0" e "1", que são associadas a dois estados quânticos $|\phi_0\rangle$ e $|\phi_1\rangle$, respectivamente. Para que Bob consiga distinguir perfeitamente entre esses dois estados, eles devem ser ortogonais, i.e., $\langle\phi_i|\phi_j\rangle = \delta_{ij}$, com $i, j \in \{0, 1\}$. Ainda, esses estados devem conter, como qualquer outro sistema de autenticação, alguma etiqueta (*tag*) que permita a Bob checar a autenticidade da mensagem. Os autores consideram que os estados $|\phi_i\rangle$ pertencem a um espaço de estado de dois qubits (um espaço de Hilbert de dimensão quatro) $\mathcal{H}$, em que o primeiro qubit transporta informação sobre a mensagem e o segundo sobre a etiqueta.

A chave secreta consiste de um estado quântico de dois qubits emaranhados, a serem compartilhados por Alice e Bob. Dessa forma, Alice guarda o primeiro qubit do estado $|\psi\rangle_{AB}$ e Bob o segundo qubit, em que

$$|\psi\rangle_{AB} = \frac{|01\rangle_{AB} - |10\rangle_{AB}}{\sqrt{2}}. \tag{3.1}$$

O procedimento de autenticação é descrito a seguir. Quando Alice deseja enviar um bit certificado i, ela prepara dois qubits no estado $|\phi_i\rangle$ (lembre-se que $|\phi_i\rangle$ já é um estado quântico de dois qubits) e aplica a operação de codificação

$$E_{A\mathcal{H}} = |0\rangle\langle 0|_A I_{\mathcal{H}} + |1\rangle\langle 1|_A U_{\mathcal{H}} \tag{3.2}$$

na sua parte de $|\psi\rangle_{AB}$ e na mensagem, em que $U_{\mathcal{H}}$ é um operador unitário de conhecimento público. Basicamente, o resultado desta operação de codificação é criar um estado de superposição, no qual o operador unitário $U_{\mathcal{H}}$ é aplicado no estado $|\phi_i\rangle$ [segundo termo da Equação (3.2)] ou não [primeiro termo da Equação (3.2)], dependendo do estado do qubit da chave de Alice.

Em virtude do estado quântico usado como chave secreta ser um estado emaranhado, não é possível escrever separadamente equações para os estados quânticos dos sistemas de Alice e Bob, mesmo após a operação de construção da etiqueta quântica. Entretanto, é possível escrever uma equação para o estado global do sistema $|G\rangle$ (chave de Alice + chave de Bob + mensagem). Após escolher por enviar a mensagem i, Alice cria o estado $|\phi_i\rangle$, e o estado global do sistema passa a ser $|G\rangle = |\psi\rangle_{AB}|\phi_i\rangle$. Observe que não há emaranhamento entre a chave e o recém-criado estado $|\phi_i\rangle$. Em seguida, Alice cria a etiqueta aplicando o operador unitário $E_{A\mathcal{H}}$ na sua parte do par EPR, que é o primeiro qubit de $|\psi\rangle_{AB}$, e na mensagem. O estado do sistema global passa a ser

$$\begin{aligned}
|G\rangle &\overset{(1)}{=} E_{A\mathcal{H}}\frac{1}{\sqrt{2}}(|01\rangle_{AB} - |10\rangle_{AB})\,|\phi_i\rangle \\
&\overset{(2)}{=} \frac{1}{\sqrt{2}}\{[(|0\rangle\langle 0|_A I_{\mathcal{H}} + |1\rangle\langle 1|_A U_{\mathcal{H}})|01\rangle_{AB}|\phi_i\rangle] - [(|0\rangle\langle 0|_A I_{\mathcal{H}} + |1\rangle\langle 1|_A U_{\mathcal{H}})|10\rangle_{AB}|\phi_i\rangle]\} \\
&\overset{(3)}{=} \frac{1}{\sqrt{2}}(|0\rangle\langle 0|0\rangle|1\rangle_{AB}|\phi_i\rangle + |1\rangle\langle 1|0\rangle|1\rangle_{AB}U_{\mathcal{H}}|\phi_i\rangle - |0\rangle\langle 0|1\rangle|0\rangle_{AB}|\phi_i\rangle \\
&\qquad -|1\rangle\langle 1|1\rangle|0\rangle_{AB}U_{\mathcal{H}}|\phi_i\rangle) \\
&\overset{(4)}{=} \frac{1}{\sqrt{2}}(|01\rangle_{AB}|\phi_i\rangle - |10\rangle_{AB}U_{\mathcal{H}}|\phi_i\rangle). \qquad (3.3)
\end{aligned}$$

As igualdades (3) e (4) são obtidas bastando observar que a operação unitária $E_{A\mathcal{H}}$ é somente aplicada ao primeiro qubit do par EPR e na mensagem, e que $\langle a|b\rangle = \delta_{ab}$, para $a, b \in \{0, 1\}$.

Após esse processo, Alice envia a mensagem e a etiqueta para Bob. Para obter a equação exata do estado quântico enviado, o operador de densidade referente ao estado global do sistema é calculado

$$\begin{aligned}
\rho &= |G\rangle\langle G| = \frac{1}{2}(|01\rangle_{AB}|\phi_i\rangle - |10\rangle_{AB}U_{\mathcal{H}}|\phi_i\rangle)(\langle 01|_{AB}\langle\phi_i| - \langle 10|_{AB}\langle\phi_i|U_{\mathcal{H}}^{\dagger}) \\
&= \frac{1}{2}(|01\rangle\langle 01|_{AB}|\phi_i\rangle\langle\phi_i| - |10\rangle\langle 01|_{AB}|\phi_i\rangle\langle\phi_i|U_{\mathcal{H}} - |01\rangle\langle 10|_{AB}|\phi_i\rangle\langle\phi_i|U_{\mathcal{H}}^{\dagger} \\
&\qquad + |10\rangle\langle 10|_{AB}U_{\mathcal{H}}|\phi_i\rangle\langle\phi_i|U_{\mathcal{H}}^{\dagger}). \qquad (3.4)
\end{aligned}$$

A mensagem autenticada que Alice envia a Bob é obtida de ρ através de um traço parcial sobre as variáveis de Alice e Bob, ou seja, sobre os qubits da chave secreta.

Assim,

$$\begin{aligned}\rho' &= \mathrm{tr}_{AB}(\rho) \\ &= \frac{1}{2}(|\phi_i\rangle\langle\phi_i|\mathrm{tr}(|01\rangle\langle 01|_{AB}) - |\phi_i\rangle\langle\phi_i|U_{\mathcal{H}}\mathrm{tr}(|10\rangle\langle 01|_{AB}) \\ &\quad - |\phi_i\rangle\langle\phi_i|U_{\mathcal{H}}^{\dagger}\mathrm{tr}(|01\rangle\langle 10|_{AB}) + U_{\mathcal{H}}|\phi_i\rangle\langle\phi_i|U_{\mathcal{H}}^{\dagger}\mathrm{tr}(|10\rangle\langle 10|_{AB})) \\ &= \frac{1}{2}(\rho_i + U_{\mathcal{H}}\rho_i U_{\mathcal{H}}^{\dagger}), \end{aligned} \tag{3.5}$$

em que $\rho_i = |\phi_i\rangle\langle\phi_i|$.

Ao receber a mensagem enviada através de um canal quântico perfeito e inseguro, Bob deve realizar um procedimento de verificação de autenticidade, baseado no estado quântico recebido e na chave secreta (sua parte do par EPR). Para isso, Bob aplica o operador unitário

$$D_{B\mathcal{H}} = |0\rangle\langle 0|_B U_{\mathcal{H}}^{\dagger} + |1\rangle\langle 1|_B I_{\mathcal{H}} \tag{3.6}$$

na sua parte de $|\psi\rangle_{AB}$ e na mensagem recebida. Finalmente, Bob realiza uma medição ortogonal (projetiva) no espaço $\mathcal{H}$. Desde que o espaço $\mathcal{H}$ possui dimensão quatro, e que os estados $|\phi_0\rangle$ e $|\phi_1\rangle$ são ortogonais, a medição é definida para um conjunto de projetores ortogonais

$$\mathcal{P} = \{P_i = |\phi_i\rangle\langle\phi_i|; i = 0, \ldots, 3\}, \tag{3.7}$$

em que $|\phi_2\rangle$ e $|\phi_3\rangle$ são vetores escolhidos usando o procedimento de Gram-Schmidt, de tal maneira que $\{|\phi_i\rangle\}$ forma uma base ortonormal para $\mathcal{H}$. Bob aceita que a mensagem é autêntica se o resultado da medição corresponder a um dos dois primeiros resultados (0 ou 1). Caso contrário, ele rejeita a mensagem.

Resumo do protocolo

O protocolo pode ser resumido como segue. Alice e Bob combinam publicamente em utilizar dois estados quânticos de dois qubits, $|\phi_0\rangle$ e $|\phi_1\rangle$, $\langle\phi_i|\phi_j\rangle = \delta_{ij}$, para representarem as mensagens clássicas 0 e 1, respectivamente. Para enviar a mensagem i, Alice e Bob seguem os passos abaixo:

1. Alice e Bob compartilham um estado quântico emaranhado $|\psi\rangle_{AB}$ [(Equação (3.1)], usado como chave secreta.

2. Alice cria dois qubits no estado $|\phi_i\rangle$. O primeiro qubit leva informação sobre a mensagem e o segundo sobre a etiqueta (*tag*);

3. Alice cria a etiqueta de verificação de autenticidade aplicando o operador unitário $E_{A\mathcal{H}}$ sobre a mensagem e sobre a sua parte do par EPR compartilhado;

4. Alice envia para Bob a mensagem, juntamente com a etiqueta;

5. Na recepção, Bob aplica o operador unitário $E_{B\mathcal{H}}$ na etiqueta recebida e na sua parte da chave secreta;

6. Bob realiza uma medição projetiva definida pelos operadores em $\mathcal{P}$ [(Equação (3.7)], e considera que a mensagem é autêntica se o resultado for 0 ou 1. Caso contrário, a mensagem é descartada.

3.2.1 Segurança

Os autores do trabalho analisaram a segurança do protocolo considerando alguns tipos de ataques e um canal de comunicação sem ruído para a transmissão das mensagens - canal quântico perfeito entre Alice, Eva e Bob. Os tipos de ataques foram:

Envio. Eva prepara um estado quântico não autêntico e o envia para Bob, tentando impersonalizar Alice.

Intercepta-Reenvia. Eva intercepta a mensagem enviada por Alice, aplica algum mapeamento e reenvia o resultado para Bob.

Medição. Eva intercepta a mensagem enviada por Alice e efetua medições na tentativa de obter alguma informação sobre a chave secreta.

Em linhas gerais, os autores mostraram que a segurança do protocolo depende somente da escolha do operador unitário $U_{\mathcal{H}}$. Como $U_{\mathcal{H}}$ é uma matriz de ordem quatro, é possível escrever

$$U_{\mathcal{H}} = \begin{bmatrix} M_0 & M_1 \\ M_2 & M_3 \end{bmatrix}, \tag{3.8}$$

em que M_i são matrizes complexas de ordem dois. Defina ainda $\overline{M}_i^j$ como sendo a linha j da submatriz M_i e M_i^j a coluna j da submatriz M_i de $U_{\mathcal{H}}$. Definida a probabilidade de falha do protocolo P_f e, para que se tenha $P_f < 1$, as seguintes restrições devem ser impostas ao operador unitário $U_{\mathcal{H}}$ escolhido por Alice e Bob:

1. Se $|\overline{M}_0^1 \overline{M}_0^{0\dagger}| = 0$, então $|\overline{M}_0^0|^2 < 1$ e $|\overline{M}_0^1|^2 < 1$.

2. Se $|\overline{M}_0^1 \overline{M}_0^{0\dagger}| = 0$, então

$$\frac{1}{2}x\left\{1+\left(\frac{x}{y}\right)\left[1+\left(\frac{x}{y}\right)^2\right]^{-1/2}\right\}+\frac{1}{2}y\left[1+\left(\frac{x}{y}\right)^2\right]^{1/2}+z<1, \tag{3.9}$$

em que $x = |\overline{M}_0^0|^2 - |\overline{M}_0^1|^2$, $y = 2|\overline{M}_0^1 \overline{M}_0^{0\dagger}|$, e $z = |\overline{M}_0^1|^2$.

3. $M_0^0 \neq e^{i\gamma} S(\delta)\sigma_x M_0^1$, ou $M_2^{0\dagger} M_2^1 \neq 0$ e $M_2^0 \neq e^{i\chi} M_2^1$, em que

$$S(\beta) = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\beta} \end{bmatrix} \tag{3.10}$$

é o operador de deslocamento de fase, σ_x é matriz de Pauli e γ, δ e χ são tais que $U_{\mathcal{H}}$ é um operador unitário.

4. $|M_0^0| > 0$ ou $|M_0^1| > 0$.

Para se ter uma idéia de como as restrições foram encontradas, considere o primeiro tipo de ataque, que é o de envio de mensagens.

