

INSTITUT DE
RECHERCHE
MATHÉMATIQUE
AVANCÉE

UMR 7501

Strasbourg

Thèse

présentée pour obtenir le grade de docteur de
l'Université de Strasbourg
Spécialité MATHÉMATIQUES

Kees van Schenk Brill

**Reviving Nuclear Magnetic Resonance as a viable
approach for quantum computation
and
Solving simultaneous Pell equations by quantum
computational means**

Soutenu le 3 décembre 2010
devant la commission d'examen

Edward Belaga, invité
Frits Beukers, rapporteur
Yann Bugeaud, examinateur
Daniel Grucker, co-directeur de thèse
Maurice Mignotte, directeur de thèse
Francis Taulelle, rapporteur

www-irma.u-strasbg.fr

Cette thèse contient deux parties. Je décris une approche pour construire une réalisation physique d'un ordinateur quantique par Résonance Magnétique Nucléaire (RMN). Je propose un nouveau cadre pour la RMN dans les réalisations physiques d'un ordinateur quantique. Je construis une description de la RMN à partir de la mécanique quantique avec laquelle je peux construire les opérateurs élémentaires essentiels pour le calcul quantique. Je décris les expériences pour construire ces opérateurs. Je propose un algorithme quantique en temps polynomial pour résoudre des équations de Pell simultanées comme extension de l'algorithme de Hallgren pour des équations de Pell simples.

tel-00534864, version 1 - 13 Dec 2010

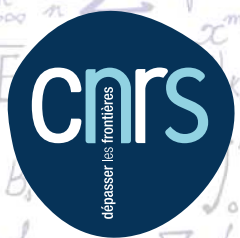
INSTITUT DE RECHERCHE MATHÉMATIQUE AVANCÉE
UMR 7501
 Université de Strasbourg et CNRS
 7 Rue René Descartes
 67 084 STRASBOURG CEDEX

Tél. 03 68 85 01 29
 Fax 03 68 85 03 28
www-irma.u-strasbg.fr
irma@math.unistra.fr

IRMA
 Institut de Recherche
 Mathématique Avancée

IRMA 2010/013
<http://tel.archives-ouvertes.fr/tel-00534864>

ISSN 0755-3390





INSTITUT DE RECHERCHE MATHÉMATIQUE AVANCÉE
Université de Strasbourg et C.N.R.S. (UMR 7501)
7 rue René Descartes
67084 STRASBOURG Cedex

LABORATOIRE D'IMAGÉRIE ET DE NEUROSCIENCES COGNITIVES
Institut de Physique Biologique (UMR 7191)
4 Rue Kirschleger
67085 STRASBOURG Cedex

**Reviving Nuclear Magnetic Resonance as a viable approach for
quantum computation
and
Solving simultaneous Pell equations by quantum computational
means**

par

Kees van Schenk Brill

Keywords: Quantum Computing, Nuclear Magnetic Resonance,
Time-dependent Schrödinger Equation, Simultaneous Pell Equations,
Diophantine Approximation, Algebraic Number Theory.

Mathematics Subject Classification: 11A51, 11A55, 11D09, 11J86,
34L40, 35Q41, 35Q60, 43A25, 68Q12, 81P68.

Voor Mina

Sommaire

Cette thèse contient deux parties qui peuvent être lues indépendamment. Dans la première partie je décris notre nouvelle approche pour construire une réalisation physique d'un ordinateur quantique par Résonance Magnétique Nucléaire (RMN). Avant de parler de RMN, je donne une introduction générale sur le calcul quantique. Je rappelle des notions de mécanique quantique nécessaires pour pouvoir décrire des algorithmes pour des ordinateurs quantiques.

Ensuite je rappelle la langage du calcul quantique. Je décris les manipulations que l'on peut faire avec des quantum bits, ou qubits, équivalents quantiques des bits pour un ordinateur ordinaire. Je détaille les avantages des ordinateurs quantiques pour des opérations du type « Transformée de Fourier » et je traite les deux algorithmes fondateurs dans le domaine: la factorisation en nombres premiers par l'algorithme de Shor et la recherche dans des bases de données par l'algorithme de Grover.

Je continue avec une description des réalisations physiques possibles pour construire un tel ordinateur.

Je parle de plusieurs approches différentes, mais celle à laquelle je consacre le plus de temps est l'approche par RMN. C'est avec cette technique que l'on a jusqu'à maintenant obtenu les résultats les plus intéressants en calcul quantique. Je discute ces succès et également pourquoi la RMN est devenue une technique obsolète.

A partir de ce point là, je propose un nouveau cadre pour la RMN dans les réalisations physiques d'un ordinateur quantique. Afin d'obtenir un tel cadre, je construis une nouvelle description de la RMN à partir de la mécanique quantique avec laquelle je peux construire les opérateurs élémentaires essentiels pour le calcul quantique. Je décris nos expériences pour construire ces opérateurs en distinguant entre des opérateurs agissant sur un qubit et des opérateurs agissant sur deux qubits. Je finis la première partie de la thèse avec une discussion sur la viabilité de cette approche pour permettre à la RMN de regagner sa place dans les techniques utilisées pour construire un ordinateur quantique.

Dans la deuxième partie de cette thèse je propose un algorithme quantique en temps polynomial pour résoudre des équations de Pell simultanées. Cette partie est inspirée d'une part de l'algorithme quantique de Hallgren pour

résoudre des équations de Pell simples en temps polynomial et d'autre part par la démonstration de Cipu et Mignotte du fait que dans le cas général, des équations de Pell simultanées ont au plus deux solutions distinctes.

Je commence cette partie avec une discussion sur l'équation de Pell simple. Je traite la résolution par fractions continues ainsi que les techniques plus modernes qui utilisent la théorie algébrique des nombres, notamment la notion du régulateur d'un corps de nombres. Je continue avec l'algorithme de Hallgren pour résoudre des équations de Pell. Cet algorithme est en temps polynomial contrairement aux méthodes décrites auparavant. C'est un algorithme quantique basé sur des extensions de techniques de Transformée de Fourier discutées dans la première partie.

Après le cas des équations de Pell simples, je m'intéresse au cas des équations de Pell simultanées. Je donne d'abord une borne supérieure pour la plus petite solution. Pour obtenir cette borne, j'utilise des résultats qui viennent de la théorie de l'approximation diophantienne pour les formes linéaires en logarithmes. Après avoir obtenu une borne supérieure, je continue avec la démonstration de Cipu et Mignotte du fait qu'il y a au plus deux solutions distinctes pour une paire d'équations de Pell simultanées. Dans cette démonstration on obtient une borne supérieure pour toutes les solutions des équations de Pell simultanées. J'utilise cette borne ensuite ainsi que l'algorithme de Hallgren pour des équations de Pell simples pour construire un algorithme qui résout en temps polynomial des équations de Pell simultanées. Cet algorithme a une partie quantique, la procédure de Hallgren pour résoudre les équations de Pell simples et obtenir les solutions fondamentales de chaque équation, et une partie « classique » de recherche de solutions à partir de ces solutions fondamentales, jusqu'à la borne supérieure. Je finis cette partie avec une discussion sur la possibilité d'étendre ces techniques pour résoudre d'autres problèmes similaires dans la théorie de nombres.

Dans les appendices je donne quelques détails supplémentaires sur la théorie des fractions continues et la théorie des nombres algébriques

Abstract

This text consists of two parts that can be read almost independently. In the first part I describe a renewed approach by Nuclear Magnetic Resonance (NMR) to build a quantum computer. I start with an introduction on quantum computing. I briefly describe the most important algorithms and the most promising physical realizations of a quantum computer. I continue with a description of NMR and the methods used earlier to build a quantum computer by NMR. I explain the shortcomings of these techniques and construct a new framework for quantum computation using NMR. For this I introduce a new quantum mechanical description of NMR with which the basic quantum gates needed for quantum computation can be built. I describe the experiments to build these gates, distinguishing between one qubit operations and two qubit operations. I conclude this part with a discussion on the practicality of this approach and whether these methods will allow for a revival of NMR as a quantum computing device.

The second part consists of the resolution and computation of simultaneous Pell equations. This part is inspired by Hallgren's quantum algorithm to solve the simple Pell equation in quantum polynomial time and by the proof of Cipu and Mignotte that in the general case, the simultaneous Pell equation has at most two solutions. I start this part with a discussion of the simple Pell equation, the classical techniques used to solve it, as well as more modern techniques. Afterwards I describe Hallgren's algorithm, for which I will need some extensions of the quantum computing techniques that I introduced in the first part. After this, I tackle simultaneous Pell equations. First I describe some classical results and solving techniques, culminating in the proof by Cipu and Mignotte that there are at most two distinct solutions for any given pair of independent Pell equations. To obtain this result, I have to introduce some Diophantine approximation theory. Finally I extend Hallgren's algorithm to simultaneous Pell equations using bounds from Diophantine approximation theory and some simple sieving techniques to compute solutions of simultaneous Pell equations in polynomial time on a quantum computer. I end this part with a discussion on extensions of these techniques to similar computational number theory problems.

In the appendices I give a short overview on continued fractions and on algebraic number theory.

Preface

No one who achieves success
does so without the help of
others. The wise and confident
acknowledge this help with
gratitude.

ALFRED NORTH WHITEHEAD

This thesis finds its origin in a chance meeting between two of my advisors, Edward Belaga and Daniel Grucker during a Mathematics and Biology seminar in the winter of 2005, where Professor Belaga gave a talk on molecular computing. During a coffee break they decided to organise another conference, this time on computing in general and on quantum computing and its physical realizations in particular. They received a research grant from the ANR (Agence Nationale de la Recherche) to continue their interdisciplinary work and they decided that it would be a good idea to look for a PhD student to assist them. I applied for this position and after two pleasant meetings they offered me the possibility to work with them. As I was not the beneficiary of a PhD grant from the French state and as the ANR grant was not sufficient to finance a full PhD position, it was difficult to begin our research. At this point it became unlikely that our collaboration would continue and I started to explore other avenues. During this time I was invited by the French embassy in the Hague to a reception for former beneficiaries of their embassy's grant to study a year in France. At this reception I explained my problems to two members of their grant committee, Jos van der Kruk and Gilbert van der Louw, who told me that one of the applicants for that year's grant had refused the embassy's offer. They then suggested me to apply for this grant. Thanks to these fine gentlemen and the swift and accurate help of Catherine Délice, I could finally begin my research on quantum computing. For this, I heartfully thank them.

My advisor, Daniel Grucker, has been a tremendous help on all fronts during the entire period of my thesis. From a financial point of view, he managed to find me a position as a technical assistant in my second and third year of research, which allowed me to continue my PhD. From an educational point of view, he taught me the basics and intricacies of Nuclear Magnetic Resonance with much clarity and great enthusiasm. As an experimentalist, he showed me how to operate the machines at our disposition and how to prepare our samples. As an advisor, he has been a driving force behind our research, pushing me to investigate our approach, showing an admirable patience for me during all these years and guiding me through the arduous process of writing a thesis. Daniel, I cannot thank you enough for all your help during my PhD. It has been a great pleasure.

My other advisor, Edward Belaga, has from the start focused on the global picture of our research, refusing to be carried away by details and always keeping in mind our ultimate goal, a functioning quantum computer combined with a well-considered architecture and well-conceived algorithms. He has personally taken my mathematical education in hand, pointing me in the right directions and providing important references for our research. He made it possible for me to attend conferences in the United States, England and Portugal, which lead to many interesting contacts. It has been impressive to see him make time for me at the most unlikely moments. While travelling between conferences he would call me to help me out with some mathematical problem, giving me just the clue that was eluding me. I consider myself lucky to have been his student and regret the fact that due to his retirement he could no longer officially be my advisor. Edward, I thank you for all the time you invested in me.

Because Edward Belaga had to retire, I needed another advisor for the mathematical contents of my research. Maurice Mignotte, who had previously supervised my Master thesis, was willing to take on this task. As my thesis was almost finished, his main contributions have been to proofread my manuscript, but this he has done with his usual modesty and expertise. Along the way, he managed to help me with the finer details on simultaneous Pell equations and diophantine approximation theory. Maurice, I thank you for accepting to be my advisor for just a year and for the pleasant discussions that usually started with Mathematics but rarely ended there.

As for the jury members, I warmly thank Frits Beukers, Francis Taulelle and Yann Bugeaud not only for having accepted to be on my thesis committee but also for the care with which they have read my manuscript and the useful suggestions they have made.

A lot of people helped me with my research during my thesis. First and foremost Tarek Khalil, who gave my work a much firmer physical grounding and who verified most of my computations. Tarek, I thank you for our heated discussions and for your insistence to correctly formulate our framework. Next, my gratitude goes to Jean Richert, who helped both Daniel and me understand how to approach the dipole-dipole interaction and who double-checked much of our work.

One of the perks of having two advisors is having two offices and therefore twice as many interesting colleagues. I would like to thank Jerome Steibel for his many fun suggestions regarding our experiments; Jerome, one day our computer will run on beer ! Many thanks also to, amongst others : Nathalie, Thierry, Renée, Laura and H el ene, who made my stay at the Institute of Physics and Biology a very pleasant one.

As to my fellow PhD-students at the Institute for Mathematics, what can I say. It was a great pleasure to share offices with Vincent, Audrey, R emi, Benjamin, Alain, Jean and Auguste. To have coffee breaks with Adrien, C edric, Camille, Alexandre, Florian, H el ene and Anne-Laure. The most pleasant

times were however during those scarce moments of extra-mathematical activity, for which a royal thank you goes to Fabien, Aurélien, Scoum, Thomas, Aurore, Ghislain, Jürgen and everybody else who contributed to the good spirit of the first floor.

During my thesis a lot of bureaucratic work was done for me by people who are far more capable than I am. I would like to thank Simone, Nathalie and Yvonne especially for all they have done for me.

A nice thing about friends is that they help you keep up when your research is desperately trying to make you feel miserable. I would like to take this opportunity to thank Alexandre and Jannes, who both greatly restored my morale when needed.

My family has been there for me during all these years and without them I never would have finished my thesis. Dick, thank you for the many hours you spent proofreading and spellchecking, for making me see how to formulate my ideas more clearly and for all the times you helped me out. Willy, thank you for supporting me throughout the entire process and helping me through the last difficult hurdles, when I felt ready to throw in the towel. I know it has been hard on both of you to have your son far away from you and I sincerely hope that in the future this will change. Maartje, Erik, Floor, Midas, Thijs and Esther, thank you for the pleasant moments, the good chocolate, the cycling, the rollercoasters and so much more.

Finally I thank my little family of my own. Julia, you had to put up with me during all those times when morale was low, when deadlines were set, when plans were altered, when dates got pushed further and further into the future, when everything seemed uncertain. I know that without you by my side, I would have given up long ago. You have been my rock, even if you think that it is the other way around. A last word goes to the smallest of my family, my lovely daughter, Mina. You have helped me realize what is important and what is secondary, you may not have known it at the time, but you have done me a great service in just being there.

Contents

Sommaire	iii
Abstract	v
Preface	vii
I Quantum Computing using NMR	1
1 Quantum Computing	3
1.1 Introduction	3
1.2 Quantum Mechanics	4
1.3 Classical and Quantum Logic	5
1.3.1 Qubits	5
1.3.2 Manipulating bits and qubits	6
1.3.3 Limitations	10
1.4 Quantum Algorithms	11
1.4.1 Discrete Fourier and Quantum Fourier Transform	11
1.4.2 Fourier Transforms over Abelian Groups	13
1.4.3 Shor's Factoring Algorithm	17
1.4.4 Grover's Search Algorithm	22
1.5 Physical Realisations	25
1.5.1 Introduction	25
1.5.2 Optical photon quantum computer	27
1.5.3 Trapped ions	29
1.5.4 Other physical realizations	31
2 NMR and Quantum Computing	33
2.1 Nuclear Magnetic Resonance	33
2.1.1 Introduction	33
2.2 Quantum computing with NMR	37
2.2.1 Ensemble system	37
2.2.2 Labeling the qubits	38
2.2.3 Unitary transformations	39
2.2.4 Ensemble measurements	40
2.3 Drawbacks	41

3	Reviving the NMR Approach	43
3.1	Introduction	43
3.2	Framework for Quantum Computing	44
3.2.1	One spin $\frac{1}{2}$	44
3.2.2	Two spins $\frac{1}{2}$	56
3.2.3	N spins	60
3.2.4	Dipole-dipole coupling	61
3.2.5	Two coupled spins	63
3.3	Five Steps to an NMR Quantum Computer	64
3.4	Experimental results	65
3.4.1	Material and methods	65
3.4.2	Results	65
3.4.3	Numerical solution of equation (3.2.18)	68
3.5	Conclusion and Perspective	71
II	Solving Simultaneous Pell Equations	73
4	Pell equations	75
4.1	Introduction	75
4.2	Classical Techniques	76
4.2.1	Chakravala Method	76
4.2.2	Continued fraction method	79
4.3	Modern Techniques	80
4.4	Quantum Computational Techniques	81
5	Simultaneous Pell equations	87
5.1	Introduction	87
5.2	A conjecture on 5 integers	88
5.3	An upper bound	89
5.3.1	Diophantine Approximation	89
5.3.2	Upper bound for smallest solution	94
5.4	Finite Number of Solutions	97
5.4.1	Introduction	97
5.4.2	Transforming the equations	98
5.4.3	Linear form in three logarithms	99
5.4.4	Gap principles	101
5.5	Quantum Algorithm	102
5.6	Conclusion and Perspective	104
A	Kronecker product and sum	107
B	Continued Fractions	109
C	Algebraic Number Theory	113
	List of Symbols and Acronyms	127
	Bibliography	129

Part I

Quantum Computing using
Nuclear Magnetic Resonance

Chapter 1

Any sufficiently advanced
technology is indistinguishable
from magic.

ARTHUR C. CLARKE

Quantum Computing

1.1 Introduction

The idea of the quantum computer has been around for some time. One of its basic elements is the notion of reversible computation, which was developed by Charles Bennett [Ben73, Ben82]. This is a model of computing that is reversible, for which a necessary condition is that the corresponding binary mapping is one-to-one. A major motivation for this type of models is that reversible computing can improve the energy efficiency of computers beyond the von Neumann-Landauer limit [Lan61, vN66] of $k_B T \log 2$ energy dissipated per irreversible bit operation.

We concentrate on logically reversible systems, which is a necessary but not a sufficient condition for a computational process to be physically reversible. Landauer's principle is the notion that the erasure of n bits of information has a cost of $n k_B T \log 2$ in thermodynamic entropy.

Poplavskii wrote in the seventies that classical computers are unable to simulate quantum mechanical systems because of the superposition principle [Pop75]. Manin added a few years later [Man80] that the exponential number of basis states of a quantum system could be exploited but that a theory of quantum computation was needed that captured the fundamental principles without committing to a physical realization.

Richard Feynmann wrote in the early eighties [Fey82] that in order to simulate the evolution of quantum systems with computers, these computers would need to have quantum mechanical properties if we wanted the simulation to be done efficiently. In 1985 David Deutsch proposed a universal quantum computer [Deu85], which can simulate any other quantum computer. In the same article he also invented a simple quantum algorithm for a decision problem, that he proved to be faster than any classical algorithm that can be constructed for this problem. Richard Josza later produced a generalization of this algorithm [DJ92]. The decision problem in question is to decide whether a given binary function is balanced or constant, given

that it has one of these properties.

Until the middle of the nineties, no serious proposal for a physical realization of a quantum computer had been made. While new quantum algorithms continued to be found, most based on the quantum computational equivalent of the Fourier Transform, nobody seemed to know how to actually build such a hypothetical computer. In 1995, Cirac and Zoller proposed to build a quantum computer from ion traps [CZ95]. From that point on, different proposals for physical realizations have slowly started to outnumber the proposals for different quantum algorithms.

In the rest of this chapter we introduce the basic elements that are needed for a quantum computer. We give a very short overview on quantum mechanics in general and a little more detail on quantum logic. We discuss the Quantum Fourier Transform and describe the two important algorithms in the domain of quantum computation. We then proceed by detailing some proposals for physical realizations.

1.2 Quantum Mechanics

Quantum computing should be seen in the framework of quantum mechanics. We give a brief overview on the basics for quantum mechanics. For a more precise review we recommend the excellent account by Nielsen and Chuang [NC00] or the standard text books on quantum mechanics [Sak94, CTDL77].

Throughout these chapters we will suppose to be working in a complex Hilbert space V of dimension N . The standard quantum mechanical notation for a vector in a vector space is $|\phi\rangle$ which is called a ket. Its vector dual $\langle\phi|$ is called a bra. An inner product between two vectors ϕ, ψ is denoted $\langle\phi|\psi\rangle$. The tensor product between two vectors is denoted as $|\phi\rangle\otimes|\psi\rangle$ but we will use the shorthand notation $|\phi\rangle|\psi\rangle$.

We will fix an orthonormal basis $\mathcal{B} = \{|0\rangle, \dots, |N-1\rangle\}$ for V . Thus we can write

$$|\phi\rangle = \sum_{i=0}^{N-1} a_i |i\rangle, \quad (1.2.1a)$$

$$\langle\phi| = \sum_{i=0}^{N-1} a_i^* \langle i|, \quad (1.2.1b)$$

where the a_i are complex numbers.

Any linear operator A on V can be written in the form

$$A = \sum_{i,j} a_{ij} |i\rangle\langle j|. \quad (1.2.2)$$

Quantum mechanics can be summarized by 4 postulates.

