Universidade Federal de Pernambuco

Departamento de Matemática

Dissertação de Mestrado:

O Algoritmo Polinomial de Shor para Fatoração em um Computador Quântico

por

Mário Sansuke Maranhão Watanabe

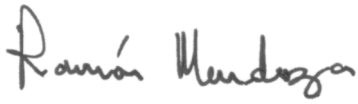
Manoel Lemos
Orientador

Este trabalho contou com o apoio financeiro da CAPES
Recife, Setembro de 2003.

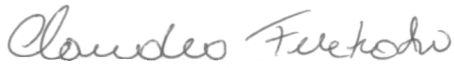
Tese submetida ao Corpo Docente do Programa de Pós-graduação do Departamento de Matemática da Universidade Federal de Pernambuco como parte dos requisitos necessários para a obtenção do Grau de Mestre em Ciências.

Aprovado:

Manoel José Machado Soares Lemos
Orientador



Ramón Oreste Mendoza Ahumada



Cláudio Benedito Silva Furtado

O ALGORITMO POLINOMIAL DE SHOR PARA FATORAÇÃO EM UM COMPUTADOR QUÂNTICO

Por

Mário Sansuke Maranhão Watanabe

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
DEPARTAMENTO DE MATEMÁTICA
Cidade Universitária - Tels. (081)3271-8410 - Fax: (081) 3271-1833
RECIFE - BRASIL

Setembro - 2003

Para Ana Teresa e Ana Júlia,
minha melhor poesia,
minha outra paixão.

Agradecimentos

Agradeço, em primeiro e muito especial lugar, à minha esposa Ana Teresa que, para dizer pouco, é perfeita, enquanto humana. A sua preciosa dedicação e incondicional amor são a paz de espírito que me permite estar nesse caminho. Agradeço à minha mãe Maurina pelo permanente incentivo a tudo aquilo que, de alguma forma, acrescente dignidade e valores imperecíveis à minha existência. Também, manifesto meus agradecimentos ao Departamento de Matemática da UFPE concretizado nas pessoas de seus professores, alunos e funcionários. Em particular, aos professores Manoel Lemos e Ramón Mendoza pela disponibilidade em compartilhar este trabalho nas condições de, respectivamente, orientador e co-orientador do projeto. Agradeço à CAPES pelo apoio financeiro.

O Algoritmo Polinomial de Shor para Fatoração em um Computador Quântico

RESUMO

Sistemas de criptografia largamente difundidos como o RSA fundamentam a sua eficiência na suposição de que, em termos práticos, é impossível fatorar números inteiros suficientemente grandes em uma escala de tempo aceitável. Mais precisamente, não existem, até o momento, algoritmos de fatoração em tempo polinomial que possam ser implementados nos atuais computadores. Dentre os algoritmos conhecidos, o mais eficiente requer um tempo computacional de ordem exponencial na quantidade de dígitos binários do número a ser fatorado.

Em 1994, baseado nos trabalhos anteriores de Benioff, Bennett, Deutsch, Feynman e Simon, dentre outros, Peter Shor apresentou um algoritmo de fatoração que requer assintoticamente uma quantidade em ordem polinomial de passos em um computador quântico para fatorar um número inteiro de tamanho arbitrário. Esse algoritmo ao invés de abordar o problema de decompor tal número em dois fatores não triviais pelo método direto de divisões sucessivas, utiliza o problema equivalente de encontrar a ordem de um certo inteiro modulo o número fatorado, onde esse inteiro é escolhido aleatoriamente relativamente primo com o número fatorado. Shor faz uso de um algoritmo quântico para calcular essa ordem.

A computação quântica revela um paradigma computacional bastante adverso da computação clássica. Enquanto esta última é realizada através de operações binárias determinísticas com base na lógica booleana clássica, a computação quântica fundamenta as suas operações nos postulados que descrevem o comportamento quântico da matéria. Portanto, é probabilística no seu *modus operandi*. Essa diferença entre os formalismos lógicos da computação clássica e da computação quântica é um reflexo direto da natureza dos sistemas físicos que são utilizados para implementar concretamente cada uma dessas computações.

Esta dissertação apresenta o algoritmo de Shor para fatoração em um computador quântico. Na seqüência, introduzimos no capítulo 1 alguns conceitos básicos da computação clássica com o objetivo de criar um ambiente de idéias favorável à apresentação da computação quântica como uma extensão, tão natural quanto possível, do modelo clássico computacional. Assim, no capítulo 2, apresentamos as bases do formalismo matemático que modela a computação quântica, atendo-nos apenas aos aspectos conceituais que são, direta ou indiretamente, aplicados na descrição do algoritmo de Shor.

Os capítulos 3 e 4 são dedicados à apresentação do algoritmo de fatoração de Shor, feita em duas partes. A primeira diz respeito a parte não quântica e aborda os aspectos algébricos do algoritmo. Também é demonstrado o teorema que assegura a viabilidade probabilística da solução desse problema. No capítulo 4, apresentamos a parte quântica do algoritmo de Shor. O ponto alto da dissertação é alcançado mostrando-se como encontrar a ordem de um inteiro módulo o número a ser fatorado relativamente primo com este, conciliando o algoritmo quântico com uma interpretação clássica de seus dados de saída, mediante o uso da expansão de um número racional em frações contínuas.

The Shor's Polynomial Algorithm for Factoring in a Quantum Computer

ABSTRACT

Cryptographic systems such as RSA are based on the assumption that, in practice, integer factoring for too large numbers is impossible in an acceptable period of time. So far, there is no polynomial algorithm running in classical computers for factoring an arbitrary size integer. The most efficient known algorithm demands an exponential computational time on the number of binary digits of the factored number.

In 1994, based on the previous works of Benioff, Bennett, Deutsch, Feynman and Simon, among others, Peter Shor presented a factoring algorithm that requests a polynomial amount of steps for factoring an arbitrary integer n on a quantum computer. That algorithm approaches the problem by splitting the n into two non trivials factors by finding the order of a certain integer a mod n. The integer a is randomly chosen among the ones relatively prime with n. Shor uses an quantum algorithm to evaluate that order.

Quantum computation uses a very different computational model when compared with classical computation. While the latter is achieved by means of deterministic binary operations based on the boolean logic, quantum computation works based on the postulates that describes the quantum behavior of matter. So, it is probabilistic in yours modus operandi. The difference between the classical and quantum computation logic formalisms is a straight reflection of the physical systems that are used to build each of these computations.

This work presents the the Shor's polynomial algorithm for factoring in a quantum computer. In the first chapter we introduce some classical computation basic concepts in view to create an ideal environment that allow us to introduce the quantum computation as an extension as natural as possible of the classical computational model. In the second chapter we give the mathematical formalism which models the quantum computation focusing our attention over the conceptual points that will be applied on the Shor's algorithm description.

Chapters 3 and 4 are devoted to presenting the Shor's factoring algorithm. The presentation divides into two parts. The first part, in Chapter 3, approaches the non quantum algebraic features of the algorithm and we also demonstrate the theorem that proves that the algorithm is probabilistically feasible. In chapter 4 we present the quantum features of Shor's algorithm. The main ponit is achieved when we show how to find the order of a mod n. For that purpose we make a classical interpretation of the quantum algorithm output data by using the expansion of a rational number in continuous fractions.

Sumário

Introdução

Sistemas de criptografia largamente difundidos como o RSA fundamentam a sua eficiência na suposição de que, em termos práticos, é impossível fatorar números inteiros suficientemente grandes como, por exemplo, da ordem de 2^{1000}. Mais precisamente, não existem até o momento algoritmos de fatoração em tempo polinomial que possam ser implementados nos atuais computadores. Dentre os algoritmos conhecidos, o mais eficiente [9, 10] requer um tempo computacional $\mathcal{O}\left(\exp\left[(\log_2 N)^{\frac{1}{3}}(\log_2 \log_2 N)^{\frac{2}{3}}\right]\right)$. Ou seja, trata-se de um algoritmo exponencial na quantidade $\mathcal{O}(\log_2 N)$ de dígitos binários de N, onde N é o número inteiro que se deseja fatorar.

Em 1994, baseado nos trabalhos anteriores de Benioff, Bennett, Deutsch, Feynman e Simon, dentre outros, Peter Shor [16] apresentou um algoritmo de fatoração que requer assintoticamente $\mathcal{O}\left((\log_2 N)^2(\log_2 \log_2 N)(\log_2 \log_2 \log_2 N)\right)$ passos em um computador quântico para fatorar um número inteiro N com $\mathcal{O}(\log_2 N)$ dígitos binários. Esse algoritmo, que opera em tempo polinomial, ao invés de abordar o problema de decompor N em dois fatores não triviais pelo método direto de divisões sucessivas, utiliza o problema equivalente de encontrar a ordem de $y^a \ (mod\ N)$, onde $1 < y < N$ é um número inteiro escolhido aleatoriamente tal que $mdc(y, N) = 1$. Shor faz uso de um algoritmo quântico para calcular essa ordem.

A computação quântica revela um paradigma computacional bastante adverso da computação clássica. Enquanto esta última é realizada através de operações binárias determinísticas com base na lógica booleana clássica, a computação quântica fundamenta as suas operações nos postulados que descrevem o comportamento quântico da matéria. Portanto, é probabilística no seu *modus operandi*. Essa diferença entre os formalismos lógicos da computação clássica e da computação quântica é um reflexo direto da natureza dos sistemas físicos que são utilizados para implementar concretamente cada uma dessas computações. De fato, enquanto os métodos computacionais clássicos são implementados fisicamente por sistemas eletrônicos descritos pelas leis do eletromagnetismo, os algoritmos quânticos devem ser fisicamente implementados por sistemas discretos de partículas como átomos, elétrons e fótons, cujo comportamento é governado pelas leis da mecânica quântica.

Esta dissertação apresenta o algoritmo de Shor para fatoração em um computador quântico e foi, na maior parte, baseada no artigo [11] de S.J. Lomonaco Jr. Em nossa abordagem, optamos por introduzir no capítulo 1 alguns conceitos básicos da computação clássica com o objetivo de criar um ambiente de idéias, de forma tal, que tornasse possível a apresentação da computação quântica como uma extensão, a mais natural quanto possível, do modelo clássico computacional. Assim, no capítulo 2, apresentamos as bases do formalismo matemático que modela a computação quântica, atendo-nos apenas aos aspectos conceituais que são, direta ou indiretamente, aplicados na descrição do algoritmo de Shor. Neste capítulo, faz-se um sistemático uso dos postulados da teoria quântica que descrevem o comportamento de um sistema computacional quântico.

Os capítulos 3 e 4 são dedicados à apresentação do algoritmo de fatoração de Shor, feita em duas partes. A primeira, contida no capítulo 3, diz respeito a parte não quântica e aborda os aspectos algébricos do algorítmo no que diz respeito à equivalência entre o problema de encontrar dois fatores não triviais de um inteiro N e o de descobrir a ordem de $y^a \ (mod\ N)$. Também é demonstrado o teorema que assegura a viabilidade probabilística do método de solução desse problema. No capítulo 4, apresentamos a parte quântica do algoritmo de Shor. O ponto alto da dissertação é alcançado mostrando-se como encontrar a ordem de $y^a \ (mod\ N)$, conciliando o algoritmo quântico com uma interpretação clássica de seus dados de saída, mediante o uso da expansão de um número racional em frações contínuas.

Capítulo 1

Computação Clássica

Neste capítulo inicial, apresentamos de forma breve alguns temas básicos da computação clássica que julgamos necessários para a introdução dos conceitos da computação quântica que virão a seguir. Optamos aqui por uma abordagem que permita a criação de uma ambiência favorável à apresentação das idéias futuras. Dessa forma, poderemos estabelecer mais adiante um instrutivo parelelo que evidencie as diferenças conceituais entre esses dois paradigmas da computação.

1.1 Bit, Registros e Espaço de Estados

Os computadores clássicos utilizam o *bit*, ou dígito binário, como o componente básico da memória. O *bit* representa a menor quantidade de informação digital capaz de ser armazenada porque pode estar, de cada vez, em apenas um de dois estados distintos e mutuamente exclusivos. A concepção desse princípio de memória justifica-se pela relativa facilidade, do ponto de vista físico, de se armazenar a informação digital através da distinção entre dois valores de uma grandeza física contínua como tensão ou corrente elétrica. Dessa forma, o estado de cada *bit* corresponde a um estado físico distinto da máquina. Essa característica de armazenamento faz com que o sistema binário de numeração seja a escolha natural para representar as informações envolvidas na computação clássica.

Assim, do ponto de vista matemático, um **bit** está associado a uma variável $x \in \{0, 1\}$, onde o conjunto $\{0, 1\}$ é dito o **espaço de estados** do *bit*. Desse modo, um *bit* x pode estar a cada instante em apenas um dentre os dois possíveis estados 0 ou 1, mas nunca em uma sobreposição simultânea de ambos.

Por sua vez, um **registro** é uma área mais extensa de memória formada pela justaposição de uma quantidade finita de k *bits*. Isto é, um registro está associado a uma k-upla $(x_{k-1}, \cdots, x_0) \in \{0,1\}^k$, onde o produto cartesiano

$$\{0,1\}^k = \underbrace{\{0,1\} \times \cdots \times \{0,1\}}_{k}$$

é dito o **espaço de estados** desse registro. Além disso, como herda o princípio de funcionamento de cada um de seus *bits* componentes, um registro pode estar a cada instante em apenas um dos 2^k diferentes estados possíveis e mutuamente exclusivos do seu espaço de estados.

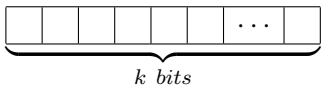


fig.1 Um registro com k *bits*

Observação 1.1.1 *Note que cada um desses* 2^k *possíveis estados corresponde biunivocamente ao armazenamento de um número inteiro* $0 \leq N \leq 2^k - 1$. *De fato, existe uma bijeção natural* ϕ *entre o conjunto das* k*-uplas* $\{(x_{k-1}, \cdots, x_0);\ x_i \in \{0,1\}\}$ *e o conjunto* $\{N \in \mathbb{Z};\ 0 \leq N \leq 2^k - 1\}$ *definida por:*

$$\phi(x_{k-1}, \cdots, x_0) = x_{k-1} \cdot 2^{k-1} + \cdots + x_1 \cdot 2^1 + x_0 \cdot 2^0,$$

que é a representação binária de N. *Além disso, os valores mínimo e máximo armazenados ocorrem quando, respectivamente,* $x_i = 0$ *e* $x_i = 1$ *para todo* i. *Mais claramente,* $\phi(0, \cdots, 0) = 0$ *e* $\phi(1, \cdots, 1) = \sum_{j=0}^{k-1} 2^j = 2^k - 1$. *Abaixo, ilustramos a representação binária do número* 13 *em um registro de cinco bits.*

0	1	1	0	1

fig.2 *Representação binária do número 13*

Em síntese, um registro é capaz de armazer somente um valor de cada vez, isto é, não é possível um mesmo registro de k *bits* clássicos armazenar simultaneamente uma sobreposição de dois, ou mais, números inteiros distintos $0 \leq N_1, N_2 \leq 2^k - 1$.

Observação 1.1.2 *Dado um número inteiro positivo* N *podemos estar interessados em avaliar a quantidade mínima* k *de bits que deverá possuir um registro para ser capaz de armazená-lo. Afirmamos que será necessário um registro com* $\lceil \log_2 N \rceil \leq k$ *bits, onde o símbolo* $\lceil\ \rceil$ *denota o menor inteiro maior que. Com efeito,* k *deverá satisfazer* $N \leq 2^k - 1 < 2^k$, *donde* $\log_2 N < k$.

1.2 Álgebra Booleana e Portas lógicas

A Álgebra Booleana, assim nomeada em homenagem ao seu descobridor o matemático inglês *George Boole* (1815-1864), caracteriza-se pelo fato de que suas variáveis e funções apenas podem operar com os valores 0 e 1. Também conhecida como **álgebra de chaveamentos**, a álgebra booleana, por suas características, é o modelo matemático adequado para descrever a dinâmica de funcionamento dos circuitos lógicos digitais.

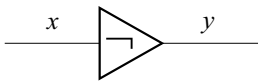
As **Funções booleanas** são da forma $f : \{0,1\}^k \longrightarrow \{0,1\}^m$, cuja entrada é um estado $x = (x_{k-1}, \cdots, x_0) \in \{0,1\}^k$ e cuja saída é um estado $y = (f_{m-1}(x), \cdots, f_0(x)) \in \{0,1\}^m$. Funções booleanas são implementadas na prática através de circuitos eletrônicos digitais chamados **portas lógicas**. Ou seja, a mesma operação que uma função booleana realiza abstratamente com os números 0 e 1, a porta lógica correspondente realiza de forma concreta com, digamos, os estados de tensão elétrica de 0 e 5 volts. Essa correspondência costuma levar ao emprego indistinto das duas expressões.

1.2.1 Portas Lógicas Elementares

Dentre as portas lógicas, ou correspondentes funções booleanas, existem três que merecem uma especial referência. Tais portas são as funções **Not**, **And** e **Or**, que funcionam como blocos elementares para a construção de qualquer outra porta lógica. Isso significa que qualquer função booleana pode ser implementada através de uma combinação adequada dessas funções ou portas básicas [21, pp 62-64], por isso mesmo denominadas de **funcões booleanas elementares** ou **portas lógicas elementares**. A seguir, apresentamos tais funções.

A porta **Not**

A porta lógica **Not**, tabém conhecida como operação de *negação* ou *inversão*, é denotada pelo símbolo "$\neg$" e implementa a correspondente função booleana $\neg : \{0,1\} \longrightarrow \{0,1\}$ definida por $\neg(x) = 1 - x$. Ou seja, essa função atua sobre um único *bit* e inverte o seu estado. Portanto, a função **Not** possui apenas uma entrada e uma saída. Esquematicamente, essa porta é representada pelo seguinte diagrama:



acompanhado de sua tabela-verdade:

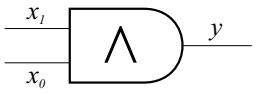
x	y
0	1
1	0

A porta **And**

A porta lógica **And**, também denominada operação de *disjunção*, é denotada pelo símbolo "$\wedge$" e implementa a correspondente função booleana $\wedge : \{0,1\}^k \longrightarrow \{0,1\}$, $k \geq 2$, definida por:

$$\wedge(x_{k-1}, \cdots, x_0) = \begin{cases} 1 & se \quad (x_{k-1}, \cdots, x_0) = (1, 1, \cdots, 1) \\ \\ 0 & se \quad (x_{k-1}, \cdots, x_0) \neq (1, 1, \cdots, 1). \end{cases}$$

A função **And** pode ter mais de duas entradas, mas possui apenas uma saída. Exibimos abaixo a representação esquemática da ocorrência mais usual da porta **And**, isto é, o caso em que $k = 2$:



cuja tabela-verdade é:

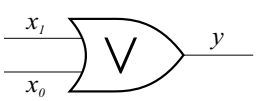
x_1	x_0	y
0	0	0
0	1	0
1	0	0
1	1	1

A porta **Or**

A porta lógica **Or**, também chamada de operação de *conjunção*, é denotada pelo símbolo "$\vee$" e implementa a correspondente função booleana $\vee : \{0,1\}^k \longrightarrow \{0,1\}$, $k \geq 2$, definida por:

$$\vee(x_{k-1}, \cdots, x_0) = \begin{cases} 0 & se \quad (x_{k-1}, \cdots, x_0) = (0, 0, \cdots, 0) \\ \\ 1 & se \quad (x_{k-1}, \cdots, x_0) \neq (0, 0, \cdots, 0). \end{cases}$$

Assim como a porta anterior, a função **Or** pode ter mais de duas entradas, mas possui apenas uma saída. Também exibimos abaixo a representação esquemática da ocorrência mais usual da porta **Or**, isto é, o caso em que $k = 2$:



cuja tabela-verdade é:

x_1	x_0	y
0	0	0
0	1	1
1	0	1
1	1	1

Observação 1.2.1 (Um comentário sobre reversibilidade) *Dentre as portas lógicas elementares, apenas a porta* **Not** *é reversível. Por uma porta reversível, queremos qualificar aquela que está associada a uma função booleana inversível. Assim, a inversa da função* $\neg(x) = y$ *é a função* $\neg^{-1}(y) = x$*, definida por* $\neg^{-1}(y) = 1 - y$*. De fato,* $(\neg^{-1} \circ \neg)(x) = x$*. Além disso, note que a função* **Not** *é auto-inversa, isto é,* $\neg^{-1} = \neg$*. Em termos simples, uma porta lógica, ou correspondente função booleana, será reversível se for possível determinar o valor de entrada (input) uma vez que se conheça o valor de saída (output). Ou seja, se for possível reverter a operação. Neste sentido, note que mesmo conhecendo o valor de saída das funções* **And** *e* **Or***, os valores de entrada permanecem indeterminados. Veremos no capítulo dois que, ao contrário do que ocorre na computação clássica, toda porta quântica é reversível.*

Como afirmamos no início desta seção, qualquer função booleana pode ser implementada como uma combinação adequada das portas lógicas elementares **Not**, **And** e **Or**. A seguir, ilustraremos esse fato escolhendo para exemplo uma função booleana que será utilizada em dois momentos no capítulo seguinte: em uma primeira aplicação da computação quântica conhecida como *Algoritmo de Deutsch-Jozsa* e na definição de um certo *operador linear unitário*. Este último, de uso essencial no algoritmo de fatoração de Shor. Assim, pretendemos torná-la, desde já, familiar.