Suponha que Eva prepara um estado puro normalizado $|\epsilon\rangle \in \mathcal{H}$ e o envia para Bob na tentativa de impersonalizar Alice. No caso mais geral, esta mensagem quântica não autêntica pode ser descrita como $|\epsilon\rangle = \sum_{i=0}^{3} e_i|\phi_i\rangle$. Quando Bob recebe esta mensagem, ele necessita aplicar o procedimento de verificação da etiqueta para ter certeza que a mensagem foi enviada por Alice. Assim, ele aplica a operação de decodificação e, em seguida, realiza uma medição definida pelo conjunto $\mathcal{P}$. Antes da medição, o estado da mensagem que Bob possui é descrito por $\rho'_E = \frac{1}{2}(\rho_E + U_{\mathcal{H}}^{\dagger}\rho_E U_{\mathcal{H}})$, em que $\rho_E = |\epsilon\rangle\langle\epsilon|$. Como mencionado, Bob rejeita a mensagem se o resultado da medição corresponde a um dos dois últimos projetores em $\mathcal{P}$; portanto, a probabilidade P_f de sucesso de Eva é

$$\begin{aligned} P_f &= \sum_{i=0}^{1} \langle\phi_i|\rho'_E|\phi_i\rangle \\ &= \frac{1}{2}\sum_{i=0}^{1}(|e_i|^2 + |\langle\epsilon|U_{\mathcal{H}}|\phi_i\rangle|^2). \end{aligned} \tag{3.11}$$

Esta quantidade depende da estratégia de Eva na criação do estado $|\epsilon\rangle$ e do operador unitário $U_{\mathcal{H}}$. O fato de que $|\epsilon\rangle$ possui norma unitária e $U_{\mathcal{H}}$ ser unitário, assegura que

ambos os dois termos do lado direito de (3.11) são menores ou iguais a 0, 5. O primeiro termo depende somente da criação de $|\epsilon\rangle$, e, para ser igual a 0,5, e_2 e e_3 devem ser nulos. A análise é feita considerando que Eva faz essa escolha. O segundo termo do lado direito da Equação (3.11) pode ser escrito como

$$\frac{1}{2}\sum_{i=0}^{1}|\langle\epsilon|U_{\mathcal{H}}|\phi_i\rangle|^2 = \frac{1}{2}\left[(|\overline{M}_0^0|^2 - |\overline{M}_0^1|^2)|e_0|^2 + 2|\overline{M}_0^1\overline{M}_0^{0\dagger}||e_0|\sqrt{1-|e_0|^2}\cos\theta_E + |\overline{M}_0^1|^2\right], \tag{3.12}$$

em que o ângulo θ_E depende somente da escolha de Eva quanto ao estado $|\epsilon\rangle$. O objetivo de Eva é tornar P_f tanto maior quanto possível. O pior caso para Alice e Bob é quando $\cos\theta_E = 2\pi k$ para k inteiro qualquer, e quando $|e_0|$ maximiza (3.12). A primeira e a segunda restrição decorrem diretamente desse fato:

1. Se $|\overline{M}_0^1\overline{M}_0^{0\dagger}| = 0$, então o máximo da Equação (3.12) é estritamente menor que 0,5 quando $|\overline{M}_0^0|^2 < 1$ e $|\overline{M}_0^1|^2 < 1$ (primeira restrição).

2. Se $|\overline{M}_0^1\overline{M}_0^{0\dagger}| \neq 0$, então o máximo da Equação (3.12) é estritamente menor que 0,5 quando a Equação (3.9) é satisfeita (segunda restrição).

3.2.2 Considerações

O protocolo descrito por Curty e Santos apresentou uma probabilidade de falha $P_f < 1$ quando a matriz acordada entre Alice e Bob satisfaz um conjunto de restrições. Com relação à implementação deste protocolo, algumas considerações são tecidas abaixo.

Como mencionado, a operação de criação da etiqueta por Alice, bem como a sua verificação por Bob, requer a aplicação de uma operação unitária genérica. Operadores unitários são implementados por meio de circuitos quânticos [24]. Atualmente, somente uma pequena classe de operadores unitários pode ser implementada usando tecnologia existente. A construção de circuitos capazes de implementar um operador unitário genérico, mesmo para um espaço de Hilbert de dimensão quatro, não é possível de acordo com o estado da tecnologia ou com tecnologias presentes no futuro próximo.

A chave secreta usada pelos participantes envolvidos é um par EPR maximamente emaranhado e compartilhado entre Alice e Bob. Embora já seja possível criar e distribuir estados emaranhados entre duas partes, o seu armazenamento por períodos

razoáveis de tempo ainda está distante dos horizontes tecnológicos [24]. Isto porque os aparatos atualmente disponíveis interagem fortemente com os estados quânticos, gerando decoerência. A construção de memórias quânticas confiáveis ainda é um desafio para os cientistas.

Ainda com relação à chave secreta, os autores não analisaram o impacto do reuso do par EPR no cálculo da probabilidade de falha do protocolo. Isto é importante, pois originalmente o sistema utiliza, para cada mensagem (bit) enviada de forma autêntica, um par EPR ou um bit clássico de chave secreta. Tal sistema está longe de ser prático, visto que desde de 1981 existem sistemas de autenticação capazes de transferir n mensagens autênticas usando chaves de comprimento $n \log(1/p)$, apresentando segurança incondicional com probabilidade de falha menor do que p, mesmo quando computadores quânticos com poder computacional ilimitado estão disponíveis para criptoanalistas [32].

Por último, os autores até hoje não descreveram uma classe ótima de operadores unitários ótimos que minimize a probabilidade de falha do protocolo para determinados estados quânticos representando as mensagens clássicas.

3.3 Outros protocolos

Em virtude de ser uma área de pesquisa relativamente recente, são poucos os trabalhos na área de autenticação quântica. Até a data da elaboração desta dissertação, somente três artigos descreviam protocolos diferentes para esse fim: o já descrito protocolo de Curty e Santos, uma generalização deste protocolo para autenticação de qubits [13] (mensagens quânticas representadas por vetores em um espaço de Hilbert de dimensão dois) e, por último, um protocolo para autenticação de mensagens quânticas de comprimento arbitrário [1].

Os protocolos destinados à autenticação de mensagens quânticas devem utilizar uma quantidade maior de recursos para serem implementados, visto que Alice e Bob manipulam, ao invés de um simples conjunto de estados quânticos ortogonais, estados quânticos genéricos a serem transmitidos de forma autêntica por meio de um canal quântico. Como o objetivo deste trabalho de dissertação é descrever um protocolo de autenticação quântica de mensagens clássicas que utilize o mínimo possível de recursos, os protocolos de autenticação de mensagens quânticas mencionados acima serão descritos de forma breve nas próximas subseções. Caso o leitor tenha interesse, são

disponibilizadas referências para cada novo conceito/recurso usado em cada protocolo.

3.3.1 Autenticação de um qubit

Em 2002, Curty, Santos *et al.* propuseram um protocolo para autenticação de mensagens quânticas [13]. O esquema é capaz de autenticar um único qubit descrito por um operador de densidade $\rho_{\mathcal{M}}$, pertencente a um espaço de mensagens $\mathcal{M}$ de dimensão dois. A diferença fundamental em relação ao caso clássico é que a etiqueta agora é dada por um operador densidade $\rho_{\mathcal{T}}$, de algum espaço de etiquetas $\mathcal{T}$. A mensagem e a etiqueta são descritas por um operador de densidade $\rho_{\mathcal{H}} = \rho_{\mathcal{M}} \otimes \rho_{\mathcal{T}}$, que atua no espaço de estado $\mathcal{H} = \mathcal{M} \otimes \mathcal{T}$. No caso em que o canal quântico é perfeito, $\mathcal{T}$ é o espaço de Hilbert de dimensão dois e $\rho_{\mathcal{T}} = |0\rangle\langle 0|_{\mathcal{T}}$.

Da mesma forma que no protocolo anterior, o par EPR $|\psi\rangle_{AB}$ é usado como chave secreta quando Alice deseja enviar para Bob uma mensagem $\rho_{\mathcal{M}}$. O estado do sistema global (chave, mensagem e etiqueta) é dado por

$$\rho_{AB\mathcal{H}} = |\psi\rangle\langle\psi|_{AB} \otimes \rho_{\mathcal{H}} = |\psi\rangle\langle\psi|_{AB} \otimes \rho_{\mathcal{M}} \otimes |0\rangle\langle 0|_{\mathcal{T}}. \tag{3.13}$$

Em seguida, Alice realiza uma operação de codificação, descrita pelo operador $E_{A\mathcal{H}}$, na sua parte da chave, na mensagem e na etiqueta, em que

$$E_{A\mathcal{H}} = |0\rangle\langle 0|_A \otimes I_{\mathcal{H}} + |1\rangle\langle 1|_A \otimes U_{\mathcal{H}} \tag{3.14}$$

e $U_{\mathcal{H}}$ é algum operador unitário em $\mathcal{H}$. Se $\rho^e_{AB\mathcal{H}}$ é o operador densidade que descreve o estado do sistema global após a codificação, i.e., $\rho^e_{AB\mathcal{H}} = E_{A\mathcal{H}}\rho_{AB\mathcal{H}}E^\dagger_{A\mathcal{H}}$, então o estado da mensagem que Alice envia a Bob através do canal quântico é dado por $\rho^e_{\mathcal{H}} = \mathrm{tr}_{AB}\left(\rho^e_{AB\mathcal{H}}\right)$, em que $\rho^e_{AB\mathcal{H}}$ pode ser escrito como

$$\begin{aligned}\rho^e_{AB\mathcal{H}} \;=\; & \frac{1}{2}(|01\rangle\langle 01|_{AB} \otimes \rho_{\mathcal{H}} + |10\rangle\langle 10|_{AB} \otimes U_{\mathcal{H}}\rho_{\mathcal{H}}U^\dagger_{\mathcal{H}} \\ & -|01\rangle\langle 10|_{AB} \otimes \rho_{\mathcal{H}}U^\dagger_{\mathcal{H}} - |10\rangle\langle 01|_{AB} \otimes U_{\mathcal{H}}\rho_{\mathcal{H}}).\end{aligned} \tag{3.15}$$

Assim,

$$\rho^e_{\mathcal{H}} = \frac{1}{2}(\rho_{\mathcal{H}} + U_{\mathcal{H}}\rho_{\mathcal{H}}U^\dagger_{\mathcal{H}}). \tag{3.16}$$

No lado da recepção, Bob decodifica a informação enviada por Alice aplicando a operação unitária

$$D_{B\mathcal{H}} = |0\rangle\langle 0|_B U^\dagger_{\mathcal{H}} + |1\rangle\langle 1|_B I_{\mathcal{H}} \tag{3.17}$$

ao seu qubit da chave secreta e à mensagem autenticada recebida, obtendo $\rho^d_{AB\mathcal{H}} = D_{B\mathcal{H}}\rho^e_{AB\mathcal{H}}D^\dagger_{B\mathcal{H}}$. Finalmente, Bob verifica o estado da etiqueta através de uma medição ortogonal $\{|0\rangle\langle 0|_\mathcal{T}, |1\rangle\langle 1|_\mathcal{T}\}$ sobre a porção da etiqueta de $\rho^d_\mathcal{H} = \mathrm{tr}_{AB}\left(\rho^d_{AB\mathcal{H}}\right)$, em que $|1\rangle_\mathcal{T}$ é o estado em $\mathcal{T}$ ortogonal a $|0\rangle_\mathcal{T}$. Se o resultado desta medição for $|0\rangle_\mathcal{T}$, Bob deve assumir que o estado quântico recebido é autêntico. Caso contrário, ele o descarta.

3.3.2 Autenticação de mensagens quânticas de comprimento arbitrário

Recentemente, Barnum *et al.* [1] descreveram um esquema não iterativo para autenticação de mensagens quânticas de comprimento m. O protocolo faz uso de códigos estabilizadores [18] para codificar a mensagem usando $m + O(s)$ qubits, em que a probabilidade de falha de autenticação decresce exponencialmente com o parâmetro de segurança s. O protocolo requer a criação de pares EPR por Alice e o envio de uma das metades a Bob. Foi demonstrado também que, para alcançar tal segurança, Alice e Bob devem compartilhar uma chave secreta clássica de comprimento maior ou igual a $2m$ bits, para cada mensagem quântica de comprimento m. Este protocolo requer circuitos quânticos para codificação e decodificação das mensagens.

Devido a quantidade e a complexidade de conceitos envolvidos, a descrição detalhada deste protocolo foge ao escopo deste trabalho.

3.4 Conclusões

Neste capítulo foram apresentados os protocolos presentes na literatura para autenticação quântica de mensagens. Como visto, o protocolo de Curty e Santos para autenticação quântica de mensagens clássicas exige a aplicação de um operador unitário genérico na criação e na verificação da etiqueta de autenticação. A implementação de tal operador unitário requer a construção de circuitos quânticos que não são passíveis de implementação usando o estado da arte da tecnologia.

No próximo capítulo, são descritos dois sistemas clássicos de autenticação que serão usados na composição do protocolo proposto nesta dissertação. Ainda, será apresentado um gerador de números pseudo-aleatórios baseado em exponenciação modular. De início, serão abordados os conceitos de segurança incondicional e computacional em criptografia e autenticação.