1. To an isolated physical system we associate a Hilbert space with inner product which is the state space of the system. The system is completely described by its state vector which is a unit vector in the state space.
2. The evolution of a closed quantum system is described by a unitary transformation.
3. Quantum measurements are described by a collection $\{M_m\}$ of measurement operators. These operators satisfy the completeness relation

$$\sum_m M_m^\dagger M_m = I. \quad (1.2.3)$$

4. The state space of a composite physical system is the tensor product of the state spaces of the component systems.

1.3 Classical and Quantum Logic

1.3.1 Qubits

Bits are the basic elements in classical computing. As a physical entity they can be considered as electronic switches that are either switched ON or switched OFF. In a computational sense they have either the value 0 or 1. The quantum mechanical analogue of bits are qubits, which is shorthand for quantum bits. As a physical entity they can be a multitude of objects. They could be the two different polarizations of a photon, the alignment of a nuclear spin in a uniform magnetic field or something else entirely. In a mathematical sense they are simply unit vectors in \mathbb{C}^2 . The standard orthonormal basis for qubits is denoted as $|0\rangle, |1\rangle$. These vectors correspond to the column vectors $(1, 0)^T, (0, 1)^T$. An arbitrary qubit $|\psi\rangle$ can be written as

$$|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle, \quad (1.3.1)$$

with $\alpha_0, \alpha_1 \in \mathbb{C}$ and $\alpha_0^2 + \alpha_1^2 = 1$. Measuring the qubit $|\psi\rangle$ will give $|0\rangle$ with probability $|\alpha_0|^2$ and $|1\rangle$ with probability $|\alpha_1|^2$. It is possible to rewrite equation (1.3.1) as

$$|\psi\rangle = e^{i\gamma} \left(\cos \frac{\theta}{2} |0\rangle + e^{i\phi} \sin \frac{\theta}{2} |1\rangle \right), \quad (1.3.2)$$

where θ, ϕ and γ are real numbers. The factor $e^{i\gamma}$ can be ignored as it has no observable effect. This leads to

$$|\psi\rangle = \cos \frac{\theta}{2} |0\rangle + e^{i\phi} \sin \frac{\theta}{2} |1\rangle. \quad (1.3.3)$$

The qubit $|\psi\rangle$ can be considered as a point on the three-dimensional unit sphere. This sphere is called the Bloch-sphere.

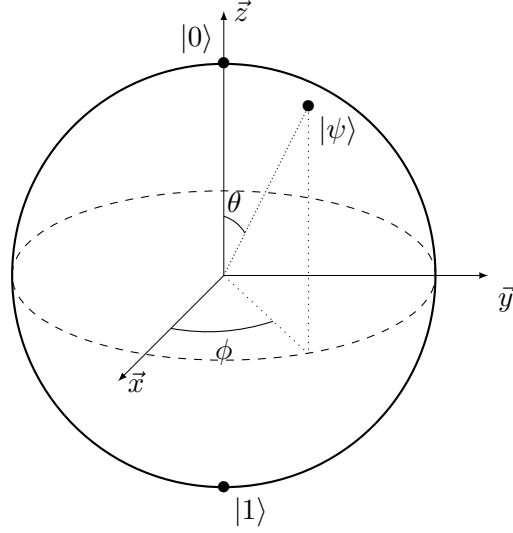


Figure 1.1: Bloch sphere representation of a qubit $|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle$.

We can use the fourth postulate in order to combine several qubits. The vectors $\{|0\rangle \otimes \cdots \otimes |0\rangle, \dots, |1\rangle \otimes \cdots \otimes |1\rangle\}$ form a set of n qubits that span a space of dimension 2^n . We will denote by $|n\rangle$ the qubit $|z_0\rangle \otimes \cdots \otimes |z_k\rangle$ with $z_i \in \{0, 1\}$ and $n = \sum_{i=0}^k z_i 2^i$.

An arbitrary qubit $|\psi\rangle = \sum_{i=0}^{2^n-1} \alpha_i |i\rangle$ is a unit vector in \mathbb{C}^{2^n} . When measured it returns the state $|j\rangle$ with probability $|\alpha_j|^2$. After measuring the state $|\psi\rangle$ becomes $|\psi'\rangle = |j\rangle$. This process is called the collapse of the waveform.

1.3.2 Manipulating bits and qubits

Classical bits

In order to compute with classical bits we use logical gates. A logical gate is a function $f: \{0, 1\}^k \rightarrow \{0, 1\}^l$ with k input bits and l output bits. The following seven gates are well-known.

$$\neg = \text{NOT}: \{0, 1\} \rightarrow \{0, 1\} \\ x \mapsto x + 1 \pmod{2} \quad (1.3.4a)$$

$$\vee = \text{OR}: \{0, 1\} \rightarrow \{0, 1\} \\ (x_1, x_2) \mapsto x_1 x_2 + x_1 + x_2 \pmod{2} \quad (1.3.4b)$$

$$\oplus = \text{XOR}: \{0, 1\}^2 \rightarrow \{0, 1\} \\ (x_1, x_2) \mapsto x_1 + x_2 \pmod{2} \quad (1.3.4c)$$

$$\wedge = \text{AND}: \{0, 1\}^2 \rightarrow \{0, 1\} \\ (x_1, x_2) \mapsto x_1 x_2 \pmod{2} \quad (1.3.4d)$$

$$\begin{aligned} \uparrow = \text{NAND}: \{0, 1\}^2 &\longrightarrow \{0, 1\} \\ (x_1, x_2) &\longmapsto x_1 x_2 + 1 \pmod{2} \end{aligned} \quad (1.3.4e)$$

$$\begin{aligned} \text{FAN}: \{0, 1\} &\longrightarrow \{0, 1\}^2 \\ x &\longmapsto (x, x) \end{aligned} \quad (1.3.4f)$$

$$\begin{aligned} \text{SWAP}: \{0, 1\}^2 &\longrightarrow \{0, 1\}^2 \\ (x_1, x_2) &\longmapsto (x_2, x_1) \end{aligned} \quad (1.3.4g)$$

With these gates we can compute any function.

Theorem 1.1. *An arbitrary function $f: \{0, 1\}^n \longrightarrow \{0, 1\}$ can be simulated with the logical gates NOT, AND, XOR, FAN and SWAP.*

Proof. We use induction on n . For $n = 1$ there are four possible functions:

1. The identity function, which does not need any gate.
2. The NOT-function, which is one of the five gates that can be used.
3. The constant function 0, which we can produce by using the following gates:

$$0 = 0(x) = \wedge \left(\text{FAN}_1(x), \neg(\text{FAN}_2(x)) \right), \quad (1.3.5)$$

where FAN_i is the i -th output bit of the FAN-function.

4. We can obtain the constant function 1 by taking the NOT of the previous function:

$$1 = \neg(0(x)). \quad (1.3.6)$$

Suppose now that any function on n bits can be computed and let f be a function on $n + 1$ bits. Define the n -bit functions f_0 and f_1 by

$$f_i(x_1, \dots, x_n) = f(i, x_1, \dots, x_n). \quad (1.3.7)$$

Then we have

$$f(x_0, \dots, x_n) = \oplus \left(\wedge (f_0(x_1, \dots, x_n), \neg(x_0)), \wedge (f_1(x_1, \dots, x_n), x_0) \right). \quad (1.3.8)$$

□

Alternative proof without induction. The function f can be written as

$$\begin{aligned} f &= \sum_x f(x) \chi_x, \\ &= \sum_{x|f(x)=1} \chi_x \end{aligned} \quad (1.3.9)$$

where

$$\chi_x(y) = \delta_{xy} = \begin{cases} 1, & \text{if } x = y, \\ 0, & \text{otherwise.} \end{cases} \quad (1.3.10)$$

So that we can write

$$f = \bigvee_{x|f(x)=1} \chi_x, \quad (1.3.11)$$

where χ_x is a product of z_i or \bar{z}_i and

$$z_i(y) = \begin{cases} 1, & \text{if } y_i = 1, \\ 0, & \text{otherwise.} \end{cases} \quad (1.3.12)$$

□

We actually need only three gates.

Theorem 1.2. *The NAND-function together with the FAN-function can simulate the functions NOT, AND and XOR.*

Proof.

$$\neg(x) = \uparrow (\text{FAN}(x)) \quad (1.3.13a)$$

$$\wedge(x_1, x_2) = \uparrow (\text{FAN}(\uparrow(x_1, x_2))) \quad (1.3.13b)$$

$$\oplus(x_1, x_2) = \uparrow \left(\uparrow \left(\uparrow (\text{FAN}(x_1)), x_2 \right), \uparrow \left(x_1, \uparrow (\text{FAN}(x_2)) \right) \right) \quad (1.3.13c)$$

□

So the NAND-gate together with the FAN-gate and the SWAP-gate allows us to compute any function. However, the NAND-gate is not reversible, nor can it be made reversible by adding an extra bit with information on the input. There are logical gates on three bits that are reversible and can compute any function. For instance the Toffoli-gate

$$\text{TOF}(x_1, x_2, x_3) = (x_1, x_2, x_1x_2 + x_3), \quad (1.3.14)$$

and the Fredkin-gate

$$\text{FRE}(x_1, x_2, x_3) = (\text{SWAP}(x_1, x_2)x_3 + \text{Id}(x_1, x_2)(x_3 + 1), x_3), \quad (1.3.15)$$

which swaps the first two bits if and only if the third bit is set to 1.

Manipulating qubits

The quantum equivalent of logical gates on bits are unitary transforms on qubits. Given a 2^n -dimensional vector space V with basis \mathcal{B} and a $2^m \times 2^m$ matrix U with $m \leq n$, an expansion of U relative to \mathcal{B} is any matrix of the form

$$G(U \otimes I_{2^{n-m}})G^{-1}, \quad (1.3.16)$$

where G permutes the basis and I_k is the $k \times k$ identity matrix.

Let $\mathcal{U} = \{U_1, \dots, U_k\}$ be a set of unitary matrices of dimension dividing 2^n . Then $(\mathcal{B}, \mathcal{U})$ is the set of all expansions of the U_i relative to \mathcal{B} .

We define the following matrices, which are respectively called the Hadamard operator, the rotation operator of angle θ , the control-Not operator and the control-control-Not operator:

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad P(\theta) = \begin{pmatrix} e^{\frac{i\theta}{2}} & 0 \\ 0 & e^{-\frac{i\theta}{2}} \end{pmatrix}, \quad (1.3.17ab)$$

$$CNOT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad (1.3.17c)$$

$$CCNOT = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}. \quad (1.3.17d)$$

The control-NOT operator is a special case of the general class of controlled operators. These operators act on two registers of qubits in a very specific manner. If the first register of qubits is in a specified control state, usually $|1\rangle \cdots |1\rangle$, then an operator U is applied to the second register of qubits. If the first register is not in the specified control state, the identity operator is applied to the second register. For any $n > 2$ and θ , such that $P(\theta)$ is not idempotent, the set $U_\tau = \{H, CNOT, CCNOT, P(\theta)\}$ generates a group GU_τ that is dense in $U(2^n)$. To be a little bit more precise we define the norm of a vector

$$\|\phi\rangle\| = \sqrt{\langle\phi|\phi\rangle}. \quad (1.3.18)$$

The norm of an operator U is defined as

$$\|U\| = \sup_{|\phi\rangle \neq 0} \frac{\|U|\phi\rangle\|}{\|\phi\rangle\|}. \quad (1.3.19)$$

We say that an operator \tilde{U} represents an operator U with precision ϵ if

$$\|\tilde{U} - U\| \leq \epsilon. \quad (1.3.20)$$

With this definition we can say that the group GU_τ represents $U(2^n)$ with precision ϵ for any $\epsilon > 0$.

A quantum circuit is a unitary matrix built by composing elementary operations from U_τ . The size of a quantum circuit will be the minimal number of operations composed to obtain it. A register in a quantum computer is a subset of the total set of qubits. Writing $|\phi_1\rangle|\phi_2\rangle$ means that the first register is in state $|\phi_1\rangle$ and the second in $|\phi_2\rangle$.

1.3.3 Limitations

The most important limitation for qubits is the following theorem.

Theorem 1.3 (No Cloning Theorem). *It is not possible to copy any given quantum state*

Proof. Suppose we have two qubits. The qubit to be copied is in state $|\phi_1\rangle$ and the other qubit in some state $|s\rangle$. Suppose that we have a copying machine, using a unitary operation U . Then

$$|\phi_1\rangle \otimes |s\rangle \xrightarrow{U} U(|\phi_1\rangle \otimes |s\rangle) = |\phi_1\rangle \otimes |\phi_1\rangle. \quad (1.3.21)$$

For another quantum state $|\phi_2\rangle$ we have the same relation. We now take inner products to get the following.

$$(\langle\phi_1| \otimes \langle s|)U^\dagger U(|\phi_2\rangle \otimes |s\rangle) = (\langle\phi_1| \otimes \langle\phi_1|)(|\phi_2\rangle \otimes |\phi_2\rangle). \quad (1.3.22a)$$

$$\langle\phi_1|\phi_2\rangle\langle s|s\rangle = \langle\phi_1|\phi_2\rangle\langle\phi_1|\phi_2\rangle. \quad (1.3.22b)$$

$$\langle\phi_1|\phi_2\rangle = \langle\phi_1|\phi_2\rangle^2. \quad (1.3.22c)$$

This equation has solutions if and only if $\langle\phi_1|\phi_2\rangle$ is 0 or 1. So copying cannot be done for general states. \square

The consequences of this negative result are clear. Even for simple operations like switching two bits we would like to make a copy of one of the bits before overwriting it. In quantum computing we need to design algorithms in such a way that we never need to store an intermediate result, which is a fundamentally different approach than what we are used to on classical computers. So in a sense we need to develop a quantum mechanical way of algorithmic thinking to design algorithms for quantum computers.

1.4 Quantum Algorithms

1.4.1 Discrete Fourier and Quantum Fourier Transform

Let x_0, \dots, x_{N-1} be a vector of complex numbers. The Discrete Fourier Transform is defined by:

$$y_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j e^{\frac{2\pi i j k}{N}}. \quad (1.4.1)$$

The Coole-Tukey algorithm [CT65] for Discrete Fourier Transforms reduced the complexity from $O(e^{n^2})$ to $O(e^{n \log n})$. Let $|k\rangle$ be a vector in a complex Hilbert space V of dimension N and let $|0\rangle, \dots, |N-1\rangle$ be an orthonormal basis for V . The Quantum Fourier Transform (QFT) is defined in the same way as the Discrete Fourier Transform:

$$|k\rangle \mapsto \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} e^{\frac{2\pi i j k}{N}} |j\rangle. \quad (1.4.2)$$

It is possible to give a matrix notation for the QFT. Let $\xi = e^{\frac{2\pi i}{2^N}}$, then the unitary $2^N \times 2^N$ matrix, given by

$$a_{jk} = \frac{1}{\sqrt{2^N}} \xi^{(j-1)(k-1)}, \quad (1.4.3)$$

is the Quantum Fourier Transform. An example for $N = 3$ and $\xi^8 = 1$:

$$\begin{aligned} \text{QFT}_{N=3} &= \frac{1}{\sqrt{8}} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \xi & \xi^2 & \xi^3 & \xi^4 & \xi^5 & \xi^6 & \xi^7 \\ 1 & \xi^2 & \xi^4 & \xi^6 & \xi^8 & \xi^{10} & \xi^{12} & \xi^{14} \\ 1 & \xi^3 & \xi^6 & \xi^9 & \xi^{12} & \xi^{15} & \xi^{18} & \xi^{21} \\ 1 & \xi^4 & \xi^8 & \xi^{12} & \xi^{16} & \xi^{20} & \xi^{24} & \xi^{28} \\ 1 & \xi^5 & \xi^{10} & \xi^{15} & \xi^{20} & \xi^{25} & \xi^{30} & \xi^{35} \\ 1 & \xi^6 & \xi^{12} & \xi^{18} & \xi^{24} & \xi^{30} & \xi^{36} & \xi^{42} \\ 1 & \xi^7 & \xi^{14} & \xi^{21} & \xi^{28} & \xi^{35} & \xi^{42} & \xi^{49} \end{pmatrix} \\ &= \frac{1}{\sqrt{8}} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \xi & \xi^2 & \xi^3 & \xi^4 & \xi^5 & \xi^6 & \xi^7 \\ 1 & \xi^2 & \xi^4 & \xi^6 & 1 & \xi^2 & \xi^4 & \xi^6 \\ 1 & \xi^3 & \xi^6 & \xi & \xi^4 & \xi^7 & \xi^2 & \xi^5 \\ 1 & \xi^4 & 1 & \xi^4 & 1 & \xi^4 & 1 & \xi^4 \\ 1 & \xi^5 & \xi^2 & \xi^7 & \xi^4 & \xi & \xi^6 & \xi^3 \\ 1 & \xi^6 & \xi^4 & \xi^2 & 1 & \xi^6 & \xi^4 & \xi^2 \\ 1 & \xi^7 & \xi^6 & \xi^5 & \xi^4 & \xi^3 & \xi^2 & \xi \end{pmatrix}. \quad (1.4.4) \end{aligned}$$

The QFT is useful because the complexity of the DFT is $O(e^{n \log n})$ whereas the complexity of the QFT is $O(n^2)$. It is exactly this gain which will allow

us to solve classically infeasible problems with quantum algorithms by using the QFT. The following example, of which Shor's algorithm is a special case, clearly shows how the QFT can be used in quantum algorithms.

Let $N > 1$ be a positive integer, $G = \mathbb{Z}/N\mathbb{Z}$ the additive group of integers modulo N and X a finite set. Suppose that we have a function $f: G \rightarrow X$, such that for some subgroup $H = \langle d \rangle$ of G , f is constant on H and separates cosets of H . Suppose that we do not know d . We want to find a generator for H . To do so we start with two registers in the zero state $|0\rangle|0\rangle$ and we apply the QFT to the first register to obtain

$$\frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} |j\rangle|0\rangle. \quad (1.4.5)$$

We then apply f to the second register to get

$$\frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} |j\rangle|f(j)\rangle. \quad (1.4.6)$$

We now measure the second register and obtain $f(j_0)$ for some j_0 . The effect of measuring the second register is that all registers that do not have $f(j_0)$ in the second register collapse. As f separates cosets of H this means that only the coset $H + j_0$ remains in the first register. If $|H| = M$, the first register can be described as

$$\frac{1}{\sqrt{M}} \sum_{s=0}^{M-1} |j_0 + sd\rangle. \quad (1.4.7)$$

We apply the QFT to this register to obtain

$$\frac{1}{\sqrt{MN}} \sum_{k=0}^{N-1} e^{\frac{2\pi i j_0 k}{N}} |k\rangle \sum_{s=0}^{M-1} e^{\frac{2\pi i s d k}{N}}. \quad (1.4.8)$$

Using the fact that $N = dM$ and evaluating the second sum as a geometric series, only the values of $|k\rangle$ that are multiples of M remain, giving

$$\frac{1}{\sqrt{d}} \sum_{t=0}^{d-1} e^{\frac{2\pi i j_0 t M}{N}} |tM\rangle. \quad (1.4.9)$$

Measuring the first register gives a multiple of M . Repeating this procedure we get several multiples of M . Using the Euclidean algorithm we obtain M with high probability.

1.4.2 Fourier Transforms over Abelian Groups

The above example works well because it was straightforward to identify the elements of the group $\mathbb{Z}/N\mathbb{Z}$ with the qubits $|0\rangle, \dots, |N-1\rangle$. For general finite abelian groups, this identification is not that simple and we will need to define a more general form of Fourier transform. To do so we need to introduce some basic representation and character theory. We follow the description of Chris Lomont [Lom]. Every finite Abelian group G can be written as the direct sum of cyclic groups, so

$$G = \mathbb{Z}/N_1\mathbb{Z} \oplus \dots \oplus \mathbb{Z}/N_k\mathbb{Z}. \quad (1.4.10)$$

We suppose that we have a function f from G to a finite set X , such that f separates cosets of a subgroup H of G . We will write elements of G as k -tuples (g_1, \dots, g_k) , with $g_i \in \{0, \dots, N_i - 1\}$. Define

$$\beta_i = (0, \dots, 0_{i-1}, 1_i, 0_{i+1}, \dots, 0). \quad (1.4.11)$$

A character of G is a group homomorphism χ from G to the multiplicative group of nonzero complex numbers \mathbb{C}^* . For every character χ and every element $g = (g_1, \dots, g_k)$ we have

$$\chi(g) = \chi\left(\sum_{i=1}^k g_i \beta_i\right) = \prod_{i=1}^k \chi(\beta_i)^{g_i}. \quad (1.4.12)$$

So every character χ is determined by its action on the β_i . As the order of β_i is N_i , the order of $\chi(\beta_i)$ must divide N_i . Therefore

$$\chi(\beta_i) = e^{\frac{2\pi i h_i}{N_i}}, \quad (1.4.13)$$

for some $h_i \in \{0, \dots, N_i - 1\}$. So we can determine a character by a k -tuple (h_1, \dots, h_k) , which can be seen as an element $h \in G$. This leads to the following definition for characters. For every $g \in G$, we define

$$\begin{aligned} \chi_g: G &\longrightarrow \mathbb{C}^* \\ h &\longmapsto \prod_{j=1}^k e^{\frac{2\pi i g_j h_j}{N_j}}. \end{aligned} \quad (1.4.14)$$

A useful theorem on characters is the following.

Theorem 1.4. *Let G be a finite Abelian group and χ a character. Then*

$$\sum_{g \in G} \chi(g) = \begin{cases} |G| & \text{if } \chi = \chi_e, \\ 0 & \text{otherwise.} \end{cases} \quad (1.4.15)$$

Here χ_e is the identity character sending every element of the group to 1.