Exemplo 1.2.1 *A função* **Xor** *(***Or** *exclusivo)*

A porta lógica **Xor** *é denotada pelo símbolo "*$\oplus$*" e implementa a correspondente função booleana* $\oplus : \{0,1\}^2 \longrightarrow \{0,1\}$ *definida por:*

$$\oplus(x_1, x_0) = \begin{cases} x_0 & se \quad x_1 = 0 \\ 1 - x_0 & se \quad x_1 = 1. \end{cases}$$

Em tempo, esclarecemos que o símbolo $\oplus$*, aqui empregado, em nada se relaciona com a sua usual denotação de soma direta. Isso posto, apresentamos abaixo uma implementação da função* **Xor** *como uma composição das funções elementares* **Not***,* **And** *e* **Or***.*

$$\boxed{\oplus(x_1, x_0) = \vee(\wedge(x_1, \neg(x_0)), \wedge(x_0, \neg(x_1)))}$$

É imediata a verificação que o resultado destas operações lógicas satisfaz a tabela-verdade:

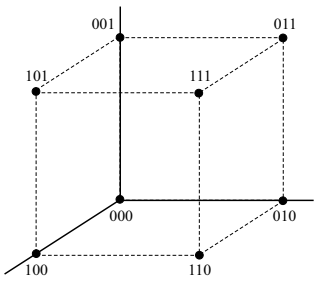
x_1	x_0	y
0	0	0
0	1	1
1	0	1
1	1	0

1.3 Estado, Evolução e Medição de um Registro

Nosso propósito nesta seção é apresentar os conceitos de *estado de um sistema*, *evolução de um sistema* e *medição desse sistema* aplicados ao caso específico de nosso interesse, onde esse sistema é um registro clássico de k *bits*.

Definição 1.3.1 *O* **estado** *de um registro de k-bits é uma k-upla* $(x_{k-1}, \cdots, x_0) \in \{0,1\}^k$*, dito o seu espaço de estados.*

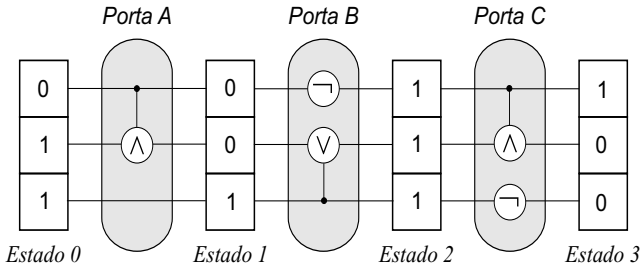
Considere, por exemplo, um registro composto por 3 bits. Nesse caso, o espaço de estados desse registro é formado por oito estados distintos, isto é, existem somente oito combinações distintas para a tripla $(x_2, x_1, x_0),\ x_i \in \{0,1\}$. De modo ilustrativo, podemos visualizar os diferentes estados desse registro ternário como sendo os oito vértices de um cubo de aresta unitária, conforme a figura abaixo:



Assim, o *espaço de estados* do registro é o conjunto formado pela união disjunta desses oito vértices. De modo geral, o espaço de estados de um registro de k *bits* pode ser visualizado como os 2^k vértices de um cubo de aresta unitária k dimensional.

Definição 1.3.2 *Considere um registro de k bits que se encontra inicialmente no estado* $x = (x_{k-1}, \cdots, x_0) \in \{0,1\}^k$. *Dizemos que o registro* **evolui** *de x para y quando passa do estado x para o estado* $y = f(x)$, *onde* $f : \{0,1\}^k \longrightarrow \{0,1\}^k$ *é uma função booleana.*

Um registro evolui de um estado para outro sob a ação de uma porta lógica, ou função booleana. Assim, considere o exemplo anterior com o registro de 3 *bits* no estado inicial $(0,1,1)$. A aplicação sucessiva de três portas lógicas poderia levar o registro a evoluir para, digamos, os estados sucessivos $(0,0,1)$, $(1,1,1)$ e, finalmente, $(1,1,0)$. Tais funções booleanas que provocariam essas evoluções poderiam ser, por exemplo, as *portas lógicas* A, B e C indicadas no diagrama abaixo:



Note que cada porta lógica é composta internamente pela portas elementares **Not**, **And** e **Or**. Neste diagrama, o tempo flui da esquerda para a direita e os estados de 0 a 3 constituem a evolução sequenciada do registro pela ação das sucessivas portas.

A Medição de um registro

Medir um registro significa observar o estado no qual se encontra através de algum processo computacional adequado. Na computação clássica, observar o estado de um registro num certo instante equivale a observar o estado individual de cada um de seus *bits* componentes. Isto é, se sabemos, por exemplo, que após a aplicação de uma sequência de portas lógicas um registro de 3 *bits* deverá ter armazenado em si o número 6 em dígitos binários, então, inequivocamente, o estado do primeiro *bit* será 0 e o estado dos segundo e terceiro *bits* será 1. Ou seja, a medição de um registro clássico é um processo determinístico. Mais ainda, a observação do estado de um registro não altera esse estado. Fatos muitos adversos ocorrem na computação quântica como veremos a seguir.

Capítulo 2

Computação Quântica

2.1 Matrizes e Transformações Unitárias

Apresentamos na sequência alguns conceitos básicos da álgebra linear que serão utilizados nos postulados que descrevem o comportamento dos sistemas quânticos nos quais estamos interessados, a saber, os sistemas de *bits* e registros quânticos. Dessa forma, seja $\mathbb{M}_{m,n}(\mathbb{C})$ o espaço das matrizes com entradas $(a_{ij}) \in \mathbb{C}$, onde $1 \leq i \leq m$ e $1 \leq j \leq n$. Agora, considere as seguintes definições:

Definição 2.1.1 *Seja $A \in \mathbb{M}_{m,n}(\mathbb{C})$. Definimos a matriz* **adjunta** *de A como sendo a matriz $A^\dagger \in \mathbb{M}_{n,m}(\mathbb{C})$ tal que $A^\dagger = \overline{A^t}$. Isto é, a adjunta da matriz A é a matriz conjugada de sua transposta.*

Definição 2.1.2 *Dizemos que uma matriz $A \in \mathbb{M}_{m,n}(\mathbb{C})$ é uma matriz* **unitária** *se e somente se $A^\dagger A = I_n$, onde I_n denota a matriz identidade $n \times n$.*

Em particular, note que se $m = n$ então uma matriz $A \in \mathbb{M}_n(\mathbb{C})$ é unitária se e somente se $A^{-1} = A^\dagger$. Agora, considere o conjunto das matrizes coluna $C \in \mathbb{M}_{n,1}(\mathbb{C})$. Decorre da definição 2.1.2 a seguinte propriedade:

Propriedade 2.1.1 *Uma matriz coluna $C = (c_{i1})$ $n \times 1$ é unitária se e somente se:*

$$|c_{11}|^2 + |c_{21}|^2 + \cdots + |c_{n1}|^2 = 1$$

Demonstração. Suponha que C é unitária. Então $C^\dagger C = I_1 = 1$. Isto é,

$$\overline{c_{11}} \cdot c_{11} + \overline{c_{21}} \cdot c_{21} + \cdots + \overline{c_{n1}} \cdot c_{n1} = 1 \tag{2.1}$$

e a conclusão segue de imediato. Reciprocamente, seja C uma matriz coluna tal que

$$|c_{11}|^2 + |c_{21}|^2 + \cdots + |c_{n1}|^2 = 1 \tag{2.2}$$

Assim, podemos escrever (2.2) na forma (2.1). Ocorre que o lado esquerdo da igualdade (2.1) é exatamente $C^\dagger C$. Isto é:

$$C^\dagger C = 1 = I_1$$

e, portanto, C é unitária. □

Finalmente, sejam V e W dois espaços vetoriais complexos com produto interno hermitiano e com bases ortonormais, respectivamente, $\alpha = \{e_1, e_2, \cdots, e_n\}$ e $\beta = \{f_1, f_2, \cdots, f_m\}$. Além disso, sejam $\mathsf{U} : V \longrightarrow W$ uma transformação linear e $A \in \mathbb{M}_{m,n}(\mathbb{C})$ a matriz de representação de U em relação às bases α e β. Nessas condições, temos a seguinte definição:

Definição 2.1.3 *Uma transformação linear $\mathsf{U} : V \longrightarrow W$ é dita* **unitária** *se e somente se A é uma matriz unitária. Além disso, $A^\dagger$ é a matriz de representação do operador $\mathsf{U}^\dagger : W \longrightarrow V$, adjunto de U, em relação às bases β e α. Mais ainda, se $V = W$ então $\mathsf{U}^\dagger\mathsf{U} = \mathsf{U}\mathsf{U}^\dagger = \mathsf{I}$, onde I é o operador identidade.*

2.2 Sistemas Quânticos

Na computação clássica, o formalismo matemático que descreve o funcionamento dos *bits* e registros está de acordo com a maneira como se comportam os sistemas físicos que implementam concretamente esses sistemas lógicos. Neste caso, os sistemas físicos são diminutos chaveamentos elétricos que controlam a passagem ou interrupção de corrente mediante a distinção entre os dois valores de tensão que concretizam os estados abstratos 0 e 1.

Em contrapartida, o formalismo matemático que descreve o funcionamento dos *bits* e registros quânticos deverá estar de acordo com os sistemas físicos que podem implementar concretamente essa computação. Tais sistemas quânticos são usualmente sistemas isolados e diminutos como átomos, elétrons, fótons etc. Assim, apresentaremos a seguir três postulados oriundos da *mecânica quântica* e que são suficientes para descrever o comportamento dos sistemas quânticos com os quais iremos lidar, a saber, *qubits* e registros quânticos. O formalismo matemático que será apresentado na próxima seção estará de acordo com esses postulados.

2.2.1 Postulados da Teoria Quântica

Dizemos que um sistema é *isolado* quando não está interagindo com o ambiente no qual está inserido. Ademais, passaremos a usar a notação **ket** $|\ \rangle$, devida a *Dirac*, para denotar o vetor de estado de um sistema quântico.

Postulado 2.2.1 (Estado de um sistema) *O Estado de um sistema quântico isolado é completamente descrito por uma matriz unitária* $|C\rangle = (c_{i1}) \in \mathbb{M}_{n,1}(\mathbb{C})$.

Postulado 2.2.2 (Evolução de um sistema) *Um sistema quântico isolado no estado* $|C_1\rangle$ *evoluirá para um novo estado* $|C_2\rangle$*, após um certo intervalo de tempo, de acordo com*

$$|C_2\rangle = U|C_1\rangle$$

onde $U \in \mathbb{M}_n(\mathbb{C})$ *é uma matriz unitária.*

Comentário. Note que este postulado é consistente com o primeiro. De fato, como $|C_1\rangle \in \mathbb{M}_{n,1}(\mathbb{C})$ e $U \in \mathbb{M}_n(\mathbb{C})$, temos que $|C_2\rangle \in \mathbb{M}_{n,1}(\mathbb{C})$. Além disso, $|C_2\rangle$ é unitária. Com efeito,

$$\begin{aligned} |C_2\rangle^\dagger |C_2\rangle &= (U|C_1\rangle)^\dagger (U|C_1\rangle) \\ &= |C_1\rangle^\dagger U^\dagger U |C_1\rangle \\ &= |C_1\rangle^\dagger |C_1\rangle \\ &= I_1 \end{aligned}$$

Portanto, $|C_2\rangle$ é um estado quântico válido.

Postulado 2.2.3 (Medição de um sistema) *Quando um sistema quântico no estado*

$$|C\rangle = \begin{pmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{n1} \end{pmatrix}$$

é **medido**, *ele colapsa com probabilidade* $Prob(i) = |c_{i1}|^2$ *para o estado*

$$|i\rangle \quad = \quad \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \quad \leftarrow \quad \textit{i-ésima posição}$$

fornecendo como resultado da medição o valor i*, onde* $0 \leq i \leq n-1$.

2.3 Qubits

O espaço de estados de um *bit* clássico era simplesmente o conjunto $\{0, 1\}$. Ou seja, os únicos e mutuamente exclusivos estados nos quais um *bit* poderia estar eram os valores 0 e 1. Nesta seção apresentaremos o conceito de **qubit**, ou *bit quântico*, cujo espaço de estados é algo muito adverso. Mais ainda, o próprio conceito de estado de um *qubit* possui uma natureza bem diferente dos estados 0 e 1 do *bit* clássico. Assim, nosso próximo objetivo será responder às perguntas: ondem vivem os *qubits* e como são definidos.

Seja $\mathcal{H}$ um espaço vetorial complexo com um produto interno hermitiano $\langle\ ,\ \rangle$. Tal produto interno induz uma norma $\|\ \|$ nesse espaço. Dizemos que $\mathcal{H}$ é um espaço **completo** se toda *sequência de Cauchy* de vetores de $\mathcal{H}$ for convergente para um vetor pertencente ao próprio espaço com relaçao a essa norma.

Definição 2.3.1 *Dizemos que um espaço vetorial complexo $\mathcal{H}$ com produto interno hermitiano é um* **Espaço de Hilbert** *se for completo com relação à norma induzida por esse produto interno.*

Um **qubit** é um sistema quântico cujo espaço de estados é um *Espaço de Hilbert* bidimensional. Mais adiante, veremos que *registros quânticos* constituídos por dois ou mais *qubits* também são sistemas cujo espaço de estados também é um Espaço de Hilbert com dimensão, embora maior do que dois, finita. Com isso estamos dizendo que todos os espaços de Hilbert com os quais lidamos na computação quântica são espaços vetoriais de dimensão finita.

Agora, como todo espaço vetorial de dimensão finita possui uma base ortonormal, sempre escolheremos tais bases para os espaços de Hilbert com os quais trabalharemos. O motivo dessa escolha é manter a coerência com a definição 2.1.3 e com os postulados da seção 2.2.1. Além disso, denotaremos por **estados puros** os vetores da base do espaço de Hilbert que for tomado como espaço de estados do sistema em questão.

Considere o espaço de Hilbert $\mathbb{C}^2$ com a base ortonormal canônica dada pelo conjunto das matrizes coluna unitárias $\mathcal{C} = \{(1\ 0)^t, (0\ 1)^t\}$. Por definição, $\mathbb{C}^2$ é o **espaço de estados de um qubit**. Utilizaremos a notação **Ket** de Dirac para representar os vetores desse espaço. Em particular, denotamos os vetores da base canônica $\mathcal{C}$ pelos símbolos abaixo:

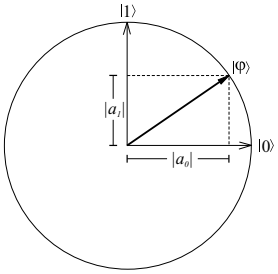
$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad\qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{2.3}$$

Definição 2.3.2 *Um* **qubit** *é um sistema quântico cujo vetor de estado $|\varphi\rangle \in \mathbb{C}^2$ é dado por*

$$|\varphi\rangle = a_0|0\rangle + a_1|1\rangle$$

onde $a_i \in \mathbb{C}$, $|a_0|^2 + |a_1|^2 = 1$ e $\mathbb{C}^2$ é um espaço de Hilbert de dimensão dois.

Note que esta definição de qubit satisfaz a propriedade 2.1.1 e, portamto, é coerente com o postulado 2.2.1. Ainda, os estados quânticos $|\varphi\rangle$ são vetores unitários de $\mathbb{C}^2$. De forma ilustrativa, podemos representar um qubit e seu espaço de estados pelo seguinte gráfico:



Como formam uma base ortonormal de $\mathbb{C}^2$, dizemos que os estados $|0\rangle$ e $|1\rangle$ são os **estados puros** do qubit. Fora esses, qualquer outro vetor de estado $|\varphi\rangle$ de um qubit é uma combinação linear de $|0\rangle$ e $|1\rangle$, conforme nos diz a definição 2.3.2. Também se diz que um estado $|\varphi\rangle$ é uma **sobreposição** desses estados puros. Diante disso, uma primeira grande diferença entre *bits* e qubits torna-se evidente. Enquanto um *bit* pode existir em apenas dois estados 0 e 1, um qubit pode existir em infinitos estados que são as infinitas possibilidades de combinações lineares dos estados puros que satisfazem a definição 2.3.2.

Agora, suponha que um qubit esteja no estado $|\varphi\rangle = a_0|0\rangle + a_1|1\rangle$ no momento em que é observado. Pelo postulado 2.2.3, esse vetor de estado $|\varphi\rangle$ deverá colapsar, no momento da observação, para um dos estados puros $|0\rangle$ ou $|1\rangle$ fornecendo como resultado os valores, respectivamente, 0 com probabilidade $|a_0|^2$ ou 1 com probabilidade $|a_1|^2$. Por essa razão, os coeficientes a_0 e a_1 são ditos **amplitudes de probabilidade**. Também usamos a notação:

$$Prob(0) = |a_0|^2 \qquad Prob(1) = |a_1|^2$$

Dessa forma, embora um qubit, enquanto estado quântico isolado, possa existir em infinitas sobreposições dos estados puros, somente é possível extrair deles informações equivalentes a valores de *bits* clássicos, isto é, 0 ou 1. A razão disso é o fato de que para extraírmos informação de um qubit é necessário observá-lo, ou seja, é necessário interagir com o sistema. Mas, quando isso ocorre, o qubit deixa de ser um sistema quântico isolado e colapsa em um dos estados puros. Contudo, o critério que o sistema quântico utiliza para escolher em qual desses dois estados puros irá colapsar são as amplitudes de probabilidade de cada um no momento da observação. Em síntese, o postulado 2.2.3 diz que:

A medição de um estado quântico não puro altera esse estado.

Esse fato atesta outra surpeendente característica da computação quântica que a diferencia largamente da computação clássica. A saber, o estado de um *bit* clássico não se altera quando o medimos ou observamos. Portanto, enquanto o ato de medir um sistema clássico é um processo determinístico, medir um sistema quântico é um processo probabilístico.

Exemplo 2.3.1 *Suponha que temos 100 computadores quânticos e em cada um deles exista um qubit no estado* $|\varphi\rangle = \sqrt{0.7}\,|0\rangle + \sqrt{0.3}\,|1\rangle$ *no exato momento em que decidimos medir esses qubits simultaneamente nos 100 computadores. O que a teoria quântica nos diz através do postulado 2.2.3 é que em aproximadamente 70 desses computadores deveremos obter como resultado o valor* 0*, pois em* 70% *dos casos o vetor de estado* $|\varphi\rangle$ *irá colapsar para o estado puro* $|0\rangle$ *e em aproximadamente 30 deles deveremos obter o valor* 1*, pois em* 30% *dos casos o vetor de estado* $|\varphi\rangle$ *deverá colapsar para o estado puro* $|1\rangle$*.*

Tendo em vista esse processo probabilístico, note que uma situação muito especial ocorre no seguinte caso. Suponha que temos um qubit no estado quântico:

$$|\varphi\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \tag{2.4}$$

Isso significa que $Prob(0) = Prob(1) = \frac{1}{2}$. Ou seja, em 50% dos casos obteremos o valor 0 como resultado da medição e nos demais 50% obteremos o valor 1. Quando isso ocorre, dizemos que o vetor de estado $|\varphi\rangle$ do qubit está em uma **sobreposição uniforme** de estados puros. Mais adiante, estenderemos esse conceito a registros quânticos e exploraremos as vantagens computacionais desse tipo de sobreposição.

2.4 Registros Quânticos

Semelhante ao que ocorre na computação clássica, um **registro quântico** é uma justaposição ordenada de um número finito de qubits e, portanto, também é um sistema quântico. Assim, nossa próxima tarefa será apresentar o espaço de Hilbert onde vivem os registros quânticos e como são definidos. Para isso, lembramos que um registro clássico de k *bits* tem como espaço de estados o produto cartesiano $\{0,1\}^k$ dos k espaços de estado $\{0,1\}$ correspondentes a cada um de seus *bits* componentes.

Por sua vez, o espaço de estados de um registro quântico de m qubits é o espaço de Hilbert dado por m vezes o produto tensorial de $\mathbb{C}^2$, isto é:

$$\underbrace{\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2}_{m}$$

Inicialmente, considere o espaço de estados $\mathbb{C}^2 \otimes \mathbb{C}^2$. Sabemos que se $\{e_1, \cdots, e_m\}$ e $\{f_1, \cdots, f_n\}$ são bases dos espaços vetoriais, respectivamente, $\mathcal{H}_1$ e $\mathcal{H}_2$, então o conjunto

$\{e_i \otimes f_j |\ 1 \leq i \leq m \ e\ 1 \leq j \leq n\}$ é uma base de $\mathcal{H}_1 \otimes \mathcal{H}_2$. Em particular, temos que os vetores

$$|0\rangle \otimes |0\rangle,\ |0\rangle \otimes |1\rangle,\ |1\rangle \otimes |0\rangle,\ |1\rangle \otimes |1\rangle \tag{2.5}$$

constituem uma base de $\mathbb{C}^2 \otimes \mathbb{C}^2$, um espaço de Hilbert de dimensão quatro.