Capítulo 4

Sistemas Clássicos de Autenticação

4.1 Introdução

Um sistema de autenticação é um conjunto de procedimentos que permite o envio de mensagens através de um canal de comunicações não confiável, permitindo ao destinatário (Bob) identificar a identidade do remetente (Alice), garantido ainda que a mensagem não foi modificada por uma terceira parte (Eva).

A autenticação de mensagens pode ser feita usando dois métodos diferentes: a assinatura digital e os códigos de autenticação de mensagens, os chamados MAC (do inglês, *Message Authentication Codes*).

Os códigos de autenticação de mensagens provêem um método para assegurar a Bob que a mensagem foi enviada por Alice ou alguém autorizada por ela, e que a mensagem não foi alterada.

A assinatura digital consiste de uma seqüência de bits que é concatenada a mensagem. Somente Alice conhece a função de assinatura que é usada para gerar a seqüência. Entretanto, ela anuncia publicamente uma função que é usada para verificar a assinatura. Essa função permite que qualquer um teste se a assinatura é válida para uma mensagem em particular. Somente com o conhecimento da função de verificação deve ser difícil para Eva determinar uma assinatura válida para uma mensagem qualquer. Para todas as mensagens devem existir assinaturas válidas [14].

Diffie e Hellman [14] e Rivest *et al.* [26] apresentaram sistemas de autenticação por chave pública que provêem ainda um tipo de criptografia na mensagem.

Atualmente, diversos sistemas são utilizados para implementar autenticação de mensagens clássicas. Dentre eles, destaque para o *Message Digest Algorithm* (MD5),

o *Secure Hash Algorithm* (SHA-1), o HMAC e o algoritmo DSS. O livro de Stallings (1998) [31] apresenta uma discussão sobre esses e outros esquemas clássicos de autenticação de mensagens.

4.2 Segurança computacional e incondicional

Dois conceitos importantes no estudo da criptografia quântica, introduzidos inicialmente por Shannon [27], são os conceitos de esquemas de criptografia ou de autenticação que apresentam segurança computacional e os que apresentam segurança incondicional.

Definição 12 (Segurança computacional) *Um esquema de criptografia ou autenticação é dito possuir segurança computacional quando a segurança do sistema depende de limitações de tempo e de recursos computacionais, que inviabilizam a resolução de uma determinada classe de problemas.*

Os sistemas de criptografia e autenticação que apresentam segurança computacional são utilizados nas mais diversas aplicações. Os mais conhecidos são os sistemas de criptografia por chave pública, a exemplo dos algoritmos RSA, MD5 e, mais recentemente, os algoritmos SHA-1 e HMAC. Suas aplicações incluem transações financeiras, comerciais, militares, autenticação de usuários e sistemas, proteção de arquivos, transmissão segura de dados, etc. A segurança desses esquemas de criptografia e autenticação é baseada em problemas da teoria dos números que são atualmente intratáveis via algoritmos clássicos, como a fatoração de números em produto de primos, o problema do logaritmo discreto e o problema dos resíduos quadráticos. Esses problemas serão discutidos em mais detalhes ainda neste capítulo.

Com relação aos sistemas de criptografia e autenticação por chave pública, o seguinte resultado pode ser provado:

Teorema 4 ([32]) *Todo sistema de criptografia por chave pública apresenta segurança computacional. Isto é, desde que um criptoanalista disponha de recursos computacionais ilimitados, as mensagens podem ser forjadas.*

Definição 13 (Segurança incondicional) *Um esquema de criptografia ou autenticação é dito possuir segurança incondicional quando a segurança do sistema independe do poder computacional de criptoanalistas.*

O exemplo mais conhecido de sistemas que apresentam segurança incondicional é a cifra de Vernam. A cifra pode ser sucintamente descrita como

$$C_i = b_i \oplus k_i, \tag{4.1}$$

em que C_i representa o bit criptografado, b_i o bit em claro (a ser criptografado), $\oplus$ é a operação ou-exclusivo e k_i é o bit de chave secreta compartilhado entre as duas partes, que deve ser escolhido de forma aleatória, uniforme e independente do bit b_i. Além disso, para cada bit de mensagem b_i, um bit de chave k_i deve ser usado e descartado em seguida. Se uma mensagem binária de comprimento n é criptografada usando a cifra de Vernam, é fácil verificar que a probabilidade de um criptoanalista decifrá-la é 2^{-n}, independentemente do poder computacional que ele possua. O grande problema da cifra de Vernam é o tamanho da chave secreta em relação ao tamanho da mensagem.

4.3 Códigos de autenticação de mensagens

Um código de autenticação de mensagens, ou código MAC, é uma técnica de autenticação que envolve o uso de uma chave secreta para gerar um bloco pequeno de tamanho fixo, conhecido como etiqueta MAC, que é concatenado a mensagem a ser enviada por um canal inseguro. Na recepção, é gerada uma segunda etiqueta, que é função da mensagem recebida e da chave secreta. As duas etiquetas são comparadas e a mensagem é considerada autêntica se as etiquetas coincidirem. Caso contrário, a mensagem é descartada.

Existem diversos códigos MAC cada um apresentando um determinado nível de segurança [31]. É possível projetar um esquema MAC que alcance segurança incondicional, de tal forma que uma terceira parte maliciosa tenha uma probabilidade arbitrariamente pequena de criar uma mensagem que o destinatário aceitaria como autêntica.

Um código de autenticação de mensagens pode ser formalizado como segue [32]. Defina M como sendo o conjunto de mensagens possíveis e T como sendo o conjunto de etiquetas de autenticação possíveis. É definido também um conjunto de funções F, publicamente conhecido, em que cada função em F mapeia uma mensagem de M em uma etiqueta de T. Para usar o sistema, Alice e Bob devem compartilhar uma chave secreta que especifica uma função $f \in F$. Para transmitir uma determinada mensagem $m \in M$, Alice calcula a função f na mensagem m, e concatena o resultado a mensagem, transmitindo ambos em seguida. Na recepção, Bob aplica a mesma função f

na mensagem recebida e compara com a etiqueta recebida. Etiquetas idênticas indicam que a mensagem é autêntica. Neste esquema, a probabilidade de Eva encontrar a função f, conhecendo apenas a mensagem e a etiqueta, deve ser arbitrariamente pequena.

Wegman e Carter provaram que, dada uma chave secreta compartilhada entre Alice e Bob, qualquer esquema de autenticação incondicionalmente seguro pode ser usado somente um número finito de vezes, e tal número depende do tamanho da chave secreta e da probabilidade que um adversário qualquer possui de inferir a função de geração da etiqueta.

Definição 14 ([32]) *Um código de autenticação de mensagens é incondicionalmente seguro com probabilidade p se, dadas uma mensagem m e a etiqueta correspondente $f(m)$, a probabilidade de se encontrar uma mensagem diferente m' tal que $f(m') = f(m)$ é sempre menor do que p.*

O que a definição afirma é que a segurança do esquema de autenticação é condicionada somente a manutenção em segredo da chave secreta por Alice e Bob. Gilbert *et al.* [16] propuseram um código MAC considerando as premissas acima. A dificuldade de utilizar tal esquema reside no tamanho da chave secreta, que deveria ter comprimento mínimo igual ao tamanho da mensagem a ser enviada. Outro problema é que a chave secreta só poderia ser usada para criar uma única etiqueta, devendo ser descartada em seguida. Na Seção 4.3.2 será apresentado um código de autenticação de mensagens incondicionalmente seguro, proposto por Wegman e Carter em 1981, permitindo o reuso da chave secreta na autenticação de n mensagens, em que o número de bits da chave é proporcional ao logaritmo do tamanho da mensagem. Antes, porém, se faz necessário definir uma importante classe de funções bastante utilizadas em criptografia: as funções *hash*.

4.3.1 Funções *hash*

O conceito de funções *hash* em autenticação foi introduzido por Carter e Wegman [10]. As funções *hash* mapeiam domínios extensos em intervalos pequenos. Elas podem ser vistas como uma forma de atribuir uma abreviação para um nome. As funções *hash* desempenham um papel fundamental em criptografia e autenticação [31].

Definição 15 ([10]) *Uma classe H de funções hash $A \longrightarrow B$ é universal$_2$ se, para quaisquer $a_1, a_2 \in A$, $a_1 \neq a_2$, a probabilidade de se ter $f(a_1) = f(a_2)$ é, no máximo*

$1/|B|$, *quando* f *é escolhida do conjunto* H *de maneira aleatória e uniforme.*

Um exemplo de uma classe universal$_2$ é a classe de todas as funções lineares de $\{0,1\}^n$ para $\{0,1\}^r$, apresentadas por Carter e Wegman [10]. Estas funções podem ser descritas por matrizes M de dimensões $r \times n$ sobre $GF(2)$, ou seja, por rn bits. Outras classes de funções universais$_2$ que são mais econômicas em termos do número de bits necessários para especificá-las são discutidas pelos mesmos autores em dois trabalhos interessantes [32, 10]. A classe de funções universal$_2$ definida no Lema a seguir é baseada em $GF(2^n)$, que é um corpo algébrico formado a partir de um polinômio primitivo de grau n em $GF(2)[x]$.

Lemma 16 ([10]) *Seja* a *um elemento de* $GF(2^n)$. *Considere a função* $\{0,1\}^n \longrightarrow \{0,1\}^r$, *atribuindo a um argumento* $x \in GF(2^n)$ *os* r *primeiros bits do elemento* $ax \in GF(2^n)$. *A classe de todas as funções para* $a \in GF(2^n)$ *é uma classe universal*$_2$ *de funções para* $1 \leq r \leq n$.

A classe de funções descrita no Lema acima necessita de apenas n bits para especificar qualquer função que pertença a mesma.

Como visto, para ser universal$_2$, um conjunto de funções de A para B deve somente satisfazer a restrição na probabilidade de que uma escolha aleatória de uma função do conjunto mapeie dois pontos distintos em A no mesmo valor.

Definição 17 (Conjunto de funções *hash* fortemente universal$_n$ [32]) *Suponha que* H *é um conjunto de funções hash, em que cada elemento de* H *é uma função que mapeia um elemento de* A *em um elemento de* B. H *é fortemente universal*$_n$ *se, dados* n *elementos distintos de* A, $a_1, \ldots, a_n$, *e* n *elementos (não necessariamente distintos) de* B, $b_1, \ldots, b_n$, *então* $|H|/|B|^n$ *funções devem levar* a_1 *em* b_1, a_2 *em* b_2, *etc. Um conjunto de funções é fortemente universal*$_\omega$ *se ele é fortemente universal*$_n$ *para todos os valores de* n.

A definição afirma que um conjunto de funções fortemente universal$_n$ deve mapear, com igual probabilidade, quaisquer n elementos distintos de A em quaisquer n elementos de B. Em outras palavras, as funções do conjunto devem distribuir aleatoriamente quaisquer n elementos de A em B.

Os mesmos autores mostraram que a classe de funções universal$_2$ do Lema 16 é também uma classe de funções fortemente universal$_2$. De um modo geral, é possível criar classes de funções fortemente universal$_n$ usando polinômios definidos em algum

corpo finito. Em particular, considere que A e B são campos de Galois idênticos. Seja H um conjunto de polinômios de grau menor do que n. Então H é fortemente universal$_n$, desde que dados quaisquer n elementos distintos de A e elementos correspondentes de B, existe exatamente um polinômio de grau menor a n que "interpola" os pares designados. Isto pode ser provado usando o fato de que a inversibilidade da matriz de Vandermonde também se aplica a campos finitos.

Pode parecer peculiar definir um conjunto de funções *hash* com A e B possuindo o mesmo tamanho. Entretanto, isso pode ser facilmente contornado para fazer B menor, bastando para isso considerar os últimos bits do valor da função – no caso dos polinômios, os coeficientes dos termos de menor grau do polinômio resultante. Se o corpo finito tiver ordem que é potência de dois, o resultado será ainda uma classe fortemente universal$_n$ [32].

4.3.2 O esquema de Wegman e Carter e funções *hash*

Esta seção apresenta um código de autenticação de mensagens incondicionalmente seguro com probabilidade p, proposto por Wegman e Carter[32], e que segue a formulação descrita no início da Seção 4.3.

Inicialmente, escolha um conjunto de etiquetas T que possua pelo menos $1/p$ elementos. Seja F uma classe de funções *hash* fortemente univeral$_2$, mapeando elementos do conjunto de mensagens M em T. A definição de classe fortemente universal$_2$ implica que, se $m, m' \in M$, e para qualquer mensagem $m' \neq m$, a proporção de funções em F que mapeia m' em alguma etiqueta t', em particular, é $1/|T|$. Desde que $|T| \geq 1/p$, qualquer escolha de Eva possui sempre uma probabilidade menor ou igual a p de estar correta. A classe de funções utilizada na criação da etiqueta é descrita a seguir.