Proof. We have

$$G = \mathbb{Z}/N_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/N_k\mathbb{Z}. \quad (1.4.16)$$

Choose $h \in G$. Then

$$\begin{aligned} \sum_{g \in G} \chi_h(g) &= \sum_{\substack{g_j \in \mathbb{Z}/N_j\mathbb{Z} \\ j \in \{1, \dots, k\}}} \left(\prod_{j=1}^k e^{2\pi i h_j g_j / N_j} \right) \\ &= \prod_{j=1}^k \sum_{g_j \in \mathbb{Z}/N_j\mathbb{Z}} e^{2\pi i h_j g_j / N_j}. \end{aligned} \quad (1.4.17)$$

If for some j we have $e^{2\pi i h_j / N_j} \neq 1$, then the geometric series

$$\sum_{g_j \in \mathbb{Z}/N_j\mathbb{Z}} e^{\frac{2\pi i h_j}{N_j} g_j} = 0. \quad (1.4.18)$$

The only time this does not happen is when for all j we have

$$e^{\frac{2\pi i h_j}{N_j}} = 1. \quad (1.4.19)$$

This is the identity character. In this case the result is $\prod_{j=1}^k N_j = |G|$. \square

We can now define the notion of an orthogonal subgroup. Let H be a subgroup of G . The orthogonal subgroup of H is

$$H^\perp = \{g \in G \mid \chi_g(h) = 1, \text{ for all } h \in H\}. \quad (1.4.20)$$

While the cyclic QFT returns multiples of the generator of H , the general finite abelian QFT returns elements of the orthogonal subgroup of H . It is defined as

$$F_G = \frac{1}{\sqrt{|G|}} \sum_{g, h \in G} \chi_g(h) |g\rangle \langle h|. \quad (1.4.21)$$

We also define a translation operator

$$\tau_t = \sum_{g \in G} |t + g\rangle \langle g|, \quad (1.4.22)$$

and a phase-change operator

$$\phi_h = \sum_{g \in G} \chi_g(h) |g\rangle \langle g|. \quad (1.4.23)$$

We first show that the Fourier transform of a subgroup is its orthogonal subgroup.

Theorem 1.5. *We have the following relation between subgroups and Fourier transforms:*

$$F_G|H\rangle = |H^\perp\rangle. \quad (1.4.24)$$

Proof. By definition, we have

$$|H\rangle = \frac{1}{\sqrt{|H|}} \sum_{h \in H} |h\rangle. \quad (1.4.25)$$

We then have:

$$F_G|H\rangle = \frac{1}{\sqrt{|G|}} \sum_{g, h' \in G} \chi_g(h') |g\rangle \langle h'| \frac{1}{\sqrt{|H|}} \sum_{h \in H} |h\rangle. \quad (1.4.26)$$

Using the fact that $\langle h|h'\rangle = 1$, if $h = h'$ and zero otherwise, the above expression can be simplified to

$$\frac{1}{\sqrt{|G||H|}} \sum_{g \in G} \left(\sum_{h \in H} \chi_g(h) \right) |g\rangle. \quad (1.4.27)$$

The character χ_g of G is also a character of H , therefore $\sum_{h \in H} \chi_g(h)$ is zero unless the character is the identity on H , in which case the sum is equal to $|H|$. That is exactly the definition of the orthogonal subgroup, therefore we can reduce the equation to

$$\frac{1}{\sqrt{|G||H|}} \sum_{g \in H^\perp} |H||g\rangle. \quad (1.4.28)$$

As $|H||H^\perp| = |G|$, this is equal to

$$\frac{1}{\sqrt{|H^\perp|}} \sum_{g \in H^\perp} |g\rangle = |H^\perp\rangle. \quad (1.4.29)$$

□

In a similar way the following three identities can be proved.

Theorem 1.6. *For all elements $h, t \in G$ we have*

$$\chi_h(t) \tau_t \phi_h = \phi_h \tau_t, \quad (1.4.30a)$$

$$F_G \phi_h = \tau_{-h} F_G, \quad (1.4.30b)$$

$$F_G \tau_t = \phi_t F_G. \quad (1.4.30c)$$

We can now give the algorithm for the hidden subgroup problem for general finite abelian groups. As in the cyclic case we start with two registers of

qubits in the zero state and we apply the Fourier transform to the first register.

$$|0\rangle|0\rangle \mapsto \frac{1}{\sqrt{|G|}} \sum_{g \in G} |g\rangle|0\rangle. \quad (1.4.31)$$

We then apply the coset separating function f to the second register, which leads to

$$\frac{1}{\sqrt{|G|}} \sum_{g \in G} |g\rangle|f(g)\rangle. \quad (1.4.32)$$

Define $T = (t_1, \dots, t_m)$ as a set of coset representatives for H in G . We obviously have $|T||H| = |G|$. Using the separation property of f we can simplify the above expression to

$$\frac{1}{\sqrt{|T|}} \sum_{t \in T} |t + H\rangle|f(t)\rangle. \quad (1.4.33)$$

This is equal to

$$\frac{1}{\sqrt{|T|}} \sum_{t \in T} \tau_t|H\rangle|f(t)\rangle. \quad (1.4.34)$$

We apply the Fourier transform to the first register and use the above theorems to obtain the following result.

$$\begin{aligned} \frac{1}{\sqrt{|T|}} \sum_{t \in T} \tau_t|H\rangle|f(t)\rangle &\xrightarrow{F_G} \frac{1}{\sqrt{|T|}} \sum_{t \in T} F_G \tau_t|H\rangle|f(t)\rangle \\ &= \frac{1}{\sqrt{|T|}} \sum_{t \in T} \phi_t F_G|H\rangle|f(t)\rangle \\ &= \frac{1}{\sqrt{|H^\perp|}} \sum_{t \in T} \phi_t|H^\perp\rangle|f(t)\rangle. \end{aligned} \quad (1.4.35)$$

We now measure the first register and obtain a random element of the orthogonal subgroup of H . Since $(H^\perp)^\perp = H$, determining a generating set for the orthogonal subgroup determines H completely. This does however not mean that it is an easy task to get a generating set for H starting with a generating set for H^\perp . Suppose that we have a generating set g_1, \dots, g_t for H^\perp . As $H = H^{\perp\perp}$, we have $h \in H$ if and only if

$$\chi_h(g_j) = 1, \quad \text{for all } j = 1, \dots, t. \quad (1.4.36)$$

Let $d = \text{LCM}(N_1, \dots, N_k)$ and $\alpha_i = \frac{d}{N_i}$. Then

$$\chi_h(g_j) = \prod_{l=1}^k e^{\frac{2\pi i \alpha_l h_l g_{jl}}{d}} = 1, \quad (1.4.37)$$

if and only if

$$\sum_{l=1}^k \alpha_l h_l g_{jl} \equiv 0 \pmod{d}. \quad (1.4.38)$$

So to find elements of H we have to solve this system of t linear equations. This is a simple linear algebra problem that can be efficiently solved with the use of Smith normal forms. Solving this equation gives an element

$$h = (h_1, \dots, h_k) \in H. \quad (1.4.39)$$

Repeating the procedure will lead to a set of generators for H .

1.4.3 Shor's Factoring Algorithm

Let N be an integer. We want to find an integer $1 < p < N$, such that $p \mid N$. By repeating this process for the integers p and $q = \frac{N}{p}$ we will eventually find a factorization

$$N = \prod_{i=1}^n p_i^{e_i}, \quad (1.4.40)$$

where p_i are prime numbers and e_i are positive integers. The fundamental theorem of arithmetic tells us that this factorization is unique. The problem is to find integers p_i that divide N . The factoring algorithm proposed by Shor [Sho97] is designed to find the order r of an element x modulo N , which is the smallest positive integer, such that

$$x^r \equiv 1 \pmod{N}. \quad (1.4.41)$$

If we can find such an element, then we verify whether

$$x^{\frac{r}{2}} \not\equiv -1 \pmod{N}. \quad (1.4.42)$$

If this is the case we compute

$$\text{GCD}(x^{\frac{r}{2}} \pm 1, N), \quad (1.4.43)$$

and we might find a non-trivial factor of N . The quantum part of this algorithm revolves around the Quantum Fourier Transform and Quantum Phase Estimation.

Quantum Phase Estimation

Let U be a unitary operator and let $|u\rangle$ be an eigenvector of U with eigenvalue $e^{2\pi i\phi}$. So

$$U|u\rangle = e^{2\pi i\phi}|u\rangle. \quad (1.4.44)$$

The purpose of phase estimation is to find an approximation $\tilde{\phi}$ for the unknown value $0 \leq \phi < 1$. The quantum algorithm for phase estimation uses

two registers of qubits. The first register $|0\rangle_k$ consists of k qubits initialized in the state $|0\rangle$. The number k depends on the desired accuracy of the approximation $\tilde{\phi}$ and on the desired success probability of the algorithm. The second register is initialized as $|u\rangle$ and takes as many qubits as are needed to describe $|u\rangle$. On each of the qubits of the first register a Hadamard operator is applied:

$$|0\rangle \mapsto \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle). \quad (1.4.45)$$

Then on each qubit

$$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)_{j+1} \quad (1.4.46)$$

of the first register a controlled- U^{2^j} gate is applied, where the integer j ranges from 0 to $k-1$:

$$\begin{aligned} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|u\rangle &\mapsto \frac{1}{\sqrt{2}}(|0\rangle|u\rangle + |1\rangle U^{2^j}|u\rangle) \\ &= \frac{1}{\sqrt{2}}(|0\rangle|u\rangle + |1\rangle e^{2\pi i \phi 2^j}|u\rangle) \\ &= \frac{1}{\sqrt{2}}(|0\rangle + e^{2\pi i \phi 2^j}|1\rangle)|u\rangle. \end{aligned} \quad (1.4.47)$$

Doing this operation on each of the k qubits of the first register, we obtain the following state:

$$\begin{aligned} |0\rangle_k &\mapsto \frac{1}{\sqrt{2^k}} \left((|0\rangle + e^{2\pi i \phi 2^{k-1}}|1\rangle) \cdots (|0\rangle + e^{2\pi i \phi 2^0}|1\rangle) \right) |u\rangle \\ &= \frac{1}{\sqrt{2^k}} \sum_{j=0}^{2^k-1} e^{2\pi i \phi j} |j\rangle, \end{aligned} \quad (1.4.48)$$

where we use the convention that if

$$j = a_0 \cdot 2^0 + \cdots + a_n 2^n, \quad (1.4.49)$$

with $a_i \in \{0, 1\}$, then $|j\rangle$ indicates the qubits $|a_0\rangle \cdots |a_n\rangle$. We can write

$$\phi = \left(\frac{a}{2^k} + \delta \right), \quad (1.4.50)$$

where $a = a_{k-1} \dots a_0$ is in binary notation,

$$|\delta| \leq \frac{1}{2^{k+1}}, \quad (1.4.51)$$

and $\frac{a}{2^k}$ is the best k -bit approximation of ϕ . This gives

$$\frac{1}{\sqrt{2^k}} \sum_{j=0}^{2^k-1} e^{2\pi i j \left(\frac{a}{2^k} + \delta \right)} |j\rangle. \quad (1.4.52)$$

We apply the inverse Fourier Transform on the first register, sending $|j\rangle$ to

$$\frac{1}{\sqrt{2^k}} \sum_{l=0}^{2^k-1} e^{-\frac{2\pi ijl}{2^k}} |l\rangle. \quad (1.4.53)$$

Putting this into the equation we obtain:

$$\begin{aligned} |0\rangle_k |u\rangle &\mapsto \frac{1}{\sqrt{2^k}} \sum_{j=0}^{2^k-1} e^{2\pi ij \left(\frac{a}{2^k} + \delta\right)} |j\rangle |u\rangle \\ &\mapsto \frac{1}{\sqrt{2^k}} \left(\sum_{j=0}^{2^k-1} e^{2\pi ij \left(\frac{a}{2^k} + \delta\right)} \frac{1}{\sqrt{2^k}} \sum_{l=0}^{2^k-1} e^{-\frac{2\pi ijl}{2^k}} |l\rangle \right) |u\rangle \\ &= \frac{1}{2^k} \sum_{j,l=0}^{2^k-1} e^{-\frac{2\pi ijl}{2^k}} e^{2\pi ij \left(\frac{a}{2^k} + \delta\right)} |l\rangle |u\rangle \\ &= \frac{1}{2^k} \sum_{j,l=0}^{2^k-1} e^{\frac{2\pi ij(a-l)}{2^k}} e^{2\pi ij\delta} |l\rangle |u\rangle. \end{aligned} \quad (1.4.54)$$

Now the first register is measured. There are two cases to consider. If $\delta = 0$, then we will measure exactly $|a\rangle = |\phi\rangle$. If $\delta \neq 0$, we will measure $|a\rangle$, the best k -bit approximation of ϕ with probability $p_a = |c_a|^2$, where

$$c_a = \frac{1}{2^k} \sum_{j=0}^{2^k-1} (e^{2\pi i\delta})^j. \quad (1.4.55)$$

This is a geometric series which can be bounded with some trigonometric manipulations to obtain

$$p_a \geq \frac{4}{\pi^2} \geq 0.4. \quad (1.4.56)$$

Order finding

We use quantum phase estimation to find the order of an element x modulo N . The quantum algorithm for finding the order of x uses the unitary operator U_x that acts in the following way:

$$U_x |y\rangle = |xy \pmod{N}\rangle. \quad (1.4.57)$$

The eigenstates of this operator are

$$|u_s\rangle = \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} e^{-\frac{2\pi isk}{r}} |x^k \pmod{N}\rangle, \quad (1.4.58)$$

with $0 \leq s \leq r - 1$ an integer. Indeed we have that

$$\begin{aligned} U_x |u_s\rangle &= \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} e^{-\frac{2\pi i s k}{r}} |x^{k+1} \pmod{N}\rangle \\ &= e^{\frac{2\pi i s}{r}} |u_s\rangle. \end{aligned} \quad (1.4.59)$$

So the eigenvalues of U_x are $e^{\frac{2\pi i s}{r}}$, with $0 \leq s \leq r - 1$ an integer.

We apply the quantum phase estimation algorithm on U_x to obtain approximations of $\phi = \frac{s}{r}$. There are two problems that need to be solved to execute this algorithm. We have to efficiently implement controlled- U^{2^j} operators for integers j and we need to prepare an eigenstate $|u_s\rangle$ with a non-trivial eigenvalue. The first of these problems can be overcome by modular exponentiation.

Modular Exponentiation Modular exponentiation means computing the remainder when dividing a positive integer x^k by a positive integer N . That is, we want to compute x' , such that:

$$x' \equiv x^k \pmod{N}. \quad (1.4.60)$$

If we compute this value by first calculating x^k and then computing the remainder modulo N , then this would require $O(k)$ multiplications to complete. This method can be slightly improved by using the following relation:

$$a \cdot b \pmod{m} \equiv (a \pmod{m}) \cdot (b \pmod{m}) \pmod{m}. \quad (1.4.61)$$

So after each multiplication by x we compute the remainder modulo N . This will reduce the size of the numbers that need to be multiplied, saving memory, but this still requires $O(k)$ multiplications.

A third method reduces both the number of operations and the memory required to perform modular exponentiation. It is a combination of the previous method and a more general principle called binary exponentiation. We first convert k to a binary number:

$$k = \sum_{i=0}^{n-1} a_i 2^i, \quad (1.4.62)$$

where a_i is either 0 or 1. We can then write x^k in binary form:

$$x^k = x^{\sum_{i=0}^{n-1} a_i 2^i} = \prod_{i=0}^{n-1} (x^{2^i})^{a_i}. \quad (1.4.63)$$

Therefore x' is equal to:

$$x' \equiv \prod_{i=0}^{n-1} (x^{2^i})^{a_i} \pmod{m}. \quad (1.4.64)$$

The running time of this algorithm is $O(\log k)$.

Eigenstate Preparation The second problem that needed to be overcome was the preparation of an eigenstate $|u_s\rangle$ without the knowledge of the order r . It is relatively straightforward to prove that

$$\frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} e^{-\frac{2\pi i s k}{r}} |u_s\rangle = |x^k \pmod{N}\rangle. \quad (1.4.65)$$

Using this result with $k = 0$, we obtain

$$\frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} |u_s\rangle = |1\rangle. \quad (1.4.66)$$

The quantum state we produce before applying the inverse QFT is

$$|\phi\rangle_1 |\phi\rangle_2 = \sum_{j=0}^{2^n-1} |j\rangle U^j |1\rangle = \sum_{j=0}^{2^n-1} |j\rangle |x^j \pmod{N}\rangle, \quad (1.4.67)$$

where n is the size of the first register of qubits and is of size $O(\log N)$. In the end we have an n -bit approximation of $\phi = \frac{s}{r}$. We would like to find r from this result and we can do this by using the continued fraction algorithm.

Theorem 1.7. *Let $\frac{s}{r} \in \mathbb{Q}$ be such that*

$$\left| \phi - \frac{s}{r} \right| \leq \frac{1}{2r^2}. \quad (1.4.68)$$

Then $\frac{s}{r}$ is a convergent of the continued fraction of ϕ and can be computed by the continued fraction algorithm.

This algorithm produces numbers r', s' with no common factor, such that

$$\frac{s'}{r'} = \frac{s}{r}. \quad (1.4.69)$$

There are two ways for the algorithm to fail. The phase estimation algorithm may produce a bad estimate of $\frac{s}{r}$ in which case the above theorem no longer applies. The probability of this event depends on the size of the first register and can be made negligibly small. The second problem is that s will be randomly chosen by the quantum algorithm, when we measure, and there is always the possibility that it is a divisor of r . In that case r' will be a divisor of r and not r itself. If this happens, then

$$x^{r'} \not\equiv 1 \pmod{N}. \quad (1.4.70)$$

We repeat the algorithm to obtain r'', s'' . If $r'' \neq r$ and $\text{GCD}(s'', s') = 1$, then

$$r = \text{LCM}(r'', r'). \quad (1.4.71)$$

The probability that $\text{GCD}(s'', s') = 1$ is at least $\frac{1}{4}$.

Reducing factoring to order finding

To reduce factoring a number N to computing the order of an element x modulo N we need the following theorems:

Theorem 1.8. *Let N be a composite positive integer and $x \neq \pm 1$ a non-trivial solution to the equation $x^2 \equiv 1 \pmod{N}$. Then at least one of $\text{GCD}(x-1, N)$ and $\text{GCD}(x+1, N)$ is a non-trivial factor of N .*

Theorem 1.9. *Suppose $N = \prod_{i=1}^n p_i^{\alpha_i}$ is the prime factorization of an odd composite positive integer. Let $1 \leq x \leq N-1$ be chosen at random. Let r be the order of x modulo N . Then the probability that r is even and that*

$$x^{\frac{r}{2}} \not\equiv -1 \pmod{N}, \quad (1.4.72)$$

is at least $1 - \frac{1}{2^n}$.

So in order to factor a number N we randomly choose a positive integer x smaller than N . We use the order finding algorithm to find the order r of x modulo N . If r is even, we compute $y \equiv x^{\frac{r}{2}} \pmod{N}$ and check whether $y \not\equiv -1 \pmod{N}$. If this is the case we compute $\text{GCD}(y \pm 1, N)$ and test whether either of these is a non-trivial factor of N . The performance of this algorithm is $O(\log^3 N)$ if we use simple multiplication and $O(\log^2 N \log \log N \log \log \log N)$ if we use fast multiplication.

1.4.4 Grover's Search Algorithm

Grover's algorithm is a quantum algorithm to search an unsorted database with N entries in $O(\sqrt{N})$ time and using $O(\log N)$ storage space [Gro97]. In classical computation searching an unsorted database cannot be done in less than linear time $O(N)$. Grover's algorithm provides a quadratic speedup, unlike other quantum algorithms, which may provide exponential speedup over their classical counterparts. Consider an unsorted database with N entries. The algorithm requires an N -dimensional state space \mathcal{H} , which can be supplied by $\log N$ qubits. For simplicity we will assume that $N = 2^n$ and that the search problem has exactly one solution. It is possible to generalize Grover's algorithm to search problems with M solutions, but we will not do so here. The database entries are $1, 2, \dots, N$. We call this set V . We suppose that i_0 is the solution to the search problem. Let $f: V \rightarrow \{0, 1\}$ be a function, such that $f(x) = 1$ if and only if x is the solution to the search problem. Suppose we have a unitary operator O , such that

$$O|x\rangle|y\rangle = |x\rangle|y \oplus f(x)\rangle. \quad (1.4.73)$$

If we put the second register $|y\rangle$ in the superposition

$$|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle), \quad (1.4.74)$$

then we have that

$$O|x\rangle|-\rangle = (-1)^{f(x)}|x\rangle|-\rangle. \quad (1.4.75)$$

This operator is called the oracle. Grover's algorithm uses two registers. The first register consists of n qubits initialized in state $|0\rangle_n$. The second register has one qubit and is initialized in state $|1\rangle$. We start by applying the Hadamard operator on the first register

$$H^{\otimes n}|0\rangle_n = \frac{1}{\sqrt{N}} \sum_{i=1}^N |i\rangle = |\phi\rangle_n, \quad (1.4.76)$$

and on the second register

$$H|1\rangle = |-\rangle. \quad (1.4.77)$$

We apply the oracle operator O to the first register and obtain

$$\begin{aligned} O(|\phi\rangle_n|-\rangle) &= \frac{1}{\sqrt{N}} \sum_{i=1}^N (-1)^{f(i)} |i\rangle|-\rangle \\ &= |\phi_1\rangle_n|-\rangle. \end{aligned} \quad (1.4.78)$$

The first register is a superposition of states, but the searched element has negative amplitude while all other elements have positive amplitude. The next steps of Grover's algorithm slowly increase this negative amplitude, while decreasing the positive amplitudes, making it more likely that a measurement of the first register results in the searched element. We have the following equality:

$$|\phi_1\rangle_n = |\phi\rangle_n - \frac{2}{\sqrt{N}}|i_0\rangle. \quad (1.4.79)$$

Moreover, we have $\langle\phi|\phi\rangle = 1$ and $\langle\phi|i_0\rangle = \frac{1}{\sqrt{N}}$. We apply the operator

$$R = 2|\phi\rangle\langle\phi| - I \quad (1.4.80)$$

on the first register and get

$$\begin{aligned} R|\phi_1\rangle_n &= (2|\phi\rangle\langle\phi| - I)(|\phi\rangle_n - \frac{2}{\sqrt{N}}|i_0\rangle) \\ &= (1 - \frac{4}{N})|\phi\rangle_n + \frac{2}{\sqrt{N}}|i_0\rangle \\ &= |\phi_G\rangle_n. \end{aligned} \quad (1.4.81)$$

Grover's algorithm consists of repeatedly applying the operator

$$G = R \circ O = (2|\phi\rangle\langle\phi| - I) \circ O \quad (1.4.82)$$

on the qubits. We have the following geometric interpretation.