Notação 2.4.1 *Com o objetivo de simplificar a escrita dos vetores (2.5), convenciona-se usar a seguinte notação:*

$$|a\rangle \otimes |b\rangle = |ab\rangle,\ a, b \in \{0, 1\}$$

Dessa forma, os vetores (2.5) passam a ser escritos como $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$. Agora, como $\mathbb{C}^2 \otimes \mathbb{C}^2 \sim \mathbb{C}^4$, estes símbolos podem ser identificados com a base ortonormal canônica de $\mathbb{C}^4$ cujos vetores passam a representar os estados puros de $\mathbb{C}^2 \otimes \mathbb{C}^2$. Ou seja,

$$|00\rangle \sim \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad |01\rangle \sim \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad |10\rangle \sim \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad |11\rangle \sim \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{2.6}$$

Diante disso, considere um registro quântico formado por dois qubits com vetores de estado $|\varphi_1\rangle = a_0|0\rangle + a_1|1\rangle \in \mathbb{C}^2$ e $|\varphi_2\rangle = b_0|0\rangle + b_1|1\rangle \in \mathbb{C}^2$. Dessa forma, o estado quântico do registro $|\psi\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$ será o produto tensorial dos estados $|\varphi_1\rangle$ e $|\varphi_2\rangle$. Ou seja,

$$\begin{aligned} |\psi\rangle &= |\varphi_1\rangle \otimes |\varphi_2\rangle \\ &= (a_0|0\rangle + a_1|1\rangle) \otimes (b_0|0\rangle + b_1|1\rangle) \\ &= a_0b_0|00\rangle + a_0b_1|01\rangle + a_1b_0|10\rangle + a_1b_1|11\rangle. \end{aligned} \tag{2.7}$$

Ademais, note que

$$\sum_{i,j=0}^{1} |a_i b_j|^2 = 1 \tag{2.8}$$

pois, $|a_0|^2 + |a_1|^2 = 1$ e $|b_0|^2 + |b_1|^2 = 1$. Isso nos diz que o vetor de estado de um registro de dois qubits é um vetor unitário $|\psi\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$. Assim, $|\psi\rangle$ satisfaz a propriedade 2.1.1 e, portanto é consistente com o postulado 2.2.1. Logo, $|\psi\rangle$ é um estado quântico válido. Dessa forma, podemos apresentar a seguinte definição:

Definição 2.4.1 *Um* **registro de dois qubits** *é um sistema quântico cujo vetor de estado $|\psi\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$ é dado por*

$$|\psi\rangle = c_o|00\rangle + c_1|01\rangle + c_2|10\rangle + c_3|11\rangle$$

onde $c_i \in \mathbb{C}$, $\sum_{k=0}^{3} |c_k|^2 = 1$ e $\mathbb{C}^2 \otimes \mathbb{C}^2$ é um espaço de Hilbert de dimensão quatro.

Agora, suponha que um registro de dois qubits encontra-se no estado

$$|\psi\rangle = c_o|00\rangle + c_1|01\rangle + c_2|10\rangle + c_3|11\rangle \tag{2.9}$$

no momento em que é medido. Pelo postulado 2.2.3 tal estado $|\psi\rangle$ deverá colapsar para um dos estados puros $|00\rangle$, $|01\rangle$, $|10\rangle$ ou $|11\rangle$, fornecendo como resutado um dos valores, respectivamente, 0, 1, 2 ou 3 com probabilidade $Prob(i) = |c_i|^2$. Assim, embora um registro de dois qubits possa existir, enquanto estado isolado, em infinitas sobreposições de estados dada pela equação (2.9), a sua medição apenas permite extrair informações equivalentes aos valores possíveis de serem armazenados em um registro clássico de dois bits. Ou seja, $0 \sim (0,0)$, $1 \sim (0,1)$, $2 \sim (1,0)$ e $3 \sim (1,1)$. Confira a observação 1.1.1 do capítulo 1.

Também no caso de um registro quântico de dois qubits, podemos ter um estado de sobreposição uniforme dado por:

$$|\psi\rangle = \frac{|00\rangle + |01\rangle + |10\rangle + |11\rangle}{2} \tag{2.10}$$

Ao ser medido nesse estado, o registro quântico deverá fornecer um dos resultados 0, 1, 2 ou 3 com iguais probabilidades de 25%.

Faremos agora a generalização do conceito de registro quântico para o caso em que esse registro é composto por m qubits. Contudo, cabe antes uma ressalva quanto à notação mais adequada para representar os vetores da base canônica do espaço vetorial

$$\underbrace{\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2}_{m}.$$

Assim, note que, sendo $\{|0\rangle, |1\rangle\}$ uma base de $\mathbb{C}^2$, temos que um elemento da base de $\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$ é dado genericamente por:

$$|b_{m-1}\rangle \otimes \cdots \otimes |b_1\rangle \otimes |b_0\rangle \tag{2.11}$$

onde $b_i \in \{0,1\}$. Se usarmos a notação 2.4.1 então o vetor acima pode ser escrito como $|b_{m-1} \cdots b_1 b_0\rangle$. Ocorre que para registros com um número m suficientemente grande de qubits, mesmo esta notação torna-se inadequada. Dessa forma, note que a sequência de números $b_{m-1} \cdots b_1 b_0$ é a representação binária do número inteiro não negativo (cf. observação 1.1.1)

$$c = b_{m-1} \cdot 2^{m-1} + \cdots + b_1 \cdot 2^1 + b_0 \cdot 2^0 \tag{2.12}$$

Com isso, podemos adotar a seguinte notação:

Notação 2.4.2 *Seja $b_{m-1}\cdots b_1 b_0$ a representação binária do número inteiro não negativo c dada pela equação (2.12). Assim, denotamos:*

$$|c\rangle = |b_{m-1}\cdots b_1 b_0\rangle = |b_{m-1}\rangle \otimes \cdots \otimes |b_1\rangle \otimes |b_0\rangle$$

No caso de um registro quântico de três qubits, por exemplo, os oito vetores da base canônica de $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ podem ser representados da seguinte forma:

$$\begin{array}{lcllcl} |0\rangle & = & |000\rangle & |4\rangle & = & |100\rangle \\ |1\rangle & = & |001\rangle & |5\rangle & = & |101\rangle \\ |2\rangle & = & |010\rangle & |6\rangle & = & |110\rangle \\ |3\rangle & = & |011\rangle & |7\rangle & = & |111\rangle \end{array}$$

Em princípio, o emprego dos símbolos $|0\rangle$ e $|1\rangle$ poderia causar alguma confusão. Contudo, o contexto em que forem empregados deixará sempre muito claro se estão sendo aplicados no sentido da notação 2.4.2, acima, ou se denotam os estados puros de um qubit simples dados pelas igualdades (2.3).

Após essas considerações sobre a notação, faremos a generalização dos conceitos para um registro quântico de m qubits. Para isso, adotaremos, *mutatis mutandis*, as construções e definições feitas para o caso $m = 2$. Dessa forma, temos a seguinte definição:

Definição 2.4.2 *Um* **registro quântico** *de m qubits é um sistema quântico cujo vetor de estado $|\psi\rangle \in \underbrace{\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2}_{m}$ é dado por:*

$$|\psi\rangle = \sum_{x=0}^{2^m-1} c_x |x\rangle$$

onde $c_x \in \mathbb{C}$, $\sum_{x=0}^{2^m-1} |c_x|^2 = 1$ e $\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$ é um espaço de Hilbert de dimensão 2^m com base canônica $\{|0\rangle,\ |1\rangle,\ |2\rangle, \cdots,\ |2^m - 1\rangle\}$.

Como $\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2 \sim \mathbb{C}^{2^m}$, essa base pode ser identificada com a base ortonormal canônica de $\mathbb{C}^{2^m}$ cujos vetores passam a representar os estados puros de $\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$. Assim, para $0 \leq x \leq 2^m - 1$, temos a seguinte identificação:

$$|x\rangle \sim \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow x\text{-ésima posição}$$

Agora, suponha que um registro quântico de m qubits encontre-se no estado

$$|\psi\rangle = c_0|0\rangle + c_1|1\rangle + c_2|2\rangle + \cdots + c_{2^m-1}|2^m - 1\rangle \tag{2.13}$$

no exato momento em que é medido. Pelo postulado 2.2.3, esse estado $|\psi\rangle$ deverá colapsar para um dos estados puros $|0\rangle,\ |1\rangle,\ |2\rangle, \cdots,\ |2^m - 2\rangle$ ou $|2^m - 1\rangle$ e fornecer, respectivamente, como resultado dessa medição um dos valores $0,\ 1,\ 2, \cdots, 2^m - 2$ ou $2^m - 1$ com probabilidade $Prob(x) = |c_x|^2$, para $0 \leq x \leq 2^m - 1$. Note que tais valores são os 2^m números inteiros não negativos possíveis de serem armazenados em um registro clássico de m *bits*. Mais uma vez, confira a observação 1.1.1.

Como um caso particular da sobreposição de estados dada pela equação (2.13), temos a generalização da equação (2.10) dada pela seguinte definição:

Definição 2.4.3 *Dizemos que o vetor de estado $|\varphi\rangle$ de um registro de m qubits está em* **sobreposição uniforme** *se*

$$|\varphi\rangle = \frac{1}{\sqrt{2^m}} \sum_{x=0}^{2^m-1} |x\rangle$$

Nesse caso, ao se medir um registro em estado de sobreposição uniforme, o resultado será um dos valores $0,\ 1,\ 2,\ \cdots,\ 2^m - 2$ ou $2^m - 1$ com iguais probabilidades $Prob(x) = \frac{1}{2^m}$.

2.4.1 Estados Correlacionados

O estado de um sistema físico clássico sempre pode ser descrito em termos dos estados de suas partes componentes. Assim, na computação clássica, o estado de um registro de k *bits* sempre pode ser apresentado na forma de uma k-upla $(x_{k-1}, \cdots, x_0) \in \{0,1\}^k$ onde cada coordenada $x_i \in \{0,1\}$ é inequivocamente o estado de cada um dos k *bits* individuais que compõem o resgistro. Por sua vez, o estado de um sistema quântico nem sempre pode ser descrito em termos dos estados individuais de suas partes. Esse fato não possui contrapartida no mundo clássico de nossa experiência direta e, por isso, é não intuitivo. Dessa forma, na computação quântica, pode não ser possível expressar o estado de um registro quântico de m qubits através dos estados separados de cada um desses qubits. Formalmente, resumimos tal fato na seguinte definição:

Definição 2.4.4 *Seja $|\psi\rangle$ o vetor de estado de um registro quântico de m qubits. Dizemos que $|\psi\rangle$ é um estado* **correlacionado** *se não puder ser escrito como o produto tensorial dos estados individuais de seus m qubits. Caso contrário, os m 1-qubits são ditos* **independentes**.

Assim, considere, por exemplo, um registro de dois qubits cujo vetor de estado $|\psi\rangle$ é dado por:

$$|\psi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}} \tag{2.14}$$

Das equações (2.7) temos que, de modo geral, um registro de dois qubits dado pelo produto tensorial de seus qubits componentes é da forma:

$$|\psi\rangle = a_0b_0|00\rangle + a_0b_1|01\rangle + a_1b_0|10\rangle + a_1b_1|11\rangle \tag{2.15}$$

Logo, ao igualarmos as equações (2.14) e (2.15), caímos no seguinte sistema de equações:

$$\begin{cases} a_0b_0 &= \frac{1}{\sqrt{2}} \\ a_0b_1 &= 0 \\ a_1b_0 &= 0 \\ a_1b_1 &= \frac{1}{\sqrt{2}} \end{cases}$$

que, claramente, não possui solução. De fato, se assumimos que $a_i \neq 0$ então $b_{1-i} = 0$. Por outro lado, se fazemos $b_i \neq 0$ então $a_{1-i} = 0$ e em qualquer dos casos concluímos que $a_ib_i = 0$, uma contradição. Portanto, não existem números $a_0,\ a_1,\ b_0$ e b_1 tais que o vetor de estado $|\psi\rangle$ possa ser expresso como o produto tensorial de estados individuais dados por $|\varphi_1\rangle = a_0|0\rangle + a_1|1\rangle$ e $|\varphi_2\rangle = b_0|0\rangle + b_1|1\rangle$. Em outras palavras, não existem tais estados individuais, mas apenas o estado correlacionado $|\psi\rangle$ de ambos qubits.

Medição de qubits independentes e correlacionados

Inicialmente, mostraremos que quando um registro quântico é formado por qubits independentes, a medição de um desses qubits não provoca o colapso de todo o sistema. Para ver isso, considere um registro quântico formado por dois qubits independentes cujo vetor de estado seja inicialmente

$$|\psi_1\rangle = c_0|00\rangle + c_1|01\rangle + c_2|10\rangle + c_3|11\rangle$$

Como os qubits são independentes, então existem $a_0,\ a_1,\ b_0$ e $b_1 \in \mathbb{C}$ tais que o estado $|\psi_1\rangle$ pode ser escrito como o produto tensorial dos estados de seus qubits dados por $|\varphi_1\rangle = a_0|0\rangle + a_1|1\rangle$ e $|\varphi_2\rangle = b_0|0\rangle + b_1|1\rangle$. Ou seja, inicialmente temos:

$$|\psi_1\rangle = |\varphi_1\rangle \otimes |\varphi_2\rangle$$

Agora, suponha, sem perda de generalidade, que ao realizarmos a medição do primeiro qubit, este colapse para o estado puro $|0\rangle \in \mathbb{C}^2$. Dessa forma, o registro passa do estado $|\psi_1\rangle$ para

$$|\psi_2\rangle = |0\rangle \otimes |\varphi_2\rangle.$$

Afirmamos, que $|\psi_2\rangle$ ainda é um estado quântico. De, fato:

$$\begin{aligned} |\psi_2\rangle &= |0\rangle \otimes (b_0|0\rangle + b_1|1\rangle) \\ &= b_0|0\rangle \otimes |0\rangle + b_1|0\rangle \otimes |1\rangle \\ &= b_0|00\rangle + b_1|01\rangle \end{aligned}$$

Logo, como $|b_0|^2+|b_1|^2 = 1$, temos que $|\psi_2\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$ é um estado quântico válido. Assim, embora o registro quântico tenha passado do estado $|\psi_1\rangle$ para $|\psi_2\rangle$, este último ainda é uma sobreposição de estados e não um estado puro. Isso mostra que a medição de um dos qubits independentes não provocou o colapso do registro como um todo.

Por outro lado, se um registro encontra-se em um estado quântico correlacionado, então a medição de um dos qubits afetará todos os demais, provocando o colapso de todo o sistema. Dessa forma, considere o exemplo do registro de dois qubits no estado correlacionado $|\psi\rangle$ dado pela equação (2.14). Ao medirmos esse registro o vetor de estado $|\psi\rangle$ poderá colapsar, com iguais probabilidades de $\frac{1}{2}$, para um dos estados puros $|00\rangle$ ou $|11\rangle$. Contudo, nunca irá colapsar para os estados $|01\rangle$ ou $|10\rangle$, pois $Prob(1) = Prob(2) = 0$.

Com isso, suponha que realizamos a medição de um dos qubits. Se este colapsar para o estado puro $|0\rangle$ então, necessariamente, o outro será forçado a colapsar também para o estado puro $|0\rangle$ e o sistema como um todo irá colapsar para o estado $|00\rangle$, já que não é possível ocorrer $|01\rangle$. Similarmente, se o resultado da medição de um dos qubits for o estado puro $|1\rangle$ então o outro necessariamente também irá colapsar para o estado $|1\rangle$, levando todo o sistema para o estado $|11\rangle$. Assim, a medição de qualquer dos qubits correlacionados provoca o colapso de todo o registro quântico.

2.5 Transformações Unitárias e Portas Quânticas

Até agora, vimos os aspectos referentes ao vetor de estado de um qubit e sua medição que comportam-se de acordo com os postulados 2.2.1 e 2.2.3. A partir desta seção, passaremos a abordar o comportamento dinâmico dos qubits, isto é, a maneira como evoluem de um estado a outro de acordo com o postulado 2.2.2.

No capítulo um, vimos que a evolução dos estados de um registro clássico dá-se pela ação de uma porta lógica. Ocorre que, naquele caso, os espaços de estados de um registro de k bits são produtos cartesianos da forma $\{0,1\}^k$. Por isso, exige-se que as portas lógicas que provocam tais evoluções sejam, naturalmente, as funções booleanas $f : \{0,1\}^k \longrightarrow \{0,1\}^k$, onde as entradas são k-uplas $x = (x_{k-1}, \cdots, x_1, x_0)$ e as saídas são k-uplas $f(x) = (f_{k-1}(x), \cdots, f_1(x), f_0(x))$, ambas constituídas de 0's e 1's.

Na computação quântica, os espaços de estados dos qubits são espaços vetoriais. Dessa forma, é natural, em primeiro lugar, que as funções que provocam as mudanças de estado sejam as *transformações lineares* do espaço nele mesmo. Mais ainda, como os estados

são vetores unitários e as bases são ortonormais, também é natural que, dentre as transformações lineares, sejam escolhidas as *unitárias*. De fato, estas preservam a norma e a ortogonalidade dos vetores sobre os quais atuam. Por essa razão, o postulado 2.2.2 exige que as transformações que provocam a evolução dos vetores de estado dos qubits sejam aquelas cuja matriz de representação na base canônica ortonormal do espaço sejam matrizes unitárias. Assim, pela definição 2.1.3, dado um registro quântico formado por m qubits, as transformações unitárias U que provocam sua evolução são os operadores lineares tais que

$$\mathsf{U}^\dagger\mathsf{U} = \mathsf{U}\mathsf{U}^\dagger = \mathsf{I} \tag{2.16}$$

onde I é o operador identidade do espaço vetorial de Hilbert $\underbrace{\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2}_{m}$.

Assim, uma consequência imediata do postulado 2.2.2 é que os operadores lineares que provocam mudanças nos estados dos qubits são operadores *inversíveis*. De fato, $\mathsf{U}^{-1} = \mathsf{U}^\dagger$. Agora, da mesma forma como na computação clássica as funções booleanas são implementadas por portas lógicas, na compução quântica os operadores unitários são implementados por portas quânticas. Ora, pelo que afirmamos no início deste parágrafo, temos que

Toda porta quântica é reversível.

Tal comportamento das portas quânticas é mais uma das notáveis diferenças em relação à computação clássica. Como vimos na observação 1.2.1, a maioria das funções booleanas elementares clássicas são irreversíveis. Também, como no caso clássico, existem portas quânticas ditas *elementares*. Neste caso, consideram-se **portas quânticas elementares** aquelas que atuam sobre registros de um, dois ou até três qubits por serem as de implementação física mais simples. Dessa forma, toda porta quântica pode ser construída a partir de um número finito de portas elementares [17, Seção 2].

2.5.1 Portas Quânticas Elementares

Neste seção daremos quatro exemplos de portas quânticas elementares. As duas primeiras atuam sobre registros simples de apenas um qubit, a terceira atua sobre registros de dois qubits e a quarta sobre registros de três qubits.

A porta quântica $\mathsf{U}_{\mathsf{NOT}}$

A porta $\mathsf{U}_{\mathsf{NOT}}$ desempenha na computação quântica uma função similar àquela realizada pela função **Not** da computação clássica. Ou seja, se o estado de um qubit é $|0\rangle$ a porta $\mathsf{U}_{\mathsf{NOT}}$ converte para $|1\rangle$ e vice versa. Como é um operador linear, basta definir a porta $\mathsf{U}_{\mathsf{NOT}} : \mathbb{C}^2 \longrightarrow \mathbb{C}^2$ em relação aos vetores da base. Assim, temos:

$$\begin{aligned} \mathsf{U}_{\mathsf{NOT}}|0\rangle &= 0 \cdot |0\rangle + 1 \cdot |1\rangle \\ \mathsf{U}_{\mathsf{NOT}}|1\rangle &= 1 \cdot |0\rangle + 0 \cdot |1\rangle \end{aligned}$$

Logo, a matriz de transformação de $\mathsf{U_{NOT}}$ em relação à base canônica ortonormal de $\mathbb{C}^2$ é a matriz unitária $A \in \mathbb{M}_{2,2}(\mathbb{C})$ dada por:

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

A Transformação de Hadamard - H

A transformação de Hadamard é um operador unitário $\mathsf{H} : \mathbb{C}^2 \longrightarrow \mathbb{C}^2$, assim definido nos vetores da base:

$$\begin{aligned} \mathsf{H}|0\rangle &= \tfrac{1}{\sqrt{2}} \cdot |0\rangle + \tfrac{1}{\sqrt{2}} \cdot |1\rangle \\ \mathsf{H}|1\rangle &= \tfrac{1}{\sqrt{2}} \cdot |0\rangle - \tfrac{1}{\sqrt{2}} \cdot |1\rangle \end{aligned}$$

Sua matriz de representação em relação à base canônica ortonormal de $\mathbb{C}^2$ é a matriz unitária $A \in \mathbb{M}_{2,2}(\mathbb{C})$ dada por:

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

Uma das importantes aplicações dessa porta quântica elementar é a capacidade de criar um estado de *sobreposição uniforme* quando atua sobre um qubit no estado $|0\rangle$. De fato, note que $\mathsf{H}|0\rangle = \frac{1}{\sqrt{2}} \cdot |0\rangle + \frac{1}{\sqrt{2}} \cdot |1\rangle$ é exatamente o estado dado pela equação (2.4).