A classe de funções *hash* desejável para um código MAC deve mapear elementos de um conjunto extenso A', que é o conjunto de todas as mensagens possíveis, em elementos de um conjunto menor B', o conjunto das etiquetas de autenticação. Defina a' e b' como sendo o comprimento das mensagens e das etiquetas, respectivamente, e seja $s = b' + \log_2 \log_2(a')$. Agora considere H uma classe de funções fortemente universal$_2$ que mapeia seqüências de bits de comprimento $2s$ em seqüências de comprimento s[1]. Defina H' como sendo uma classe de funções do conjunto de mensagens A' para o conjunto das etiquetas B'. Cada função h' de H' será construída a partir de

[1]Tal classe pode ser facilmente implementada usando multiplicação de polinômios em $GF(2^{2s})$ [10].

uma seqüência de $\log_2 a' - \log_2 b'$ funções de H. Suponha que $h_1, h_2, \ldots, h_{\log_2 a' - \log_2 b'}$ é uma dessas seqüências. Para aplicar uma função h' de H' à mensagem, o seguinte procedimento é realizado. A mensagem é dividida em subseqüências de comprimento $2s$. Se necessário, a última subseqüência é completada com zeros. Assim, a mensagem irá ser dividida em $\lceil a'/2s \rceil$ subseqüências. Em seguida, h_1 é aplicada a todas as $\lceil a'/2s \rceil$ subseqüências e o resultado é concatenado para formar uma nova seqüência, cujo comprimento é aproximadamente igual a metade do comprimento da mensagem original. Este processo é repetido usando $h_2, h_3, \ldots$. O esquema multiplicativo sugerido em Carter e Wegman [10] utiliza uma chave de aproximadamente duas vezes o tamanho do argumento das funções. Desta forma, se a classe H for usada na composição das funções de H', o tamanho da chave secreta a ser usada para identificar uma função em H' será da ordem de $4s \log_2(a')$. Tal esquema representa uma redução brusca no tamanho da chave secreta, que seria da ordem de $2a'$ caso o esquema multiplicativo fosse usado diretamente na definição das funções de H'.

A classe de funções H' é, no sentido do Teorema abaixo, "quase" fortemente universal$_2$. Ao mesmo tempo, o Teorema garante que o código MAC é incondicionalmente seguro com probabilidade p.

Teorema 5 ([32]) *Dadas duas mensagens distintas, m_1 e m_2, e duas etiquetas quaisquer, t_1 e t_2, o número de funções que levam m_1 em t_1 é $1/|B'|$ vezes o número total de funções. Entretanto, menos do que $2/|B'|$ dessas funções levarão m_2 em t_2.*

Prova A cada vez que o tamanho das mensagens (m_1 e m_2) é dividido por dois (via a aplicação da seqüência de funções de H), existe uma probabilidade de $1/2^s$ de que as duas subseqüências resultantes sejam idênticas. Desde que esse processo é repetido $\log_2 a' - \log_2 b'$ vezes, a probabilidade de que as duas subseqüências sejam idênticas até o último passo é, no máximo, $\log_2 a'/2^s$, que é igual a $1/2^{b'}$. O fato de que a função que faz a última redução é escolhida a partir de uma classe fortemente universal$_2$ é usado para mostrar que m_1 irá ser levada em qualquer etiqueta com igual probabilidade. Como a penúltima subseqüência também é diferente, m_2 é levada em uma etiqueta qualquer com probabilidade $1/|B'|$. Assim, se $t_1 \neq t_2$, então $1/|B'|$ das funções irão levar m_2 em t_2 ou, caso contrário, menos de $2/|B'|$ o irão. ■

O teorema acima pode parecer contrastante com a definição de classe fortemente universal$_2$, que diz que $1/|B'|$ das funções deve levar m_1 em t_1, e que a mesma proporção também deve levar m_2 em t_2. Em se tratando de um código de autenticação, o teorema

afirma que se Eva conhece um par mensagem-etiqueta, então ela não consegue substituir a mensagem e calcular uma nova etiqueta válida com uma probabilidade maior do que $2/|B'|$. Assim, o sistema é seguro com probabilidade $2/|B'|$, e essa probabilidade pode ser feita tão pequena quanto desejada.

4.3.3 Autenticação de múltiplas mensagens

O código de autenticação de mensagens descrito acima não permite criar mais de uma etiqueta de autenticação usando a mesma função *hash*. Uma maneira de solucionar esse problema seria usar funções de uma classe universal$_n$, que permitiriam a autenticação de $n-1$ mensagens. Como a complexidade dessas classes é maior, Wegman e Carter, no mesmo trabalho, propuseram um método mais eficiente para autenticar múltiplas mensagens que é descrito abaixo.

Seja F uma classe de funções *hash* universal$_2$ de M para B, em que B é o conjunto das seqüências de bits de comprimento k. Cada mensagem em M deve possuir um número entre 1 e n. A chave secreta compartilhada por Alice e Bob consiste agora de duas partes. A primeira parte especifica uma função f em F. A segunda parte da chave é uma seqüência $(b_i, \ldots, b_n)$ de elementos de B. Alice deve assegurar que nunca irá enviar duas mensagens com o mesmo número. Para criar a etiqueta de autenticação t_i para a mensagem m_i, Alice primeiro calcula $f(m_i)$ e então realiza uma operação ou-exclusivo usando a seqüência b_i. A etiqueta de autenticação será portanto $t_i = f(m_i) \oplus b_i$. Desde que Bob conhece a seqüência b_i para a $i-$ésima mensagem enviada, ele repete a operação para obter o valor da função *hash* e assim proceder com o processo de autenticação.

Na realidade, o efeito da aplicação da seqüência b_i é o de mascarar o valor da função *hash*, de tal forma que Eva, ao ler a etiqueta, não obtenha qualquer informação acerca da função utilizada. O teorema seguinte mostra que o sistema descrito acima é seguro com probabilidade $1 - 1/2^k$.

Teorema 6 ([32]) *Suponha que uma chave secreta $(f, (b_i, \ldots, b_n))$ é escolhida aleatoriamente a partir do conjunto de chaves possíveis. Seja $m_i, \ldots, m_n$ mensagens quaisquer, com a restrição de que os índices das mensagens sejam todos diferentes. Suponha também que Eva conheça somente o conjunto F e as n mensagens com suas respectivas etiquetas, $t_i = f(m_i) \oplus b_i$. Então, não existe nenhuma mensagem, não importando o índice, para a qual Eva consiga inferir a etiqueta correta com probabilidade maior do*

que $1/2^k$.

Os autores mostraram que o teorema acima é válido mesmo quando uma classe de funções "quase" fortemente universal$_2$ é usada. Por último, será mostrado que o número de bits da chave secreta é assintoticamente ótimo, quando o número de mensagens n tende para infinito.

Suponha inicialmente que Eva escolhe uma função do conjunto F. Em seguida, ela escolhe uma mensagem m_i e tenta inferir o valor da etiqueta, que imagina ser t_i. O processo é repetido para todo $1 \leq i \leq n$.

Teorema 7 ([32]) *No cenário acima, se a probabilidade de sucesso de Eva na* $i-$*ésima tentativa for menor do que* p_i*, então* F *deve conter pelo menos* $1/(p_1 p_2 \ldots p_n)$ *funções.*

Prova Seja $F_0 = F$ e $F_k = \{f \in F | f(m_i) = t_i \ \forall \ i = 1, \ldots, k\}$. Eva deve usar a estratégia a seguir para inferir a etiqueta. Após escolher a $i-$ésima mensagem, ela enumera o conjunto F_{i-1}, escolhe aleatoriamente um membro deste conjunto e tenta inferir a etiqueta $f(m_i)$. Desde que Eva tem uma probabilidade de sucesso menor do que p_i, então $|\{f \in F_{i-1}; f(m_i) = t_i\}| \leq p_i |F_{i-1}|$. O conjunto do lado esquerdo é F_i, de forma que $|F_{i-1}| \geq (1/p_i)|F_i|$. Isto é válido para cada i, de tal forma que $|F_0| \geq (1/p_i)(1/p_2)\ldots(1/p_n)|F_n|$. O teorema é assim verificado, desde que $F_0 = F$ e $|F_n| \geq 1$. ■

Corolário 18 *No caso em que* $p_1 = p_2 = \ldots p_n = p$*, então são necessários pelo menos* $n \log_2(1/p)$ *bits para especificar um membro de* F *escolhido de forma aleatória, para que o esquema seja incondicionalmente seguro com probabilidade* p *e capaz de enviar* n *mensagens. Note que o código MAC apresentado nesta seção requer* $n \log_2(1/p) + K$ *bits, em que* K *é o número de bits necessários para especificar um elemento de uma classe de funções hash fortemente universal*$_2$.

A seção seguinte aborda a geração de números pseudo-aleatórios, que será usada mais adiante para compor um código de autenticação de mensagens baseado no esquema apresentado nesta seção.

4.4 Geração de números pseudo-aleatórios

A geração de números aleatórios desempenha um papel fundamental em praticamente todas as áreas da criptografia e segurança de dados em geral. Idealmente, é desejável

que um gerador de seqüências pseudo-aleatórias rapidamente produza seqüências longas de bits a partir de uma seqüência curta e pré-determinada, de forma que os bits gerados, de todas as formas, aparentem ter sido obtidos aleatóriamente.

Obviamente, a idéia de tal mecanismo determinístico gerando seqüências de comportamento aleatório parece ser contraditório: observando diversas de suas saídas deve ser possível, pelo menos em princípio, determinar seus parâmetros de forma a simular o gerador.

A solução usual é projetar o gerador de tal forma que as seqüências produzidas não sejam reprovadas por um conjunto de testes estatísticos. Esses testes incluem o cálculo da freqüência de zeros e uns, comprimento de surtos e distribuição dos bits na seqüência. Existem testes para determinar se uma seqüência segue uma determinada distribuição, incluindo a uniforme, mas não existe nenhum teste que prove a independência entre os bits da mesma. Ao contrário, existem testes que podem ser aplicados para demonstrar que a seqüência não possui independência [31].

Uma característica importante dos geradores de seqüências pseudo-aleatórias é a imprevisibilidade computacional, essencial para a maioria das aplicações em criptografia clássica.

Definição 19 (Imprevisibilidade em tempo polinomial [7]) *Um gerador de seqüência pseudo-aleatória é imprevisível em tempo polinomial (imprevisível à esquerda e imprevisível à direita), ou computacionalmente imprevisível, se e somente se as seqüências geradas não são previsíveis (à esquerda e à direita) em tempo polinomial. Isto é, dada uma seqüência finita produzida pelo gerador, $x_k, x_{k+1}, \ldots, x_{n-1}, x_n$, o melhor algoritmo que determina, em tempo polinomial, o termo anterior à seqüência, x_{k-1}, ou o termo posterior, x_{n+1}, consegue fazê-lo somente com uma probabilidade arbitrariamente pequena, do que se os termos fossem obtidos a partir de lançamentos sucessivos de uma moeda não viciada.*

A definição acima afirma que seqüências produzidas por geradores que satisfazem os requisitos da Definição 19 resistem a todos os testes estatísticos em tempo polinomial, ou seja, essas seqüências não podem ser distinguidas por nenhum teste estatístico em tempo polinomial, com uma probabilidade arbitrariamente pequena, de seqüências produzidas por lançamentos sucessivos de uma moeda não viciada.

A segurança dos geradores computacionalmente seguros é geralmente baseada em problemas computacionalmente intratáveis da teoria dos números. Na seção seguinte,

será apresentado um dos primeiros geradores descritos na literatura, bem como os fundamentos em que a sua segurança é baseada.

4.4.1 O gerador de Blum e Micali

O gerador de Blum-Micali [8] (BM) é um gerador simples, baseado na intratabilidade computacional (clássica) do problema do logaritmo discreto. Considere a equação

$$y = g^x \mod p, \tag{4.2}$$

em que p é um primo, g é um gerador g para $GF(p)$ e $x \in GF(p)$. Dados g, x e p, o cálculo de y é direto. Entretanto, dados y, g e p, o problema de encontrar x, que é o logaritmo discreto de y módulo p na base g, é considerado computacionalmente intratável. Acredita-se que a dificuldade de calcular o logaritmo discreto é da mesma ordem de magnitude da fatoração em produtos de primos. Formalmente, o problema do logaritmo discreto é resumido como segue:

O problema do logaritmo discreto [7]. Seja p um número primo,e seja g um gerador para o grupo $\mathbb{Z}_p^*$. A função $f_{g,p} : \mathbb{Z}_p^* \longrightarrow \mathbb{Z}_p^*$ definida por $f_{b,P} = b^x \mod P$ é uma permutação de $\mathbb{Z}_p^*$ calculada em tempo $O(|p|^3)$. O problema do logaritmo discreto com parâmetros g e p consiste em encontrar, para cada $y \in \mathbb{Z}_p^*$, o índice $x \in \mathbb{Z}_p^*$ tal que $g^x \mod p = y$. Um algoritmo probabilístico $\mathbf{P}[g,p,y]$ resolve o problema do logaritmo discreto se para todo primo p, para todos os geradores g de $\mathbb{Z}_p^*$ e para todo $y \in \mathbb{Z}_p^*$, $\mathbf{P}[g,p,y] = x$, $x \in \mathbb{Z}_p^*$, tal que $g^x \mod p = y$. O problema do logaritmo discreto é comumente chamado de problema de encontrar o índice.