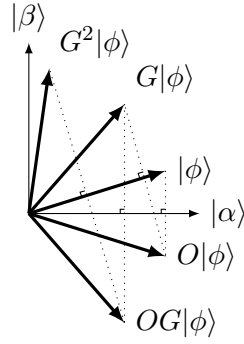


Figure 1.2: A geometrical interpretation of Grover's search algorithm: successive reflections around the axes $|\alpha\rangle$ and $|\phi\rangle$.

Let

$$|\alpha\rangle = \frac{1}{\sqrt{N-1}} \sum_{i \neq i_0} |i\rangle, \quad (1.4.83)$$

and $|\beta\rangle = |i_0\rangle$. We can write

$$|\phi\rangle = \sqrt{\frac{N-1}{N}} |\alpha\rangle + \sqrt{\frac{1}{N}} |\beta\rangle. \quad (1.4.84)$$

Let

$$\cos \frac{\theta}{2} = \sqrt{\frac{N-1}{N}}, \quad (1.4.85)$$

then

$$|\phi\rangle = \cos \frac{\theta}{2} |\alpha\rangle + \sin \frac{\theta}{2} |\beta\rangle. \quad (1.4.86)$$

After straightforward computation we find that

$$G|\phi\rangle = \cos \frac{3\theta}{2} |\alpha\rangle + \sin \frac{3\theta}{2} |\beta\rangle, \quad (1.4.87)$$

and more generally

$$G^k |\phi\rangle = \cos \frac{(2k+1)\theta}{2} |\alpha\rangle + \sin \frac{(2k+1)\theta}{2} |\beta\rangle. \quad (1.4.88)$$

We have that

$$\theta = 2 \arccos \left(\sqrt{\frac{N-1}{N}} \right), \quad (1.4.89)$$

so the number of times we need to apply G verifies the equation

$$k\theta + \frac{\theta}{2} = \frac{\pi}{2}. \quad (1.4.90)$$

So

$$k = \left\lfloor \frac{\pi - \theta}{2\theta} \right\rfloor. \quad (1.4.91)$$

Setting

$$\theta = 2 \arccos \left(\sqrt{\frac{N-1}{N}} \right), \quad (1.4.92)$$

and using the Taylor expansion for arccos we get

$$k = \left\lceil \frac{\pi\sqrt{N}}{4} \right\rceil. \quad (1.4.93)$$

If we apply the Grover operator k times and we measure the first register, then the probability of obtaining i_0 is close to 1.

1.5 Physical Realisations

1.5.1 Introduction

There are several conditions that a physical system needs to verify to be a good candidate for a physical realization of a quantum computer. The qubits need a robust physical representation where they retain their quantum mechanical properties. The system itself must allow us to perform a universal family of unitary transformations. It should be possible to prepare the qubits in a specified set of initial states and it should be possible to measure the final output states of the qubits.

The difficulty with physical realizations for quantum computation is that these requirements are often only partially met. An important obstacle for quantum computers is decoherence, which are processes that corrupt the desired evolution of the system. Every physical realization has a decoherence time τ_Q . Operations on qubits need to be performed in this time, because after a time τ_Q , the evolution becomes unreliable. An operation on a qubit usually takes some predefined time τ_{op} , depending on the physical system that is chosen. The ratio $\frac{\tau_Q}{\tau_{op}}$ indicates the maximum number of operations that can be performed on the system before it becomes decoherent.

Representing qubits

Quantum computation is based on unitary transformations on quantum states. Qubits are two-level quantum systems and provide a useful method of labeling for pairs of states. For instance a spin $\frac{3}{2}$ particle has four states. We could make the following correspondence:

$$|m = \frac{3}{2}\rangle = |00\rangle, \quad |m = \frac{1}{2}\rangle = |01\rangle, \quad (1.5.1ab)$$

$$|m = -\frac{1}{2}\rangle = |10\rangle, \quad |m = -\frac{3}{2}\rangle = |11\rangle. \quad (1.5.1cd)$$

So we could use such a particle to represent two qubits. It is important to make a good choice to represent qubits. A poor representation results in general in a quantum system with a short decoherence time.

A good measure of decoherence for single qubits is the minimum lifetime of an arbitrary superposition of the ground states. This measure is called T_2 , the transverse relaxation time. As the name suggests, there exists also another measure for decoherence. The longitudinal relaxation time T_1 is the relaxation time of the higher energy state $|1\rangle$.

Performing unitary transformations

A natural goal for experimental quantum computation is to be able to perform arbitrary unitary transforms on a single qubit and a CNOT transform on two qubits. If the system allows us to perform these operations, then, in theory, we can perform any arbitrary unitary transform on more than one qubit. There are some issues that need to be made clear. In order to have such an arbitrary unitary operation, we need to be able to address individual qubits and arbitrary pairs of qubits, without disturbing the other qubits. When there is an error in a unitary transform, this error will propagate, causing decoherence.

State preparation

If we want to make a quantum computation, we need to be able to initialize the qubits to represent the input of the computation. In classical computing, the initialization rarely poses any serious problems, but in quantum computing this is no longer true. Depending on the physical realization it may be very difficult to interact with the qubits. There is one positive point to make though. If we have any arbitrary one qubit transformation at our disposal, then we will only need to produce one initialized state with high fidelity. All other starting states can be obtained from this state by applying a unitary transform on it. In many physical realizations, the initialization of choice is the ground state $|0 \dots 0\rangle$. There are two measures that indicate the quality of initial state preparation. The first one is the minimum fidelity of the quantum gate needed to transform the ground state to an input state $|x_0 \dots x_n\rangle$. The second one is the entropy of the initial state. In general, input states that have non-zero entropy reduce the accessibility of the answer from the output state.

Measurement

We can consider the measurement of the qubits as a process where the qubits are coupled to a classical system, which permits to read the state of the qubits. An important characteristic of the process of measurement is the collapse of the wave function in case of projective measurement. Quantum algorithms need to be designed in such a way that when the output is measured, a useful result is found with high probability. Measuring qubits is not a simple process. Projective measurements can be difficult to implement as they need a large coupling between the quantum system and the classical system. Furthermore, we only want to make measurements when we choose to do so. Unwanted measurements can be considered as a decoherence process and are therefore undesirable. So the coupling between the quantum and classical systems should not be too large either. The signal to noise ratio is usually a good indicator of the measurement capability of a system.

1.5.2 Optical photon quantum computer

Physical description

Photons are particles without charge that do not interact strongly with each other. It is possible to guide photons along long distances in optical fibers with low loss. They can be manipulated in several ways. It is possible to delay photons with phase shifters and to combine them with beamsplitters. A photon can be represented as a qubit in the following way. The energy in an electromagnetic cavity is quantized in units of $\hbar\omega$. Each such quantum is called a photon. Consider two cavities whose total energy equals $\hbar\omega$.

We can then describe the states of the qubit as being the cavity in which the photon is located. That is state $|0\rangle$ corresponds to a photon in the first cavity and $|1\rangle$ to a photon in the second cavity. Single photons can be detected for a wide range of wavelengths.

There are several devices to manipulate qubits. Mirrors with high reflectivity reflect photons and change their propagation direction in space. Phase shifters, which are just transparent media with a different refraction index than the vacuum. Propagation of photons through such a medium will result in a phase shift. Beamsplitters, which are partially silvered pieces of glass, reflect a fraction R of the incident photons and transmit a fraction $1 - R$ of the incident photons.

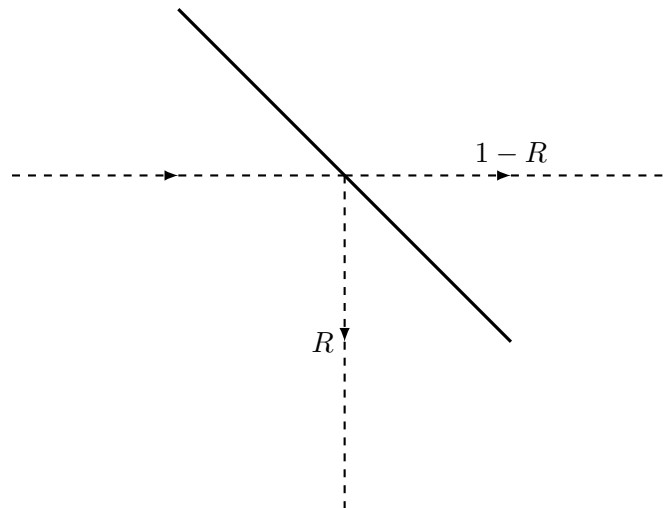


Figure 1.3: A beamsplitter that reflects a fraction R of incident photons and transmits a fraction $1 - R$.

A material that has a refraction index that is proportional to the total intensity I of light going through it is called a non-linear Kerr medium. This medium has a non-linear effect on the qubits and is used for interaction between photons.

Quantum computing

The three key elements for quantum computing are the phase shifter, the beamsplitter and a non-linear Kerr medium.

The phase shifter P acts on a qubit

$$|q\rangle = \alpha|0\rangle + \beta|1\rangle \quad (1.5.2)$$

in the following way:

$$P|q\rangle = \alpha e^{-\frac{i\Delta}{2}}|0\rangle + \beta e^{\frac{i\Delta}{2}}|1\rangle, \quad (1.5.3)$$

where

$$\Delta = \frac{(n - n_0)L}{c_0}, \quad (1.5.4)$$

with n the refraction index of light through the medium of the phase shifter, n_0 that through vacuum, L the distance the light travels through the medium and c_0 the speed of light in the medium. So the phase shifter acts as a rotation around the z -axis on a single qubit.

The beamsplitter B acts on a qubit $|q\rangle$ in the following way:

$$B|q\rangle = (\alpha \cos \theta - \beta \sin \theta)|0\rangle + (\alpha \sin \theta + \beta \cos \theta)|1\rangle, \quad (1.5.5)$$

where the angle θ of the beamsplitter verifies the equation

$$R = \cos \theta, \quad (1.5.6)$$

with R the fraction of incident light on the beamsplitter that is reflected. The beamsplitter acts as a rotation around the y -axis. The beamsplitter and the phase shifter together allow us to make arbitrary single qubit operations. The non-linear Kerr medium K is used for operations on two qubits:

$$K = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\chi L} \end{pmatrix}, \quad (1.5.7)$$

where L is the distance the light travels through the medium and χ is a characteristic coefficient of the Kerr medium. If the length L is set, such that

$$\chi L = \pi, \quad (1.5.8)$$

then the matrix for K becomes:

$$K = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (1.5.9)$$

We have the following relation:

$$\begin{aligned}
CNOT &= (I \otimes H)K(I \otimes H) \\
&= \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \tag{1.5.10}
\end{aligned}$$

where H is the Hadamard operator. The gate for $\chi L = \pi$ is also called the CZ or ControlZ gate. We can combine the three basic operations on the qubits to make a CNOT operator. This, combined with arbitrary operations on single qubits is in theory sufficient for any quantum operator.

Drawbacks

While single photons are easily generated and measured, it is difficult to make photons interact. The best non-linear Kerr media available are very weak and cannot provide a cross phase modulation of π between single photon states. Moreover, there is usually absorption associated with the non-linearity of a Kerr medium and it is estimated that nearly 50 photons need to be absorbed in order to experience a π cross phase modulation on a single photon. Therefore, the decoherence of the system will be very large.

1.5.3 Trapped ions

Physical description

An ion trap quantum computer consists of an electromagnetic trap with lasers and photodetectors, and ions. The electromagnetic trap is constructed from four cylindrical electrodes, with the end segments biased at a different voltage U_0 than the middle. Therefore, the ions are axially confined by a static potential

$$\Phi_{\text{St}} = \frac{\kappa U_0}{2} (z^2 - x^2 - y^2) \tag{1.5.11}$$

along the z -axis, where κ is a geometrical factor. A charge cannot be confined in three dimensions by static potentials and therefore two of the electrodes are grounded while the other two electrodes are driven by a fast oscillating voltage which creates a radiofrequency potential

$$\Phi_{\text{RF}} = \frac{(U_0 \cos \omega t + U_r)(1 - \frac{x^2 - y^2}{R^2})}{2}, \tag{1.5.12}$$

where R is a geometrical factor. The combination of these two potentials creates a harmonic potential. The motion of the electromagnetically confined ion becomes quantized when it is sufficiently well isolated. The purpose of the electromagnetic trap is to allow ions to be cooled to the extent that their vibrational state is close to having zero phonons. This will be the qubit state $|0\rangle$. The internal atomic states of a trapped ion form a qubit representation. These states are a combination of electron spin S and nuclear spin I , giving a total spin $F = S + I$. Suppose that an ion has an electron spin $\frac{1}{2}$ and a nuclear spin $\frac{1}{2}$. Each of these spins could be either $\frac{1}{2}$ or $-\frac{1}{2}$. This would give the obvious computational basis \mathcal{B} :

$$\mathcal{B} = \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}, \quad (1.5.13)$$

where $|ij\rangle$ might correspond to a trapped ion with electron spin $(-1)^i \cdot \frac{1}{2}$ and nuclear spin $(-1)^j \cdot \frac{1}{2}$. In physics, a basis consisting of eigenstates of the total momentum operator is preferred. This operator is defined by the Pauli operators:

$$\sigma^X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (1.5.14abc)$$

and the directional momentum operators:

$$J_x = \frac{\sigma_1^X + \sigma_2^X}{2}, \quad J_y = \frac{\sigma_1^Y + \sigma_2^Y}{2}, \quad (1.5.15ab)$$

$$J_z = \frac{\sigma_1^Z + \sigma_2^Z}{2}, \quad J^2 = J_x^2 + J_y^2 + J_z^2, \quad (1.5.15cd)$$

where the subscripts indicate whether the operator acts on the electron or on the nuclear spin. The operator J^2 has the following eigenstates:

$$|0, 0\rangle_J = \frac{|01\rangle - |10\rangle}{\sqrt{2}}, \quad |1, -1\rangle_J = |00\rangle, \quad (1.5.16ab)$$

$$|1, 0\rangle_J = \frac{|01\rangle + |10\rangle}{\sqrt{2}}, \quad |1, 1\rangle_J = |11\rangle. \quad (1.5.16cd)$$

These eigenstates are described as $|j, m_j\rangle_J$, which are eigenstates of the operator J^2 with eigenvalue $j(j+1)$ and of the operator J_z with eigenvalue m_j .

Quantum computing

The key element for quantum computing with spins is an electromagnetic field. If we apply an electromagnetic field of frequency ω_0 with the right angle and duration we can construct arbitrary single qubit operations.

Drawbacks

While the scaling of ion traps to a large number of qubits is conceptionally viable, there are two limitations to ion trap quantum computers. Phonon lifetimes are short, therefore the decoherence of a trapped ion is large. Moreover, it is not easy to prepare these ions in their motional ground states.

1.5.4 Other physical realizations

Several other physical implementation schemes for quantum computers are possible. We will describe a few of those. Quantum computing by nuclear magnetic resonance will be treated in much greater detail in the next chapter.

Quantum dots

A fundamental quantum unit that could serve as qubit representation is electric charge. It is possible with modern electronics to manipulate charges at the level of a single electron. Quantum dots are three-dimensional boxes with electrostatic potentials that confine electric charge quanta. Unlike photons, net charge cannot be destroyed and therefore it is necessary to use two boxes with only one charge quantum to represent a qubit. Single qubit operations can be performed by electrostatic gates and single mode waveguide couplers for moving electrons and tunnel junctions for quantum dots. The long-range Coulomb interaction of the electric charge can be used to perform operations on two qubits. It is simple to measure single electron charges using modern field effect transistors. Decoherence occurs through uncontrolled distant charge motion.

Superconductors

At low temperature in certain metals two electrons can bind together through a phonon interaction to form a Cooper pair, with charge $2e$. These pairs can be confined within an electrostatic box. A qubit is represented by one Cooper pair in two boxes. Single qubit gates are realized by electrostatic gates to modulate the box potential and Josephson junctions between coupled boxes. Josephson junctions are also used to couple different qubits, where an external magnetic field coupled to the superconducting interferometric loops is used. Qubits are measured by measuring the electric charge in a box. Cooper pairs are relatively robust and therefore the main decoherence factor is spontaneous emission of electromagnetic photons.

Chapter 2

All things are difficult before
they are easy.

THOMAS FULLER

Nuclear Magnetic Resonance and Quantum Computing

2.1 Nuclear Magnetic Resonance

2.1.1 Introduction

A much longer introduction to Nuclear Magnetic Resonance can be found in the books of Shaw and Slichter [Sha76, Sli80]. A magnetic system that possesses both magnetic moments and angular momentum can exhibit a phenomenon called magnetic resonance. If the magnetic system is a nucleus we speak of nuclear magnetic resonance. The fact that nuclei can have magnetic moments was first suggested in 1924 by Pauli, while studying the hyperfine structure of atomic spectra [Pau24].

The angular momentum of nuclei is quantized and nuclei have a quantum number I which can be any half integer value. A nucleus with quantum number I has an angular momentum of $I\hbar$.

The quantization of atomic magnetic moments was already demonstrated in 1921 by Stern and Gerlach [Ste21]. Their techniques to distinguish various quantum states of atoms were refined to measure transition energies of nuclei. In 1945 two groups simultaneously discovered resonant absorption in bulk matter. Bloch et al. detected resonance absorption in water protons [BHP46] and Purcell et al. detected resonance absorption in paraffin wax [PTP46].

The nucleus possesses a total magnetic moment $\vec{\mu}$ and a total angular momentum \vec{J} . We can take these two vectors parallel and have the following equation:

$$\vec{\mu} = \gamma \vec{J}, \quad (2.1.1)$$

where γ is a scalar constant. This constant is called the gyromagnetic ratio. A classical first order approximation will give an estimate for γ . Consider a particle of mass m and charge e moving in a circular path of radius r with

period T . The angular momentum of this particle is:

$$\begin{aligned}\vec{\mathbf{J}} &= mvr \\ &= \frac{2\pi r^2 m}{T}.\end{aligned}\tag{2.1.2}$$

The magnetic moment of the particle can be computed if we treat the rotating particle as a current loop of area A with current i :

$$\begin{aligned}\vec{\mu} &= iA \\ &= \frac{e\pi r^2}{cT}.\end{aligned}\tag{2.1.3}$$

As we have the equation

$$\vec{\mu} = \gamma \vec{\mathbf{J}},\tag{2.1.4}$$

it follows that the gyromagnetic ratio verifies the following equation:

$$\gamma = \frac{e}{2mc}.\tag{2.1.5}$$

We now consider the consequences of placing a nucleus with a magnetic moment in a magnetic field $\vec{\mathbf{B}}_0$. We first consider this from a classical point of view. The nucleus is a magnetic dipole and will acquire an energy:

$$E = -\vec{\mu} \cdot \vec{\mathbf{B}}_0.\tag{2.1.6}$$

As the nucleus has angular momentum, it will not only align itself with the magnetic field $\vec{\mathbf{B}}_0$, but it will also precess with a frequency ω_0 at an angle θ about this field. This effect is caused by the interaction of the torque generated by rotational motion of the nucleus and the magnetic field of the nuclear magnetic moment. The torque between the magnetic moment of the nucleus and the field is

$$\tau = \vec{\mu} \times \vec{\mathbf{B}}_0.\tag{2.1.7}$$

The torque is equal to the rate of change of angular momentum:

$$\begin{aligned}\tau &= \frac{d\vec{\mathbf{J}}}{dt} \\ &= \omega_0 \vec{\mathbf{J}}.\end{aligned}\tag{2.1.8}$$

The frequency of this precession is therefore

$$\omega_0 = \gamma |B_0|,\tag{2.1.9}$$

which is called the Larmor frequency. It is the basic phenomenon of NMR. The magnetic field is proportional to the precession frequency and the proportionality constant is the gyromagnetic ratio.

We now consider the basic properties of NMR from a quantum mechanical point of view. First we define the dimensionless angular momentum operator $\vec{\mathbf{I}}$ by:

$$\vec{\mathbf{J}} = \hbar\vec{\mathbf{I}}. \quad (2.1.10)$$

The operator \mathbf{I}^2 has eigenvalues I , which are either integer or half-integer. All components of $\vec{\mathbf{I}}$ commute with \mathbf{I}^2 . The operator \mathbf{I}_z has eigenvalues m , where m can be any of the $2I + 1$ values $-I, \dots, I$.

The application of a magnetic field $\vec{\mathbf{B}}$ produces an interaction energy of the nucleus of amount $-\vec{\boldsymbol{\mu}} \cdot \vec{\mathbf{B}}$. If we take the magnetic field to be B_0 along the z -direction we have the following Hamiltonian:

$$\mathcal{H} = -\gamma\hbar B_0 \mathbf{I}_z. \quad (2.1.11)$$

The eigenvalues of this Hamiltonian are multiples $\gamma\hbar B_0$ of the eigenvalues of \mathbf{I}_z and therefore the allowed energies are:

$$E = -\gamma\hbar B_0 m, \quad \text{with } m = -I, \dots, I. \quad (2.1.12)$$

We want to detect such a set of energy levels by spectral absorption. Therefore an interaction is needed that causes transitions between energy levels. Such an interaction must be time dependent and of angular frequency ω , such that:

$$\hbar\omega = \Delta E, \quad (2.1.13)$$

where ΔE is the difference of energy between two levels of the spectrum.