A porta quântica $\mathsf{U_{CNOT}}$

A porta $\mathsf{U_{CNOT}}$ (*controlled-NOT*) atua sobre um registro de dois qubits e o seu efeito é alterar o estado do segundo qubit, controlado pelo estado do primeiro. Especificamente, a porta $\mathsf{U_{CNOT}}$ nega o estado do segundo qubit se o estado do primeiro é $|1\rangle$ e deixa inalterado se o estado do primeiro for $|0\rangle$. Assim, a porta quântica $\mathsf{U_{CNOT}} : \mathbb{C}^2 \otimes \mathbb{C}^2 \longrightarrow \mathbb{C}^2 \otimes \mathbb{C}^2$ é definida nos vetores da base canônica da seguinte forma:

$$\begin{aligned} \mathsf{U_{CNOT}}|00\rangle &= |00\rangle \\ \mathsf{U_{CNOT}}|01\rangle &= |01\rangle \\ \mathsf{U_{CNOT}}|10\rangle &= |11\rangle \\ \mathsf{U_{CNOT}}|11\rangle &= |10\rangle \end{aligned}$$

e sua matriz de representação em relação à base canônica ortonormal de $\mathbb{C}^2 \otimes \mathbb{C}^2$ é a matriz unitária $A \in \mathbb{M}_{4,4}(\mathbb{C})$ dada por:

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

A porta quântica $\mathsf{U}_{\mathsf{CCNOT}}$

Também denominada *operador de Toffoli*, a porta $\mathsf{U}_{\mathsf{CCNOT}}$ (*controlled-controlled-NOT*) atua sobre um registro quântico de três qubits e o seu efeito é negar o estado do terceiro qubit se e somente se o estado dos dois primeiros é 11. O operador linear de Toffoli $\mathsf{U}_{\mathsf{CCNOT}} : \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 \longrightarrow \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ é assim definido em relação aos vetores da base:

$$\begin{array}{lcllcl} \mathsf{U}_{\mathsf{CCNOT}}|000\rangle & = & |000\rangle & \mathsf{U}_{\mathsf{CCNOT}}|100\rangle & = & |100\rangle \\ \mathsf{U}_{\mathsf{CCNOT}}|001\rangle & = & |001\rangle & \mathsf{U}_{\mathsf{CCNOT}}|101\rangle & = & |101\rangle \\ \mathsf{U}_{\mathsf{CCNOT}}|010\rangle & = & |010\rangle & \mathsf{U}_{\mathsf{CCNOT}}|110\rangle & = & |111\rangle \\ \mathsf{U}_{\mathsf{CCNOT}}|011\rangle & = & |011\rangle & \mathsf{U}_{\mathsf{CCNOT}}|111\rangle & = & |110\rangle \end{array}$$

Sua matriz de representação em relação à base canônica ortonormal de $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ é a matriz unitária $A \in \mathbb{M}_{8,8}(\mathbb{C})$ dada por:

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

2.5.2 O Operador Unitário U_f

Sabemos que a maioria das portas lógicas clássicas são irreversíveis. Por outro lado, toda porta quântica é reversível. Assim, é natural perguntarmos se os algorítmos clássicos podem ser quanticamente implementados. Ou seja, será que sempre existirão operadores lineares que possam implementar uma computação classicamente exeqüível?

Felizmente, Deutsch mostrou [5] que sempre será possível construir portas quânticas reversíveis para toda função que possa ser classicamente implementada. Assim, apresentamos nesta seção o operador linear U_f que desempenhará um papel fundamental na parte quântica do algoritmo de fatoração de Shor. Dessa forma, considere uma função booleana clássica

$$f : \{0,1\}^k \longrightarrow \{0,1\}^m.$$

Isto é, as entradas dessa função são k-uplas $(x_{k-1},\ \cdots,\ x_1,\ x_0) = x$ e as saídas são m-uplas $(f_{m-1}(x), \cdots,\ f_1(x),\ f_0(x)) = f(x)$. Antes de mais nada, note que, pela observação 1.1.1, temos $0 \leq x \leq 2^k - 1$ e $0 \leq f(x) \leq 2^m - 1$, onde x e $f(x)$ são números inteiros não

negativos cujas representações binárias são as ênupas, respectivamente, de entrada e de saída.

Agora, sendo f uma função possível de ser classicamente computada, assumiremos a existência de uma certa transformação linear unitária U_f que implementa f quanticamente. Para apresentar essa transformação, denotamos

$$\mathcal{H}_1 = \underbrace{\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2}_{k} \; e \; \mathcal{H}_2 = \underbrace{\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2}_{m}.$$

Assim $\mathsf{U}_f : \mathcal{H}_1 \otimes \mathcal{H}_2 \longrightarrow \mathcal{H}_1 \otimes \mathcal{H}_2$ é um operador linear definido nos vetores da base por:

$$\mathsf{U}_f \; |x, y\rangle = |x, y \oplus f(x)\rangle \tag{2.17}$$

onde $\mathcal{H}_1 \otimes \mathcal{H}_2$ é um espaço de Hilbert de dimensão 2^{k+m}. A notação $|x, y\rangle$ significa um **estado puro** de $\mathcal{H}_1 \otimes \mathcal{H}_2$ formado pela justaposição das representações binárias dos números x e y com, respectivamente, k e m dígitos, de tal forma que $|x\rangle$ é um estado puro de $\mathcal{H}_1$ e $|y\rangle$ é um estado puro de $\mathcal{H}_2$. Isto é,

$$|xy\rangle = |x_{k-1} \; \cdots \; x_1 \; x_0 \; y_{m-1} \; \cdots \; y_1 \; y_0\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2.$$

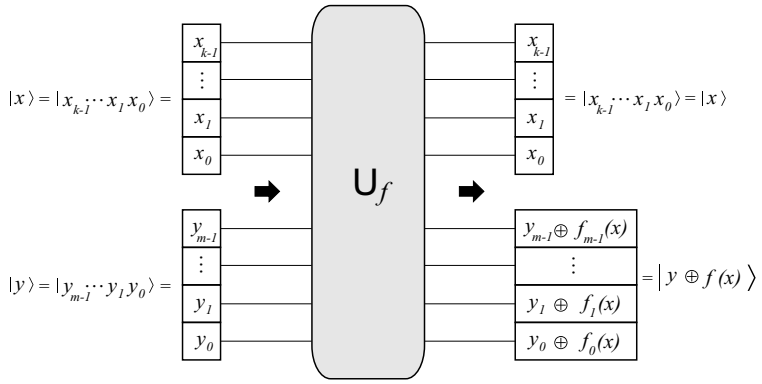
Similarmente, $|x, y \oplus f(x)\rangle$ também é um estado puro de $\mathcal{H}_1 \otimes \mathcal{H}_2$, pois

$$|x, y \oplus f(x)\rangle = |x_{k-1} \; \cdots \; x_1 \; x_0 \; (y_{m-1} \oplus f_{m-1}(x)) \cdots (y_1 \oplus f_1(x)) \; (y_0 \oplus f_0(x))\rangle,$$

onde,

$$y_j \oplus f_j(x) = \begin{cases} f_j(x) & se \quad y_j = 0 \\ \\ 1 - f_j(x) & se \quad y_j = 1 \end{cases} \tag{2.18}$$

é a função **Xor**, apresentada no exemplo 1.2.1 do capítulo um. Graficamente, o operador U_f pode ser representado pelo seguinte diagrama:



Para ver que U_f é um operador unitário, mostraremos o caso mais simples em que f é uma função booleana do tipo $f : \{0,1\} \longrightarrow \{0,1\}$ e, por conseguinte, $\mathsf{U}_f : \mathbb{C}^2 \otimes \mathbb{C}^2 \longrightarrow \mathbb{C}^2 \otimes \mathbb{C}^2$ é um operador definido nos vetores da base canônica $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ de $\mathbb{C}^2 \otimes \mathbb{C}^2$ por $\mathsf{U}_f\ |x, y\rangle = |x, y \oplus f(x)\rangle$. Ou seja,

$$\begin{array}{rcccl} \mathsf{U}_f\ |0,0\rangle & = & |0, 0 \oplus f(0)\rangle & = & |0\ f(0)\rangle \\ \mathsf{U}_f\ |0,1\rangle & = & |0, 1 \oplus f(0)\rangle & = & |0\ (1 - f(0))\rangle \\ \mathsf{U}_f\ |1,0\rangle & = & |1, 0 \oplus f(1)\rangle & = & |1\ f(1)\rangle \\ \mathsf{U}_f\ |1,1\rangle & = & |1, 1 \oplus f(1)\rangle & = & |1\ (1 - f(1))\rangle \end{array}$$

onde usamos que:

$$y \oplus f(x) = \begin{cases} f(x) & se \quad y = 0 \\ \\ 1 - f(x) & se \quad y = 1 \end{cases}$$

Com isso, seja $A \in \mathbb{M}_{4,4}(\mathbb{C})$ a matriz de representação de U_f em relação à base canônica de $\mathbb{C}^2 \otimes \mathbb{C}^2$. Para cada uma das quatro possibilidades de definição de uma função $f : \{0,1\} \longrightarrow \{0,1\}$ existirá uma matriz A associada. Consideremos esses quatro casos:

1. Se $f(0) = 0$ e $f(1) = 0$ então:

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

2. Se $f(0) = 0$ e $f(1) = 1$ então:

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

3. Se $f(0) = 1$ e $f(1) = 0$ então:

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

4. Se $f(0) = 1$ e $f(1) = 1$ então:

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

Logo, qualquer que seja f, A é uma matriz unitária.

Agora, voltando à definição geral de U_f dada pela equação (2.17), note que para computar $f(x)$ através dessa porta quântica é suficiente aplicar U_f a um estado inicial da forma $|x, 0\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$. De fato,

$$\mathsf{U}_f|x, 0\rangle = |x, f(x)\rangle \tag{2.19}$$

pois, conforme a definição dada pela equação (2.18), se $y = 0$ então $y_j = 0$ para todo j. Finalmente, temos que $\mathsf{U}_f\mathsf{U}_f = \mathsf{I}$, onde I é o operador identidade de $\mathcal{H}_1 \otimes \mathcal{H}_2$. Com efeito, $\mathsf{U}_f|x, y \oplus f(x)\rangle = |x, y \oplus f(x) \oplus f(x)\rangle = |x, y\rangle$, pois $f(x) \oplus f(x) = 0$.

2.5.3 A Transformada Quântica de Fourier $\mathsf{U}_{\mathcal{QF}}$

Seja $\mathcal{H} = \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$ um espaço de Hilbert de dimensão $Q = 2^m$ e base canônica ortonormal $\mathcal{C} = \{|0\rangle,\ |1\rangle,\ |2\rangle,\ \cdots,\ |Q-1\rangle\}$. Agora, considere a transformação linear $\mathsf{U}_{\mathcal{QF}} : \mathcal{H} \longrightarrow \mathcal{H}$, assim definida nos vetores da base $\mathcal{C}$:

$$\mathsf{U}_{\mathcal{QF}}\ |x\rangle = \frac{1}{\sqrt{Q}} \sum_{y=0}^{Q-1} \omega^{xy} |y\rangle \tag{2.20}$$

onde $\omega = e^{\frac{2\pi i}{Q}}$ é a Q-ésima raiz complexa da unidade. Queremos mostrar que $\mathsf{U}_{\mathcal{QF}}$ é unitária. Com efeito, considere a matriz $A \in \mathbb{M}_{Q,Q}(\mathbb{C})$ de representação de $\mathsf{U}_{\mathcal{QF}}$ em relação à base $\mathcal{C}$:

$$A = \frac{1}{\sqrt{Q}} \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega^1 & \omega^2 & \omega^3 & \cdots & \omega^{Q-1} \\ 1 & \omega^2 & \omega^4 & \omega^6 & \cdots & \omega^{2(Q-1)} \\ 1 & \omega^3 & \omega^6 & \omega^9 & \cdots & \omega^{3(Q-1)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{Q-1} & \omega^{2(Q-1)} & \omega^{3(Q-1)} & \cdots & \omega^{(Q-1)(Q-1)} \end{pmatrix}$$

Como A é uma matriz simétrica, pois $\omega^{xy} = \omega^{yx}$, temos que $A^\dagger = \overline{A^t} = \overline{A}$. Portanto, basta mostrar que $A\overline{A} = I$, onde I é a matriz identidade $Q \times Q$. Dessa forma, temos que $A\overline{A} = \frac{1}{Q}(a_{xy})_{Q\times Q}$, onde

$$a_{xy} = \sum_{k=0}^{Q-1} \omega^{xk}\overline{\omega}^{yk} = \sum_{k=0}^{Q-1} (\omega^x \overline{\omega}^y)^k.$$

Afirmamos que

$$a_{xy} = \begin{cases} Q & se \quad x = y \\ \\ 0 & se \quad x \neq y \end{cases}$$

De fato, se $x = y$ temos que $a_{xy} = \sum_{k=0}^{Q-1} 1 = Q$. Por outro lado, se $x \neq y$ então temos dois casos a considerar:

1. Se $x > y$, podemos escrever

$$a_{xy} = \sum_{k=0}^{Q-1} (\omega^{x-y}(\omega\overline{\omega})^y)^k = \sum_{k=0}^{Q-1} \omega^{k(x-y)} = \frac{1-\omega^{Q(x-y)}}{1-\omega^{x-y}}$$

onde usamos que $\omega\overline{\omega} = |\omega|^2 = 1$. Além disso, como $0 < x - y \leq Q - 1$, temos que $\omega^{x-y} \neq 1$. Assim, $\omega^{Q(x-y)} = e^{2(x-y)\pi i} = 1$. Logo $a_{xy} = 0$.

2. Se $y > x$, podemos escrever

$$a_{xy} = \sum_{k=0}^{Q-1} ((\omega\overline{\omega})^x \overline{\omega}^{y-x})^k = \sum_{k=0}^{Q-1} \overline{\omega}^{k(y-x)} = \frac{1 - \overline{\omega}^{Q(y-x)}}{1 - \overline{\omega}^{y-x}}$$

onde usamos que $\omega\overline{\omega} = |\omega|^2 = 1$. Além disso, como $0 < y - x \leq Q - 1$, temos que $\overline{\omega}^{y-x} \neq 1$. Assim, $\overline{\omega}^{Q(x-y)} = \overline{e^{2(y-x)\pi i}} = 1$. Logo $a_{xy} = 0$.

Portanto, $A\overline{A} = I$ e isso mostra que A é unitária e, por conseguinte, $\mathsf{U}_{\mathcal{QF}}$ é uma transformação linear unitária.

2.6 O Algoritmo de Deutsch-Jozsa

Um algoritmo quântico sempre cumpre as seguintes etapas:

1. `Preparar um registro quântico com um número finito de qubits;`
2. `Colocar este registro em um estado puro conveniente;`
3. `Aplicar sucessivas e adequadas transformações unitárias de modo a deixar o registro em uma sobreposição de estados desejada;`
4. `Realizar a medição do registro`.

Portanto, o êxito de um algoritmo quântico está na razão direta da habilidade em escolher as transformações unitárias que devem atuar sucessivamente sobre o registro. Para ilustrar o poder de alcance da computação quântica e demonstrar o seu impacto sobre os métodos clássicos, escolhemos como aplicação inicial o algorimto de Deutsch-Jozsa por ser muito mais simples, para uma primeira abordagem, do que o algoritmo de fatoração de Shor, objeto desta dissertação. Assim, considere o seguinte problema:

Problema 2.6.1 *Seja $f : \{0,1\} \longrightarrow \{0,1\}$ uma função booleana que a cada valor de $x \in \{0,1\}$ associa um valor $f(x) \in \{0,1\}$. Suponha agora que desejamos calcular $f(0) \oplus f(1)$, onde $\oplus : \{0,1\}^2 \longrightarrow \{0,1\}$ é a função* **Xor** *apresentada no exemplo 1.2.1. Contudo, só é possível aplicar a função f apenas uma vez. Ou seja, se escolhermos conhecer o valor de $f(0)$, ficaremos sem saber o valor de $f(1)$ ou vice versa.*

Classicamente, esse é um problema impossível de ser solucionado, pois a função **Xor** não é reversível. Antes de prosseguir, apresentamos uma definição equivalente da função **Xor** que será mais útil para os nossos propósitos atuais.

Dessa forma, sejam $a, b \in \{0, 1\}$. Definimos:

$$a \oplus b = \begin{cases} 0 & se \quad a = b \\ \\ 1 & se \quad a \neq b \end{cases}$$

Esse problema, impossível de ser solucionado classicamente, é solúvel através de um algoritmo quântico. De acordo com a definição da função **Xor** dada acima, tudo que precisamos para resolver o problema é perguntar uma única vez se $f(0)$ é igual ou diferente de $f(1)$. Note que na computação clássica, tal pergunta é equivalente a usar a função f duas vezes, uma para cada valor de x. Mas isso é proibido pelas condições do problema. Portanto, quanticamente, resolveremos esse impasse fazendo a pergunta acima apenas uma vez, não para 0 e 1 separadamente, mas para uma sobreposição de estados: utilizaremos um registro quântico de dois qubits onde, inicialmente, o primeiro estará no estado $|0\rangle$ e o segundo no estado $|1\rangle$.

Para isso, considere as seguintes matrizes:

$$U = \begin{pmatrix} 1 - f(0) & f(0) & 0 & 0 \\ f(0) & 1 - f(0) & 0 & 0 \\ 0 & 0 & 1 - f(1) & f(1) \\ 0 & 0 & f(1) & 1 - f(1) \end{pmatrix}$$

$$H = \begin{pmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \end{pmatrix}$$

Note que para quaisquer escolhas de $f(0)$ e $f(1)$ a matriz U é unitária. Além disso, a matriz H também é unitária, pois $H^t H = I$. O algoritmo quântico que apresentaremos para solucionar o problema 2.6.1 utilizará os operadores unitários representados pelas matrizes acima e que agem sobre o espaço de Hilbert $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$. Isto é, trabalharemos com registros quânticosde dois qubits. Em linhas gerais, tal algoritmo terá o seguinte formato:

$$|\varphi_0\rangle = |01\rangle \xrightarrow{H} |\varphi_1\rangle = H|\varphi_0\rangle \xrightarrow{U} |\varphi_2\rangle = UH|\varphi_0\rangle \xrightarrow{H} |\varphi_3\rangle = HUH|\varphi_0\rangle.$$

Ao realizarmos a medição do estado $|\varphi_3\rangle$ obteremos a resposta do problema. Note que esta sequência está de acordo com as etapas de um algoritmo quântico descritas no início desta seção.

Algoritmo 2.6.1 (Algoritmo de Deutsch-Jozsa) *Considere os seguintes passos:*

1. `Inicialize com um registro quântico no estado puro` $|\varphi_0\rangle = |01\rangle$`;`
2. `Faça o registro evoluir sob a ação do operador` H`;`
3. `Faça o registro evoluir sob a ação do operador` U`;`
4. `Faça o registro evoluir sob a ação do operador` H`;`
5. `Realize a medição do sistema`.