O problema do logaritmo discreto é considerado intratável em tempo polinomial. Isto porque ainda não foi descrito um algoritmo clássico que resolva tal problema eficientemente. Formalmente,

Consideração acerca do problema do logaritmo discreto [7]: Afirma que existe uma fração fixa de tempo em que o problema do logaritmo discreto não pode ser resolvido eficientemente. Seja $\mathbf{P}[g,p,y]$ um procedimento probabilístico para resolver o problema do logaritmo discreto, que executado em um computador clássico. Seja $0 \leq c \leq 1$ uma constante fixa e *poly* um polinômio também fixo.

Então, para todo número n suficientemente grande e para todos, com exceção de uma fração ϵ, dos números primos p de n bits, para todos os geradores g de $\mathbb{Z}_p^*$ e para pelo menos uma fração ϵ de números $y \in \mathbb{Z}_p^*$, $\mathbf{P}[g, p, y]$ gasta um tempo (esperado) maior do que $poly(n)$ para calcular um elemento particular x tal que $g^x \mod p = y$.

Segundo Stallings (1998) [31], o algoritmo mas eficiente para calcular o logaritmo discreto módulo um primo p tem ordem

$$e^{\left((\ln p)^{(1/3)} \ln(\ln p)\right)^{2/3}}, \tag{4.3}$$

que é inviável para primos grandes.

Definição 20 (Resíduos quadráticos.) *Um número x é dito ser quadrado* mod n, *também chamado de resíduo quadrático, se e somente se existe $y^2 \equiv x \mod n$. O número y é uma raiz de x* mod n. *O conjunto de todos os quadrados* mod n *é chamado RQ_n.*

Note que se y é uma raiz de x, então $-y$ também o é, pois $(-y)^2 = (-1)^2 y^2 = x$.

A idéia de construção do gerador BM é bastante simples: a partir de uma semente são feitas sucessivas exponenciações, tendo os resultados intermediários usados como expoentes nas exponenciações seguintes. Para cada resultado intermediário, o gerador emite um bit para formar a seqüência pseudo-aleatória. Mais especificamente:

Definição do gerador Seja p um número primo tal que $p \equiv 3 \mod 4$.

1. Escolha g como sendo um gerador para o grupo $\mathbb{Z}_p^*$;
2. Escolha $x_0 \in RQ_p$.

O gerador é definido para

$$\text{Gerador BM} := (p, g, x_0). \tag{4.4}$$

Os valores de p e g podem ser publicamente conhecidos e usados seguidas vezes. x_0 é a semente, a chave secreta do gerador.

Geração dos bits. Para o $i-$ésimo bit b_i, começando com $i = 1$, seja

$$x_i \equiv g^{x_{i-1}} \mod p; \tag{4.5}$$

$$b_i = \delta_{x_i > (p-1)/2}. \tag{4.6}$$

A Equação (4.6) significa que $b_i = 1$ se e somente se $x_i > (p-1)/2$.

A prova de que o gerador BM é computacionalmente seguro, considerando a intratabilidade computacional do problema do logaritmo discreto, não será mostrada aqui. Entretanto, um passo fundamental da prova é mostrado abaixo, de extrema importância para o entendimento do Lema 23 no Capítulo 5, que é uma das contribuições deste trabalho.

Blum e Micali provaram a redução do problema de inferir um bit anterior ou posterior de uma dada seqüência do gerador a um problema conhecido como o bit central (tradução do inglês, *hard-core bit problem*). Além disso, foi mostrado que qualquer algoritmo probabilístico capaz de resolver o problema do bit central em tempo polinomial, também resolve o problema do logaritmo discreto eficientemente.

O problema do bit central consiste em encontrar o bit $b_i = \delta_{x_i > (p-1)/2}$ a partir do valor de $x_{i+1} = g^{x_i} \mod p$.

Antes, porém, são enunciados alguns resultados bem conhecidos da teoria dos números, dispostos nos dois lemas a seguir e cujas provas podem ser encontradas em textos sobre teoria dos números [25].

Lema 21 (Quadrados e raízes $\mod p$**. Critério de Euler.)** *As regras seguintes são válidas para quadrados e raízes módulo um primo p em $\mathbb{Z}_p^*$:*

(a). *Todo número $x \in RQ_p$ possui exatamente duas raízes* $\mod p$*, sendo da forma $\{y, -y\}$. A única exceção é* $1 \mod 2$*, que tem somente uma raiz,* $1 \equiv -1$.

(b). *Precisamente metade dos números* $\mod p$ *são resíduos quadráticos, i.e.,*

$$|RQ_p| = \frac{p-1}{2}. \tag{4.7}$$

(c). *Um algoritmo eficiente para testar se um número é um quadrado para $p > 2$ é dado pelo critério de Euler: para todo $x \in \mathbb{Z}_p^*$,*

$$x \in RQ_p \iff x^{(p-1)/2} = 1 \mod p. \tag{4.8}$$

(d). *Para* $x_{i+1} = g^{x_i} \mod p$ *tem-se*

$$x_{i+1} \in RQ_p \iff x_i \text{ é par.} \tag{4.9}$$

Lema 22 *Se $x \in RQ_p$ para* $p \equiv 3 \mod 4$*, então uma raiz quadrada de x pode ser facilmente calculada como*

$$w = x^{(p+1)/4} \mod p. \tag{4.10}$$

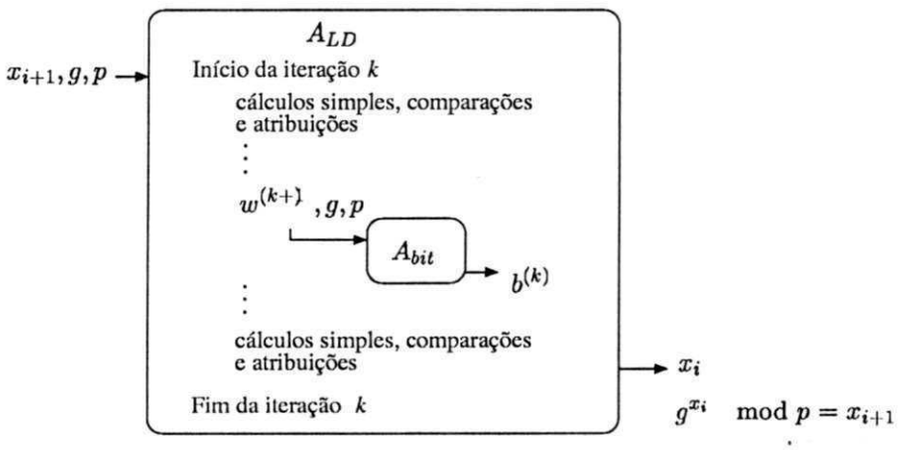


Figura 4.1: Visão geral da prova do teorema do bit central.

Teorema 8 (Problema do bit central para o logaritmo discreto [8].) *Considerando a intratabilidade computacional do problema do logaritmo discreto, não existe nenhum procedimento probabilístico em tempo polinomial, A_{bit}, capaz de resolver o problema do bit central.*

A prova do teorema é bem interessante, feita por contradição. É admitido que tal algoritmo A_{bit} existe e então é construído um algoritmo A_{LD} que resolve eficientemente o problema do logaritmo discreto.

A Figura 4.1 dá uma idéia da prova. As entradas para o algoritmo que calcula o logaritmo discreto são x_{i+1}, g e p, em que $x_{i+1} = g^{x_i} \mod p$ e x_i é o índice procurado, ou seja, a saída do algoritmo. Como será mostrado, o algoritmo A_{LD} realiza $l + 1$ iterações e, ao final de cada iteração, retorna um bit da representação binária de x_i. Ainda, para cada iteração, todos os cálculos antes e depois da chamada ao algoritmo A_{bit} são realizados de forma eficiente, de forma que se A_{bit} pode ser realizado eficientemente, também será o algoritmo A_{LD}. Isto contradiz a hipótese de intratabilidade computacional do cálculo do logaritmo discreto feita anteriormente.

Prova A prova do teorema é feita por contradição. Considere que o algoritmo A_{bit} contradiz o teorema, e então é construído um algoritmo A_{LD} capaz de resolver o problema do logaritmo discreto eficientemente. A idéia básica é usar o algoritmo A_{bit} como a principal sub-rotina do algoritmo A_{LD}, que faz, em cada iteração, uma chamada ao algoritmo A_{bit}. Em cada chamada, A_{bit} decide entre duas raízes. Dependendo desta escolha, e ao final da iteração, o algoritmo A_{LD} retorna um bit da representação

binária do logaritmo discreto desejado, ou seja, um bit de x_i com representação binária $(x_{i_l}, \ldots, x_{i_0})$. A construção é descrita a seguir.

Seja $x_{i+1}^{(0)} = x_{i+1}$, g e p as entradas para o algoritmo A_{LD} e seja x_i o logaritmo discreto procurado, com representação binária $(x_{i_l}, \ldots, x_{i_0})$.

1. Primeiro, A_{LD} calcula o bit menos significativo de x_i, x_{i_0}, i.e., ele decide se x_i é par ou ímpar. Pelo Lema 21(d), é necessário somente verificar se $x_{i+1}^{(0)}$ é um quadrado, e A_{LD} pode decidir isto eficientemente usando o Lema 21(c).

2. Para usar o procedimento 1 novamente e calcular o próximo bit, o algoritmo A_{LD} necessita deslocar x_i para a direita. Para isso, é necessário primeiro definir

$$x_{i+1}^{(0)'} = x_{i+1}^{(0)} / g^{x_{i_0}} \mod p. \tag{4.11}$$

   Dessa forma, $x_{i+1}^{(0)'} = g^{x_i'} \mod p$, em que os bits de x_i' são $(x_{i_l}, \ldots, x_{i_1}, 0)$.

3. O deslocamento à direita significa dividir o expoente por 2, o que equivale a calcular a raiz quadrada $x_{i+1}^{(0)''} = g^{x_i''} \mod p$ de $x_{i+1}^{(0)'}$, em que x_i'' possui uma representação binária dada pelos bits $(x_{i_l}, \ldots, x_{i_1})$. Para isso, A_{LD} usa o resultado do Lema 22. Entretanto, não é possível saber de antemão se o resultado w é $x_{i+1}^{(0)''}$ ou $-x_{i+1}^{(0)''}$.

   Neste passo, o algoritmo A_{bit} é usado, pois a raiz correta possui um expoente $x_i'' \leq (p-1)/2$ (isto porque $x_i'' \leq x_i/2$). É necessário mostrar que a outra raiz tem expoente $> (p-1)/2$. Isto porque

$$-x_{i+1}^{(0)''} = (-1)x_{i+1}^{(0)''} = g^{(p-1)/2} g^{x_i''} = g^{(p-1)/2 + x_i''}. \tag{4.12}$$

   Portanto, dada a raiz w, o algoritmo A_{bit} é acionado. Se a saída for igual a zero, é usado $x_{i+1}^{(0)''} = w$. Caso contrário, $x_{i+1}^{(0)''} = -w$.

O algoritmo A_{LD} segue fazendo $x_{i+1}^{(1)} = x_{i+1}^{(0)''}$, com logaritmo discreto dado por $(x_{i_l}, \ldots, x_{i_1})$. Os três primeiros passos acima são repetidos para encontrar x_{i_1} e, em seguida, removê-lo. Na j–ésima iteração, o valor de $x_{i+1}^{(j)}$ com logaritmo discreto dado por $(x_{i_l}, \ldots, x_{i_j})$, e a cada iteração o bit x_{i_j} é encontrado. Ao final, todos os bits de x_i estarão determinados. ■

4.5 Códigos de autenticação de mensagens computacionalmente seguros

Como visto na Seção 4.3.2 (página 62), o código de autenticação de mensagens proposto por Wegman e Carter permite a Bob identificar que uma mensagem recebida é autêntica com uma probabilidade de falha p que pode ser feita tão pequena quanto desejada. Tal segurança independe do poder computacional de Eva. Além disso, o protocolo permite o envio de n mensagens, fazendo uso de uma chave secreta de comprimento $n\log(1/p) + K$ bits, em que K é o número de bits necessários para identificar uma função específica de uma classe fortemente universal$_2$.

É possível reduzir consideravelmente o tamanho da chave secreta se, ao invés de fazer uso da segurança incondicional prevista pelo protocolo, Alice e Bob utilizarem um esquema cuja segurança é baseada na complexidade computacional de um problema qualquer. Isto quer dizer que Alice e Bob abdicariam da segurança incondicional em troca da redução drástica do tamanho da chave secreta a ser compartilhada.

Brassard [9] propôs uma modificação simples no protocolo de Wegman e Carter permitindo a autenticação de múltiplas mensagens usando uma chave secreta consideravelmente menor. Para isso, foi sugerida a substituição das seqüências $b_i, \ldots, b_n$ do protocolo original por um gerador de seqüências binárias pseudo-aleatórias criptograficamente seguro. O protocolo é descrito a seguir.