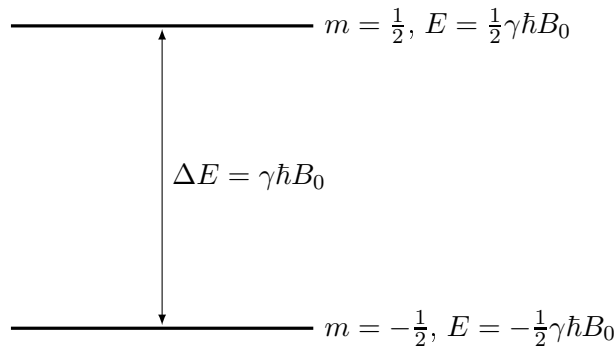


Figure 2.1: Energy levels for a spin $\frac{1}{2}$ particle

The coupling used to produce magnetic resonance is an alternating magnetic field of amplitude $\vec{\mathbf{B}}_1$ perpendicular to the static magnetic field. The Hamiltonian of this alternating field is:

$$\mathcal{H} = -\gamma\hbar\vec{\mathbf{B}}_1 \cdot \vec{\mathbf{I}}_x \cos(\omega t). \quad (2.1.14)$$

The allowed transitions are between adjacent energy levels and therefore:

$$\begin{aligned}\hbar\omega &= \Delta E \\ &= \gamma\hbar B_0. \\ &\Leftrightarrow \\ \omega &= \gamma B_0.\end{aligned}\tag{2.1.15}$$

We see that Planck's constant has disappeared from the resonance equation. If we can estimate γ , we can compute the frequency that produces a magnetic resonance.

We now consider a macroscopic sample of nuclei with spin $\frac{1}{2}$. Let N_+ be the number of nuclei in the state $m = \frac{1}{2}$ and N_- the number of nuclei in the state $m = -\frac{1}{2}$. Obviously, the total number of nuclei N verifies:

$$N = N_+ + N_-. \tag{2.1.16}$$

Moreover, the equilibrium populations N_+^0 and N_-^0 verify the equation:

$$\frac{N_-^0}{N_+^0} = e^{-\frac{\gamma\hbar B_0}{k_B T}}, \tag{2.1.17}$$

where k_B is the Boltzmann constant.

If we apply an alternating magnetic field, the total number of nuclei will remain constant, but N_+ and N_- will vary because of the energy transitions induced by the field. The probability per second of inducing a transition from $m = \frac{1}{2}$ to $m = -\frac{1}{2}$ is equal to P_\downarrow and the probability per second of inducing a transition in the other direction is P_\uparrow . This leads to the following differential equation:

$$\frac{dN_+}{dt} = P_\uparrow N_- - P_\downarrow N_+. \tag{2.1.18}$$

We can rewrite this equation as the difference between the two populations:

$$n = N_+ - N_-, \tag{2.1.19}$$

and obtain the following differential equation:

$$\frac{dn}{dt} = \frac{n_0 - n}{T_1}, \tag{2.1.20}$$

where we have:

$$n_0 = N \left(\frac{P_\uparrow - P_\downarrow}{P_\uparrow + P_\downarrow} \right), \tag{2.1.21a}$$

$$\frac{1}{T_1} = P_\uparrow + P_\downarrow. \tag{2.1.21b}$$

The solution of this differential equation is:

$$n(t) = n_0 + Ce^{-\frac{t}{T_1}}, \quad (2.1.22)$$

with C a constant that depends on n , n_0 the thermal equilibrium population difference and T_1 a characteristic time associated with the approach to thermal equilibrium. This characteristic time T_1 is called the spin-lattice relaxation time.

2.2 Quantum computing with NMR

2.2.1 Ensemble system

NMR differs from other physical realizations of a quantum computer in the sense that instead of a single photon or other physical entity it uses an ensemble of systems as single qubit representation. As a direct consequence, the measurement is also an ensemble average. Furthermore, it is technically infeasible to prepare the ensemble in a special state such as the ground state, therefore the initial state will be the thermal equilibrium state:

$$|\rho\rangle = \frac{e^{-\beta\mathcal{H}}}{\mathcal{Z}}, \quad (2.2.1)$$

where \mathcal{H} is the Hamiltonian of the system, $\beta = \frac{1}{k_B T}$ and $\mathcal{Z} = \text{Trace}(e^{\beta\mathcal{H}})$ is the partition function normalisation to ensure that the trace of ρ is equal to 1. For modest fields at room temperature we can use the approximation:

$$|\rho\rangle \approx 2^{-n} (1 - \beta\mathcal{H}), \quad (2.2.2)$$

where the system has n spins. As spin-spin couplings are small compared to the precession frequencies, we can interpret the thermal state density matrix as a mixture of the pure states $|00\dots 0\rangle, \dots, |11\dots 1\rangle$.

The principal output of an experiment is the free induction decay signal:

$$V(t) = V_0 \text{Tr} \left(e^{-i\mathcal{H}t} \rho e^{i\mathcal{H}t} (iX_k + Y_k) \right), \quad (2.2.3)$$

where X_k and Y_k operate only on the spin k , and V_0 is a constant that depends on the coil, the quality factor and the sample volume. This induction signal has an exponential decay, which is caused by several factors. The inhomogeneity of the static magnetic field, spin-spin coupling resulting in phase randomisation and thermalisation of the spins to their equilibrium are all contributing to the exponential decay of the signal.

For successful quantum computation we need to perform unitary transformations to a properly initialized qubit and to measure the output. In the ensemble approach of NMR quantum computing several problems need to be

addressed. First, how can we use the thermal state (2.2.1) to initialize our system? How can we perform arbitrary unitary transforms on this state? Most important of all, how can an ensemble average measurement produce the same results as projective quantum measurements?

2.2.2 Labeling the qubits

The initial state of our system is the thermal state. In order to perform quantum computation, we want to have an initial state of qubits $|0 \cdots 0\rangle$. There are several techniques to obtain this initial state from the thermal state. These techniques are called labeling techniques. We consider the temporal labeling technique, which is based on the fact that quantum operations are linear and that observables measured in NMR are traceless. Suppose that our initial thermal state for a two spin system is the density matrix:

$$\rho_0 = \begin{pmatrix} a_1 & 0 & 0 & 0 \\ 0 & a_2 & 0 & 0 \\ 0 & 0 & a_3 & 0 \\ 0 & 0 & 0 & a_4 \end{pmatrix}, \quad (2.2.4)$$

where the a_i are positive real numbers that sum to 1. Supposing furthermore that we can overcome our second problem of performing unitary transformations, we use SWAP-gates to obtain states with permuted populations:

$$\rho_1 = \begin{pmatrix} a_1 & 0 & 0 & 0 \\ 0 & a_3 & 0 & 0 \\ 0 & 0 & a_4 & 0 \\ 0 & 0 & 0 & a_2 \end{pmatrix}, \quad \rho_2 = \begin{pmatrix} a_1 & 0 & 0 & 0 \\ 0 & a_4 & 0 & 0 \\ 0 & 0 & a_2 & 0 \\ 0 & 0 & 0 & a_3 \end{pmatrix}. \quad (2.2.5ab)$$

A unitary quantum computation U is applied to each of these three thermal states in three separate experiments at different times, resulting in three different outcomes C_k :

$$C_k = U \rho_k U^{-1}. \quad (2.2.6)$$

We take the sum of these three outcomes to obtain the following result:

$$\begin{aligned} \sum_k C_k &= \sum_k U \rho_k U^{-1} \\ &= U \left(\sum_k \rho_k \right) U^{-1} \\ &= (4a_1 - 1)U \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} U^{-1} + (1 - a_1) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned} \quad (2.2.7)$$

In NMR the only observables that are measured, are traceless observables. Let M be such an observable. We have:

$$\begin{aligned} \text{Tr} \left(\sum_k C_k M \right) &= \sum_k \text{Tr} (C_k M) \\ &= (4a_1 - 1) \text{Tr} \left(U \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} U^{-1} M \right) \\ &= (4a_1 - 1) \text{Tr} (U|00\rangle\langle 00|U^{-1}). \end{aligned} \quad (2.2.8)$$

Therefore the sum of the three outcomes is proportional to the outcome of an initial state $|00\rangle$. This technique can always be accomplished if the decoherence time is sufficiently long. It is also possible to perform these different experiments at the same time but at a different space, using for instance magnetic field gradients. In that case we call the technique spatial labeling.

2.2.3 Unitary transformations

In order to perform arbitrary single qubit operations it is sufficient to apply a large RF at the correct frequency. We consider the following three rotation operators:

$$\begin{aligned} \mathbf{R}_x(\theta) &= e^{-\frac{i\sigma^X\theta}{2}} \\ &= \begin{pmatrix} \cos \frac{\theta}{2} & -i \sin \frac{\theta}{2} \\ -i \sin \frac{\theta}{2} & \cos \frac{\theta}{2} \end{pmatrix}, \end{aligned} \quad (2.2.9a)$$

$$\begin{aligned} \mathbf{R}_y(\theta) &= e^{-\frac{i\sigma^Y\theta}{2}} \\ &= \begin{pmatrix} \cos \frac{\theta}{2} & -\sin \frac{\theta}{2} \\ \sin \frac{\theta}{2} & \cos \frac{\theta}{2} \end{pmatrix}, \end{aligned} \quad (2.2.9b)$$

$$\begin{aligned} \mathbf{R}_z(\theta) &= e^{-\frac{i\sigma^Z\theta}{2}} \\ &= \begin{pmatrix} e^{-\frac{i\theta}{2}} & 0 \\ 0 & e^{\frac{i\theta}{2}} \end{pmatrix}. \end{aligned} \quad (2.2.9c)$$

In the Bloch sphere notation of qubits, these operators define rotations of an angle θ around the three coordinate axes. A rotation around an arbitrary axe $\hat{u} = (u_x, u_y, u_z)$ is given by:

$$\begin{aligned} \mathbf{R}_{\hat{u}}(\theta) &= e^{-\frac{i\theta\vec{\sigma}\cdot\hat{u}}{2}} \\ &= \cos \frac{\theta}{2} I - i \sin \frac{\theta}{2} (\sigma^X u_x + \sigma^Y u_y + \sigma^Z u_z). \end{aligned} \quad (2.2.10)$$

Let $\mathbf{R}_1 = \mathbf{R}_{x_1}(\frac{\pi}{2})$ be a rotation of $\frac{\pi}{2}$ around the x -axis of the first qubit and define \mathbf{R}_2 likewise for the second qubit. We have the identity:

$$\mathbf{R}_i^2 e^{-ia\sigma_i^Z t} \mathbf{R}_i^2 = e^{ia\sigma_i^Z t}. \quad (2.2.11)$$

This property is called the refocusing property and it is used as a technique to remove time evolution.

The ControlNot gate is built from a ControlZ gate, just as in other physical realizations. This CZ gate is built by using the scalar \mathbf{J} -coupling between qubits, which are indirect interactions, mediated by electrons shared through a chemical bond. We have the following identity:

$$\sqrt{i} e^{\frac{i\pi\sigma_1^Z\sigma_2^Z}{4}} e^{-\frac{i\pi\sigma_1^Z}{4}} e^{-\frac{i\pi\sigma_2^Z}{4}} = CZ. \quad (2.2.12)$$

So we can build the ControlZ gate and from equation (1.5.10) we can construct a ControlNot gate. We therefore have the basic operators to do quantum computation.

2.2.4 Ensemble measurements

Ensemble measurements are fundamentally different from measurements of a single series of qubits. Quantum algorithms are designed, such that no matter what state the qubit collapses to, the resulting measured amplitude will tell us something meaningful. In Shor's factoring algorithm for instance, we obtain a random fraction $\frac{p}{q}$, with p a random integer and q the outcome that will be extracted in the classical postprocessing phase. In an ensemble measurement we will not obtain this random fraction $\frac{p}{q}$, but rather an average over a large number of these kind of fractions. The problem is that this average does not contain any meaningful information that can be extracted. This difficulty can be overcome under certain conditions. If we are able to build quantum gates that can do the classical postprocessing part, then it is possible to have meaningful ensemble measurements. The idea is to wait for the measurement until after the postprocessing part is done in a quantum computational way and only then measure the outcome. In the example above, if we apply the continued fraction algorithm as a quantum algorithm, then our outcome would always be q . The average would therefore also always be q . This technique does beg the following question: if we can do the postprocessing on a quantum computer, is it not a better idea to always do the postprocessing in this fashion? There are several reasons not to do so. While Fourier Transformations have exponential speedup on a quantum computer, other algorithms do not have this advantage. Moreover, the decoherence on quantum computers is much more important than the decoherence on classical computers, where there is hardly decoherence at all. If we are obliged to do the postprocessing also quantum computationally, that effectively reduces the number of gate operations we can use for the

main part of a quantum algorithm before decoherence sets in. It is therefore preferable to have a classical postprocessing part of a quantum algorithm. Nevertheless, ensemble measurement can be given a useful meaning, but we have to adapt the quantum algorithms in order for the outcome to be useful.

2.3 Drawbacks

The physical realization of an NMR quantum computer by labeling atoms of molecules as qubits has met with impressive successes. The factoring of the number 15 by using Shor's factoring algorithm on 7 qubits can be considered the high mark of NMR as a quantum computer [VSB⁺01]. No other physical realization has so far been able to repeat this result. The NMR approach has nevertheless met with severe criticism.

From the point of view of long term development, physical realizations of quantum computers need to have several nice properties. One of them is scalability. If we can realize an N -qubit quantum computer in some sort of physical realization, it should be reasonable to hope that an $(N + 1)$ -qubit quantum computer can be realized by just slightly widening the physical constraints and some small additional effort. In classical computers the analog is clear: if we are able to place N chips on a circuit board, we expect that placing $(N + 1)$ chips would require some architectural effort, some designing constraints, but no fundamental problem whatsoever. The problem with the current approach of NMR quantum computing is that an N -qubit quantum computer would be some kind of complicated molecule, with each qubit some properly labeled atom in this molecule. If we would like to build an $(N + 1)$ -qubit quantum computer we cannot simply add another qubit to the system. We would have to design a new molecule altogether. Therefore, the NMR approach to quantum computing lacks scalability.

Another difficulty with using atoms of specifically designed molecules for quantum computing is the inherent architecture of the qubits. As we use the atoms of a molecule for qubits, some qubits will have quite some distance between them. The scalar coupling between these qubits, which is needed to make a CNOT operator, will be rather weak. Therefore it will be difficult to have direct operations between these qubits. It is possible to circumvent this problem by using a cellular automata style architecture, where an operation on distant qubits will be executed by a series of local operations moving from one qubit to the other qubit. While this approach may be possible, it will certainly come with an additional cost of extra operations which slow down the algorithms to be executed.

A third difficulty is the weak signal because of the labeling techniques used. By repeating experiments in a permutation, such that all other ground states except the initialization ground state cancel out, we may achieve initialization, but the probability of the initialization state will not be increased. If we

want to initialize our system in the ground state $|0 \cdots 0\rangle$, then the probability of this state is:

$$p_{0 \cdots 0} = \frac{1}{Z} \langle 0 \cdots 0 | e^{-\beta \mathcal{H}} | 0 \cdots 0 \rangle. \quad (2.3.1)$$

This probability is proportional to $n2^{-n}$, if we have a molecule with n qubits. Therefore the signal will decrease exponentially if the number of qubits increases. This problem might also be overcome by improving the labeling techniques and by using optical pumping methods, but there will always be a decrease in signal if the number of qubits increases.

The last criticism to the NMR approach for quantum computing is the most severe. It starts with the remark that for quantum computing to be efficient we need to be able to have entangled states [LP01]. That is to say states of the form

$$|ab\rangle, \quad (2.3.2)$$

that cannot be separated into two separate states

$$|a\rangle \otimes |b\rangle. \quad (2.3.3)$$

The mixed thermal state that we use in NMR quantum computing does not exhibit an entangled nature [BCJ+99] and it can therefore be argued that no real quantum computing takes place in an NMR quantum computer. This objection does not put into question the NMR approach in itself, but rather the use of thermal initialization states. As NMR quantum computing seems to need these thermal states, this seems like an insurmountable problem.

So what are the problems that NMR quantum computing needs to overcome? Scalability, architecture, signal loss in case of lots of qubits and the lack of entanglement in the so called thermal state.

In the next chapter we will try a different NMR approach where all these problems can be addressed. Our approach has of course problems of its own and whether any successful physical implementation of our scheme will be realized remains to be seen.

Chapter 3

Curiosity killed the cat, but for
a while I was a suspect.

STEVEN WRIGHT

Reviving the Nuclear Magnetic Resonance Approach

3.1 Introduction

As we noted at the end of the previous chapter, the NMR approach to quantum computing has lately met with rather severe criticism and has slowly been fading from the field of physical realizations. In 2001, bulk liquid NMR was the hotbed of quantum computation and physical realizations of quantum computers, but in the last years no major publication has appeared that continues to propose this approach for a physical realization. In order for bulk liquid NMR to be made viable again, at least three of the following problems need to be solved:

1. **Scalability:** A major objection to the NMR approach is the fact that it has no scalability whatsoever. Even augmenting the number of qubits by one would demand an entirely different molecule on which the qubits are labeled.
2. **Decoherence:** The thermal approach as initialization scheme for quantum computation has as a direct consequence that the signal decreases exponentially if the number of qubits increases.
3. **Entanglement:** The thermal approach does not exhibit entangled quantum states. These states are essential in the sense that without them, quantum computing cannot be faster than classical computing.
4. **Architecture:** The ideal architecture for a quantum computer is one where every qubit can communicate directly with every other qubit. In molecules this architecture is naturally unachievable and we have to use indirect interactions between distant qubits.

It is clear that the mixed thermal state is a major problem. If we could work with pure states, than we would not need to have a labeling scheme which

decreases the signal exponentially and we would not need to worry about not having entangled quantum states. We therefore reconsider bulk liquid NMR, but instead of using atoms on a specifically designed molecule as qubits, we will use magnetic field gradients to create different resonance frequencies for spatially separated parts of the liquid. We then proceed to assign to specific resonance frequencies the value of a logical qubit. We build a framework around this approach in which we show that we can properly initialize this system. This in itself lifts three of the major objections: we no longer need to use a mixed thermal state and using magnetic field gradients allows us to have an easily scalable quantum system. The architecture objection remains for the moment unaddressed as this problem only becomes an issue if we have a working system of more than some qubits. It is however conceptually not an insurmountable problem. For starters, via an indirect approach with interaction via nearest neighbour qubits, there will be some loss of efficiency, but it can be shown [Wat95, Llo93] that this still represents a universal quantum computer. Another reason why the architecture need not be an issue is the fact that we can potentially use magnetic field gradients in three directions in order to obtain more neighbours for each qubit.

The main problem with our approach is that while molecules have an obvious interaction for qubits by using the scalar interaction via the shared electron cloud, we do not have such an obvious interaction. We show that we cannot directly use the dipole moment between qubits as this is averaged away to zero, but we may use the long-range dipolar effect, which is not averaged to zero because of the geometrical constraints of the sample. This is still work in progress and it is as of yet unclear whether this approach will actually result in a useful interaction between qubits.

In the rest of this chapter we first describe the framework in which our computations are executed. Via this framework we obtain the methods to make single qubit gates as well as how to initialize these qubits. We conclude with a roadmap which if followed successfully should lead to a working NMR quantum computer. Those steps in this roadmap which have already been executed will be given together with the experimental data to support them.

3.2 Framework for Quantum Computing by Nuclear Magnetic Resonance

3.2.1 One spin $\frac{1}{2}$

Static field

For one nuclear spin $\frac{1}{2}$ in a \vec{B}_0 magnetic field, the Hamiltonian is:

$$\begin{aligned}\mathcal{H}_1 &= -\vec{\mu} \cdot \vec{B}_0 \\ &= -\gamma \vec{I}_1 \cdot \vec{B}_0,\end{aligned}\tag{3.2.1}$$

where γ is the gyromagnetic ratio of the nuclear spin and \vec{I}_1 is the spin that verifies the following equation:

$$\vec{I}_1 = \frac{1}{2}\hbar\vec{\sigma}, \quad (3.2.2)$$

with $\vec{\sigma}$ a vector defined by the following Pauli matrices:

$$\sigma^X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (3.2.3abc)$$

The matrix notation of equation (3.2.1) is given by:

$$\mathcal{H}_1 = -\frac{1}{2}\gamma\hbar \begin{pmatrix} B_Z & B_X - iB_Y \\ B_X + iB_Y & -B_Z \end{pmatrix}. \quad (3.2.4)$$

By convention, the vector $\vec{B}_0 = (B_X, B_Y, B_Z)$ is placed in the Oz -direction, that is $\vec{B}_0 = (0, 0, B_0)$ and the xOy -plane is called the transverse plane. Therefore the matrix form of the Hamiltonian reduces to:

$$\mathcal{H}_1 = -\frac{1}{2}\gamma\hbar \begin{pmatrix} B_0 & 0 \\ 0 & -B_0 \end{pmatrix}. \quad (3.2.5)$$

The two eigenvalues of the Hamiltonian \mathcal{H}_1 , which give the energy of the quantum states, are:

$$E_+ = -\frac{1}{2}\gamma\hbar B_0, \quad E_- = \frac{1}{2}\gamma\hbar B_0, \quad (3.2.6ab)$$

which have the following two corresponding eigenvectors:

$$|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (3.2.7ab)$$

We will use these eigenvectors as our canonical basis for computation. If we have more than one spin our canonical basis will not necessarily be the basis of eigenvectors. Sometimes we will use the qubit notation $|0\rangle, |1\rangle$ in stead of $|+\rangle, |-\rangle$.