Se $f(0) \oplus f(1) = 0$ *então o valor medido será sempre igual* 1 *pois o sistema irá colapsar necessariamente para o estado puro* $|01\rangle$*. Por outro lado, se* $f(0) \oplus f(1) = 1$ *então o valor medido será sempre igual a* 3 *pois o sistema irá colapsar necessariamente para o estado puro* $|11\rangle$*. Assim o algoritmo será sempre capaz de determinar o valor de* $f(0) \oplus f(1)$ *com apenas uma utilização do operador* U *ao qual está associada a função* f*.*

Demonstração. No passo 1, o sistema encontra-se no estado puro:

$$|\varphi_0\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Após o passo 2, o registro evolui para a sobreposição de estados:

$$H|\varphi_0\rangle = \begin{pmatrix} 1/2 \\ -1/2 \\ 1/2 \\ -1/2 \end{pmatrix}$$

Após o passo 3, o registro evolui para a seguinte sobreposição de estados, que depende de $f(0)$ e $f(1)$ em iguais proporções:

$$UH|\varphi_0\rangle = \begin{pmatrix} 1/2 - f(0) \\ -1/2 + f(0) \\ 1/2 - f(1) \\ -1/2 + f(1) \end{pmatrix}$$

Após o passo 4, o registro encontra-se no estado desejado para realizarmos a medição do sistema. Esse estado é:

$$HUH|\varphi_0\rangle = \begin{pmatrix} 0 \\ 1 - (f(0) + f(1)) \\ 0 \\ f(1) - f(0) \end{pmatrix}$$

Ou seja,

$$HUH|\varphi_0\rangle = 0 \cdot |00\rangle + [1 - (f(0) + f(1))] \cdot |01\rangle + 0 \cdot |10\rangle + [f(1) - f(0)] \cdot |11\rangle$$

De acordo com o postulado 2.2.3, ao realizarmos a medição de tal estado no passo 5, devemos obter apenas um dos seguintes resultados abaixo com respectivas probabilidades:

$$\begin{array}{lll} 0 & \text{com probabilidade} & Prob(0) = 0 \\ 1 & \text{com probabilidade} & Prob(1) = [1 - (f(0) + f(1))]^2 \\ 2 & \text{com probabilidade} & Prob(2) = 0 \\ 3 & \text{com probabilidade} & Prob(3) = (f(1) - f(0))^2 \end{array}$$

Ora, se $f(0) \oplus f(1) = 0$ significa que $f(0) = f(1)$. Logo, $Prob(3) = 0$ e o sistema deverá colapsar necessariamente para o estado puro $|01\rangle$ fornecendo o valor 1 com $Prob(1) = 1$. Por outro lado, se $f(0) \oplus f(1) = 1$ teremos $f(0) \neq f(1)$. Portanto, $Prob(1) = 0$ e o registro deverá colapsar necessariamente para o estado puro $|11\rangle$ fornecendo o valor 3 com $Prob(3) = 1$. Assim, em qualquer das duas possibilidades o computador quântico será capaz de solucionar o problema com probabilidade de acerto igual a 1. □

Capítulo 3

Algoritmo de Fatoração de Shor - parte I

Dados um número composto ímpar positivo N e um inteiro $1 < y < N$, escolhido aleatoriamente com $mdc\ (y, N) = 1$, a tarefa de encontrar um fator não trivial de N está em conexão com a de determinar o período r da função $f_N(a) = y^a (mod\ N)$, desde que r seja par e $y^{\frac{r}{2}} \not\equiv -1 (mod\ N)$. Mas a probabilidade dessa ocorrência é maior ou igual a $1 - \frac{1}{2^{k-1}}$, onde k é o número de fatores primos distintos de N.

Neste capítulo, abordamos o Algoritmo de Fatoração de Shor, a menos da parte quântica. O objetivo de tal abordagem é fornecer uma visão geral dos seus passos e consolidar a compreensão dos aspectos algébricos envolvidos nas demonstrações do seu funcionamento e da sua eficiência probabilística. Os passos do algoritmo são descritas na seção 3.2, após a apresentação dos conceitos iniciais. Com a finalidade de preservar a essencial compreensão do funcionamento do algoritmo, reservamos para o final a demonstração da sua eficiência probabilística, na qual estão presentes a maioria dos detalhes técnicos.

3.1 Conceitos Iniciais

Dado um número ímpar composto positivo N queremos fatorá-lo em uma decomposição da forma $N = n_1 \cdot n_2$, com $1 < n_i < N$. Ou seja, estamos interessados em encontrar um fator não trivial de N, onde denominamos de fator não trivial um divisor d de N diferente da unidade e do próprio N.

Por outro lado, considere a seguinte definição:

Definição 3.1.1 *Sejam y e N inteiros tais que $1 < y < N$ e $mdc\ (y, N) = 1$. Denomina-se a ordem de y módulo N ao menor inteiro positivo r tal que $y^r \equiv 1\ (mod\ N)$.*

Uma maneira equivalente de apresentar a ordem de um inteiro y módulo N é através do período de uma certa função definida sobre o conjunto dos números naturais. Assim, tendo em vista nossos objetivos imediatos, considere esta segunda maneira de apresentar a definição da ordem de um inteiro y.

Definição 3.1.2 *Sejam y e N inteiros tais que $1 < y < N$ e $mdc\ (y, N) = 1$. Considere agora a seguinte função:*

$$\begin{array}{rccc} f_N: & \mathbb{N} & \longrightarrow & \mathbb{N} \\ & a & \longmapsto & y^a (mod\ N). \end{array} \tag{3.1}$$

Definimos a ordem de y módulo N como o menor inteiro positivo r tal que $f_N(a+r) = f_N(a)$, para todo $a \in \mathbb{N}$. Isto é, r é o menor inteiro positivo tal que $f_N(r) = 1$.

Observação 3.1.1 *Note que embora o contradomínio de f_N seja $\mathbb{N}$, o seu conjunto imagem é finito e possui cardinalidade menor ou igual a N. Mais claramente, os elementos do conjunto imagem de f_N serão, no máximo, todos os restos $0, 1, ..., N-1$ da divisão por N.*

Agora, voltemos ao problema proposto no primeiro parágrafo desta seção, ou seja, encontrar um fator não trivial de um número ímpar composto positivo N. Especificamente, queremos mostrar a conexão existente entre o problema de encontrar um fator não trivial de N e o de encontrar o período da função definida em 3.1.2.

3.2 O Algoritmo de Fatoração de Shor

O Algoritmo de Fatoração de Shor aborda o problema de encontrar um fator não trivial de um número composto impar positivo N através de uma adequada utilização do período da função definida em 3.1.2. Em linhas gerais, o algoritmo substitui a fatoração direta de N pelo seguinte procedimento: escolha aleatoriamente um número inteiro $1 < y < N$ de modo que $mdc\ (y, N) = 1$ e calcule o período r da função $f_N(a) = y^a (mod\ N)$. Se esse período satisfizer ambas as condições:

$$\begin{array}{ll} (P1) & r\ par \\ (P2) & y^{\frac{r}{2}} \not\equiv -1\ (mod\ N) \end{array} \tag{3.2}$$

então será possível encontrar um fator não trivial de N.

Aqui, esclarecemos que o Algoritmo de Fatoração de Shor é composto essencialmente de 5 passos. Dentre esses, apenas o segundo passo, exatamente o que calcula o período de f_N, envolve computação de natureza quântica. Os quatro outros requerem computação clássica em tempo polinomial. A entrada do algoritmo é um número N inteiro, positivo, ímpar e composto. A saída são dois fatores não triviais de N. Para verificar se um número é primo ou composto existe um algoritmo polinomial [1], [2]. A seguir, apresentamos os passos do algoritmo de Shor.

3.2.1 Os Passos do Algoritmo de Shor

△ Entrada: um número inteiro positivo N composto e ímpar.
▽ Saída: um fator não trivial d de N.

Passo um

Escolha aleatoriamente $1 < y < N$ e calcule $mdc\ (y, N)$ através do algoritmo polinomial de Euclides.

Se $mdc\ (y, N) \neq 1$ **então** finalize com $d = mdc\ (y, N)$ sendo um fator não trivial de N.

Senão, vá para o passo dois.

Passo dois

Use uma computação quântica para calcular o período r da função

$$f_N(a) = y^a (mod\ N).$$

Passo três

Se r é ímpar, **então** volte ao passo um.

Senão, prossiga para o passo quatro.

Passo quatro

Como r é par, temos que

$$y^r - 1 = (y^{\frac{r}{2}} - 1)(y^{\frac{r}{2}} + 1) \equiv 0\ (mod\ N). \tag{3.3}$$

Observe que necessariamente $(y^{\frac{r}{2}} - 1) \not\equiv 0\ (mod\ N)$, pois r é a ordem de y módulo N. Dessa forma,

Se $(y^{\frac{r}{2}} + 1) \equiv 0\ (mod\ N)$, **então** volte ao passo um.

Senão, vá para o passo cinco.

Passo cinco

Desde que $(y^{\frac{r}{2}} + 1) \not\equiv 0\ (mod\ N)$, use o algoritmo polinomial de Euclides para calcular $d = mdc(y^{\frac{r}{2}} + 1, N)$. **Finalize** o algoritmo com d um fator não trivial de N.

3.2.2 Comentários sobre o algoritmo

Em primeiro lugar, observamos que a eficiência do algoritmo depende fortemente de que o inteiro $1 < y < N$ escolhido aleatoriamente possua uma ordem que satisfaça simultaneamente as condições (P1) e (P2). Do contrário, os passos terceiro ou quarto serão sistematicamente interrompidos, forçando a reinicialização do processo a partir do passo inicial. Felizmente, desde que N é um número ímpar, existe um resultado garantindo que para qualquer escolha aleatória do y com ordem respectiva igual a r, a probabilidade de r ser um número ímpar ou, sendo par, ocorrer $y^{\frac{r}{2}} \equiv -1(mod\ N)$ é menor ou igual a $\frac{1}{2^{k-1}}$, onde k é o número de fatores primos distintos de N. Tal resultado será demonstrado na próxima seção.

Uma vez tendo chegado ao Quinto passo, note que da equação (3.3) temos que N divide o produto $(y^{\frac{r}{2}}-1)(y^{\frac{r}{2}}+1)$ sem dividir qualquer dos dois fatores. Esta situação ideal permite que o inteiro ímpar possa ser decomposto em dois fatores $N = n_1 \cdot n_2$, isto é, N possui um fator não trivial. Essa afirmação está demonstrada a seguir.

Proposição 3.2.1 *Sejam N, a_1 e a_2 inteiros tais que $a_1 \cdot a_2 \equiv 0\ (mod\ N)$. Se $a_i \not\equiv 0\ (mod\ N)$ para $i = 1, 2$, então $mdc(a_i, N)$ é um fator não trivial de N.*

Demonstração. É imediato que $mdc\ (a_i, N) \neq N$, pois $a_i \not\equiv 0\ (mod\ N)$ para $i = 1, 2$. Agora, sem perda de generalidade, suponha que $mdc\ (a_1, N) = 1$. Como $N \mid a_1 \cdot a_2$ então necessariamente $N \mid a_2$, isto é, $a_2 \equiv 0\ (mod\ N)$, o que é uma contradição. E isso demonstra a proposição. □

3.3 Eficiência do Algoritmo

Nesta seção, demonstraremos o resultado mencionado no primeiro parágrafo da seção 3.2.2, essencialmente contido no teorema 3.3.1. Mas para demonstrá-lo, será necessário provar alguns resultados anteriores.

Lema 3.3.1 *Seja $G = \langle g \rangle$ um grupo cíclico de ordem n. Se $i \in \{1, 2, \cdots, n\}$, então,*

$$ord(g^i) = \frac{n}{mdc(i, n)}$$

Demonstração. Seja $r = ord(g^i)$. Então $(g^i)^r = e$, onde e denota o elemento neutro de G. Assim, $g^{ir} = e$ e isso implica que $n \mid ir$, pois $n = ord(g)$. Agora,

$$\frac{n}{mdc(i, n)} \mid \left(\frac{i}{mdc(i, n)}\right) r,$$

donde

$$\frac{n}{mdc(i,n)} \mid r.$$

Reciprocamente,

$$(g^i)^{\frac{n}{mdc(i,n)}} = (g^n)^{\frac{i}{mdc(i,n)}} = e.$$

Por conseguinte,

$$r \mid \frac{n}{mdc(i,n)}$$

e o resultado segue. □

Proposição 3.3.1 *Seja G um grupo cíclico com ordem n par. Considere o conjunto X_0 de números escritos na forma fatorada $2^t s$, onde $t \geq 0$ é um inteiro fixado e s um número ímpar qualquer. Considere agora o conjunto*

$$E = \{y \in G;\ ord(y) \in X_0\}.$$

Então, $\mid E \mid \leq \frac{n}{2}$, onde $\mid E \mid$ denota a cardinalidade de E.

Demonstração. Seja g um gerador de G. Como o grupo é cíclico então qualquer elemento $y \in G$ pode ser escrito como uma potência de g, isto é, $y = g^i$, para algum $i \in \{1, 2, ..., n\}$. Mais explicitamente, $G = \{g, g^2, ..., g^n\}$. Pelo lema 3.3.1,

$$ord(g^i) = \frac{n}{mdc(i,n)}$$

Agora, escreva a ordem de G na forma $n = 2^\tau \sigma$ onde σ é um número ímpar e τ é um inteiro positivo, isto é, $\tau > 0$ pois, por hipótese, n é um número par. Similarmente, escreva o expoente de g^i na forma $i = 2^\beta b$, onde $\beta \geq 0$ é um inteiro e b um número ímpar. Desse modo, $mdc(i,n) = 2^{min\{\beta,\tau\}} mdc(\sigma, b)$. Portanto,

$$ord(g^i) = 2^{\tau - min\{\beta,\tau\}} \frac{\sigma}{mdc(\sigma, b)} \tag{3.4}$$

onde $\frac{\sigma}{mdc(\sigma,b)}$ é um número ímpar. Por outro lado, estamos interessados nos elementos $y \in G$ que possuem ordem da forma $2^t s$. Mais claramente, queremos determinar os expoentes $i's$ que satisfazem:

$$ord(g^i) = 2^t s. \tag{3.5}$$

Ora, das equações (3.4) e (3.5) concluímos que tais valores $i's$ são aqueles cujo expoente β satisfaz a relação

$$\tau - min\{\beta, \tau\} = t \tag{3.6}$$

Desse modo, temos três casos distintos a considerar:

Caso 1: $t > \tau$.

Note que $\tau - min\{\beta, \tau\} \leq \tau$. Logo, por (3.6), $t \leq \tau$. Assim, $E = \emptyset$ se $t > \tau$ e, neste caso, $\mid E \mid = 0 \leq \frac{n}{2}$.

Caso 2: $0 < t \leq \tau$.

Decorre da condição (3.6) que devemos ter $\tau - min\{\beta, \tau\} > 0$, isto é, $\tau > min\{\beta, \tau\}$. Neste caso, necessariamente $\beta = min\{\beta, \tau\}$. Logo, os valores $i's$ para os quais $ord(g^i) = 2^t s$ são aqueles cujo expoente $\beta = \tau - t$, ou seja,

$$i = 2^{\tau - t} b \tag{3.7}$$

onde b é um número ímpar. Dessa forma, note que avaliar a cardinalidade do conjunto

$$E = \{y \in G;\ ord(y) \in X_0\}$$

reduz-se à tarefa de contar quantos elementos do conjunto de expoentes

$$\{1, 2, 3, ..., n = 2^\tau \sigma\} \tag{3.8}$$

possuem a forma dada por (3.7). Isto é, procuramos no conjunto (3.8) os múltiplos de $2^{\tau - t}$ cujo fator de multiplicação é um número ímpar. Tais elementos são, explicitamente:

$$2^{\tau - t} \cdot 1,\ 2^{\tau - t} \cdot 3,\ 2^{\tau - t} \cdot 5,\ 2^{\tau - t} \cdot 7,\ ...\ ,\ 2^{\tau - t} \cdot (2^t \sigma - 1) \tag{3.9}$$

onde afirmamos que o maior número ímpar b que multiplica a potência $2^{\tau - t}$ é igual a $(2^t \sigma - 1)$. Com efeito, podemos escrever $2^\tau \sigma = 2^t \sigma \cdot 2^{\tau - t}$, isto é, $n = 2^\tau \sigma$ é, ele próprio, o maior múltiplo de $2^{\tau - t}$ dentro do conjunto (3.8). Neste caso, note que o fator de multiplicação $2^t \sigma$ é par. Ora, o múltiplo antecessor deverá ter fator de multiplicação ímpar. Isso mostra que $(2^t \sigma - 1)$ é o maior número ímpar que multiplicado por $2^{\tau - t}$ permanece dentro do conjunto (3.8).

Agora, observe que a quantidade de elementos relacionados em (3.9) é exatamente a cardinalidade do conjunto

$$\{1, 3, 5, 7, ..., (2^t \sigma - 1)\}. \tag{3.10}$$

Finalmente, é imediata a verificação de que o número de elementos de (3.10) é igual a $\frac{2^t \sigma}{2}$. Portanto, como $0 \leq \beta = \tau - t$, segue que

$$\mid E \mid = 2^{t-1} \sigma = \frac{2^\tau \sigma}{2^{\beta + 1}} \leq \frac{n}{2}$$

CASO 3: $t = 0$.

Com esta condição, segue de (3.6) que $\tau = min\{\beta, \tau\}$ e isso implica que $\beta \geq \tau$. Assim, estamos interessados nos expoentes $i = 2^{\beta}b \in \{1, 2, ..., n = 2^{\tau}\sigma\}$, onde b é ímpar e $\beta \geq \tau > 0$, pois n é par por hipótese. Desse modo, note que a quantidade de tais expoentes será, no máximo, a quantidade de números pares entre 1 e $n = 2^{\tau}\sigma$, pois todos os expoentes ímpares estão excluídos. Portanto,

$$| E | \leq \frac{n}{2}$$

como queríamos demonstrar. □

Lema 3.3.2 *Seja $N > 1$ um inteiro com fatoração prima $N = p_1^{\alpha_1} \ ... \ p_k^{\alpha_k}$. Sejam ainda y, a, e b inteiros, com $a \geq 0$. Então $y^a \equiv b \pmod{N}$ se e somente se $y^a \equiv b \pmod{p_i^{\alpha_i}}$, para cada $i = 1, ..., k$.*

DEMONSTRAÇÃO. Se $y^a \equiv b \pmod{N}$ então $N \mid y^a - b$ e isso implica que $p_i^{\alpha_i} \mid y^a - b$, pois $p_i^{\alpha_i} \mid N$. Logo, $y^a \equiv b \pmod{p_i^{\alpha_i}}$, para cada $i = 1, ..., k$. Reciprocamente, suponha que $y^a \equiv b \pmod{p_i^{\alpha_i}}$, para todo $i = 1, ..., k$. Então $p_i^{\alpha_i} \mid y^a - b$. Ocorre que $mdc(p_i, p_j) = 1$ se $i \neq j$. Logo, $N = p_1^{\alpha_1} \ ... \ p_k^{\alpha_k} \mid y^a - b$, ou seja, $y^a \equiv b \pmod{N}$. □

Teorema 3.3.1 *Seja N um inteiro positivo composto ímpar com fatoração*

$$N = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \ ... \ p_k^{\alpha_k},$$

onde $3 \leq p_1 < ... < p_k$ são números primos e $\alpha_1, ..., \alpha_k$ inteiros positivos. Suponha que y é um inteiro escolhido aleatoriamente no intervalo $1 < y < N$ tal que $mdc(y, N) = 1$. Seja r a ordem de y módulo N. Então,

$$Prob\ (r\ \acute{e}\ \acute{i}mpar \ \ ou \ \ y^{\frac{r}{2}} \equiv -1 \pmod{N}) \leq \frac{1}{2^{k-1}}$$

.

DEMONSTRAÇÃO. Como $mdc(y, N) = 1$ temos que $mdc(y, p_i^{\alpha_i}) = 1$ para cada $i = 1, ..., k$. Seja r_i a ordem de y módulo $p_i^{\alpha_i}$, isto é, r_i é o menor inteiro positivo que satisfaz:

$$y^{r_i} \equiv 1 \pmod{p_i^{\alpha_i}}. \tag{3.11}$$

Afirmamos que

$$r = mmc\ (r_1,\ r_2,\ ...\ ,\ r_k). \tag{3.12}$$

De fato, como $y^r \equiv 1 \pmod{N}$ então, pelo lema 3.3.2, $y^r \equiv 1 \pmod{p_i^{\alpha_i}}$. Logo, $r_i \mid r$ para cada $i = 1, ..., k$, isto é, r é um múltiplo comum de $r_1,\ r_2,\ ...\ ,\ r_k$. Por conseguinte, $mmc\ (r_1,\ r_2,\ ...\ ,\ r_k) \mid r$. Reciprocamente, $y^{mmc\ (r_1,\ r_2,\ ...\ ,\ r_k)} \equiv 1 \pmod{p_i^{\alpha_i}}$, pois

$r_i \mid mmc\ (r_1,\ r_2,\ \ldots\ ,\ r_k)$ para cada i. Assim, novamente pelo lema 3.3.2, temos que $y^{mmc\ (r_1,\ r_2,\ \ldots\ ,\ r_k)} \equiv 1 \pmod N$ e isso implica que $r \mid mmc\ (r_1,\ r_2,\ \ldots\ ,\ r_k)$. Portanto, (3.12) segue.

Agora, escreva as ordens de y módulo $p_i^{\alpha_i}$ na forma:

$$r_i = s_i \cdot 2^{t_i} \tag{3.13}$$

onde s_i é um número ímpar e $t_i \geq 0$. Além disso, tome $s = mmc(s_1,\ s_2,\ \ldots,\ s_k)$ e $M = max\ (t_1,\ t_2,\ \ldots,\ t_k)$. Assim, decorre de (3.12) que

$$r = s \cdot 2^M \tag{3.14}$$

Mais ainda, note que r é ímpar se e somente se $t_i = 0$, para todo $i = 1, \ldots, k$. Isto é,

$$Prob\ (r\ impar)\ =\ Prob\ (t_1 = t_2 = \cdots = t_k = 0) \tag{3.15}$$

Por outro lado, suponha que r seja par. Em primeiro lugar, afirmamos que

$$y^{\frac{r}{2}} \equiv -1 \pmod{p_i^{\alpha_i}}\ implica\ que\ t_i = M. \tag{3.16}$$

De fato, suponha por absurdo que $t_i < M$. Isso implica, por (3.13) e (3.14), que $r_i \mid \frac{r}{2}$ e como r_i é a ordem de y módulo $p_i^{\alpha_i}$ então $y^{\frac{r}{2}} \equiv 1 \pmod{p_i^{\alpha_i}}$, o que é uma contradição.