Seja P_f a probabilidade de falha aceitável que Alice e Bob estão dispostos a aceitar, considerando o esquema incondicionalmente seguro de Wegman e Carter. Defina k como sendo um inteiro maior do que $\log(1/P_f)$; seja M o espaço de mensagens e B o conjunto de todas as seqüências de bits de comprimento k, $B = \{0,1\}^k$. Seja também H um conjunto de funções fortemente universal$_2$ de M para B. Desta forma, Alice e Bob devem compartilhar uma chave secreta composta de duas partes: uma seqüência de bits que identifica uma função *hash* $h \in H$ e uma semente x_0 para o gerador pseudo-aleatório criptograficamente seguro. Para a $n-$ésima mensagem $m \in M$ trocada entre eles, a etiqueta de autenticação será

$$a(m,n) = h(m) \oplus x_0(n), \tag{4.13}$$

em que $x_0(n) = x_0[(n-1)k+1, \ldots, nk]$ denota k bits seqüenciais, a partir do $((n-1)k+1)-$ésimo bit gerado a partir da semente x_0. O efeito é o de prover pseudo-criptografia de uso único (*one-time-pad*), de forma a proteger a identificação da função *hash* usada.

A segurança do esquema acima depende da segurança do gerador de seqüências pseudo-aleatórias. Note que, desde que Eva não consiga distinguir a seqüência pseudo-aleatória de uma seqüência realmente aleatória, o protocolo de Brassard alcança o nível de segurança do protocolo de Wegman e Carter. No trabalho original, Brassard sugere o uso do gerador de Blum e Micali, por se tratar de um gerador criptograficamente seguro.

Em um cenário clássico, a extensão de Brassard seria uma alternativa interessante ao esquema de Wegman e Carter, especialmente no caso em que uma grande quantidade de mensagens deve ser trocada entre Alice e Bob. A segurança do esquema de Brassard é baseada na inexistência de um algoritmo clássico eficiente que resolva o problema do logaritmo discreto.

4.6 Conclusões

Este capítulo apresentou uma introdução aos sistemas clássicos de autenticação de mensagens, especialmente os que utilizam códigos de autenticação de mensagens (MAC). Inicialmente, foram definidos os conceitos de segurança computacional e incondicional. Em seguida, foi apresentado um código de autenticação de mensagens, proposto por Wegman e Carter em 1981, e que apresenta segurança incondicional. Tal esquema é baseado no uso de classes de funções *hash* fortemente universais$_2$ para criação de etiquetas de autenticação. Ele permite ainda a autenticação de múltiplas mensagens. Os autores mostraram que o tamanho da chave secreta utilizada para autenticar n mensagens com segurança incondicional tende a ser ótimo quando n tende para infinito.

Em seguida, foi apresentado um gerador de números pseudo-aleatórios proposto por Blum e Micali. Tal gerador é provado ser criptograficamente seguro. Isto significa que, dada uma seqüência de bits gerados a partir de uma semente (secreta) x_0, não existe um procedimento probabilístico clássico em tempo e recursos polinomiais capaz de predizer o bit anterior ou o bit posterior da seqüência. A segurança de tal gerador é baseada na intratabilidade computacional (clássica) do problema do logaritmo discreto.

Por último, a modificação de Brassard para o protocolo de Wegman e Carter foi discutida. O autor sugeriu a introdução do gerador pseudo-aleatório com o objetivo de diminuir o tamanho da chave utilizada no código MAC de Wegman e Carter. A redução da chave, neste caso, fez com que a segurança do protocolo original se reduzisse à segurança do gerador utilizado.

No próximo capítulo, será proposto um novo protocolo para autenticação quântica de mensagens clássicas. O esquema apresenta uma segurança incondicional, mesmo quando Eva dispõe de computadores quânticos e clássicos, sem restrição de tempo de processamento. Além disso, o protocolo proposto utiliza uma chave secreta significantemente menor do que os sistemas de autenticação quânticos apresentados no Capítulo 3, além de exigir recursos físicos que o torna passível de implementação de acordo com o estado da arte da tecnologia quântica.

Capítulo 5

Um Protocolo para Autenticação Quântica de Mensagens Clássicas

5.1 Introdução

O protocolo de Wegman e Carter apresentado no capítulo anterior demonstra segurança incondicional mesmo quando Eva possui recursos computacionais quânticos e clássicos infinitos. Isto porque a segurança do esquema é baseada no caráter unidirecional (do inglês, *one way*) das funções *hash* fortemente universais$_2$.

A extensão de Brassard, entretanto, apresenta somente segurança computacional. Quando o gerador de Blum e Micali é usado no esquema de Brassard, a descrição de um algoritmo clássico capaz de resolver eficientemente o problema do logaritmo discreto seria suficiente para que Eva pudesse forjar mensagens e enganar Bob.

A seção seguinte apresenta um breve comentário sobre algoritmo de Shor. Tal algoritmo deve ser executado em um computador quântico e tem como finalidade resolver, em tempo e recursos polinomiais, o problema de encontrar a ordem de um elemento em um corpo finito. O algoritmo de Shor funciona como uma sub-rotina para algoritmos que resolvem problemas da teoria dos números, que estão imersos em uma classe de problemas denominada de subgrupo escondido. Fazem parte desta classe a fatoração em produtos de primos, o problema do logaritmo discreto, o problema dos resíduos quadráticos, entre outros.

Em um cenário onde Eva dispõe de computadores quânticos, pode-se concluir que o esquema de Brassard pode ser facilmente quebrado.

5.2 O problema de encontrar a ordem e o problema do subgrupo escondido

Considere x e N inteiros positivos que não possuem fatores comuns, e tais que $x < N$. A *ordem* de x módulo N é definida como sendo o menor inteiro positivo r tal que $x^r \equiv 1 \mod N$. O problema de encontrar a ordem é o de determinar a ordem para dois inteiros x e N. Este problema é intratável em um computador clássico. Shor [28] propôs um algoritmo para resolver o problema de encontrar a ordem, que é executado em um computador quântico usando tempo e recursos polinomiais. Este algoritmo faz uso de um procedimento conhecido como *estimação de fase*, que consiste em estimar a fase φ de um autovalor $e^{2\pi i\varphi}$ associado com um autovetor particular $|u\rangle$ de um operador unitário U [24].

Shor também demonstrou que o *problema do subgrupo escondido* (HSP) para um grupo abeliano finito G também pode ser resolvido por um computador quântico usando um número de operações que é polinomial em $\log |G|$. Ainda, ele mostrou que o HSP pode ser reduzido ao problema de encontrar a ordem. O problema da fatoração, do logaritmo discreto, de encontrar o período, e diversos outros problemas da teoria dos números são instâncias do problema do subgrupo escondido [25].

O aspecto principal a ser destacado aqui é que, para qualquer um dos problemas a ser resolvido com um computador quântico é necessário ter, explicitamente, os valores dos parâmetros (argumentos) clássicos. Por exemplo, para resolver o problema do logaritmo discreto, faz-se necessário conhecer os inteiros a e $b = a^s$, que são usados para construir um operador unitário usado pelo algoritmo de busca da ordem.

5.3 Descrição do protocolo

O protocolo descrito aqui [22] é uma extensão do esquema proposto por Brassard. A idéia principal é usar as propriedades da mecânica quântica para "mascarar" a seqüência pseudo-aleatória de bits de forma que Eva, mesmo possuindo recursos computacionais clássicos e quânticos infinitos, não consiga distinguir a seqüência aleatória falsa de uma seqüência verdadeiramente aleatória. Na prática, isto quer dizer que as propriedades intrínsecas da mecânica quântica aumentam a segurança computacional do protocolo de Brassard para uma segurança incondicional do protocolo de Wegman e Carter, com a vantagem de usar uma chave secreta consideravelmente menor, princi-

palmente no caso em que um grande número de mensagens deve ser enviado.

Considere P_f, k, M, B e x_0 como definidos na Seção 4.5. O protocolo de Brassard faz uso de uma chave secreta composta de duas partes, uma função *hash* particular $h \in H$ e uma semente x_0 para o gerador de Blum e Micali. O protocolo descrito aqui faz uso de outra chave secreta y_0, que é uma semente para o mesmo gerador de seqüências pseudo-aleatórias. Quando Alice deseja enviar uma mensagem certificada para Bob, ela realiza todos os passos descritos pelo protocolo de Brassard. Para a $n-$ésima mensagem $m \in M$ a ser enviada, Alice prepara uma etiqueta $a(m,n)$ de k bits, dada pela Equação (4.13). Em seguida, vem a parte quântica do protocolo.

Considere que Alice e Bob combinem de usar duas bases ortonormais para o espaço de Hilbert de dimensão dois,

$$\mathcal{Z} = \{|0\rangle, |1\rangle\} \tag{5.1}$$

$$\mathcal{X} = \{|+\rangle, |-\rangle\}, \tag{5.2}$$

em que

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \qquad \text{e} \qquad |-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \tag{5.3}$$

Estas são as mesmas bases ortonormais usadas para criar os quatro estados quânticos no protocolo BB84 [3]. Para cada bit de $a(m,n)$, Alice prepara um estado quântico não emaranhado $|\psi_{n_j}\rangle$, que é baseado no bit correspondente emitido pelo gerador BM com semente y_0. Dessa forma, se o $j-$ésimo bit de $y_0(n)$ é zero, Alice prepara $|\psi_{n_j}\rangle$ usando a base $\mathcal{Z}$ da seguinte maneira:

$$|\psi_{n_j}\rangle = \begin{cases} |0\rangle & \text{se o } j-\text{ésimo bit de } a(m,n) \text{ é } 0 \\ |1\rangle & \text{se o } j-\text{ésimo bit de } a(m,n) \text{ é } 1. \end{cases} \tag{5.4}$$

Analogamente, se o $j-$ésimo bit de $y_0(n)$ é 1, Alice prepara $|\psi_{n_j}\rangle$ usando a base $\mathcal{X}$, em que

$$|\psi_{n_j}\rangle = \begin{cases} |+\rangle & \text{se o } j-\text{ésimo bit de } a(m,n) \text{ é } 0 \\ |-\rangle & \text{se o } j-\text{ésimo bit de } a(m,n) \text{ é } 1. \end{cases} \tag{5.5}$$

Após a geração dos qubits, Alice envia o estado $|\psi_{n_j}\rangle^{\otimes k}$ para Bob usando um canal quântico sem ruído. A mensagem m pode ser enviada usando um canal inseguro, seja ele clássico ou quântico.

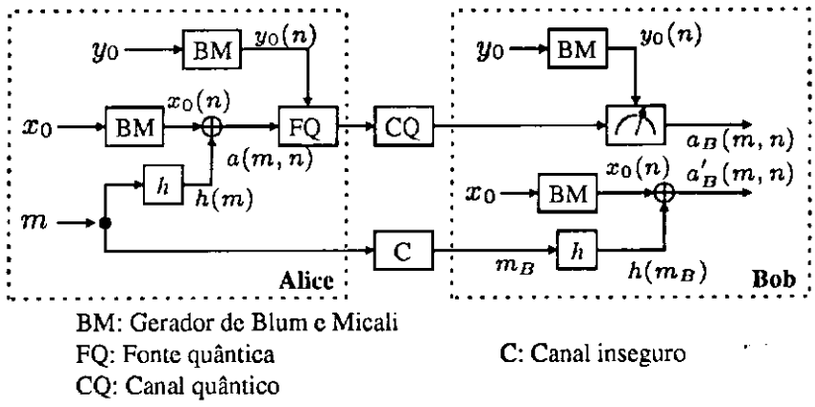


Figura 5.1: Diagrama de blocos para o protocolo proposto.

Na recepção, Bob realiza medições POVMs nas bases $\mathcal{Z}$ e $\mathcal{X}$, definidas pelos seguintes conjuntos

$$E_Z = \{E_0 = |0\rangle\langle 0|, E_1 = |1\rangle\langle 1|\} \tag{5.6}$$

e

$$E_{\mathcal{X}} = \{E_+ = |+\rangle\langle +|, E_- = |-\rangle\langle -|\}. \tag{5.7}$$

Para o $j-$ésimo qubit $|\psi_{n_j}\rangle$ recebido, Bob realiza uma medição usando o conjunto E_Z ou $E_{\mathcal{X}}$, dependendo se o $j-$ésimo bit de $y_0(n)$ é 0 ou 1, respectivamente. Dessa forma, Bob considera que o $j-$ésimo bit da etiqueta de Alice é 0 quando ele obtém as saídas E_0 ou E_+. Se a saída for E_1 ou E_-, Bob considera que o bit j de $a(m,n)$ é 1. Ao final de k medições, Bob dispõe de uma seqüência $a_B(m,n)$ de k bits clássicos. Além disso, Bob dispõe também da mensagem recebida, denotada por m_B, que pode estar modificada ou não.

O próximo passo de Bob é calcular uma etiqueta local baseado na função *hash* e na seqüência gerada pela semente x_0, obtendo $a'_B(m,n) = h(m_B) \oplus x_0(n)$. Como o canal quântico é perfeito, Bob considera que a mensagem é autêntica se $a_B(m,n) = a'_B(m,n)$. Caso contrário, ele descarta a mensagem recebida.

A afirmação acima pode ser feita porque no caso em que Eva não interfere na transmissão quântica da etiqueta, Bob obterá na medição a mesma etiqueta enviada por Alice, ou seja, $a_B(m,n) = a(m,n)$. Isto porque o gerador cuja semente é y_0 indica em qual das bases Alice deve criar os qubits, ao mesmo tempo em que diz a Bob qual

dos conjuntos POVMs ele deve escolher para medi-los. Se a medição é feita sempre na mesma base em que os qubits são criados, Bob sempre interpreta corretamente o bit enviado por Alice. A Figura 5.1 resume o código de autenticação de mensagens proposto acima. Realizações físicas dos diversos dispositivos quânticos, como fontes quânticas e medidores, são detalhadas, por exemplo, no paper de Zbinder *et. al.* [17].