The probability for a spin $\frac{1}{2}$ to be in either of these states is equal to $\frac{1}{2}$ at $T = 0$. At higher temperatures the probability to occupy a state depends on the temperature T . We can describe the wave function as:

$$|\psi(0)\rangle = a|+\rangle + b|-\rangle, \quad (3.2.8a)$$

$$|a|^2 + |b|^2 = 1. \quad (3.2.8b)$$

The wave function for a spin $\frac{1}{2}$ particle can be written as:

$$\begin{aligned} |\psi(0)\rangle &= \frac{1}{\sqrt{2}} \left(e^{-\frac{i\phi}{2}} |+\rangle + e^{\frac{i\phi}{2}} |-\rangle \right) \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-\frac{i\phi}{2}} \\ e^{\frac{i\phi}{2}} \end{pmatrix}. \end{aligned} \quad (3.2.9)$$

The time evolution far from the speed of light \vec{c} is given by the Schrödinger equation:

$$\frac{i\hbar\partial}{\partial t}|\psi_1(t)\rangle = \mathcal{H}|\psi_1(t)\rangle. \quad (3.2.10)$$

If we write

$$|\psi_1(t)\rangle = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \quad (3.2.11)$$

then we have the following differential equations:

$$\begin{aligned} i\hbar \begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix} &= -\frac{1}{2}\gamma\hbar \begin{pmatrix} B_0 & 0 \\ 0 & -B_0 \end{pmatrix} \cdot \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \\ &= \frac{1}{2}\hbar \begin{pmatrix} \omega_0 x(t) \\ -\omega_0 y(t) \end{pmatrix}, \end{aligned} \quad (3.2.12)$$

where

$$\omega_0 = -\gamma B_0 \quad (3.2.13)$$

is the resonance or Larmor frequency. The following time dependent wave function is the obvious solution of this system of equations:

$$|\psi_1(t)\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-\frac{i(\omega_0 t + \phi)}{2}} \\ e^{\frac{i(\omega_0 t + \phi)}{2}} \end{pmatrix}. \quad (3.2.14)$$

The effect of an RF magnetic field

In NMR a transition between the two states $|+\rangle$ and $|-\rangle$ is obtained by a B_1 magnetic field rotating in the transverse plane.

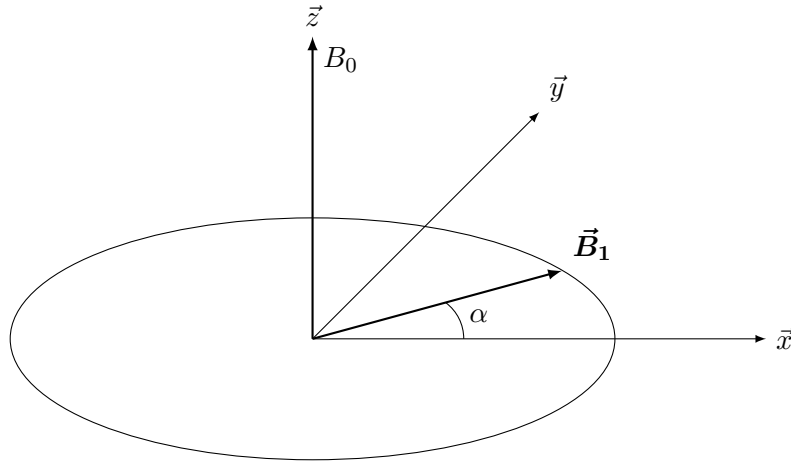


Figure 3.1: Magnetic field \vec{B}_1 rotating in the transverse plane.

This field is generated by an electromagnetic RF wave obtained by an oscillating current in a solenoid surrounding the spin system. In NMR spectroscopy the phase α of the RF field can be controlled. In this case the Hamiltonian has non-diagonal elements due to the RF magnetic field \vec{B}_1 rotating around \vec{B}_0 with an angular velocity ω :

$$\mathcal{H}_{\text{RF}}(t) = -\frac{1}{2}\gamma\hbar \begin{pmatrix} B_0 & B_1 e^{-i(\omega t + \alpha)} \\ B_1 e^{i(\omega t + \alpha)} & -B_0 \end{pmatrix}. \quad (3.2.15)$$

In this case the time evolution is no longer trivial. We have the following differential equations:

$$\begin{aligned} \frac{i\hbar\partial}{\partial t} |\psi_{\text{RF}}(t)\rangle &= \mathcal{H}_{\text{RF}}(t) \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \\ &= \frac{1}{2}\hbar \begin{pmatrix} \omega_0 & \omega_1 e^{-i(\omega t + \alpha)} \\ \omega_1 e^{i(\omega t + \alpha)} & -\omega_0 \end{pmatrix} \cdot \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \end{aligned} \quad (3.2.16)$$

where

$$\omega_1 = -\gamma B_1. \quad (3.2.17)$$

In the time-independent case these equations obviously reduce to the equations (3.2.9). For the time-dependent case we need to solve the following differential equations:

$$\begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix} = -\frac{i}{2} \begin{pmatrix} \omega_0 x(t) + \omega_1 e^{-i(\omega t + \alpha)} y(t) \\ \omega_1 e^{i(\omega t + \alpha)} x(t) - \omega_0 y(t) \end{pmatrix}. \quad (3.2.18)$$

To solve these two equations we make the following substitutions:

$$p(t) = x(t) e^{\frac{i\omega_0 t}{2}}, \quad (3.2.19a)$$

$$q(t) = y(t) e^{-\frac{i\omega_0 t}{2}}. \quad (3.2.19b)$$

This leads to the following equations:

$$\dot{p}(t) = \left(\dot{x}(t) + \frac{i\omega_0}{2} x(t) \right) e^{\frac{i\omega_0 t}{2}}, \quad (3.2.20a)$$

$$\dot{q}(t) = \left(\dot{y}(t) - \frac{i\omega_0}{2} y(t) \right) e^{-\frac{i\omega_0 t}{2}}. \quad (3.2.20b)$$

Therefore we have that

$$\begin{aligned} \dot{p}(t) &= \left(-\frac{i}{2} \left(\omega_0 x(t) + \omega_1 e^{-i(\omega t + \alpha)} y(t) \right) + \frac{i\omega_0}{2} x(t) \right) e^{\frac{i\omega_0 t}{2}} \\ &= -\frac{i\omega_1}{2} e^{-i(\omega t + \alpha)} y(t) e^{\frac{i\omega_0 t}{2}} \\ &= -\frac{i\omega_1}{2} q(t) e^{i((\omega_0 - \omega)t - \alpha)}, \end{aligned} \quad (3.2.21a)$$

and

$$\begin{aligned}
 \dot{q}(t) &= \left(-\frac{i}{2} \left(\omega_1 e^{i(\omega t + \alpha)} x(t) - \omega_0 y(t) \right) - \frac{i\omega_0}{2} y(t) \right) e^{-\frac{i\omega_0 t}{2}} \\
 &= -\frac{i\omega_1}{2} e^{i(\omega t + \alpha)} x(t) e^{-\frac{i\omega_0 t}{2}} \\
 &= -\frac{i\omega_1}{2} p(t) e^{-i((\omega_0 - \omega)t - \alpha)}.
 \end{aligned} \tag{3.2.21b}$$

When we take the second derivative of $p(t)$ we obtain the following second order differential equation:

$$\begin{aligned}
 \ddot{p}(t) &= -\frac{i\omega_1}{2} \dot{q}(t) e^{i((\omega_0 - \omega)t - \alpha)} + i(\omega_0 - \omega) \dot{p}(t) \\
 &= -\frac{i\omega_1}{2} \left(-\frac{i\omega_1}{2} p(t) e^{-i((\omega_0 - \omega)t - \alpha)} e^{i((\omega_0 - \omega)t - \alpha)} \right) + i(\omega_0 - \omega) \dot{p}(t) \\
 &= i(\omega_0 - \omega) \dot{p}(t) - \frac{\omega_1^2}{4} p(t).
 \end{aligned} \tag{3.2.22}$$

This is equivalent to

$$\ddot{p}(t) - i(\omega_0 - \omega) \dot{p}(t) + \frac{\omega_1^2}{4} p(t) = 0. \tag{3.2.23}$$

Let λ_{\pm} be the solutions of the equation

$$\lambda^2 - i(\omega_0 - \omega)\lambda + \frac{\omega_1^2}{4} = 0. \tag{3.2.24}$$

We have

$$\lambda_{\pm} = \frac{i \left((\omega_0 - \omega) \pm \sqrt{(\omega_0 - \omega)^2 + \omega_1^2} \right)}{2}. \tag{3.2.25}$$

We have that

$$p(t) = C_1 e^{\lambda_+ t} + C_2 e^{\lambda_- t}, \tag{3.2.26}$$

and

$$\begin{aligned}
 x(t) &= p(t) e^{-\frac{i\omega_0 t}{2}} \\
 &= \left(C_1 e^{\lambda_+ t} + C_2 e^{\lambda_- t} \right) e^{-\frac{i\omega_0 t}{2}} \\
 &= e^{-\frac{i\omega t}{2}} \left(C_1 e^{\frac{i\sqrt{(\omega_0 - \omega)^2 + \omega_1^2} t}{2}} + C_2 e^{-\frac{i\sqrt{(\omega_0 - \omega)^2 + \omega_1^2} t}{2}} \right),
 \end{aligned} \tag{3.2.27}$$

with initial value

$$x(0) = \frac{1}{\sqrt{2}}e^{-\frac{i\phi}{2}}. \quad (3.2.28)$$

A similar computation for $y(t)$ leads to:

$$y(t) = e^{\frac{i\omega t}{2}} \left(C_3 e^{\frac{i\sqrt{(\omega_0-\omega)^2+\omega_1^2}t}{2}} + C_4 e^{-\frac{i\sqrt{(\omega_0-\omega)^2+\omega_1^2}t}{2}} \right), \quad (3.2.29a)$$

$$y(0) = \frac{1}{\sqrt{2}}e^{\frac{i\phi}{2}}. \quad (3.2.29b)$$

As $x(t), y(t)$ verify the differential equations:

$$\dot{x}(t) = -\frac{i}{2} \left(\omega_0 x(t) + \omega_1 e^{-i(\omega t + \alpha)} y(t) \right), \quad (3.2.30a)$$

$$\dot{y}(t) = -\frac{i}{2} \left(\omega_1 e^{i(\omega t + \alpha)} x(t) - \omega_0 y(t) \right), \quad (3.2.30b)$$

we obtain the following equations for the constants C_i :

$$C_1 + C_2 = \frac{1}{\sqrt{2}}e^{-\frac{i\phi}{2}}, \quad C_3 + C_4 = \frac{1}{\sqrt{2}}e^{\frac{i\phi}{2}}, \quad (3.2.31ab)$$

$$(\Delta + R)C_1 + \omega_1 e^{-i\alpha} C_3 = 0, \quad (\Delta - R)C_2 - \omega_1 e^{-i\alpha} C_4 = 0, \quad (3.2.31cd)$$

where

$$\Delta = \sqrt{(\omega_0 - \omega)^2 + \omega_1^2}, \quad R = \omega_0 - \omega. \quad (3.2.31ef)$$

The computation of the constants C_i is now straightforward. We have

$$C_1 = \frac{1}{2\sqrt{2}} \left(\left(1 - \frac{R}{\Delta}\right) e^{-\frac{i\phi}{2}} - \frac{\omega_1 e^{-i\alpha}}{\Delta} e^{\frac{i\phi}{2}} \right), \quad (3.2.32a)$$

$$C_2 = \frac{1}{2\sqrt{2}} \left(\left(1 + \frac{R}{\Delta}\right) e^{-\frac{i\phi}{2}} + \frac{\omega_1 e^{-i\alpha}}{\Delta} e^{\frac{i\phi}{2}} \right), \quad (3.2.32b)$$

$$C_3 = \frac{1}{2\sqrt{2}} \left(-\frac{\omega_1 e^{i\alpha}}{\Delta} e^{-\frac{i\phi}{2}} + \left(1 + \frac{R}{\Delta}\right) e^{\frac{i\phi}{2}} \right), \quad (3.2.32c)$$

$$C_4 = \frac{1}{2\sqrt{2}} \left(\frac{\omega_1 e^{i\alpha}}{\Delta} e^{-\frac{i\phi}{2}} + \left(1 - \frac{R}{\Delta}\right) e^{\frac{i\phi}{2}} \right). \quad (3.2.32d)$$

These expressions are simplified if the angular velocity ω of the RF magnetic field verifies the resonance condition:

$$\omega = \omega_0. \quad (3.2.33)$$

In that case we have

$$\Delta = \omega_1, \quad R = 0, \quad (3.2.34ab)$$

and the coefficients C_i simplify to:

$$C_1 = \frac{1}{2\sqrt{2}} \left(e^{-\frac{i\phi}{2}} - e^{\frac{i(\phi-2\alpha)}{2}} \right), \quad (3.2.35a)$$

$$C_2 = \frac{1}{2\sqrt{2}} \left(e^{-\frac{i\phi}{2}} + e^{\frac{i(\phi-2\alpha)}{2}} \right), \quad (3.2.35b)$$

$$C_3 = \frac{1}{2\sqrt{2}} \left(-e^{-\frac{i(\phi-2\alpha)}{2}} + e^{\frac{i\phi}{2}} \right), \quad (3.2.35c)$$

$$C_4 = \frac{1}{2\sqrt{2}} \left(e^{-\frac{i(\phi-2\alpha)}{2}} + e^{\frac{i\phi}{2}} \right). \quad (3.2.35d)$$

This leads to the following equations:

$$x(t) = \frac{e^{-\frac{i\omega_0 t}{2}}}{2\sqrt{2}} \left(\left(e^{-\frac{i\phi}{2}} - e^{\frac{i(\phi-2\alpha)}{2}} \right) e^{\frac{i\omega_1 t}{2}} + \left(e^{-\frac{i\phi}{2}} + e^{\frac{i(\phi-2\alpha)}{2}} \right) e^{-\frac{i\omega_1 t}{2}} \right), \quad (3.2.36a)$$

$$y(t) = \frac{e^{\frac{i\omega_0 t}{2}}}{2\sqrt{2}} \left(\left(-e^{-\frac{i(\phi-2\alpha)}{2}} + e^{\frac{i\phi}{2}} \right) e^{\frac{i\omega_1 t}{2}} + \left(e^{-\frac{i(\phi-2\alpha)}{2}} + e^{\frac{i\phi}{2}} \right) e^{-\frac{i\omega_1 t}{2}} \right). \quad (3.2.36b)$$

We can write the evolution in matrix notation:

$$|\psi_{\text{RF}}(t)\rangle = A(\omega, \omega_0, \omega_1, \alpha) \cdot |\psi_{\text{RF}}(0)\rangle, \quad (3.2.37)$$

where

$$A(\omega, \omega_0, \omega_1, \alpha) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (3.2.38)$$

is a rotation in the complex plane. So in order to compute the coefficients of this matrix we need to solve the equation

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x(0) \\ y(0) \end{pmatrix}. \quad (3.2.39)$$

This leads to the following equation:

$$\begin{pmatrix} e^{-\frac{i\omega t}{2}} \left(C_1 e^{\frac{i\Delta t}{2}} + C_2 e^{-\frac{i\Delta t}{2}} \right) \\ e^{\frac{i\omega t}{2}} \left(C_3 e^{\frac{i\Delta t}{2}} + C_4 e^{-\frac{i\Delta t}{2}} \right) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} e^{-\frac{i\phi}{2}} \\ e^{\frac{i\phi}{2}} \end{pmatrix}. \quad (3.2.40)$$

The solution of this matrix equation is:

$$A(\omega, \omega_0, \omega_1, \alpha) = \begin{pmatrix} e^{-\frac{i\omega t}{2}} \left(\cos \frac{\Delta t}{2} - \frac{iR}{\Delta} \sin \frac{\Delta t}{2} \right) & -\frac{i\omega_1 e^{-\frac{i(\omega t + 2\alpha)}{2}}}{\Delta} \sin \frac{\Delta t}{2} \\ -\frac{i\omega_1 e^{\frac{i(\omega t + 2\alpha)}{2}}}{\Delta} \sin \frac{\Delta t}{2} & e^{\frac{i\omega t}{2}} \left(\cos \frac{\Delta t}{2} + \frac{iR}{\Delta} \sin \frac{\Delta t}{2} \right) \end{pmatrix}. \quad (3.2.41)$$

It is possible to separate the static field evolution from this equation. We then obtain:

$$A(\omega, \omega_0, \omega_1, \alpha) = E(\omega) \cdot R(\omega, \omega_0, \omega_1, \alpha), \quad (3.2.42)$$

which leads to

$$E(\omega) = \begin{pmatrix} e^{-\frac{i\omega t}{2}} & 0 \\ 0 & e^{\frac{i\omega t}{2}} \end{pmatrix}, \quad (3.2.43a)$$

$$R(\omega, \omega_0, \omega_1, \alpha) = \begin{pmatrix} \cos \frac{\Delta t}{2} - \frac{iR}{\Delta} \sin \frac{\Delta t}{2} & -\frac{i\omega_1 e^{-i\alpha}}{\Delta} \sin \frac{\Delta t}{2} \\ -\frac{i\omega_1 e^{i\alpha}}{\Delta} \sin \frac{\Delta t}{2} & \cos \frac{\Delta t}{2} + \frac{iR}{\Delta} \sin \frac{\Delta t}{2} \end{pmatrix}. \quad (3.2.43b)$$

At the resonance frequency, this matrix reduces to:

$$A(\omega_0, \omega_0, \omega_1, \alpha) = \begin{pmatrix} e^{-\frac{i\omega_0 t}{2}} & 0 \\ 0 & e^{\frac{i\omega_0 t}{2}} \end{pmatrix} \cdot \begin{pmatrix} \cos \frac{\omega_1 t}{2} & -ie^{-i\alpha} \sin \frac{\omega_1 t}{2} \\ -ie^{i\alpha} \sin \frac{\omega_1 t}{2} & \cos \frac{\omega_1 t}{2} \end{pmatrix}. \quad (3.2.44)$$

The effect of an RF pulse is usually described [CTDL77] as a rotation of angle $\theta_1 = \omega_1 t$ in the spin space around the vector $\vec{u} = (u_x, u_y, u_z)$:

$$\mathbf{R}_{\vec{u}, \theta_1}^{\frac{1}{2}} = \begin{pmatrix} \cos \frac{\theta_1}{2} - iu_z \sin \frac{\theta_1}{2} & (-iu_x - u_y) \sin \frac{\theta_1}{2} \\ (-iu_x + u_y) \sin \frac{\theta_1}{2} & \cos \frac{\theta_1}{2} + iu_z \sin \frac{\theta_1}{2} \end{pmatrix}. \quad (3.2.45)$$

In NMR \vec{u} lies in the transverse plane: $\vec{u} = (\cos \alpha, \sin \alpha, 0)$ and therefore rotations induced in NMR are restricted to:

$$\mathbf{R}_{\vec{u}, \theta_1}^{\frac{1}{2}} = \begin{pmatrix} \cos \frac{\theta_1}{2} & -ie^{-i\alpha} \sin \frac{\theta_1}{2} \\ -ie^{i\alpha} \sin \frac{\theta_1}{2} & \cos \frac{\theta_1}{2} \end{pmatrix}. \quad (3.2.46)$$

We notice that this is exactly the transformation matrix that we have computed, except for the fact that the time evolution of the static field \vec{B}_0 is missing in this equation.