Agora, caso $y^{\frac{r}{2}} \equiv -1 \pmod N$, então, pelo lema 3.3.2, $y^{\frac{r}{2}} \equiv -1 \pmod{p_i^{\alpha_i}}$ para $i = 1, \ldots, k$. Como a implicação (3.16) vale para cada i, segue que

$$Prob\ (y^{\frac{r}{2}} \equiv -1 \pmod N)\ \leq\ Prob\ (t_1 = t_2 = \cdots = t_k = M) \tag{3.17}$$

Portanto, das expressões (3.15) e (3.17), decorre que

$$Prob\ (r\ impar\ ou\ y^{\frac{r}{2}} \equiv -1 \pmod N)\ \leq\ Prob\ (t_i's = 0\ ou\ t_i's = M).$$

Resta mostrar que

$$Prob\ (t_i's = 0\ ou\ t_i's = M) \leq \frac{1}{2^{k-1}}$$

Para isso, precisamos verificar que

$$Prob(t_i = t) \leq \frac{1}{2}$$

onde $t \geq 0$ é um inteiro fixo. Com efeito, considere o grupo multiplicativo

$$\mathbb{Z}^*_{p_i^{\alpha_i}} = \{\overline{y} \in \mathbb{Z}_{p_i^{\alpha_i}};\ mdc\ (y, p_i^{\alpha_i}) = 1\}$$

Encontramos em [15](p.124,125) uma demonstração de que $\mathbb{Z}^*_{p_i^{\alpha_i}}$ é um grupo cíclico pois, por hipótese, $p_i^{\alpha_i}$ é necessariamente uma potência de ímpar. Além disso, a ordem de $\mathbb{Z}^*_{p_i^{\alpha_i}}$ é par. De fato,

$$\mid \mathbb{Z}^*_{p_i^{\alpha_i}} \mid = \phi(p_i^{\alpha_i}) = p_i^{\alpha_i - 1} \cdot (p_i - 1)$$

é um número par, pois $3 \leq p_i$ é ímpar; onde ϕ é a função de Euler, ou seja, $\phi(p_i^{\alpha_i})$ é o número de inteiros positivos menores que $p_i^{\alpha_i}$ relativamente primos com $p_i^{\alpha_i}$.

Agora, note que se $mdc\ (y, N) = 1$, com $1 < y < N$, então $\overline{y} \in \mathbb{Z}^*_{p_i^{\alpha_i}}$. Mais ainda, se $r_i = s_i \cdot 2^{t_i}$ é a ordem de y módulo $p_i^{\alpha_i}$, temos que no grupo $\mathbb{Z}^*_{p_i^{\alpha_i}}$:

$$ord(\overline{y}) = r_i$$

Ademais, considere o conjunto

$$E = \{\overline{y} \in \mathbb{Z}^*_{p_i^{\alpha_i}}\ ;\ ord\ (\overline{y})\ tem\ a\ forma\ 2^t \cdot s\}$$

Dessa forma,

$$Prob\ (t_i = t) = \frac{\mid E \mid}{\phi(p_i^{\alpha_i})} \tag{3.18}$$

onde $\mid E \mid$ denota a cardinalidade de E.

Ocorre que pela proposição 3.3.1:

$$\mid E \mid \leq \frac{\phi(p_i^{\alpha_i})}{2}$$

Portanto, $Prob\ (t_i = t) \leq \frac{1}{2}$. Em particular, $Prob\ (t_i = 0) \leq \frac{1}{2}$ e $Prob\ (t_i = M) \leq \frac{1}{2}$.

Em resumo, dado um inteiro y escolhido aleatoriamente no intervalo $1 < y < N$ tal que $mdc(y, N) = 1$, temos que $r_i = s_i \cdot 2^{t_i}$ é a ordem de y módulo $p_i^{\alpha_i}$. Além disso, $\overline{y} \in \mathbb{Z}^*_{p_i^{\alpha_i}}$ e $ord(\overline{y}) = ord(y)$. Dessa forma, a probabilidade de $t_i = t$ nada mais é do que a probabilidade da ordem de $\overline{y} \in \mathbb{Z}^*_{p_i^{\alpha_i}}$ ser da forma $2^t \cdot s$ que, como vimos, é dada por (3.18).

Assim,

$$Prob\ (t_i's = 0\ ou\ t_i's = M) = \prod_{i=1}^{k} Prob(t_i = 0) + \prod_{i=1}^{k} Prob(t_i = M) \leq \frac{1}{2^{k-1}}$$

e isso conclui a demonstração do teorema. $\square$

Capítulo 4

Algoritmo de Fatoração de Shor - parte II

Como vimos no capítulo três, dado um número inteiro positivo N composto e ímpar, o problema de encontrar um fator não trivial de N reduz-se ao problema de descobrir o período r da função f_N, definida em 3.1.2. Por conveniência, passaremos a denotar o período dessa função por P ao invés de r como fizemos até agora. Assim, neste capítulo, solucionaremos o problema referente ao passo 2 do algoritmo de Shor, aqui apresentado na seguinte formulação:

Problema 4.0.1 *Considere uma função periódica*

$$\begin{array}{rcl} f : \mathbb{N} & \longrightarrow & \mathbb{N} \\ x & \longmapsto & f(x) \end{array}$$

tal que $0 \leq f(x) \leq N-1$*, onde* N *é um inteiro positivo, composto e ímpar . Encontre o período* P *da função* f*. Isto é,* P *é o menor inteiro positivo tal que* $f(x+P) = f(x)$*, para todo* $x \in \mathbb{N}$*.*

No caso em que $f = f_N$, note que $P < N$. De fato, P é o período da função f_N ou, de modo equivalente, é a ordem de $y^a \ (mod\ N)$, onde $mdc\ (y, N) = 1$. Assim, $P \leq \phi(N)$, onde ϕ é a função de Euler, ou seja, $\phi(N)$ é o número de inteiros positivos menores que N relativamente primos com N. Ocorre que $\phi(N) < N$, donde segue que $P < N$.

A solução deste problema será realizada em duas partes. Na primeira, utilizaremos um algoritmo quântico cuja finalidade será fornecer como resultado um certo número natural y, ao qual estará associada um certa probabilidade $Prob\ (y)$. Este resultado carregará intrinsecamente o fato de que f é uma função periódica de período P. Na segunda parte, voltaremos a utilizar um algorítmo clássico polinomial para extrair o valor de P a partir do resultado y.

4.1 A Parte Quântica do Algoritmo de Shor

O algoritmo quântico que soluciona o problema 4.0.1 seguirá as mesmas quatro etapas descritas no início da seção 2.6. Desse modo, apresentamos na sequência cada uma dessas etapas, enumerando-as como desdobramentos do passo 2 do algoritmo de Shor.

Passo 2.1 `Preparar um registro quântico com um número finito de qubits`

Neste algoritmo, usaremos dois registros quânticos: o primeiro deverá armazenar os valores de x sobre os quais aplicaremos a função periódica f e o segundo deverá armazenar os valores de $f(x)$. Assim, precisamos calcular a quantidade de qubits que deverá possuir cada um desses registros. Como $x \in \mathbb{N}$ devemos estabelecer um subconjunto finito de $\mathbb{N}$ de tal forma que a restrição de f a esse subconjunto seja suficiente para a avaliação do período P.

A priori, tome k um inteiro positivo tal que $N^2 \leq 2^k < 2N^2$. Afirmamos que tal k sempre existe e é único. De fato, considere o conjunto $X = \{x \in \mathbb{R};\ x > 1\}$, o qual pode ser escrito como a união disjunta de intervalos

$$\bigcup_{k=1}^{\infty} (2^{k-1},\ 2^k].$$

Como $N > 1$, então existe um único k tal que $2^{k-1} < N^2 \leq 2^k$. Agora, multiplicando esta desigualdade por 2, obtemos $2^k < 2N^2 \leq 2^{k+1}$. Portanto, $N^2 \leq 2^k < 2N^2$. Com isso, faça $Q = 2^k$ e considere o conjunto

$$S_Q = \{0,\ 1,\ 2, \cdots,\ Q-1\}. \tag{4.1}$$

Afirmamos que a restrição de f ao conjunto finito $S_Q \subset \mathbb{N}$ é suficiente para os propósitos deste algoritmo. Entretanto, somente nas demonstrações dos resultados da seção 4.2, poderemos justificar a escolha de um tal Q, satisfazendo

$$N^2 \leq Q < 2N^2. \tag{4.2}$$

Assim, como o primeiro registro quântico deverá ser capaz de conter os inteiros não negativos entre 0 e $Q-1$, inclusive, então serão necessários $k = \log_2 Q$ qubits. (cf. observação 1.1.2).

Por sua vez, o segundo registro deverá possuir $m = \lceil \log_2 N \rceil$ qubits para conter os valores de $f(x)$. De fato, um tal registro será capaz de armazenar os inteiros não negativos $0,\ 1,\ \cdots,\ N-1$, satisfazendo a condição $0 \leq f(x) \leq N-1$.

Portanto, com tais valores k e m, sejam

$$\mathcal{H}_1 = \underbrace{\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2}_{k} \text{ e } \mathcal{H}_2 = \underbrace{\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2}_{m}$$

com bases canônicas ortonormais, respectivamente, $\mathcal{C}_1 = \{|0\rangle,\ |1\rangle,\ \cdots,\ |2^k - 1\rangle\}$ e $\mathcal{C}_2 = \{|0\rangle,\ |1\rangle,\ \cdots,\ |2^m - 1\rangle\}$. Dessa forma, o espaço de estados sobre os quais aplicaremos os operadores unitários do algoritmo será:

$$\mathcal{H}_1 \otimes \mathcal{H}_2$$

cujos elementos serão representados pelas seguintes notações equivalentes:

$$|x\rangle|y\rangle = |x, y\rangle = |x\rangle \otimes |y\rangle$$

onde $|x\rangle \in \mathcal{H}_1$ é o vetor de estado de um registro quântico de k qubits e $|y\rangle \in \mathcal{H}_2$ é o vetor de estado de um registro quântico de m qubits.

Passo 2.2 `Colocar este registro em um estado puro conveniente`

Para alcançar o resultado que pretendemos com a aplicação dos operadores apresentados na próxima etapa, o estado puro favorável com qual devemos iniciar o algorimto é o estado $|\psi_0\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ dado por:

$$|\psi_0\rangle = |0\rangle|0\rangle$$

Passo 2.3 `Aplicar sucessivas e adequadas transformações unitárias de modo a deixar o registro em uma sobreposição de estados desejada`

A escolha do estado puro inicial $|\psi_0\rangle = |0\rangle|0\rangle$ vem ao encontro dos nossos objetivos ao utilizar a transformada quântica de Fourier $\mathsf{U}_{\mathcal{QF}}$ para criar um estado de sobreposição uniforme e em seguida o operador U_f para computar a função f através de uma porta quântica. Denotando por I o operador identidade de $\mathcal{H}_2$, o diagrama abaixo representa a sequência de operações que realizaremos nesta etapa:

$$|\psi_0\rangle \overset{\mathsf{U}_{\mathcal{QF}}\otimes\mathsf{I}}{\longrightarrow} |\psi_1\rangle \overset{\mathsf{U}_f}{\longrightarrow} |\psi_2\rangle \overset{\mathsf{U}_{\mathcal{QF}}\otimes\mathsf{I}}{\longrightarrow} |\psi_3\rangle$$

Assim, em primeiro lugar, aplicamos a transformada quântica de Fourier apenas ao primeiro registro com a finalidade de criar neste um estado de sobreposição uniforme, ou seja:

$$|\psi_0\rangle = |0\rangle|0\rangle \quad \overset{\mathsf{U}_{\mathcal{QF}}\otimes\mathsf{I}}{\longmapsto} \quad \mathsf{U}_{\mathcal{QF}}|0\rangle \otimes \mathsf{I}|0\rangle = \left(\frac{1}{\sqrt{Q}} \sum_{x=0}^{Q-1} |x\rangle\right)|0\rangle = |\psi_1\rangle.$$

Neste momento, o primeiro registro contém a sobreposição uniforme de todos os inteiros $0,\ 1,\ ,\cdots,\ Q-1$. Note que tal sobreposição só foi alcançada porque o primeiro registro, sobre o qual agiu o operador $\mathsf{U}_{\mathcal{QF}}$, estava no estado $|0\rangle \in \mathcal{H}_1$.

Em seguida, aplicamos ao estado $|\psi_1\rangle$ o operador unitário U_f definido da secão 2.17, fazendo o sistema evoluir para o estado $|\psi_2\rangle$, isto é:

$$\begin{aligned}
|\psi_1\rangle \overset{\mathsf{U}_f}{\longmapsto} |\psi_2\rangle &= \mathsf{U}_f\left(\left(\tfrac{1}{\sqrt{Q}}\sum_{x=0}^{Q-1}|x\rangle\right)|0\rangle\right)\\
&= \mathsf{U}_f\left(\tfrac{1}{\sqrt{Q}}\sum_{x=0}^{Q-1}|x\rangle|0\rangle\right)\\
&= \tfrac{1}{\sqrt{Q}}\sum_{x=0}^{Q-1}\mathsf{U}_f(|x\rangle|0\rangle)\\
&= \tfrac{1}{\sqrt{Q}}\sum_{x=0}^{Q-1}|x\rangle|f(x)\rangle
\end{aligned}$$

onde usamos a linearidade do operador U_f e a equação 2.19. Note que nesta passagem foi indispensável a nossa escolha inicial do segundo registro também no estado puro $|0\rangle \in \mathcal{H}_2$. Finalmente, aplicamos a transformada quântica de Fourier ao primeiro registro obtendo o estado $|\psi_3\rangle$ da seguinte forma:

$$\begin{aligned}
|\psi_2\rangle \overset{\mathsf{U}_{\mathcal{QF}}\otimes\mathsf{I}}{\longmapsto} |\psi_3\rangle &= (\mathsf{U}_{\mathcal{QF}}\otimes\mathsf{I})\left(\tfrac{1}{\sqrt{Q}}\sum_{x=0}^{Q-1}|x\rangle|f(x)\rangle\right)\\
&= \tfrac{1}{\sqrt{Q}}\sum_{x=0}^{Q-1}(\mathsf{U}_{\mathcal{QF}}\otimes\mathsf{I})(|x\rangle|f(x)\rangle)\\
&= \tfrac{1}{\sqrt{Q}}\sum_{x=0}^{Q-1}\mathsf{U}_{\mathcal{QF}}|x\rangle\otimes\mathsf{I}|f(x)\rangle\\
&= \tfrac{1}{\sqrt{Q}}\sum_{x=0}^{Q-1}\left(\frac{1}{\sqrt{Q}}\sum_{y=0}^{Q-1}\omega^{xy}|y\rangle\right)|f(x)\rangle\\
&= \tfrac{1}{Q}\sum_{x=0}^{Q-1}\sum_{y=0}^{Q-1}\omega^{xy}|y\rangle|f(x)\rangle\\
&= \tfrac{1}{Q}\sum_{y=0}^{Q-1}\left(|y\rangle\sum_{x=0}^{Q-1}\omega^{xy}|f(x)\rangle\right)\\
&= \tfrac{1}{Q}\sum_{y=0}^{Q-1}|y\rangle|\varphi(y)\rangle
\end{aligned}$$

onde $|\varphi(y)\rangle = \sum_{x=0}^{Q-1} \omega^{xy} |f(x)\rangle$, com $\omega = e^{\frac{2\pi i}{Q}}$.

Agora, quando $|\varphi(y)\rangle$ é diferente do vetor nulo, podemos reescrever o estado $|\psi_3\rangle$ da seguinte forma:

$$|\psi_3\rangle = \sum_{y=0}^{Q-1} \frac{\| \, |\varphi(y)\rangle \, \|}{Q} |y\rangle \frac{|\varphi(y)\rangle}{\| \, |\varphi(y)\rangle \, \|} \tag{4.3}$$

Com isso, note que $\frac{|\varphi(y)\rangle}{\| \, |\varphi(y)\rangle \, \|}$ pertence a $\mathcal{H}_2$ e, consequentemente, $\frac{\| \, |\varphi(y)\rangle \, \|}{Q} |y\rangle$ pertence a $\mathcal{H}_1$.

Passo 2.4 `Realizar a medição do registro`

Ao finalizar a etapa anterior, o sistema quântico formado pelos dois registros encontra-se no estado dado pela equação 4.3. Neste ponto, estamos apenas interessados na medição do primeiro registro. Assim, de acordo com o postulado 2.2.3, ao ser medido, este deverá colapsar para um dos estados puros $|y_0\rangle \in \mathcal{H}_1$, tal que

$$y_0 \in \{0, \ 1, \ 2, \ \cdots, \ Q-1\}$$

fornecendo como resultado o valor y_0 com probabilidade:

$$Prob \ (y_0) = \frac{\| \, |\varphi(y_0)\rangle \, \|^2}{Q^2}.$$

Embora não possua utilidade para nossos propósitos, o segundo registro evoluirá com a medição do primeiro para a sobreposição $\frac{|\varphi(y_0)\rangle}{\| \, |\varphi(y_0)\rangle \, \|}$ de tal forma que todo o sistema será levado ao estado final $|\psi_3'\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ dado por:

$$|\psi_3'\rangle = |y_0\rangle \frac{|\varphi(y_0)\rangle}{\| \, |\varphi(y_0)\rangle \, \|}.$$

Com isso, a parte quântica do algoritmo de Shor cumpre a sua finalidade. Ou seja, o objetivo dos passos 2.1 a 2.4 foi criar uma distribuição de probabilidade no espaço amostral

$$S_Q = \{0, \ 1, \ 2, \cdots, \ Q-1\}$$

onde o evento no qual estamos interessados é a ocorrência de um número $y \in S_Q$ com probabilidade:

$$Prob \ (y) = \frac{\| \, |\varphi(y)\rangle \, \|^2}{Q^2}, \tag{4.4}$$

onde

$$|\varphi(y)\rangle = \sum_{x=0}^{Q-1} \omega^{xy} |f(x)\rangle$$

com $\omega = e^{\frac{2\pi i}{Q}}$. Assim, damos por encerrada a parte quântica do algoritmo de Shor. Nossa próxima tarefa será mostrar sob quais condições podemos extrair o período P da função f a partir da ocorrência desse valor y, utilizando para isso métodos clássicos em tempo polinomial.

4.2 O Cálculo do Período P

Nem todos os valores de $y \in S_Q$ permitirão calcular o período P da função f. Na verdade, existirá um subconjunto $Y \subset S_Q$ constituído de P elementos tal que apenas será possível calcular o período P da função f se o valor y obtido na etapa anterior satisfizer $y \in Y$. Diante disso, temos duas tarefas bem distintas a cumprir:

1. Apresentar o conjunto Y e, supondo que $y \in Y$, mostrar **como** é possível calcular o valor P usando um algoritmo polinomial;

2. Mostrar a **probabilidade** de ocorrer $y \in Y$. Em outras palavras, mostrar que sendo f uma função períodica de período P, então a probabilidade dada pela equação 4.4 assegura que ocorrerá um resultado $y \in Y$ se a parte quântica do algorimo for repetida uma quantidade polinomial $\mathcal{O}(\log_2 log_2 N)$ de vezes, onde N é o número que desejamos decompor em dois fatores não triviais.

Cumpriremos cada uma dessas tarefas, respectivamente, nas duas próximas seções.

4.2.1 A Extração do Período por Frações Contínuas

Para que possamos definir o conjunto Y, acima referido, será necessário apresentar o seguinte resultado:

Lema 4.2.1 *Sejam $x, n \in \mathbb{N}$, com $n > 0$. Então existe um único número inteiro, denotado por $[x]_n$, que satisfaz:*

$$\begin{cases} x \equiv [x]_n \ (mod\ n) \\ \\ -\frac{n}{2} \leq [x]_n < \frac{n}{2} \end{cases} \tag{4.5}$$

Demonstração. Sejam q e r os únicos inteiros positivos tais que $x = qn + r$, onde $0 \leq r < n$. Temos dois casos a considerar.

1. Se $0 \leq r < \frac{n}{2}$, então $[x]_n = r$ satisfaz as condições (4.5). De fato, $x \equiv r \ (mod\ n)$ e $-\frac{n}{2} \leq 0 \leq r < \frac{n}{2}$.

2. Se $\frac{n}{2} \leq r < n$, podemos subtrair n dos membros desta desigualdade, obtendo $-\frac{n}{2} \leq r - n < 0$. Agora, faça
$$[x]_n = r - n.$$
Assim, temos que $x \equiv r - n \ (mod\ n)$, pois $r - n \equiv r \ (mod\ n)$ e $x \equiv r \ (mod\ n)$. Além disso, $-\frac{n}{2} \leq r - n < 0 < \frac{n}{2}$.