Para um melhor entendimento do esquema de autenticação proposto, o exemplo hipotético a seguir mostra a aplicação do protocolo para o caso em que o envio da mensagem é feito sem a interferência de Eva.

5.3.1 Exemplo do uso do protocolo

Considere que Alice e Bob compartilham uma chave secreta composta por

1. uma função $h \in H$, em que $H : M \rightarrow B$, e
2. um par de sementes x_0 e y_0 para um gerador de Blum e Micali.

Considere ainda que

- as mensagens são de comprimento 32, e as etiquetas de comprimento 8;
- Alice deseja enviar a mensagem $n = 1$ dada por $m = DA\ \ 18\ \ FF\ \ 09$ (em hexadecimal);
- a função h mapeia (secretamente) $h(m) = 1010\ \ 1100 = AC$;
- os geradores BM forneçam $x_0(1) = 0011\ \ 1010$ e $y_0(1) = 1011\ \ 0011$.

Lembre-se que tanto Alice quanto Bob são capazes de gerar as seqüências $x_0(1)$ e $y_0(1)$. Inicialmente, Alice cria a etiqueta, fazendo uma operação ou-exclusivo entre $h(m)$ e $x_0(1)$:

$$a(m,1) = (1010\ \ 1100) \oplus (0011\ \ 1010) = 1001\ \ 0110. \tag{5.8}$$

Alice então usa os bits de $a(m,1)$, em conjunto com os bits de $y_0(1)$, para criar os qubits a serem enviados pelo canal quântico:

$$
\begin{array}{lcccccccc}
y_0(1): & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\
 & \Downarrow & \Downarrow & \Downarrow & \Downarrow & \Downarrow & \Downarrow & \Downarrow & \Downarrow \\
Bases: & \mathcal{X} & \mathcal{Z} & \mathcal{X} & \mathcal{X} & \mathcal{Z} & \mathcal{Z} & \mathcal{X} & \mathcal{X} \\
a(m,1): & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\
|\psi_{1_j}\rangle^{\otimes 8}: & |-\rangle & |0\rangle & |+\rangle & |-\rangle & |0\rangle & |1\rangle & |-\rangle & |+\rangle
\end{array}
\tag{5.9}
$$

Alice envia o estado produto $|\psi_{1_j}\rangle^{\otimes 8}$ por um canal quântico perfeito. Note que a suposição é que o canal seja perfeito, mas não seguro. A mensagem m é enviada por um canal quântico ou clássico não confiável.

Na recepção, Bob usa a seqüência $y_0(1)$ para decidir em que base ele deve efetuar a medição de cada qubit:

$$
\begin{array}{lcccccccc}
y_0(1): & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\
 & \Downarrow & \Downarrow & \Downarrow & \Downarrow & \Downarrow & \Downarrow & \Downarrow & \Downarrow \\
POVM: & E_{\mathcal{X}} & E_{\mathcal{Z}} & E_{\mathcal{X}} & E_{\mathcal{X}} & E_{\mathcal{Z}} & E_{\mathcal{Z}} & E_{\mathcal{X}} & E_{\mathcal{X}} \\
|\psi_{1_j}\rangle^{\otimes 8}: & |-\rangle & |0\rangle & |+\rangle & |-\rangle & |0\rangle & |1\rangle & |-\rangle & |+\rangle
\end{array}
\tag{5.10}
$$

Um cálculo simples mostra que, neste caso, Bob obterá na medição a seqüência binária $a_B(m,n) = 1001\ 0110$. Para ver isso, considere as medições do primeiro e do segundo qubit, $|\psi_{1_1}\rangle$ e $|\psi_{1_2}\rangle$, respectivamente.

Na medição do primeiro qubit, usando o POVM $E_{\mathcal{X}}$, Bob tem as seguintes probabilidades de medição:

$$
\begin{aligned}
p(E_+) &= \mathrm{tr}\,(E_+|-\rangle\langle -|) \\
&= \mathrm{tr}\left(\frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\right) \\
&= 0.
\end{aligned}
\tag{5.11}
$$

Conseqüentemente,

$$
\begin{aligned}
p(E_-) &= \mathrm{tr}\,(E_-|-\rangle\langle -|) \\
&= \mathrm{tr}\left(\frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\right) \\
&= 1.
\end{aligned}
\tag{5.12}
$$

Isto significa que Bob irá obter E_- na medição com probabilidade máxima. Logo ele interpreta que o primeiro bit da etiqueta enviada por Alice é 1.

Analogamente, para o segundo qubit $|\psi_{1_2}\rangle$ e usando a base $E_{\mathcal{Z}}$, Bob terá as seguintes probabilidades de leitura:

$$\begin{aligned} p(E_0) &= \mathrm{tr}\,(E_0|0\rangle\langle 0|) \\ &= \mathrm{tr}\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) \\ &= 1 \end{aligned} \tag{5.13}$$

e

$$\begin{aligned} p(E_1) &= \mathrm{tr}\,(E_1|0\rangle\langle 0|) \\ &= \mathrm{tr}\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) \\ &= 0. \end{aligned} \tag{5.14}$$

Pela sistemática adotada, Bob conclui que o segundo bit da etiqueta de Alice é 0. Após 8 medições, Bob obtém a seqüência $a_B(m,1) = 1001\ \ 0110$.

O próximo passo é criar uma etiqueta local $a'_B(m,1)$, baseada na mensagem recebida m_B e na seqüência $x_0(1)$, que é gerada localmente. Considerando que a mensagem não foi modificada, tem-se que

$$a'_B(m,1) = h(m_B) \oplus x_0(1) = 1001\ \ 0110. \tag{5.15}$$

Por último, Bob testa se a etiqueta enviada por Alice pelo canal quântico, $a_B(m,1)$, é igual a etiqueta gerada a partir da mensagem m_B recebida, $a'_B(m,1)$. Neste caso, tem-se que $a_B(m,1) = a'_B(m,1)$ e Bob conclui que a mensagem é autêntica.

Se Alice e Bob não dispõem de um canal clássico, uma alternativa seria escolher uma base, por exemplo $\mathcal{Z}$, que Alice usaria para criar os estados quânticos. Dessa forma, Alice criaria o estado $|m\rangle$ e o enviaria a Bob, que utilizaria o conjunto $E_{\mathcal{Z}}$ para efetuar a medição de $|m\rangle$.

5.4 Análise da segurança

A análise da segurança do protocolo descrito na seção anterior será feita considerando criptoanalistas com recursos computacionais infinitos, tanto clássicos como quânticos. Os argumentos usados aqui para provar a segurança incondicional do protocolo são

baseados nos algoritmos quânticos para a resolução de problemas da teoria dos números e na estratégia adotada na criação dos qubits da etiqueta.

O protocolo de Wegman e Carter faz uso de seqüências de bits verdadeiramente aleatórias, $b_1, \ldots, b_n$, com o objetivo de criptografar o valor da função *hash*. Com isso, ele garante que Eva não consegue obter nenhuma informação sobre a função *hash* utilizada. Este protocolo apresenta segurança incondicional, mesmo quando Eva dispõe de um computador quântico com recursos ilimitados.

A segurança do protocolo proposto reside na segurança do gerador de seqüências pseudo-aleatórias. Será provado nesta seção que, de acordo com a metodologia de criação dos qubits, a probabilidade de Eva distinguir a seqüência pseudo-aleatória de uma seqüência realmente aleatória nunca é maior do que um valor que decresce exponencialmente com o tamanho da semente do gerador em questão, mesmo que Eva possua um poder computacional quântico infinito. Isto sugere que a segurança do protocolo proposto se reduz à segurança incondicional do protocolo de Wegman e Carter.

Relembrando, Alice e Bob compartilham uma chave secreta que é composta de três partes: uma função *hash*, que é escolhida de um conjunto fortemente universal$_2$, e duas sementes x_0 e y_0 para os geradores de seqüências pseudo-aleatórias. Além disso, Alice sempre envia para Bob, via um canal quântico perfeito, um dos quatro estados quânticos: $|0\rangle$, $|1\rangle$, $|+\rangle$ e $|-\rangle$.

5.4.1 Processamento da informação quântica

Será analisado o caso em que Eva intercepta a transmissão quântica, possivelmente armazena os estados quânticos enviados por Alice e tenta processá-los de alguma maneira usando um computador quântico, na esperança de obter alguma informação sobre a chave secreta ou sobre a etiqueta.

Os estados que Eva venha a armazenar e processar pertencem a um conjunto $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$ em que nem todos os estados quânticos são ortogonais entre si. Como visto no Capítulo 2, não é possível distinguir perfeitamente entre esses estados. Por outro lado, Eva não possui nenhum conhecimento a priori sobre a função *hash* usada por Alice e Bob. Pode-se concluir, portanto, que o processamento da informação quântica não ajuda Eva na sua tarefa de burlar a segurança do protocolo. Dessa forma, Eva deve realizar medições que lhes permitirão usar seus recursos computacionais infinitos,

tanto clássicos como quânticos.

5.4.2 Ataque de medição

A essa altura está claro que a segurança do esquema proposto depende da previsibilidade do gerador de seqüências pseudo-aleatórias de Blum e Micali, quando usado em um cenário quântico. É necessário investigar, portanto, como Eva pode fazer uso dos recursos computacionais disponíveis para predizer tal gerador.

Primeiramente, é provado o seguinte resultado acerca do gerador de Blum e Micali com parâmetros p, g e x_0, em que p é um número primo ímpar tal que $p \equiv 3 \mod 4$, g é um elemento primitivo de $\mathbb{Z}_p^*$ e x_0 é uma semente secreta escolhida de RQ_p, com representação binária de l bits.

Lema 23 *Seja $d_{i+1} \ldots d_{i+s}$ uma parte de seqüência contendo s bits consecutivos produzidos pelo gerador de Blum e Micali especificado acima. O melhor algoritmo probabilístico $A_{BM}(g, p, d_{i+1}, \ldots, d_{i+s})$ capaz de predizer a seqüência completa para traz (e para frente) necessita de pelo menos $s = l$ bits para que*

$$Prob[A_{BM}(g, p, d_{i+1}, \ldots, d_{i+s}) = d_i] = 1. \tag{5.16}$$

Prova O lema é provado observando que este problema é equivalente ao problema de calcular o bit central (ver Seção 4.4.1, página 67), sendo que o último pode ser reduzido ao problema do logaritmo discreto.

Seja então $A_{LD}(x_{i+1}, g, p)$ um algoritmo executado em um computador quântico capaz de calcular o logaritmo discreto x_i, em que $d_i = \delta_{x_i > (p-1)/2}$.

O resultado segue por contradição. Suponha que tal algoritmo $A_{BM}(\cdot, \cdot)$ exista. Então, deve existir uma função $f(d_{i+1}, \ldots, d_{i+s})$ tal que

$$x_{i+1} = f(d_{i+1}, \ldots, d_{i+s}), \qquad s < l, \tag{5.17}$$

e

$$x_i = A_{LD}(x_{i+1}, g, p), \tag{5.18}$$

$$d_i = \delta_{x_i > (p-1)/2}. \tag{5.19}$$

Tal função não pode existir, pois a cardinalidade do seu domínio, (2^s), é menor do que a cardinalidade do seu contradomínio, (2^l). ■

Eva não possui inicialmente nenhuma informação sobre a composição da chave x_0, y_0 e $h \in H$, compartilhada entre Alice e Bob. A chave x_0 está associada com a criação da etiqueta, $a(m,n) = h(m) \oplus x_0(n)$, e $y_0(n)$ com a escolha das bases.

Para iniciar sua estratégia de ataque de medição, Eva tem que escolher as bases que serão usadas para ler os qubits. Eva tem duas opções: ou escolhe uma semente y_{E_0} e utiliza um gerador de BM para indicar as bases, ou escolhe aleatoriamente cada base usada em cada uma das medições. A probabilidade de sucesso para uma escolha aleatória da semente é 2^{-l}, para que se tenha $y_{E_0} = y_0$. A melhor estratégia que Eva pode usar, como provado na literatura [2], é medir cada qubit enviado por Alice usando uma única base, a base de Breidbart. A base de Breidbart para o espaço de Hilbert de dimensão dois é definida pelos estados

$$|\varphi_1\rangle = \cos(\pi/8)|0\rangle + \text{sen}(\pi/8)|1\rangle \tag{5.20}$$

$$|\varphi_2\rangle = \text{sen}(\pi/8)|0\rangle - \cos(\pi/8)|1\rangle. \tag{5.21}$$

Os autores mostraram que a probabilidade de Eva medir o mesmo bit enviado por Alice é de $p_b = \cos^2(\pi/8) \approx 0,85$. Um canal binário simétrico (BSC) clássico com probabilidade de transição $p = \cos^2(\pi/8)$ pode modelar este cenário (Figura 5.2), em que todos os qubits que representam $a(m,n) = h(m) \oplus x_0(n)$ são interceptados por Eva. Cada bit obtido por Eva após a medição é denotado por c_j^n.