Measurement

Measurement in NMR is obtained by the current induced in the same solenoid as that used for perturbing the spin system. The current is induced by the rotation of the nuclear spin magnetic momenta in the solenoid. At equilibrium there is no induced current due to the absence of nuclear magnetization. After an RF pulse (ω_1, α) the magnetization of a spin $\frac{1}{2}$ at temperature $T = 0$ is equal to:

$$\begin{aligned} M(\omega, \omega_0, \omega_1, \alpha, \phi, t) &= \vec{\mu} = \gamma \vec{I} \\ &= \frac{1}{2} \gamma \hbar (|x(t)|^2 - |y(t)|^2). \end{aligned} \quad (3.2.47)$$

To compute the magnetization, we have to compute

$$(x(t) \quad y(t)) \cdot \begin{pmatrix} \bar{x}(t) \\ -\bar{y}(t) \end{pmatrix}. \quad (3.2.48)$$

We observe the following identity for the rotation matrix $A(\omega, \omega_0, \omega_1, \alpha)$:

$$\begin{aligned} \bar{A}(\omega, \omega_0, \omega_1, \alpha) &= \begin{pmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{pmatrix} \\ &= \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \\ &= \begin{pmatrix} e^{\frac{i\omega t}{2}} & 0 \\ 0 & e^{-\frac{i\omega t}{2}} \end{pmatrix} \cdot \begin{pmatrix} \cos \frac{\Delta t}{2} + \frac{iR}{\Delta} \sin \frac{\Delta t}{2} & \frac{i\omega_1 e^{i\alpha}}{\Delta} \sin \frac{\Delta t}{2} \\ \frac{i\omega_1 e^{-i\alpha}}{\Delta} \sin \frac{\Delta t}{2} & \cos \frac{\Delta t}{2} - \frac{iR}{\Delta} \sin \frac{\Delta t}{2} \end{pmatrix}. \end{aligned} \quad (3.2.49)$$

Therefore the magnetization can be written as

$$\begin{aligned} M(\omega, \omega_0, \omega_1, \alpha, \phi, t) &= \frac{1}{2} \gamma \hbar (|x(t)|^2 - |y(t)|^2) \\ &= \frac{1}{2} \gamma \hbar (x(t) \quad y(t)) \cdot \begin{pmatrix} \bar{x}(t) \\ -\bar{y}(t) \end{pmatrix} \\ &= \frac{1}{2} \gamma \hbar (x(0) \quad y(0)) \cdot \begin{pmatrix} a & c \\ b & d \end{pmatrix} \cdot \begin{pmatrix} d & -c \\ b & -a \end{pmatrix} \cdot \begin{pmatrix} \bar{x}(0) \\ \bar{y}(0) \end{pmatrix} \\ &= \frac{\gamma \hbar}{4} \begin{pmatrix} e^{-\frac{i\phi}{2}} & e^{\frac{i\phi}{2}} \end{pmatrix} \cdot \begin{pmatrix} ad + bc & -2ac \\ 2bd & -(ad + bc) \end{pmatrix} \cdot \begin{pmatrix} e^{\frac{i\phi}{2}} \\ e^{-\frac{i\phi}{2}} \end{pmatrix}. \end{aligned} \quad (3.2.50)$$

After a few easy manipulations we get

$$M(\omega, \omega_0, \omega_1, \alpha, \phi, t) = \gamma \hbar \Re(\bar{a} b e^{-i\phi}). \quad (3.2.51)$$

If we replace the matrix coefficients a, b, c, d by their value and simplify this equation we obtain the following formula for the magnetization:

$$M(\omega, \omega_0, \omega_1, \alpha, \phi, t) = \frac{\gamma \hbar \omega_1 \sin \frac{\Delta t}{2}}{\Delta} \left(\frac{R}{\Delta} \sin \frac{\Delta t}{2} \cos(\phi - \alpha) + \cos \frac{\Delta t}{2} \sin(\phi - \alpha) \right). \quad (3.2.52)$$

At the resonance frequency the magnetization reduces to

$$M(\omega_1, \alpha, \phi, t) = \frac{\gamma \hbar}{2} \sin \omega_1 t \sin(\phi - \alpha). \quad (3.2.53)$$

Single qubit gates

The important single qubit gates are the NOT gate, the Hadamard gate and an arbitrary rotation gate. An apparently trivial gate, the identity gate, is also an essential ingredient for quantum computing. As the wave function evolves even in a static field, we cannot simply assume that not applying an RF is the same as applying the identity operator. To achieve the identity operator we need to solve the following equation:

$$\begin{pmatrix} e^{-\frac{i\omega t}{2}} & 0 \\ 0 & e^{\frac{i\omega t}{2}} \end{pmatrix} \cdot \begin{pmatrix} \cos \frac{\Delta t}{2} - \frac{iR}{\Delta} \sin \frac{\Delta t}{2} & -\frac{i\omega_1 e^{-i\alpha}}{\Delta} \sin \frac{\Delta t}{2} \\ -\frac{i\omega_1 e^{i\alpha}}{\Delta} \sin \frac{\Delta t}{2} & \cos \frac{\Delta t}{2} + \frac{iR}{\Delta} \sin \frac{\Delta t}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (3.2.54)$$

This leads to the following conditions:

$$\Delta t \equiv 0 \pmod{2\pi}, \quad (3.2.55a)$$

$$\omega t \equiv \Delta t \pmod{2\pi}. \quad (3.2.55b)$$

The NOT gate is defined as:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (3.2.56)$$

In order to perform the NOT-gate we need to have the following equality:

$$\begin{pmatrix} e^{-\frac{i\omega t}{2}} & 0 \\ 0 & e^{\frac{i\omega t}{2}} \end{pmatrix} \cdot \begin{pmatrix} \cos \frac{\Delta t}{2} - \frac{iR}{\Delta} \sin \frac{\Delta t}{2} & -\frac{i\omega_1 e^{-i\alpha}}{\Delta} \sin \frac{\Delta t}{2} \\ -\frac{i\omega_1 e^{i\alpha}}{\Delta} \sin \frac{\Delta t}{2} & \cos \frac{\Delta t}{2} + \frac{iR}{\Delta} \sin \frac{\Delta t}{2} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (3.2.57)$$

This can be reduced to

$$\begin{pmatrix} \cos \frac{\Delta t}{2} - \frac{iR}{\Delta} \sin \frac{\Delta t}{2} & -\frac{i\omega_1 e^{-i\alpha}}{\Delta} \sin \frac{\Delta t}{2} \\ -\frac{i\omega_1 e^{i\alpha}}{\Delta} \sin \frac{\Delta t}{2} & \cos \frac{\Delta t}{2} + \frac{iR}{\Delta} \sin \frac{\Delta t}{2} \end{pmatrix} = \begin{pmatrix} 0 & e^{\frac{i\omega t}{2}} \\ e^{-\frac{i\omega t}{2}} & 0 \end{pmatrix}. \quad (3.2.58)$$

Therefore we have

$$-\frac{i\omega_1 e^{-i\alpha}}{\Delta} \sin \frac{\Delta t}{2} = e^{\frac{i\omega t}{2}}, \quad (3.2.59a)$$

$$-\frac{i\omega_1 e^{i\alpha}}{\Delta} \sin \frac{\Delta t}{2} = e^{-\frac{i\omega t}{2}}. \quad (3.2.59b)$$

Taking the product of the lefthand and righthand side of these equations we get

$$-\frac{\omega_1^2 \sin^2 \frac{\Delta t}{2}}{\Delta^2} = 1. \quad (3.2.60)$$

As all parameters are reals, this has no solution. The best we can do is the following gate:

$$i \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (3.2.61)$$

by setting the parameters as follows:

$$\omega = \omega_0, \quad (3.2.62a)$$

$$\omega_1 t \equiv \pi \pmod{4\pi}, \quad (3.2.62b)$$

$$\omega t \equiv -2\alpha \pmod{2\pi}. \quad (3.2.62c)$$

The Hadamard operator has been defined as:

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (3.2.63)$$

To obtain the Hadamard gate we have to set the parameters, such that:

$$\begin{pmatrix} e^{-\frac{i\omega t}{2}} & 0 \\ 0 & e^{\frac{i\omega t}{2}} \end{pmatrix} \cdot \begin{pmatrix} \cos \frac{\Delta t}{2} - \frac{iR}{\Delta} \sin \frac{\Delta t}{2} & -\frac{i\omega_1 e^{-i\alpha}}{\Delta} \sin \frac{\Delta t}{2} \\ -\frac{i\omega_1 e^{i\alpha}}{\Delta} \sin \frac{\Delta t}{2} & \cos \frac{\Delta t}{2} + \frac{iR}{\Delta} \sin \frac{\Delta t}{2} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (3.2.64)$$

This is impossible to attain as the pair of equations:

$$e^{-\frac{i\omega t}{2}} \left(\cos \frac{\Delta t}{2} - \frac{iR}{\Delta} \sin \frac{\Delta t}{2} \right) = \frac{1}{\sqrt{2}}, \quad (3.2.65a)$$

$$e^{\frac{i\omega t}{2}} \left(\cos \frac{\Delta t}{2} + \frac{iR}{\Delta} \sin \frac{\Delta t}{2} \right) = -\frac{1}{\sqrt{2}}. \quad (3.2.65b)$$

have complex conjugates on the lefthand side but not complex conjugates on the righthand side. We can obtain the following matrix:

$$-\frac{i}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad (3.2.66)$$

by setting the parameters as follows:

$$\omega t = -2\alpha + 4k\pi, \quad (3.2.67a)$$

$$\sin \frac{\Delta t}{2} = \frac{\Delta}{\omega_1 \sqrt{2}}, \quad (3.2.67b)$$

$$\cos \frac{\omega t}{2} = \frac{R}{\omega_1}. \quad (3.2.67c)$$

At the resonance frequency, the last two conditions reduce to:

$$\sin \frac{\omega_1 t}{2} = \frac{1}{\sqrt{2}}, \quad (3.2.68a)$$

$$\cos \frac{\omega t}{2} = 0. \quad (3.2.68b)$$

The arbitrary rotation gate that we want to build is the rotation:

$$P = \begin{pmatrix} e^{\frac{i\theta}{2}} & 0 \\ 0 & e^{-\frac{i\theta}{2}} \end{pmatrix}, \quad (3.2.69)$$

where $\cos \theta = \frac{3}{5}$. This can easily be achieved by the following parameter settings:

$$\Delta t \equiv 0 \pmod{4\pi}, \quad (3.2.70a)$$

$$\omega t \equiv -\theta \pmod{4\pi}. \quad (3.2.70b)$$

We also need the gates that initialize the qubit in either the state $|0\rangle = |+\rangle$ or in the state $|1\rangle = |-\rangle$. To do so we do not look at the rotation and evolution matrix, but at the magnetization formula. For simplicity we will assume $\omega = \omega_0$. A qubit in the state $|+\rangle$ should give a magnetization of $\frac{\gamma\hbar}{2}$, while a qubit in the state $|-\rangle$ should give a magnetization of $-\frac{\gamma\hbar}{2}$. This leads to the following equations:

$$1 = \sin \omega_1 t \sin(\phi - \alpha), \quad (3.2.71a)$$

$$-1 = \sin \omega_1 t \sin(\phi - \alpha). \quad (3.2.71b)$$

for respectively qubit $|0\rangle$ or qubit $|1\rangle$. This leads to the following conditions:

$$\alpha \equiv \phi + \frac{\pi}{2} \pmod{2\pi}, \quad (3.2.72a)$$

$$\omega_1 t \equiv \frac{\pi}{2} \pmod{2\pi}, \quad (3.2.72b)$$

for initializing in the state $|0\rangle$ and

$$\alpha \equiv \phi + \frac{\pi}{2} \pmod{2\pi}, \quad (3.2.73a)$$

$$\omega_1 t \equiv -\frac{\pi}{2} \pmod{2\pi}, \quad (3.2.73b)$$

for initializing in the state $|1\rangle$.

3.2.2 Two spins $\frac{1}{2}$

For two spins we consider two cases: one for a homogeneous magnetic field \vec{B}_0 and one where the two spins are in two different magnetic fields \vec{B}_A and \vec{B}_B .

Two spins in a homogeneous magnetic field

In this case the Hamiltonian of the system is:

$$\begin{aligned}\mathcal{H}_{2,\text{Hom}} &= -\sum_{i=1}^2 \vec{\mu}_i \cdot \vec{B}_0 \\ &= -\frac{1}{2}\gamma\hbar(\vec{\sigma}_1 + \vec{\sigma}_2) \cdot \vec{B}_0 \\ &= -\frac{1}{2}\gamma\hbar(\sigma_1^x \oplus_K \sigma_2^x)B_x + (\sigma_1^y \oplus_K \sigma_2^y)B_y + (\sigma_1^z \oplus_K \sigma_2^z)B_z,\end{aligned}\tag{3.2.74}$$

where

$$\vec{B}_0 = (B_x, B_y, B_z).\tag{3.2.75}$$

We can write equation (3.2.74) in matrix form:

$$\mathcal{H}_{2,\text{Hom}} = -\frac{1}{2}\gamma\hbar \begin{pmatrix} 2B_z & B_x - iB_y & B_x - iB_y & 0 \\ B_x + iB_y & 0 & 0 & B_x - iB_y \\ B_x + iB_y & 0 & 0 & B_x - iB_y \\ 0 & B_x + iB_y & B_x + iB_y & -2B_z \end{pmatrix}.\tag{3.2.76}$$

We note that this Hamiltonian can be written as

$$\mathcal{H}_{2,\text{Hom}} = \mathcal{H}_1 \oplus_K \mathcal{H}_1,\tag{3.2.77}$$

where \oplus_K is the Kronecker sum (A.7) of two matrices.

If we place the z -axis along the homogeneous magnetic field \vec{B}_0 , this matrix becomes

$$\mathcal{H}_{2,\text{Hom}} = -\frac{1}{2}\gamma\hbar \begin{pmatrix} 2B_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2B_0 \end{pmatrix}.\tag{3.2.78}$$

We directly obtain the eigenvalues of the Hamiltonian (3.2.74):

$$E_{++} = -\gamma\hbar B_0,\tag{3.2.79a}$$

$$E_{+-} = E_{-+} = 0,\tag{3.2.79b}$$

$$E_{--} = \gamma\hbar B_0.\tag{3.2.79c}$$

The corresponding eigenvectors are:

$$|++\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |+-\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad (3.2.80ab)$$

$$|-+\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad |--\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \quad (3.2.80cd)$$

The time-dependent wave function is given by:

$$|\psi_2(t)\rangle = \frac{1}{2} \begin{pmatrix} e^{-i(\omega_0 t - \phi_{++})} \\ e^{i\phi_{+-}} \\ e^{i\phi_{-+}} \\ e^{i(\omega_0 t + \phi_{--})} \end{pmatrix}. \quad (3.2.81)$$

We note that

$$|\psi_2(t)\rangle = |\psi_1(t)\rangle \otimes |\psi_1(t)\rangle, \quad (3.2.82)$$

by properly adjusting the phase factors ϕ_i .

Two spins in different magnetic fields

If the magnetic fields experienced by spins A and B are respectively \vec{B}_A and \vec{B}_B , then the Hamiltonian of the system is:

$$\mathcal{H}_2 = - \left((\vec{\mu}_A \cdot \vec{B}_A) \oplus_K (\vec{\mu}_B \cdot \vec{B}_B) \right). \quad (3.2.83)$$

If we write this Hamiltonian in matrix form, we obtain the following:

$$\mathcal{H}_2 = -\frac{\gamma\hbar}{2} \begin{pmatrix} B_{AZ} + B_{BZ} & B_{BX} - iB_{BY} & B_{AX} - iB_{AY} & 0 \\ B_{BX} + iB_{BY} & B_{AZ} - B_{BZ} & 0 & B_{AX} - iB_{AY} \\ B_{AX} + iB_{AY} & 0 & B_{BZ} - B_{AZ} & B_{BX} - iB_{BY} \\ 0 & B_{AX} + iB_{AY} & B_{BX} + iB_{BY} & -B_{AZ} - B_{BZ} \end{pmatrix}, \quad (3.2.84)$$

where we have

$$\vec{B}_A = (B_{AX}, B_{AY}, B_{AZ}), \quad (3.2.85a)$$

$$\vec{B}_B = (B_{BX}, B_{BY}, B_{BZ}). \quad (3.2.85b)$$

If we suppose that both magnetic fields are in the Oz -direction, this matrix reduces to:

$$\mathcal{H}_2 = -\frac{\gamma\hbar}{2} \begin{pmatrix} B_A + B_B & 0 & 0 & 0 \\ 0 & B_A - B_B & 0 & 0 \\ 0 & 0 & B_B - B_A & 0 \\ 0 & 0 & 0 & -(B_A + B_B) \end{pmatrix}. \quad (3.2.86)$$

We directly obtain the eigenvalues of the Hamiltonian (3.2.86):

$$E_{++} = -\frac{1}{2}\gamma\hbar(B_A + B_B), \quad (3.2.87a)$$

$$E_{+-} = -\frac{1}{2}\gamma\hbar(B_A - B_B), \quad (3.2.87b)$$

$$E_{-+} = \frac{1}{2}\gamma\hbar(B_A - B_B), \quad (3.2.87c)$$

$$E_{--} = \frac{1}{2}\gamma\hbar(B_A + B_B). \quad (3.2.87d)$$

The corresponding eigenvectors are:

$$|++\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |+-\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad (3.2.88ab)$$

$$|-+\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad |--\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \quad (3.2.88cd)$$

This leads to the following time-dependent wave function:

$$|\psi_2(t)\rangle = \frac{1}{2} \begin{pmatrix} e^{-\frac{i}{2}((\omega_A + \omega_B)t - \phi_{++})} \\ e^{-\frac{i}{2}((\omega_A - \omega_B)t - \phi_{+-})} \\ e^{\frac{i}{2}((\omega_A - \omega_B)t - \phi_{-+})} \\ e^{\frac{i}{2}((\omega_A + \omega_B)t - \phi_{--})} \end{pmatrix}, \quad (3.2.89)$$

where:

$$\omega_A = -\gamma B_A, \quad (3.2.90a)$$

$$\omega_B = -\gamma B_B. \quad (3.2.90b)$$

As in the homogenous case, the wave function for two independent spins can be written as a tensor product of the wave function of each spin:

$$|\psi_2(t)\rangle = |\psi_1(t)\rangle_{\omega_A} \otimes |\psi_1(t)\rangle_{\omega_B}, \quad (3.2.91)$$

by properly adapting the phase factors ϕ_i . This is simply done by taking equation (3.2.91) as the proper definition for the case of two spins $\frac{1}{2}$ and using equation (3.2.14) for each single wave function to obtain the following

wave function for two spins:

$$\begin{aligned}
|\psi_2(t)\rangle &= |\psi_1(t)\rangle_{\omega_A} \otimes |\psi_1(t)\rangle_{\omega_B} \\
&= \frac{1}{2} \begin{pmatrix} e^{-\frac{i(\omega_A t + \phi_1)}{2}} \\ e^{\frac{i(\omega_A t + \phi_1)}{2}} \end{pmatrix} \otimes \begin{pmatrix} e^{-\frac{i(\omega_B t + \phi_2)}{2}} \\ e^{\frac{i(\omega_B t + \phi_2)}{2}} \end{pmatrix} \\
&= \frac{1}{2} \begin{pmatrix} e^{-\frac{i}{2}((\omega_A + \omega_B)t + (\phi_1 + \phi_2))} \\ e^{-\frac{i}{2}((\omega_A - \omega_B)t + (\phi_1 - \phi_2))} \\ e^{\frac{i}{2}((\omega_A - \omega_B)t + (\phi_1 - \phi_2))} \\ e^{\frac{i}{2}((\omega_A + \omega_B)t + (\phi_1 + \phi_2))} \end{pmatrix} \tag{3.2.92}
\end{aligned}$$

The measurement of the two spin $\frac{1}{2}$ system is given by the rotation of the magnetization of each spin in the solenoid. This can be computed in two different ways. If we write the wave function as

$$|\psi_2(t)\rangle = a_{++}|++\rangle + a_{+-}|+-\rangle + a_{-+}|-+\rangle + a_{--}|--\rangle, \tag{3.2.93}$$

then the magnetization is equal to the sum over each spin of the probability of measuring the state $|+\rangle$ minus the probability of measuring the state $|-\rangle$. For the first spin the probability of measuring $|+\rangle$ is equal to:

$$|a_{++}|^2 + |a_{+-}|^2, \tag{3.2.94}$$

and the probability of measuring $|-\rangle$ is equal to:

$$|a_{-+}|^2 + |a_{--}|^2. \tag{3.2.95}$$

A similar computation for the second spin gives the following formula for the magnetization:

$$M = \frac{1}{2}\gamma\hbar(2|a_{++}|^2 - 2|a_{--}|^2). \tag{3.2.96}$$

The second method of computing the magnetization is to consider each spin separately with its corresponding wave function for one spin. Taking the sum of these magnetizations gives the total magnetization:

$$M = M_1(\phi_1) + M_2(\phi_2). \tag{3.2.97}$$

Two spins in different magnetic fields with an RF field

Using the tensor product notation we can directly compute the wave function for two spins in an inhomogenous magnetic field with an RF magnetic field \vec{B}_1 rotating around \vec{B}_0 with an angular velocity ω :

$$\begin{aligned}
|\psi_2(t)\rangle &= |\psi_1(t)\rangle_{\omega_A} \otimes |\psi_1(t)\rangle_{\omega_B} \\
&= \left(E(\omega) \cdot R(\omega_A) \cdot |\psi_1(0)\rangle_{\omega_A} \right) \otimes \left(E(\omega) \cdot R(\omega_B) \cdot |\psi_1(0)\rangle_{\omega_B} \right) \\
&= (E(\omega) \otimes E(\omega)) \cdot (R(\omega_A) \otimes R(\omega_B)) \cdot \left(|\psi_1(0)\rangle_{\omega_A} \otimes |\psi_1(0)\rangle_{\omega_B} \right). \tag{3.2.98}
\end{aligned}$$

This description is the most general possible for two independent spins. We would like to find a set of parameters, such that this matrix becomes a ControlNOT operator. As the above equation is a tensor product of two matrices, this would imply that we can write

$$\text{CNOT} = A \otimes B. \quad (3.2.99)$$

This leads to the following equality:

$$\begin{aligned} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} &= \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \otimes \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} \\ &= \begin{pmatrix} a_1 b_1 & a_1 b_2 & a_2 b_1 & a_2 b_2 \\ a_1 b_3 & a_1 b_4 & a_2 b_3 & a_2 b_4 \\ a_3 b_1 & a_3 b_2 & a_4 b_1 & a_4 b_2 \\ a_3 b_3 & a_3 b_4 & a_4 b_3 & a_4 b_4 \end{pmatrix}. \end{aligned} \quad (3.2.100)$$

We obtain, amongst others, the following equations:

$$a_1 b_1 = 1, \quad a_4 b_1 = 0, \quad a_4 b_3 = 1. \quad (3.2.101abc)$$

These equations have no solution in \mathbb{C} and therefore it is impossible to set the parameters, such that the resulting operator on the wave function is the ControlNot operator. As we have the identity

$$(A_1 \otimes A_2) \cdot (B_1 \otimes B_2) = (A_1 B_1) \otimes (A_2 B_2), \quad (3.2.102)$$

we cannot hope to build a ControlNot operator starting with another operator obtained from an RF wave. The conclusion is that in order to build a CNOT we need an interaction between the two spins. We therefore investigate whether the dipole-dipole coupling between the two spins can be used as such an interaction. Before doing so we consider the case of N spins.

3.2.3 N spins

The description we have obtained for the wave function of two spins is easily generalized. An N spin system in an inhomogenous magnetic field has the following wave function:

$$|\psi_N(t)\rangle = \bigotimes_{i=1}^N |\psi_{i,\omega_i}(t)\rangle. \quad (3.2.103)$$

The total magnetization M is given by the sum of all N individual magnetizations:

$$M(t) = \sum_{i=1}^N M_i(t). \quad (3.2.104)$$

In this description we have not yet taken into account the population differences of the two energy levels in case of N spins at temperature T .