Para mostrar a unicidade, considere $[x]_n$ e $[x]_n^{'}$ satisfazendo as condições (4.5). Assim, temos que $[x]_n^{'} \equiv [x]_n \ (mod\ n)$, isto é, $[x]_n^{'} - [x]_n \equiv 0 \ (mod\ n)$. Por outro lado, suponha, sem perda de generalidade, que $[x]_n^{'} \geq [x]_n$. Isso implica que $0 \leq [x]_n^{'} - [x]_n < n$. Portanto, necessariamente $[x]_n^{'} - [x]_n = 0$, ou seja, $[x]_n^{'} = [x]_n$. □

Com o resultado do lema 4.2.1 podemos apresentar a definição do conjunto Y. Assim:

$$Y = \{y \in S_Q| \ -\frac{P}{2} \leq [Py]_Q < \frac{P}{2}\} \tag{4.6}$$

Note que o período P ainda é desconhecido. Entretanto, a sua existência já é suficiente para que possamos definir o conjunto Y como acima. A partir de agora, iniciaremos o trabalho de explicitá-lo. Para isso, mostraremos na sequência a seguinte proposição:

Proposição 4.2.1 *O conjunto Y possui P elementos.*

DEMONSTRAÇÃO. Inicialmente, considere os múltiplos Py de P, onde $0 \leq y \leq Q - 1$, isto é,
$$0,\ P,\ 2P,\ \cdots,\ P(Q-1)$$
Queremos determinar quantos múltiplos de Q existem entre 0 e $P(Q-1)$, inclusive. Ou seja, queremos encontrar o maior j tal que
$$jQ \leq P(Q-1)$$
Com efeito, devemos ter $jQ \leq PQ - P$, isto é, $j \leq P - \frac{P}{Q}$. Como j é inteiro, temos que
$$j \leq \left\lfloor P - \frac{P}{Q} \right\rfloor = P - 1,$$
pois $0 < \frac{P}{Q} < 1$. Assim, existem P múltiplos de Q entre 0 e $P(Q-1)$:
$$0,\ Q,\ 2Q,\ \cdots,\ Q(P-1)$$
Para ver que Y possui P elementos utilizaremos a seguinte idéia: mostraremos que existem P intervalos disjuntos I_j tais que $y \in Y$ se e somente se $Py \in I_j$ para algum $j = 0,\ 1,\ 2,\ \cdots,\ P-1$. Mais ainda, para todo j exite um e somente um $y \in Y$ tal que $Py \in I_j$.

Assim, para $j \in \{0,\ 1,\ 2, \cdots,\ P-1\}$ seja

$$I_j = \left[jQ - \frac{P}{2},\ jQ + \frac{P}{2}\right).$$

Note que $I_j \cap I_k = \emptyset$ se $j \neq k$. De fato, como $P < Q$ então $\frac{P}{2} < \frac{Q}{2}$. Também, observe que se $a, b \in I_j$ então $|a-b| < P$. Primeiro, mostraremos que $y \in Y$ se e somente se $Py \in I_j$ para algum $j \in \{0,\ 1,\ 2, \cdots,\ P-1\}$. Com efeito, suponha que $y \in Y$, ou seja

$$-\frac{P}{2} \leq [Py]_Q < \frac{P}{2} \tag{4.7}$$

Como $Py \equiv [Py]_Q\ (mod\ Q)$, então $Py - [Py]_Q = jQ$, isto é, $Py = jQ + [Py]_Q$ para algum $j \in \{0,\ 1,\ 2,\ \cdots,\ P-1\}$. Somando jQ à desigualdade (4.7), obtemos:

$$jQ - \frac{P}{2} \leq jQ + [Py]_Q < jQ + \frac{P}{2}.$$

Ou seja,

$$jQ - \frac{P}{2} \leq Py < jQ + \frac{P}{2}.$$

Portanto, se $y \in Y$ então necessariamente $Py \in I_j$ para algum $j \in \{0,\ 1,\ 2, \cdots,\ P-1\}$. Reciprocamente, suponha que para algum j, $0 \leq j \leq P-1$, exite um múltiplo Py de P tal que

$$jQ - \frac{P}{2} \leq Py < jQ + \frac{P}{2}.$$

Ou seja,

$$-\frac{P}{2} \leq Py - jQ < \frac{P}{2}.$$

Observe que $Py \equiv Py - jQ\ (mod\ Q)$ e, além disso, $-\frac{Q}{2} \leq Py - jQ < \frac{Q}{2}$, pois $\frac{P}{2} < \frac{Q}{2}$. Assim, pela unicidade do número $[Py]_Q$, temos que $Py - jQ = [Py]_Q$, isto é, $-\frac{P}{2} \leq [Py]_Q < \frac{P}{2}$. Logo, $y \in Y$.

Finalmente, mostraremos que para todo $j \in \{0,\ 1,\ 2,\ \cdots,\ P-1\}$ exite um e somente um $y \in Y$ tal que $Py \in I_j$. De fato, suponha que para algum $0 \leq j \leq P-1$ não existe $Py \in I_j$. Neste caso, seja Pk o maior múltiplo de P tal que $Pk < jQ - \frac{P}{2}$. Logo, o múltiplo consecutivo deverá satisfazer $jQ + \frac{P}{2} \leq P(k+1)$ pois estamos supondo que não existe múltiplo de P pertencente a I_j. Dessa forma, se P é um número par então teremos a seguinte sequência de desigualdades

$$Pk \leq jQ - \frac{P}{2} - 1 < jQ + \frac{P}{2} \leq P(k+1).$$

Ora, nesse caso a diferença entre estes dois multiplos consecutivos de P será maior ou igual que $P+1$, um absurdo. Por outro lado, se P é ímpar, então teremos a seguinte sequência de desigualdades

$$Pk < jQ - \frac{P}{2} < jQ + \frac{P}{2} < P(k+1),$$

haja vista que Pk e $P(k+1)$ são números inteiros. Nesse caso, a distância entre estes dois múltiplos consecutivos de P será estritamente maior que P; novamente, uma contradição. Portanto, para todo $j \in \{1,\ 2,\ \cdots,\ P-1\}$, existe $y \in Y$ tal que $Py \in I_j$.

Para mostrar a unicidade, suponha que para um mesmo j existem $y_1,\ y_2 \in Y$ com, digamos, $y_2 > y_1$ tais que $Py_1,\ Py_2 \in I_j$. Nesse caso, temos por um lado que $P \leq P(y_2 - y_1)$, mas por outro lado $Py_2 - Py_1 < P$, uma contradição. Portanto, o conjunto Y possui P elementos. □

Para extrair o período P através do uso de frações contínuas será necessário definir uma certa bijeção entre Y e o conjunto $S_P = \{0,\ 1,\ 2,\ \cdots,\ P-1\}$. Mas para apresentar tal bijeção precisamos do seguinte resultado:

Lema 4.2.2 *Considere a função arredondamento $\mathcal{A} : \mathbb{R}_+ \longrightarrow \mathbb{N}$ definida por*

$$\mathcal{A}(x) = \left\lfloor x + \frac{1}{2} \right\rfloor$$

onde $\mathbb{R}_+$ é o conjunto dos números reais não negativos. Então, para $a \in \mathbb{N}$,

$$[a]_Q = a - Q\mathcal{A}\left(\frac{a}{Q}\right)$$

Demonstração. $\mathcal{A}(x)$ é o único número natural tal que

$$\mathcal{A}(x) \leq x + \frac{1}{2} < \mathcal{A}(x) + 1.$$

Em particular, se $x = \frac{a}{Q}$ temos:

$$\mathcal{A}\left(\frac{a}{Q}\right) \leq \frac{a}{Q} + \frac{1}{2} < \mathcal{A}\left(\frac{a}{Q}\right) + 1.$$

Multiplicando esta desigualdade por Q, obtemos:

$$Q\mathcal{A}\left(\frac{a}{Q}\right) \leq a + \frac{Q}{2} < Q\mathcal{A}\left(\frac{a}{Q}\right) + Q.$$

Donde,

$$-\frac{Q}{2} \leq a - Q\mathcal{A}\left(\frac{a}{Q}\right) < \frac{Q}{2},$$

onde somamos $-Q\mathcal{A}\left(\frac{a}{Q}\right) - \frac{Q}{2}$ aos membros da desigualdade anterior. Agora, observe que

$$a \equiv a - Q\mathcal{A}\left(\frac{a}{Q}\right)\ (mod\ Q)$$

Logo, pela unicidade do número $[a]_Q$ (cf. lema 4.2.1), temos que

$$[a]_Q = a - Q\mathcal{A}\left(\frac{a}{Q}\right)$$

como queríamos demonstrar. □

Proposição 4.2.2 *Considere os conjuntos Y e $S_P = \{0,\ 1,\ 2,\ \cdots,\ P-1\}$. A função $d : Y \longrightarrow S_P$ definida por*

$$d(y) = \mathcal{A}\left(\frac{P}{Q}y\right)$$

é uma bijeção. Além disso, podemos escrever $[Py]_Q = Py - Qd(y)$.

Demonstração. Como o conjunto Y possui P elementos, basta mostrar que d é injetiva e que $0 \leq d(y) < P$. Com efeito, sejam $y_1, y_2 \in Y$ com, digamos, $y_1 < y_2$ e suponha que $d(y_1) = d(y_2)$. Pelo lema 4.2.2, $[Py]_Q = Py - Qd(y)$. Logo,

$$Py_1 - [Py_1]_Q = Py_2 - [Py_2]_Q,$$

donde $P \leq P(y_2 - y_1) = [Py_2]_Q - [Py_1]_Q < P$, uma contradição. Por conseguinte, d é injetiva. Agora, vamos mostrar que $0 \leq d(y) < P$. De fato, $d(0) = 0$. Também,

$$\begin{array}{rcl} d(y) & = & \frac{Py-[Py]_Q}{Q} \\ & \leq & \frac{P(Q-1)+\frac{P}{2}}{Q} \\ & < & \frac{P(Q-1)+P}{Q} \\ & = & P. \end{array}$$

Portanto, d é uma bijeção, como afirmamos. □

Até este momento, todo o trabalho realizado nesta seção teve como propósito apresentar dois resultados fundamentais:

1. Definir o conjunto $Y = \{y \in S_Q|\ -\frac{P}{2} \leq [Py]_Q < \frac{P}{2}\}$ com cardinalidade igual a P;
2. Definir a bijeção $d : Y \longrightarrow S_P$ e, em particular, mostrar a sua relação com o número $[Py]_Q$ através da igualdade $[Py]_Q = Py - Qd(y)$.

Para mantermos em vista os nossos objetivos principais, lembramos que no início da seção 4.2 afirmamos que nem todos os valores de $y \in S_Q$ permitiriam avaliar o período P da função f, mas somente aqueles que pertencessem ao subconjunto $Y \subset S_Q$. Note que, até agora, temos trabalhado apenas com a hipótese de que existe um período P e, por esta razão, tal período participou apenas intrinsecamente das definições e resultados. Agora, com a posse dos dois itens resumidos acima, estamos em condições de mostrar como efetivamente extrair o valor de P a partir do resultado y obtido na parte quântica do

algoritmo. Para isso, utilizaremos um resultado clássico de frações contínuas conhecido como *Teorema de Lagrange*, que pode ser encontrado em [8, seção 10.15, Teorema 184] e [15, pp 156-157]. Com esse objetivo, faremos agora uma breve apresentação das frações contínuas.

Frações contínuas

O algoritmo polinomial de Euclides para se determinar o máximo divisor comum entre dois números pode ser utilizado para se converter um número racional em uma representação denominada por fração contínua. Dessa forma, seja $\xi = \frac{p}{q}$ um número racional positivo. Este número pode ser representado pela expressão finita

$$\xi = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\cdots + \cfrac{1}{a_n}}}}}$$

onde $a_0 = \lfloor \xi \rfloor$ é um inteiro não negativo e $a_1, a_2, \cdots, a_n$ são inteiros positivos. Se além disso, acrescentamos a condição $a_n > 1$, então tal representação de ξ em frações contínuas é única. Por simplicidade de notação a expressão acima é usualmente denotada por

$$[a_0, a_1, a_2, \cdots, a_n].$$

Agora, para $0 \leq k \leq n$ a fração $\xi_k = [a_0, a_1, a_2, \cdots, a_k]$ é denominada o k-ésimo convergente da fração contínua $[a_0, a_1, a_2, \cdots, a_n]$. Evidentemente, o n-ésimo convergente é a própria fração contínua. Cada convergente pode ser expresso na forma

$$\xi_k = \frac{p_k}{q_k}$$

onde p_k e q_k são inteiros relativamente primos e, além disso, obedecem a seguinte recorrência:

$$\begin{array}{lll} p_0 = a_0, & p_1 = a_1 a_0 + 1, & p_k = a_k p_{k-1} + p_{k-2}, \\ q_0 = 1, & q_1 = a_1, & q_k = a_k q_{k-1} + q_{k-2}. \end{array}$$

Maiores detalhes sobre frações contínuas podem ser encontrados em [15, Capítulo 8], [8, Cap.X] e [12, Seção 13.7].

Teorema 4.2.1 (Lagrange) *Seja ξ um número real e sejam a e b inteiros com $b > 0$. Se*

$$\left| \xi - \frac{a}{b} \right| \leq \frac{1}{2b^2},$$

então o número racional $\frac{a}{b}$ é um convergente da expansão de ξ em frações contínuas.

Como uma corolário do Teorema de Lagrange, temos o resultado mais importante desta seção, ou seja, aquele que fornece uma condição necessária para extrair o valor de P.

Corolário 4.2.1 *Se $y \in Y$ então o número racional $\frac{d(y)}{P}$ é um convergente da expansão de $\frac{y}{Q}$ em frações contínuas.*

Demonstração. Se $y \in Y$ então $|[Py]_Q| \leq \frac{P}{2}$ e como $Py - Qd(y) = [Py]_Q$, temos que

$$|Py - Qd(y)| \leq \frac{P}{2}.$$

Dividindo ambos os membros da desigualdade acima por PQ, obtemos

$$\left|\frac{y}{Q} - \frac{d(y)}{P}\right| \leq \frac{1}{2Q}.$$

Ocorre que, pela equação 4.2, $Q \geq N^2$ e portanto

$$\left|\frac{y}{Q} - \frac{d(y)}{P}\right| \leq \frac{1}{2N^2}.$$

Agora, como $P < N$ implica $P \leq N$, podemos escrever esta última desigualdade como

$$\left|\frac{y}{Q} - \frac{d(y)}{P}\right| \leq \frac{1}{2P^2}.$$

Assim, pelo Teorema de Lagrange, $\frac{d(y)}{P}$ é um convergente da expansão de $\xi = \frac{y}{Q}$ em frações contínuas. □

Uma vez que conhecemos y e Q então através de um algoritmo polinomial, podemos calcular os convergentes $\frac{p_n}{q_n}$ da expansão de $\frac{y}{Q}$ em frações contínuas. O corolário acima assegura que para algum n devemos ter:

$$\frac{d(y)}{P} = \frac{p_n}{q_n} \tag{4.8}$$

Isso significa que para cada n devemos testar se q_n é o período da função f_N. Isto é, devemos verificar se $f_N(q_n) = 1$. Contudo, resta ainda uma condição a ser satisfeita. Se $\frac{p_n}{q_n}$ são convergentes da expansão de um número racional em frações contínuas então p_n e q_n são relativamente primos. Portanto, a condição (4.8) não é suficiente para assegurar a extração de P. Na verdade a recorrência polinomial da expansão de $\frac{y}{Q}$ em frações contínuas fornecerá o valor de P se e somente se para algum n

$$\begin{cases} p_n = d(y) \\ q_n = P \end{cases}$$

Em outras palavras, devemos ter $mdc\ (d(y), P) = 1$. Sobre esta condição trataremos na próxima seção.

4.2.2 Considerações Probabilísticas do Algoritmo Quântico

Vimos no final da seção 4.1 que a finalidade da parte quântica do algoritmo foi criar uma distribuição de probabilidade no espaço amostral

$$S_Q = \{0,\ 1,\ 2, \cdots,\ Q-1\}$$

onde o evento de interesse é a ocorrência de um número $y \in S_Q$ com probabilidade:

$$Prob\ (y) = \frac{\parallel |\varphi(y)\rangle \parallel^2}{Q^2},$$

onde

$$|\varphi(y)\rangle = \sum_{x=0}^{Q-1} \omega^{xy} |f(x)\rangle$$

com $\omega = e^{\frac{2\pi i}{Q}}$. Essa distribuição de probabilidade carrega intrinsecamente o fato de que f é uma função periódica de período P. Além disso, vimos na seção anterior que apenas será possível extrair o valor de P através da expansão de $\frac{y}{Q}$ em frações contínuas se o valor y obtido na parte quântica do algoritmo satisfizer simultaneamante:

$$\begin{cases} y \in Y \\ mdc\ (d(y), P) = 1 \end{cases}$$

Nosso principal objetivo nesta seção é mostrar que dentre os $y \in S_Q$ encontraremos um $y \in Y$ com $mdc\ (d(y), P) = 1$ se a parte quântica do algoritmo de Shor for repetida uma quantidade polinomial $\mathcal{O}(\log_2 log_2 N)$ de vezes. Para isso, começamos por demonstrar o teorema seguinte que fornece a distribuição de probabilidade em termos do período P.

Lema 4.2.3

$$\left|e^{i\theta} - 1\right|^2 = 4\sin^2\left(\frac{\theta}{2}\right)$$

Demonstração.

$$\left|e^{i\theta} - 1\right|^2 = (\cos\theta - 1)^2 + \sin^2\theta = 2 \cdot (1 - \cos\theta) = 4\sin^2\left(\frac{\theta}{2}\right)$$

onde usamos o fato que $\sin^2\left(\frac{\theta}{2}\right) = \frac{1-\cos(\theta)}{2}$. □

Teorema 4.2.2 *Sejam q e r os únicos inteiros não negativos tais que $Q = Pq + r$, onde $0 \leq r < P$; e seja $Q_0 = Pq$. Então:*

$$Prob(y) = \begin{cases} \frac{r\sin^2\left(\frac{\pi Py}{Q}\cdot\left(\frac{Q_0}{P}+1\right)\right)+(P-r)\sin^2\left(\frac{\pi Py}{Q}\cdot\frac{Q_0}{P}\right)}{Q^2\sin^2\left(\frac{\pi Py}{Q}\right)} & se \quad Py \not\equiv 0(modQ) \\ \\ \frac{r(Q_0+P)^2+(P-r)Q_0^2}{Q^2P^2} & se \quad Py \equiv 0(modQ) \end{cases}$$

DEMONSTRAÇÃO. A idéia principal desta demonstração é particionar o conjunto S_Q em P classes formadas pelos números que divididos por P deixam restos $0,\ 1,\ 2,\ \cdots,\ P-1$. Inicialmente, dividimos o seguinte somatório:

$$|\varphi(y)\rangle = \sum_{x=0}^{Q-1} \omega^{xy}|f(x)\rangle = \sum_{x=0}^{Q_0-1} \omega^{xy}|f(x)\rangle + \sum_{x=Q_0}^{Q-1} \omega^{xy}|f(x)\rangle$$

Agora, note que cada $x \in \{0,\ 1,\ 2, \cdots,\ Q-1\}$ pode ser escrito de forma única como $x = Px_1 + x_0$, onde x_1 e x_0 são, respectivamente, o quociente e o resto da divisão de x por P. Com isso, podemos escrever:

$$|\varphi(y)\rangle = \sum_{x_0=0}^{P-1} \sum_{x_1=0}^{\frac{Q_0}{P}-1} \omega^{(Px_1+x_0)y}|f(Px_1+x_0)\rangle + \sum_{x_0=0}^{r-1} \omega^{\left(P\frac{Q_0}{P}+x_0\right)y}|f(P\frac{Q_0}{P}+x_0)\rangle.$$

Usando o fato que f é periódica de período P, temos que $f(Px_1 + x_0) = f(x_0)$. Por conseguinte,

$$|\varphi(y)\rangle = \sum_{x_0=0}^{P-1} \omega^{x_0y} \left(\sum_{x_1=0}^{\frac{Q_0}{P}-1} \omega^{Pyx_1} \right) |f(x_0)\rangle + \sum_{x_0=0}^{r-1} \omega^{x_0y}\omega^{Py\left(\frac{Q_0}{P}\right)}|f(x_0)\rangle.$$

Como $0 \leq r < P$, podemos separar o primeiro somatório acima, obtendo:

$$|\varphi(y)\rangle = \sum_{x_0=0}^{r-1} \omega^{x_0y} \left(\sum_{x_1=0}^{\frac{Q_0}{P}-1} \omega^{Pyx_1} \right) |f(x_0)\rangle + \sum_{x_0=r}^{P-1} \omega^{x_0y} \left(\sum_{x_1=0}^{\frac{Q_0}{P}-1} \omega^{Pyx_1} \right) |f(x_0)\rangle +$$

$$\sum_{x_0=0}^{r-1} \omega^{x_0y}\omega^{Py\left(\frac{Q_0}{P}\right)}|f(x_0)\rangle.$$

Assim, agrupando o primeiro e terceiro somatórios, temos:

$$|\varphi(y)\rangle = \sum_{x_0=0}^{r-1} \omega^{x_0y} \left(\sum_{x_1=0}^{\frac{Q_0}{P}} \omega^{Pyx_1} \right) |f(x_0)\rangle + \sum_{x_0=r}^{P-1} \omega^{x_0y} \left(\sum_{x_1=0}^{\frac{Q_0}{P}-1} \omega^{Pyx_1} \right) |f(x_0)\rangle$$