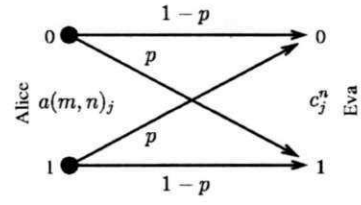


Figura 5.2: Um canal binário simétrico (BSC) clássico com $p \approx 0,15$ modelando o ataque de medição, quando Eva faz medições usando a base de Breidbart.

Lema 24 *Quantum Eva utiliza a melhor estratégia de medição (base de Breidbart) para medir os estados quânticos enviados por Alice, a seqüência de bits produzida pelo gerador de Blum e Micali com semente x_0 aparenta ser uma seqüência verdadeiramente aleatória com probabilidade*

$$P_r \geq 1 - \left(\cos^2(\pi/8)\right)^l, \tag{5.22}$$

em que l é o número de bits necessários para representar a semente x_0.

Prova O Lema 23 afirma que são necessários pelo menos l bits consecutivos, iniciando com d_{i+1}, para que o bit anterior d_i seja calculado com probabilidade um, quando o melhor procedimento probabilístico $A_{BM}(\cdot,\cdot)$ é usado. Então, um procedimento $A'_{BM}(\cdot,\cdot)$ para predizer o gerador de Blum e Micali a partir de $a(m,n)$ necessita, evidentemente, de pelo menos l bits. Mas Eva tem acesso unicamente aos bits c_j^n obtidos das medições. A prova é concluída considerando que cada bit é enviado independentemente. ■

Em outras palavras, o Lema acima afirma que a probabilidade de Eva distinguir a seqüência pseudo-aleatório produzida pelo gerador de Blum e Micali de uma seqüência verdadeiramente aleatória é limitada por $(\cos^2(\pi/8))^l$, que pode ser feita tão pequena quanto desejada. Desde que Eva não é capaz de distinguir a seqüência pseudo-aleatória de uma seqüência aleatória, a segurança do esquema proposto neste trabalho equivale à segurança do esquema de Wegman e Carter. O principal resultado desta dissertação é apresentado no teorema a seguir.

Teorema 9 *O código quântico de autenticação de mensagens clássicas proposto na Seção 5.3 é incondicionalmente seguro com probabilidade P_{WC}, mesmo que Eva disponha de recursos computacionais quânticos infinitos.*

Prova Seja P_{WC} a probabilidade de falha que Alice e Bob estão dispostos a tolerar em um código de autenticação de Wegman e Carter. Simplesmente faça $(\cos^2(\pi/8))^l$ menor que P_{WC} aumentando o tamanho l da semente secreta x_0, de forma que o nível de segurança desejado é sempre alcançado. ■

Embora a capacidade do canal BSC entre Alice e Eva seja maior do que zero, i.e., $C = 1 - H(\cos^2(\pi/8))$, em que $H(p)$ é a entropia binária de Shannon, poder-se-ia supor a existência de um esquema que necessitasse de $poly(l)$ bits c_j^n que fosse capaz de predizer a seqüência do gerador. Entretanto, o teorema da capacidade do canal afirma que se a capacidade for maior do que zero, $C > 0$, então existe um código com taxa $R \leq C$ tal que a taxa de erro decresce assintoticamente a zero. Este não é o caso, já que nenhum código é utilizado e a transmissão é feita a 1 bit por uso, que é maior do que a capacidade do canal.

5.5 Resumo do protocolo

Considerando que Alice e Bob compartilham uma chave secreta que consiste de uma função *hash* particular $h \in H$ e duas sementes x_0 e y_0, o protocolo proposto para a autenticação quântica de mensagens clássicas pode ser resumido como segue:

1. para a $n-$ésima mensagem $m \in M$, Alice gera uma ctiqueta dada por $a(m,n) = h(m) \oplus x_0(n)$, em que $x_0(n) = x_0[(n-1)k+1,\ldots,nk]$.

2. Alice cria k qubits nas bases $\mathcal{Z}$ ou $\mathcal{X}$, dependendo da etiqueta $a(m,n)$ e da seqüência $y_0(n)$. Ela envia os qubits através de um canal quântico perfeito. A mensagem é enviada por um canal inseguro, que pode ser clássico ou quântico.

3. Bob escolhe as bases (conjuntos POVMs) usadas na medição de acordo com a seqüência pseudo-aleatória $y_0(n)$. Como resultado, ele obtém uma seqüência de k bits, $a_B(m,n)$.

4. Bob calcula uma etiqueta local, $a'_B(m,n)$, baseado na mensagem recebida e na seqüência $x_0(n)$, que é comparada com $a_B(m,n)$. Se as etiquetas são idênticas, Bob considera que a mensagem é autêntica. Caso contrário ele a rejeita.

5.6 Conclusões

Este capítulo apresentou um novo protocolo para autenticação quântica de mensagens clássicas, que é a contribuição maior deste trabalho. Foi demonstrado também, através de uma prova formal, que o esquema proposto apresenta segurança incondicional, mesmo no caso em que um criptoanalista possui recursos computacionais infinitos, sejam eles clássicos ou quânticos.

O código de autenticação proposto apresenta uma série de vantagens quando comparado aos protocolos para autenticação quântica de mensagens clássicas descritos na literatura. Em primeiro lugar, ele pode ser implementado usando o estado arte da tecnologia fotônica para sistemas quânticos [17]. O tamanho da chave secreta utilizada no protocolo proposto é igual a $K + 2l$, em que K é o número de bits necessários para especificar a função *hash* e l é o comprimento da semente x_0. Isto significa que a chave é consideravelmente pequena, principalmente quando um grande número de mensagens deve ser enviado. O protocolo de Curty e Santos, por exemplo, utiliza um bit de chave para cada bit de mensagem.

Capítulo 6

Conclusões

Este trabalho de dissertação abordou o problema da autenticação quântica de mensagens clássicas. Como mencionado ao longo do texto, a construção de um computador quântico inviabilizaria todos os sistemas de criptografia e autenticação por chave pública atualmente em uso. A comunidade científica se engaja cada vez mais na busca de novos protocolos que ofereçam segurança incondicional, mesmo quando criptoanalistas dispõem de computadores quânticos e tempo de processamento infinito.

Neste trabalho, foi proposto um novo código de autenticação de mensagens que utiliza um canal quântico para a transmissão da etiqueta. Foi provado, através de um desenvolvimento matemático formal, que o esquema proposto pode ser feito tão seguro quanto o código de autenticação de mensagens de Wegman e Carter. Isto significa que um criptoanalista não pode forjar uma nova mensagem a partir de mensagens enviadas anteriormente, mesmo quando recursos computacionais infinitos, clássicos ou quânticos, estão disponíveis. A segurança do protocolo proposto é oriunda do caráter probabilístico inerente aos estados quânticos não ortogonais.

Com relação aos protocolos para autenticação quântica de mensagens clássicas definidos na literatura, o protocolo proposto apresenta uma série de vantagens. As principais dizem respeito à complexidade reduzida de implementação e ao tamanho reduzido da chave secreta que os participantes do sistema devem compartilhar, principalmente quando se deseja enviar um grande número de mensagens.

6.1 Propostas para trabalhos futuros

A seguir, serão enumeradas algumas sugestões para trabalhos futuros, que são de fundamental importância para que o protocolo de autenticação descrito neste trabalho seja implementado fisicamente. São elas:

Introduzir os efeitos do ruído do canal quântico. Pode ser feito de duas maneiras:

1. Considerar um código clássico corretor de erros para codificar a etiqueta $a(m, n)$ gerada por Alice;
2. Definir um código quântico corretor de erros, de forma a codificar os estados transmitidos por Alice;

Avaliar o impacto do ruído quântico na segurança do protocolo. Quando um código é usado, de forma a introduzir redundância na transmissão, é de se esperar que a segurança do sistema diminua.

Buscar geradores de seqüências pseudo-aleatórias mais eficientes. O gerador de Blum e Micali emite somente um bit para cada operação de exponenciação modular realizada. Existem geradores mais eficientes, capazes de gerar mais bits com operações mais simples em corpo finito.

Bibliografia

[1] H. Barnum, C. Crépeau, D. Gottesman, A. Smith, and A. Tapp. Authentication of quantum messages. *e-print quant-ph/0205128*, 2002.

[2] C. H. Bennett, F. Bessette, G. Brassard, L. Salvail, and L. Smolin. Experimental quantum cryptography. *Journal of Cryptology*, 5(1):3–28, 1992.

[3] C. H. Bennett and G. Brassard. Quantum cryptography: public-key distribution and coin tossing. *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing*, pages 175–179, 1984.

[4] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. Wootters. Teleporting an unknown quantum state via dual classical and EPR channels. *Phys. Rev. Lett.*, 70:1895–1899, 1993.

[5] C. H. Bennett and P. W. Shor. Quantum information theory. *IEEE Transactions on Information Theory*, 44(6):2724–2755, October 1998.

[6] E. Biham, M. Boyer, G. Brassard, J. Graaf, and T. Mor. Security of quantum key distribution against all collective attacks. *e-print quant-ph/9801022*, 1998.

[7] L. Blum, M. Blum, and M. Shub. Comparison of two pseudo-random number generators. In *Proceedings of Crypto'82*, pages 61–78, 1982.

[8] M. Blum and S. Micali. How to generate cryptographically strong sequences of pseudo-random bits. *SIAM J. Comput.*, 13(4):850–864, 1984.

[9] G. Brassard. On computationally secure authentication tags requiring short secret shared keys. In *Proceedings of Crypto'82*, pages 79–88, 1982.

[10] J. L. Carter and M. N. Wegman. Universal classes of hash functions. *J. Comput. Syst. Sci.*, 18:143–154, 1979.

[11] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* John Wiley & Sons Inc., New York, 1991.

[12] M. Curty and D. J. Santos. Quantum authentication of classical messages. *Phys. Rev. A*, 64:062309, 2001.

[13] M. Curty, D. J. Santos, and E. Pérez. Qubit authentication. *Phys. Rev. A*, 66:022301, 2002.

[14] W. Diffie and M. Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, 1976.

[15] M. Dusek, O. Haderka, M. Hendrych, and R. Myska. Quantum identification system. *Phys. Rev. A*, 61:149–156, 1999.

[16] E. N. Gilbert, F. J. MacWilliams, and N. J. A. Sloane. Code which detect deception. *The Bell System Tech. J.*, pages 405–424, March 1947.

[17] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden. Quantum cryptography. *e-print quant-ph/0101098*, 2001.

[18] D. Gottesman. Class of quantum error-correcting codes saturating the quantum Hamming bound. *Phys. Rev. A*, 54:1862–1868, 1996.

[19] J. G. Jensen and R. Schack. Quantum authentication and key distribution using catalysis. *e-print quant-ph/0003104*, 2000.

[20] H. K. Lo. A simple proof of the unconditional security of quantum key distribution. *e-print quant-ph/9904091*, 1999.

[21] D. Mayers. Unconditional security in quantum cryptography. *e-print quant-ph/9802025*, 1998.

[22] R. A. C. Medeiros and F. M. de Assis. A hybrid protocol for quantum authentication of classical messages. In *Proceedings of the 11th International Conference on Telecommunications*, volume 3124 of *Lecture Notes in Computer Science*, pages 1077–1082, Heidelberg, 2004. Springer-Verlag Heidelberg.

[23] A. W. Naylor and G. R. Sell. *Linear Operator Theory in Engineering and Science.* Applied Math Sciences. Springer-Verlag New York Inc., New York, 2nd. edition, 1982.

[24] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, New York, 2000.

[25] I. Niven, H. S. Zuckerman, and H. L. Montgomery. *An introduction to the theory of numbers.* John Wiley and Sons, Inc., New York, 5th. edition, 1991.

[26] R. Rivest, A. Shamir, and L. Adleman. On digital signatures and public-key cryptosystems. *MIT/LCS/TM*, 1982.

[27] C. E. Shannon. Communication theory of secrecy systems. *Bell System Technical Journal*, 1949.

[28] P. W. Shor. Algorithms for quantum computation: Discrete logarithms and factoring. In *IEEE Symposium on Foundations of Computer Science*, pages 124–134, 1994.

[29] P. W. Shor and J. Preskill. Simple proof of security of the BB84 quantum key distribution protocol. *e-print quant-ph/0003004*, 2000.

[30] T. P. Spiller. Quantum information processing: cryptography, computation, and teleportation. *Proceedings of the IEEE*, 84(12):1719–1746, December 1996.

[31] W. Stallings. *Cryptography and Network Security: Principles and Practice.* Prentice-Hall, Inc., New Jersey, 2 edition, 1998.

[32] M. N. Wegman and J. L. Carter. New hash functions and their use in authentication and set equality. *J. Comput. Syst. Sci.*, 22:265–279, 1981.

[33] S. Wiesner. Conjugate coding. *SIGACT News*, 15(1):78, 1983.

[34] G. Zeng and G. Guo. Quantum authentication protocol. *e-print quant-ph/0001046*, 2000.

[35] G. Zeng and W. Zhang. Identity verification in quantum key distribution. *Phys. Rev. A*, 61:022303, 2000.