Agora, como f é periódica de período p, então os vetores

$$|f(0)\rangle,\ |f(1)\rangle,\ |f(2)\rangle, \cdots, |f(P-1)\rangle$$

são todos distintos e, portanto, ortonormais. Assim,

$$\begin{aligned}
\||\varphi(y)\rangle\|^2 &= \sum_{x_0=0}^{r-1}\left|\omega^{x_0 y}\left(\sum_{x_1=0}^{\frac{Q_0}{P}}\omega^{Pyx_1}\right)\right|^2 + \sum_{x_0=r}^{P-1}\left|\omega^{x_0 y}\left(\sum_{x_1=0}^{\frac{Q_0}{P}-1}\omega^{Pyx_1}\right)\right|^2 \\
&= \left|\sum_{x_1=0}^{\frac{Q_0}{P}}\omega^{Pyx_1}\right|^2 \cdot \sum_{x_0=0}^{r-1}|\omega^{x_0 y}|^2 + \left|\sum_{x_1=0}^{\frac{Q_0}{P}-1}\omega^{Pyx_1}\right|^2 \cdot \sum_{x_0=r}^{P-1}|\omega^{x_0 y}|^2 \\
&= r\left|\sum_{x_1=0}^{\frac{Q_0}{P}}\omega^{Pyx_1}\right|^2 + (P-r)\left|\sum_{x_1=0}^{\frac{Q_0}{P}-1}\omega^{Pyx_1}\right|^2
\end{aligned}$$

Portanto, se $Py \equiv 0\ (mod\ Q)$ então $\omega^{Pyx_1} = 1$, pois $\omega = e^{\frac{2\pi i}{Q}}$. Logo,

$$\||\varphi(y)\rangle\|^2 = r\left(\frac{Q_0}{P}+1\right)^2 + (P-r)\left(\frac{Q_0}{P}\right)^2.$$

Donde

$$\frac{\||\varphi(y)\rangle\|^2}{Q^2} = \frac{r(Q_0+P)^2 + (P-r)Q_0^2}{Q^2P^2}$$

Por outro lado, se $Py \not\equiv 0\ (mod\ Q)$, então $\omega^{Py} \neq 1$ e com isso podemos usar a soma de uma série geométrica de modo que:

$$\begin{aligned}
\||\varphi(y)\rangle\|^2 &= r\left|\frac{\omega^{Py\cdot\left(\frac{Q_0}{P}+1\right)}-1}{\omega^{Py}-1}\right|^2 + (P-r)\left|\frac{\omega^{Py\cdot\left(\frac{Q_0}{P}\right)}-1}{\omega^{Py}-1}\right|^2 \\
& r\left|\frac{e^{\frac{2\pi i}{Q}\cdot Py\cdot\left(\frac{Q_0}{P}+1\right)}-1}{e^{\frac{2\pi i}{Q}\cdot Py}-1}\right|^2 + (P-r)\left|\frac{e^{\frac{2\pi i}{Q}\cdot Py\cdot\left(\frac{Q_0}{P}\right)}-1}{e^{\frac{2\pi i}{Q}\cdot Py}-1}\right|^2
\end{aligned}$$

Finalmente, pelo lema 4.2.3, temos:

$$\frac{\||\varphi(y)\rangle\|^2}{Q^2} = \frac{r\sin^2\left(\frac{\pi Py}{Q}\cdot\left(\frac{Q_0}{P}+1\right)\right) + (P-r)\sin^2\left(\frac{\pi Py}{Q}\cdot\frac{Q_0}{P}\right)}{Q^2\sin^2\left(\frac{\pi Py}{Q}\right)}$$

e isso demonstra o teorema. □

Corolário 4.2.2 *Se P é um divisor exato de Q, então:*

$$Prob(y) = \begin{cases} 0 & se \quad Py \not\equiv 0(modQ) \\ \\ \frac{1}{P} & se \quad Py \equiv 0(modQ) \end{cases}$$

DEMONSTRAÇÃO. Basta fazer $r = 0$ e $Q = Q_0$ na expressão do teorema. □

Proposição 4.2.3 *Seja $y \in \{0, 1, \cdots, Q-1\}$. Então*

$$Prob(y) \geq \begin{cases} \frac{4}{\pi^2} \cdot \frac{1}{P} \cdot \left(1 - \frac{1}{N}\right)^2 & se \quad 0 < |[Py]_Q| \leq \frac{P}{2} \cdot \left(1 - \frac{1}{N}\right) \\ \\ \frac{1}{P} \cdot \left(1 - \frac{1}{N}\right)^2 & se \quad [Py]_Q = 0 \end{cases}$$

DEMONSTRAÇÃO. Inicialmente, mostraremos que

$$\left| \frac{\pi[Py]_Q}{Q} \cdot \left(\frac{Q_0}{P} + 1 \right) \right| < \frac{\pi}{2}$$

Assim:

$$\begin{array}{rcl} \left| \frac{\pi[Py]_Q}{Q} \cdot \left(\frac{Q_0}{P} + 1 \right) \right| & = & \left| \frac{\pi[Py]_Q}{Q} \cdot \left(\frac{Q_0+P}{P} \right) \right| \\ & \leq & \frac{\pi}{Q} \cdot \frac{P}{2} \left(1 - \frac{1}{N}\right) \cdot \left(\frac{Q_0+P}{P} \right) \\ & = & \frac{\pi}{2} \cdot \left(1 - \frac{1}{N}\right) \cdot \left(\frac{Q_0+P}{Q} \right) \\ & \leq & \frac{\pi}{2} \cdot \left(1 - \frac{1}{N}\right) \cdot \left(\frac{Q+P}{Q} \right) \\ & = & \frac{\pi}{2} \cdot \left(1 - \frac{1}{N}\right) \cdot \left(1 + \frac{P}{Q} \right) \end{array}$$

onde usamos o fato de que $Q_0 \leq Q$. Agora, como $P \leq N$ e $N^2 \leq Q < 2N^2$, então $\frac{P}{Q} \leq \frac{N}{N^2}$ e, por conseguinte,

$$\begin{array}{rcl} \left| \frac{\pi[Py]_Q}{Q} \cdot \left(\frac{Q_0}{P} + 1 \right) \right| & = & \frac{\pi}{2} \cdot \left(1 - \frac{1}{N}\right) \cdot \left(1 + \frac{N}{N^2}\right) \\ & = & \frac{\pi}{2} \cdot \left(1 - \frac{1}{N^2}\right) \\ & < & \frac{\pi}{2} \end{array}$$

Com isso, é imediato que também

$$\left| \frac{\pi[Py]_Q}{Q} \cdot \frac{Q_0}{P} \right| < \frac{\pi}{2}$$

Agora, seja $\theta_1 = \frac{\pi Py}{Q} \cdot \left(\frac{Q_0}{P} + 1 \right)$ e $\theta_2 = \frac{\pi[Py]_Q}{Q} \cdot \left(\frac{Q_0}{P} + 1 \right)$. Dessa forma, temos que

$$\begin{array}{rcl} \theta_1 - \theta_2 & = & \frac{\pi}{Q} \cdot \left(\frac{Q_0}{P} + 1 \right) \cdot (Py - [Py]_Q) \\ & = & \frac{\pi}{Q} \cdot \left(\frac{Q_0}{P} + 1 \right) \cdot kQ \\ & = & k \left(\frac{Q_0}{P} + 1 \right) \pi \end{array}$$

Logo, $\theta_1 - \theta_2$ é um múltiplo inteiro de π, pois $k \left(\frac{Q_0}{P} + 1 \right)$ é um número inteiro, uma vez que $P \mid Q_0$ e, além disso, k é um inteiro proveniente do fato que $Py \equiv [Py]_Q \ (mod\ Q)$. Assim, como a função $\sin^2$ é periódica de período π, então se

$$0 < |[Py]_Q| \leq \frac{P}{2} \cdot \left(1 - \frac{1}{N}\right),$$

podemos escrever:

$$Prob\ (y) = \frac{r \sin^2\left(\frac{\pi[Py]_Q}{Q} \cdot \left(\frac{Q_0}{P} + 1\right)\right) + (P-r)\sin^2\left(\frac{\pi[Py]_Q}{Q} \cdot \frac{Q_0}{P}\right)}{Q^2 \sin^2\left(\frac{\pi[Py]_Q}{Q}\right)}.$$

Usando a desigualdade trigonométrica

$$\frac{4}{\pi^2}\theta^2 \leq \sin^2\theta \leq \theta^2$$

para $|\theta| \leq \frac{\pi}{2}$, obtemos:

$$\begin{aligned} Prob\ (y) &\geq \frac{r\frac{4}{\pi^2}\left(\frac{\pi[Py]_Q}{Q}\cdot\left(\frac{Q_0}{P}+1\right)\right)^2 + (P-r)\frac{4}{\pi^2}\left(\frac{\pi[Py]_Q}{Q}\cdot\frac{Q_0}{P}\right)^2}{Q^2\left(\frac{\pi[Py]_Q}{Q}\right)^2} \\ &\geq \frac{r\frac{4}{\pi^2}\left(\frac{\pi[Py]_Q}{Q}\cdot\left(\frac{Q_0}{P}\right)^2\right) + (P-r)\frac{4}{\pi^2}\left(\frac{\pi[Py]_Q}{Q}\cdot\frac{Q_0}{P}\right)^2}{Q^2\left(\frac{\pi[Py]_Q}{Q}\right)^2}. \end{aligned}$$

Ou seja,

$$\begin{aligned} Prob\ (y) &\geq \frac{P\frac{4}{\pi^2}\left(\frac{\pi[Py]_Q}{Q}\cdot\left(\frac{Q_0}{P}\right)\right)^2}{Q^2\left(\frac{\pi[Py]_Q}{Q}\right)^2} \\ &= \frac{4}{\pi^2}\frac{P\left(\frac{Q_0}{P}\right)^2}{Q^2} \\ &= \frac{4}{\pi^2}\cdot\frac{1}{P}\cdot\left(\frac{Q-r}{Q}\right)^2 \\ &= \frac{4}{\pi^2}\cdot\frac{1}{P}\cdot\left(1-\frac{r}{Q}\right)^2 \\ &\geq \frac{4}{\pi^2}\cdot\frac{1}{P}\cdot\left(1-\frac{1}{N}\right)^2 \end{aligned}$$

onde usamos que $\frac{r}{Q} \leq \frac{N}{N^2}$, pois $r < P < N$ e $N^2 \leq Q$. Agora, se $[Py]_Q = 0$ temos que

$$\begin{aligned} Prob\ (y) = \frac{r(Q_0+P)^2+(P-r)Q_0^2}{Q^2P^2} &\geq \frac{rQ_0^2+(P-r)Q_0^2}{Q^2P^2} \\ &= \frac{PQ_0^2}{Q^2P^2} \\ &= \frac{1}{P}\left(\frac{Q_0}{Q}\right)^2 \\ &= \frac{1}{P}\left(\frac{Q-r}{Q}\right)^2 \\ &= \frac{1}{P}\left(1-\frac{r}{Q}\right)^2 \\ &\geq \frac{1}{P}\left(1-\frac{N}{N^2}\right)^2 \\ &= \frac{1}{P}\left(1-\frac{1}{N}\right)^2 \end{aligned}$$

e assim concluímos a demonstração da proposição. □

O significado relevante da proposição que acabamos de demonstrar pode ser destacado da seguinte maneira: a probabilidade de obtermos como resultado da parte quântica do algoritmo um $y \in S_Q$ tal que $|[Py]_Q| < \frac{P}{2}$ é limitada inferiormente pela expressão:

$$Prob\ (y) \geq \frac{4}{\pi^2} \cdot \frac{1}{P} \cdot \left(1 - \frac{1}{N}\right)^2 \tag{4.9}$$

Note que se y satisfaz $|[Py]_Q| < \frac{P}{2}$ então $y \in Y$ e, nesse caso, asseguramos os resultados obtidos na seção anterior. Contudo, ainda é necessário acrescentar a condição que $mdc\ (d(y), P) = 1$ para que possamos extrair a o período P pela expansão em frações contínuas do número racional $\frac{y}{Q}$. Para isso, mostraremos a seguinte proposição:

Proposição 4.2.4 *A probabilidade que o valor aleatório $y \in S_Q$ produzido na parte quântica do algoritmo de Shor seja tal que $d(y)$ e P sejam relativamente primos é limitada inferiormente pela seguinte expressão:*

$$Prob\{y \in Y |\ mdc\ (d(y), P) = 1\} \geq \frac{4}{\pi^2} \cdot \frac{\phi(P)}{P} \cdot \left(1 - \frac{1}{N}\right)^2$$

onde ϕ é a função de Euler, ou seja, $\phi(P)$ é o número de inteiros positivos menores que P relativamente primos com P.

Demonstração. Para cada $y \in S_Q$ temos que $Prob\ (y)$ é limitada inferiormente pela expressão (4.9). Como o conjunto $Y \subset S_Q$ possui P elementos a probabilidade que $y \in Y$ satisfaz

$$\begin{aligned} Prob\ \{y \in Y\} &= \sum_{i=1}^{P} Prob\ (y) \\ &\geq \sum_{i=1}^{P} \frac{4}{\pi^2} \cdot \frac{1}{P} \cdot \left(1 - \frac{1}{N}\right)^2 \\ &= \tfrac{4}{\pi^2} \cdot \tfrac{1}{P} \cdot \left(1 - \tfrac{1}{N}\right)^2 \cdot \sum_{i=1}^{P} 1 \\ &= \tfrac{4}{\pi^2} \cdot \tfrac{1}{P} \cdot \left(1 - \tfrac{1}{N}\right)^2 \cdot P \\ &= \tfrac{4}{\pi^2} \cdot \left(1 - \tfrac{1}{N}\right)^2. \end{aligned}$$

Agora, cada $y \in Y$ está em bijeção com um $d(y) \in \{0, 1, \cdots, P-1\}$. Além disso,

$$Prob\ \{mdc(d(y), P) = 1\} = \frac{\phi(P)}{P}.$$

Consequentemente

$$Prob\{y \in Y |\ mdc\ (d(y), P) = 1\} \geq \frac{4}{\pi^2} \cdot \frac{\phi(P)}{P} \cdot \left(1 - \frac{1}{N}\right)^2$$

como queríamos demonstrar. □

Finalmente, mostraremos que dentre os $y \in S_Q$ encontraremos um $y \in Y$ com $mdc\,(d(y), P) = 1$ se a parte quântica do algoritmo de Shor for repetida uma quantidade polinomial $\mathcal{O}(\log_2 log_2 N)$ de vezes. Para isso, faremos uso do seguinte teorema clássico, que pode ser encontrado em [8, seção 18.4, Teorema 328].

Teorema 4.2.3

$$\liminf \frac{\phi(N)}{N/\ln\ln N} = e^{-\gamma}$$

onde $\gamma = 0,57721566490153286061...$ *denota a constante de Euler.*

Como corolário, temos:

Corolário 4.2.3

$$Prob\{y \in Y | mdc\,(d(y), P) = 1\} = \Omega\left(\frac{1}{\log_2 \log_2 N}\right)$$

Ou seja, se a medição do estado final $|\psi_3\rangle$ *for repetida* $\mathcal{O}(\log_2 log_2 N)$ *vezes, então a probabilidade de êxito é* $\Omega(1)$.

Demonstração. Pelo teorema acima, temos que

$$\frac{\phi(P)}{P/\ln\ln P} \geq e^{-\gamma} - \epsilon(P),$$

ou seja,

$$\frac{\phi(P)}{P} \geq \frac{e^{-\gamma} - \epsilon(P)}{\ln\ln P}$$

onde $\epsilon(P)$ uma sequência monotônica decrescente de números reais positivos convergindo para zero. Assim,

$$\frac{\phi(P)}{P} \geq \frac{e^{-\gamma} - \epsilon(P)}{\ln\ln P} \geq \frac{e^{-\gamma} - \epsilon(P)}{\ln\ln N} = \frac{e^{-\gamma} - \epsilon(P)}{\ln\left(\frac{\log_2 N}{\log_2 e}\right)},$$

donde

$$\frac{\phi(P)}{P} \geq \frac{e^{-\gamma} - \epsilon(P)}{\ln(\log_2 N) - \ln\log_2 e} = \frac{e^{-\gamma} - \epsilon(P)}{\frac{\log_2(\log_2 N)}{\log_2 e} - \ln\log_2 e}.$$

Isto é,

$$\frac{\phi(P)}{P} \geq \frac{e^{-\gamma} - \epsilon(P)}{\ln 2 \cdot \log_2(\log_2 N) - \ln\log_2 e} \geq \frac{e^{-\gamma} - \epsilon(P)}{\ln 2} \frac{1}{\log_2 \log_2 N}.$$

Pela proposição 4.2.4, temos que

$$Prob\{y \in Y | mdc\ (d(y), P) = 1\} \geq \frac{4(e^{-\gamma} - \epsilon(P))}{\pi^2 \ln 2} \cdot \frac{1}{\log_2 \log_2 N} \cdot \left(1 - \frac{1}{N}\right)^2.$$

Portanto, assintoticamente,

$$Prob\{y \in Y | mdc\ (d(y), P) = 1\} = \Omega\left(\frac{1}{\log_2 \log_2 N}\right),$$

onde $\Omega\left(\frac{1}{\log_2 \log_2 N}\right)$ denota o conjunto das funções de ordem mínima $\frac{1}{\log_2 \log_2 N}$. Para maiores detalhes sobre esta notação sugerimos [22, Capítulo 1, seção 4]. □

Referências Bibliográficas

[1] AGRAWAL, M., KAYAL, N., SAXENA, N. *PRIMES is in P*. Department of Computer Science & Engineering, Indian Institute of Technology Kanpur, (2002) http://www.cse.iitk.ac.in/news/primality.pdf.

[2] BERNSTEIN, D.J. *Proving Primality after Agrawal-Kayal-Saxena*. The University of Illinois at Chicago, (2003), http://cr.yp.to/papers/aks.pdf

[3] BERTHIAUME, A. *Quantum Computation*. Centrum voor Wiskunde en Informatica, Amsterdam.

[4] COUTINHO, S. C. *Números Inteiros e Criptografia RSA*. Série de Computação Matemática. Rio de Janeiro: IMPA/SBM (1997).

[5] DEUTSCH, D. *Quantum Theory, the Church-Turing principle and the universal quantum computer*. Proceedings of the Royal Society of London, Ser.A A400, 97-117 (1985).

[6] EKERT, A., JOZSA, R. *Quantum Computation and Shor's Factoring Algorithm*. Review Modern Physics, **68** (1996), pp 733-753.

[7] FENNER, S. A. *A Physics-Free Introduction to the Quantum Computation Model*. LANL e-print arXiv:cs.CC/0304008 v1 (2003).

[8] HARDY, G.H., WRIGHT, E.M. *An Introduction to the Theory of Numbers*. 4 ed. Oxford University Press, (1960).

[9] LENSTRA, A.K., LENSTRA JR., H.W., *The Development of the Number Field Sieve*. Lecture Notes in Mathematics, Vol. 1554, Springer-Velag (1993).

[10] LENSTRA, A.K., LENSTRA JR., H.W., MANASSE, M.S, POLLARD, J.M. *The Number Field Sieve*. Proc. 22nd Annual ACM Symposium on Theory of Computing. pp 564-572. New York (1990)

[11] LOMONACO JR., S.J. *Shor's Quantum Factoring Algorithm*. Symposia in Applied Mathematics (2000).

[12] LOMONACO JR., S.J. *A Rosetta Stone for Quantum Mechanics with an Introduction to Quantum Computation*. Annual Meeting of AMS, Washington (2000).

[13] LUCCHESI, C.L. *Introdução à Criptografia Computacional*. Campinas: UNICAMP (1986).

[14] RIEFFEL, E., POLAK, W. *An Introduction to Quantum Computation for Non-Physicists*. LANL e-print arXiv:quant-ph/9809016 v2 (2000).

[15] SANTOS, J.P.O. *Introdução à Teoria dos Números*. Rio de Janeiro: IMPA/SBM/CNPq (1998).

[16] SHOR, P.W. *Algorithms for Quantum Computation*. Proc. 35th Annual Symposium on Foundations os Computer Science. IEEE Computer Society Press (1994), pp 124-134.

[17] SHOR, P.W. *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*. SIAM J. on Computing, 26(5) (1997), pp 1484-1509 (quant-ph/9508027).

[18] SHOR, P.W. *Quantum Computing*. Documenta Mathematica, Extra Volume ICM (1998).

[19] STEANE, A. *Quantum Computing*. LANL e-print arXiv:quant-ph/9708022 v2 (1997)

[20] SUDBERY, A. *Quantum Mechanics and Particles of Nature: An Outline for Mathematicians*. Cambridge: Cambridge University Press (1988).

[21] TANEBAUM, A. S. *Organização Estruturada de Computadores*. 3 ed. Rio de Janeiro: Prentice-Hall do Brasil (1984).

[22] TERADA, R. *Introdução à Complexidade de Algoritmos Paralelos*. VII Escola de Computação. São Paulo: IME/USP (1990).

[23] VOLOVICH, I.V. *Quantum Computing and Shor's Factoring Algorithm*. Lectures at the Volterra-CIRM International School "Quantum Computer and Quantum Information, Trento, Italy, July 25-31, 2001. LANL e-print arXiv:quant-ph/0109004 v1 (2001).