3.2.4 Dipole-dipole coupling

The Hamiltonian of the dipole-dipole coupling is:

$$\mathcal{H}_D = \frac{\mu_0 \gamma^2}{4\pi r_{12}^3} \left(\vec{I}_1 \cdot \vec{I}_2 - 3 \frac{(\vec{I}_1 \cdot \vec{r}_{12}) \otimes (\vec{I}_2 \cdot \vec{r}_{12})}{r_{12}^2} \right). \quad (3.2.105)$$

We can write equation (3.2.105) in matrix form. First we compute $\vec{I}_1 \cdot \vec{I}_2$:

$$\begin{aligned} \vec{I}_1 \cdot \vec{I}_2 &= \left(\frac{1}{2} \hbar \vec{\sigma}_1 \right) \cdot \left(\frac{1}{2} \hbar \vec{\sigma}_2 \right) \\ &= \frac{\hbar^2}{4} \begin{pmatrix} \sigma_1^X \\ \sigma_1^Y \\ \sigma_1^Z \end{pmatrix} \cdot \begin{pmatrix} \sigma_2^X \\ \sigma_2^Y \\ \sigma_2^Z \end{pmatrix} \\ &= \frac{\hbar^2}{4} (\sigma_1^X \otimes \sigma_2^X + \sigma_1^Y \otimes \sigma_2^Y + \sigma_1^Z \otimes \sigma_2^Z) \\ &= \frac{\hbar^2}{4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned} \quad (3.2.106)$$

We proceed with $\vec{I}_1 \cdot \vec{r}_{12}$, where the distance vector \vec{r}_{12} is defined as:

$$\vec{r}_{12} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}. \quad (3.2.107)$$

This leads to:

$$\begin{aligned} \vec{I}_1 \cdot \vec{r}_{12} &= \frac{1}{2} \hbar \vec{\sigma}_1 \cdot \vec{r}_{12} \\ &= \frac{1}{2} \hbar (\sigma_1^X X + \sigma_1^Y Y + \sigma_1^Z Z) \\ &= \frac{1}{2} \hbar \begin{pmatrix} Z & X - iY \\ X + iY & -Z \end{pmatrix}. \end{aligned} \quad (3.2.108)$$

Therefore we have:

$$\begin{aligned} (\vec{I}_1 \cdot \vec{r}_{12}) \otimes (\vec{I}_2 \cdot \vec{r}_{12}) &= \frac{\hbar^2}{4} \begin{pmatrix} Z & X - iY \\ X + iY & -Z \end{pmatrix} \otimes \begin{pmatrix} Z & X - iY \\ X + iY & -Z \end{pmatrix} \\ &= \frac{\hbar^2}{4} \begin{pmatrix} Z^2 & Z(X - iY) & Z(X - iY) & X^2 - Y^2 - 2iXY \\ Z(X + iY) & -Z^2 & X^2 + Y^2 & -Z(X - iY) \\ Z(X + iY) & X^2 + Y^2 & -Z^2 & -Z(X - iY) \\ X^2 + 2iXY - Y^2 & -Z(X + iY) & -Z(X + iY) & Z^2 \end{pmatrix}, \end{aligned} \quad (3.2.109)$$

which leads to the following Hamiltonian:

$$\mathcal{H}'_D = \begin{pmatrix} r^2 - 3Z^2 & -3Z(X-iY) & -3Z(X-iY) & -3(X^2 - 2iXY - Y^2) \\ -3Z(X+iY) & -r^2 + 3Z^2 & 2r^2 - 3Z^2 & 3Z(X-iY) \\ -3Z(X+iY) & 2r^2 - 3Z^2 & -r^2 + 3Z^2 & 3Z(X-iY) \\ -3(X^2 + 2iXY - Y^2) & 3Z(X+iY) & 3Z(X+iY) & r^2 - 3Z^2 \end{pmatrix}, \quad (3.2.110)$$

where

$$\mathcal{H}'_D = \frac{r^2 \mathcal{H}_D}{\hbar K_D}, \quad r^2 = r_{12}^2, \quad K_D = \frac{\mu_0 \gamma^2 \hbar}{16\pi r_{12}^3}. \quad (3.2.111abc)$$

This Hamiltonian can also be written in matrix form with spherical coordinates. This results in:

$$\mathcal{H}_D = \hbar K_D \begin{pmatrix} 1 - 3 \cos^2 \theta & -3 \sin \theta \cos \theta e^{-i\varphi} & -3 \sin \theta \cos \theta e^{-i\varphi} & -3 \sin^2 \theta e^{-2i\varphi} \\ -3 \sin \theta \cos \theta e^{i\varphi} & -1 + 3 \cos^2 \theta & 2 - 3 \sin^2 \theta & 3 \sin \theta \cos \theta e^{-i\varphi} \\ -3 \sin \theta \cos \theta e^{i\varphi} & 2 - 3 \sin^2 \theta & -1 + 3 \cos^2 \theta & 3 \sin \theta \cos \theta e^{-i\varphi} \\ -3 \sin^2 \theta e^{2i\varphi} & 3 \sin \theta \cos \theta e^{i\varphi} & 3 \sin \theta \cos \theta e^{i\varphi} & 1 - 3 \cos^2 \theta \end{pmatrix}. \quad (3.2.112)$$

As the direction of \vec{r}_{12} is random, we should consider the mean value of each matrix element. To do so we compute the spatial average of each matrix element:

$$\bar{a}_{ij}(\theta, \varphi) = \frac{1}{2\pi^2} \int_{\theta=0}^{2\pi} \int_{\varphi=0}^{\pi} a_{ij}(\theta, \varphi) \sin \theta d\theta d\varphi. \quad (3.2.113)$$

This dramatically reduces the matrix and \mathcal{H}_D becomes:

$$\mathcal{H}_D = 0. \quad (3.2.114)$$

Therefore we cannot use the dipole-dipole coupling as the interaction between two spins to build a ControlNot operator.

In a homogeneous magnetic field it is well-known that the dipole-dipole coupling is averaged to zero by the random thermal motion in liquids, but if the two spins have two different magnetic fields this is no longer the case. This fact was first described in 1979 by Deville et al. [DBD79], and later by Botwell et al. [BBG90] in pure water. These authors have shown long-range acting dipole-dipole interactions in liquid with magnetic field gradients. Theoretical descriptions of this effect can be found in [LRVW96, JVB95], but for the moment they are difficult to use for applications to quantum computation. It is therefore necessary to find either another interaction between the spins that is not averaged to zero or to formalize their approach to long-range dipole-dipole interactions so that it is described in the same framework that we use.

3.2.5 Two coupled spins

In an inhomogeneous magnetic field, the Hamiltonian for two coupled spins is given by:

$$\mathcal{H}_3 = \mathcal{H}_2 + \mathcal{H}_c, \quad (3.2.115)$$

where \mathcal{H}_c is the Hamiltonian which describes the coupling of the two spins. For the moment we do not have a description for such a Hamiltonian, but once we do, we can use the same techniques as described earlier: in order to find the time evolution of the wave function, we have to solve the following differential equation:

$$\frac{i\hbar\partial|\psi(t)\rangle}{\partial t} = \mathcal{H}_3|\psi(t)\rangle, \quad (3.2.116)$$

where

$$|\psi(t)\rangle = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{pmatrix}. \quad (3.2.117)$$

This system of equations can also be written in matrix form:

$$\dot{X} = M \cdot X, \quad (3.2.118)$$

where \dot{X} is the vector $(\dot{x}_1(t), \dots, \dot{x}_4(t))^T$, the matrix M is equal to $\frac{\mathcal{H}_3}{i\hbar}$ and X is the vector $(x_1(t), \dots, x_4(t))^T$.

The solution of this matrix differential equation is:

$$X = e^{tM}. \quad (3.2.119)$$

To compute the exponential of the matrix M , we need to find the eigenvectors of M in order to diagonalize this matrix:

$$M = U^{-1}DU, \quad (3.2.120)$$

where U is a unitary matrix and D is a diagonal matrix:

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{pmatrix}, \quad (3.2.121)$$

with λ_i the eigenvalues of M . This will lead to a description of the wave function $|\psi(t)\rangle$, from which we can hopefully deduce the parameter settings to build a CNOT-gate.

3.3 Five Steps to an NMR Quantum Computer

We want to build an NMR quantum computer using the framework we described in the previous section. The road to a small quantum computer is essentially the same for any physical realization. For a larger quantum computer we need to take into account many other important steps such as the decoherence of the system and the fidelity of the qubit operations, but for now we concentrate on the bare necessities for a quantum computer however shortlived this computer may be. The following steps need to be followed:

1. **One qubit:** the proposed realization needs a clear description of what the physical equivalent of a logical qubit is. We need to understand how such a qubit is built and how to properly describe it. We also need to know how the qubit is initialized and how it is measured.
2. **Manipulating one qubit:** we need to be able to perform arbitrary unitary operations on a single qubit. It is not necessary to be able to perform any arbitrary unitary operation, but we at least need to have a generating set that can approximate all unitary operations. An identity operator, a NOT operator, a Hadamard operator and a phase operator are sufficient.
3. **More qubits:** we need to understand how we can have more than one logical qubit in our physical system. We have to be able to distinguish between different qubits and how we can initialize qubits simultaneously. We also have to understand how to measure individual qubits.
4. **Manipulating qubits individually:** we have to be able to perform the same generating set of unitary operations on individual qubits. It is important to have the identity operator, because while we perform an operation on a single qubit, the other qubits evolve in time. This effect needs to be undone when we do not want such an evolution.
5. **Manipulating qubits together:** the power of quantum computation lies in the entanglement of states and the natural parallelism of computation. We therefore need to have a gate which entangles two qubits. The CNOT-operator creates entanglement of qubits and is easily described. We therefore have to be able to perform a CNOT-gate on two arbitrary qubits. An equivalent entangling gate will do as well, but we concentrate on the CNOT-gate as most quantum algorithms are described with CNOT-gates.

These are the steps that we have to achieve experimentally in order to have a small scale quantum computer. From that point on, other issues such as fidelity, decoherence and error correction have to be taken into account, as well as a reasonable estimate of the real computing power of the proposed

system, but without the five steps above, it is no use to think about fidelity of gates or error correcting.

3.4 Experimental results

We describe our experimental setting as well as the results that we have obtained so far with our approach.

3.4.1 Material and methods

A sample of 10 ml of degassed water was placed at room temperature in a wide-bore magnet with a magnetic field of 4.7 T (Magnex). The NMR spectrometer (SMIS) allows a phase precision of the RF pulses of 0.25° . The RF pulses had a gaussian shaped intensity with a duration $d = 600 \mu\text{s}$, a frequency $\frac{\omega_0}{2\pi} = 200.137 \text{ MHz}$, and half-width of 3 kHz. The inter pulse delay between the ends of the first and second pulse was $\tau = 1 \text{ ms}$. The NMR signal was detected in quadrature mode with a sample frequency of 5 kHz and 8K points. The intensity of the signal is obtained as the modulus of the two parts given by the quadrature detection mode.

The homogeneity of the magnetic field was measured by the line width obtained by Fourier Transform of the free induction decay (FID) acquired after a $\frac{\pi}{2}$ pulse. The longitudinal relaxation time T_1 , measured by an inversion-recovery sequence, was 3.2 s and the transverse relaxation time T_2 , measured by a Carr-Purcell-Meiboom-Gill sequence [CP54, MG58] was 1.8 s, slightly depending on the homogeneity of the magnetic field.

The NMR spectrum of water, as for all liquid samples with no J -coupling, displays a very narrow line due to the motion averaging of the dipole-dipole coupling. Such a nuclear spin system is highly isolated from its surrounding and it is well-known that the relaxation time T_1 which characterizes the energy exchange with the lattice and the inverse of the line width which measures the decoherence time are very long in high homogeneous magnetic field.

3.4.2 Results

We first exhibit a macroscopic quantum effect in bulk liquid NMR. After that, we show a method to initialize a qubit.

Exhibiting a macroscopic quantum effect

We can show that there is a quantum interference term in bulk liquid NMR by using a $\frac{\pi}{2} - \tau - \frac{\pi}{2}$ pulse sequence at the resonance frequency. In our framework we have not given the magnetization after two pulse sequences

but it can be shown that this magnetization is proportional to

$$M(\tau) = C \sin \omega_0 \tau \sin \beta, \quad (3.4.1)$$

where C is a proportionality constant that depends on the population differences and β is the angle between the two pulse sequences.

One can easily see that in absence of free evolution, i.e. $\tau = 0$, there is no signal. This is due to the fact that in the $\tau = 0$ case, the $\frac{\pi}{2} - \tau - \frac{\pi}{2}$ sequence corresponds to a single π pulse on the sample which indeed gives no signal. In fact, according to equation (3.4.1), provided that the angle $\beta \neq 0$, we have that $M(\tau) \neq 0$, if and only if the nuclear spin state interference term $\sin \omega_0 \tau$ is different from zero. In a $\pi/2 - \tau - \pi/2$ sequence, the existence of any NMR signal is then the evidence of the occurrence of nuclear spin interferences.

Experimentally, it was impossible for us to tune τ at a time scale small enough to vary $\omega_0 \tau$ over 2π . However, it is possible to ensure over typical experimental times (a few minutes) an accurate stability of $\omega_0 \tau$, i.e. the rms magnitude of the fluctuating part of this angle $\omega_0 \tau$ remains much smaller than 2π . Under this last condition, one can then plot the NMR signal given by the $\frac{\pi}{2} - \tau - \frac{\pi}{2}$ sequence as a function of β , the relative phase of the two $\frac{\pi}{2}$ pulse fields and compare the results to that given by equation (3.4.1). If the experimental data match equation (3.4.1), then the nuclear spin interference term is revealed and also controlled.

The NMR signal (FID) after a single $\frac{\pi}{2}$ pulse is dependent on the homogeneity of the magnetic field \vec{B}_0 .

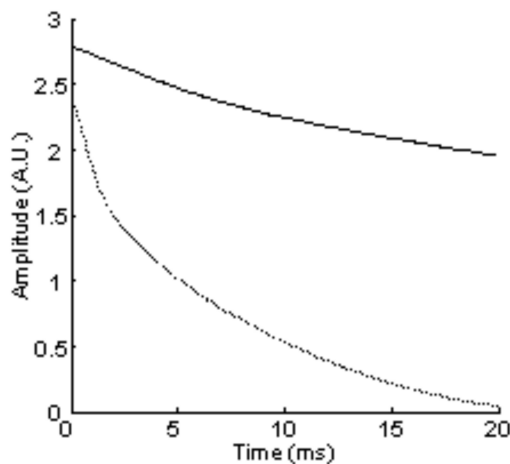


Figure 3.2: NMR signal of 10 ml of water after one $\frac{\pi}{2}$ pulse. The continuous line is obtained in a highly homogeneous magnetic field ($\frac{\Delta B_0}{B_0} = 2.0 \cdot 10^{-8}$) and the dashed line in a less homogeneous field ($\frac{\Delta B_0}{B_0} = 2.7 \cdot 10^{-7}$).

On Fig 3.2 one can see the FID recorded after a single $\frac{\pi}{2}$ pulse in a highly homogeneous field ($\frac{\Delta B_0}{B_0} = 2.0 \cdot 10^{-8}$, continuous line) compared to a less homogeneous one ($\frac{\Delta B_0}{B_0} = 2.7 \cdot 10^{-7}$, dashed line).

With a $\frac{\pi}{2} - \tau - \frac{\pi}{2}$ sequence, it is well-known that NMR gives rise to an echo at a time $t = \tau$ after the second $\frac{\pi}{2}$ pulse. This effect was described in 1950 by E. Hahn as spin echo [Hah50]. Here however, we have measured the NMR signal in a very homogeneous magnetic field and with small inter pulse delays where no spin echo is detected as seen on Fig. 3.3 (continuous line). Even in the less homogeneous magnetic field there is a modulation of

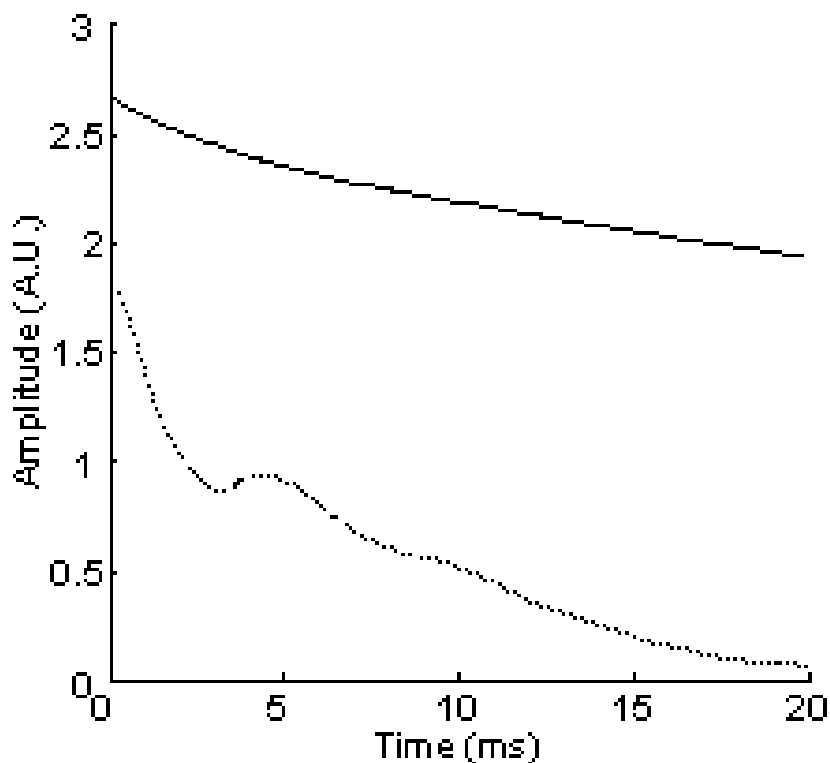


Figure 3.3: NMR signal of 10 ml of water after two $\frac{\pi}{2}$ pulses with a relative phase of $\beta = 90^\circ$. The continuous line is obtained in a highly homogeneous magnetic field ($\frac{\Delta B_0}{B_0} = 2.0 \cdot 10^{-8}$) and the dashed line in a less homogeneous field ($\frac{\Delta B_0}{B_0} = 2.7 \cdot 10^{-7}$).

the FID but no echo at 1 ms which is the delay between the two $\frac{\pi}{2}$ pulses. The absence of an echo in this case is equivalent to the absence of any echo for a homogeneous line in an Electron Spin Resonance (ESR) experiment. The FID corresponds to the magnetization in the transverse plane and therefore the signal is proportional to $\sqrt{M_x^2 + M_y^2}$. Fig. 3.4 shows the amplitude of

the NMR signal at the beginning of the FID versus the relative phase β . As

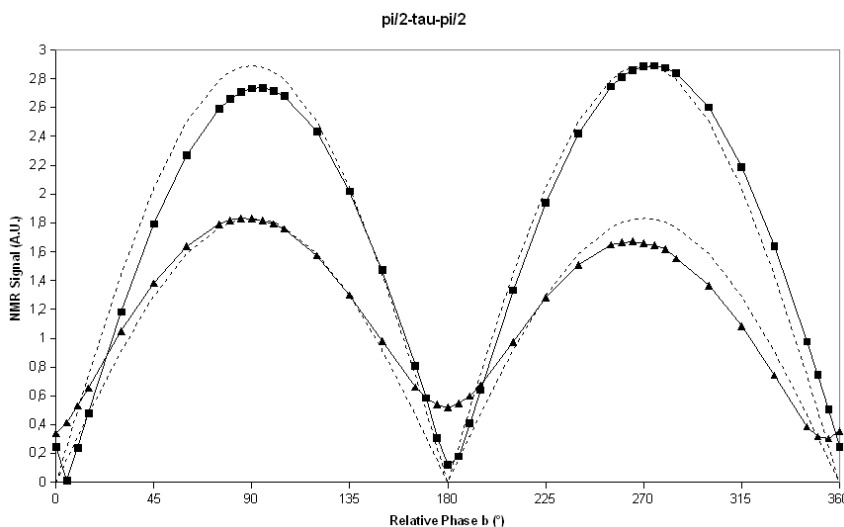


Figure 3.4: Amplitude of the NMR signal of 10 ml of water after two $\frac{\pi}{2}$ pulses versus the relative phase β of the two pulses. The continuous line (■ NMR₁) is obtained in a highly homogeneous magnetic field, the large dashed line (▲ NMR₂) in a less homogeneous field. The finely dashed lines correspond to $f(\beta) = G|\sin \beta|$ normalized to the maximum NMR signal in each case.

can be seen in Fig. 3.4, in the case of a highly homogeneous magnetic field ($\frac{\Delta B_0}{B_0} = 2.0 \cdot 10^{-8}$), the function $f(\beta) = G|\sin \beta|$, given by equation (3.4.1) for a well defined value of $\omega_0\tau$, fits the experimental data quite well. The maximum relative deviation

$$\Delta_s(\beta) = \frac{(\text{NMR}_1(\beta) - f(\beta))}{\max(\text{NMR}_1(\beta))} \quad (3.4.2)$$

between the experimental curve $\text{NMR}_1(\beta)$ and $f(\beta)$ is found to be

$$\Delta_s(15^\circ) = 9.7\%. \quad (3.4.3)$$

In the case of a less homogeneous field ($\frac{\Delta B_0}{B_0} = 2.7 \cdot 10^{-7}$), the fit is less good and the maximum relative deviation is found to be

$$\Delta_s(18^\circ) = 28.3\%. \quad (3.4.4)$$

3.4.3 Numerical solution of equation (3.2.18)

The numerical solution of equation (3.2.18) is obtained by using the ode45 subroutine of Matlab using an explicit Runge-Kutta formula for ordinary differential equations with initial values.

Reference parameters

We set the parameters of equation (3.2.18) as those used by our NMR spectrometer:

Table 3.1: Parameters of the NMR spectrometer

Parameter	Value	Unit
Resonance Frequency	200 MHz	$\omega_0 = -2 \cdot 10^8 \cdot 2\pi$ rad/s
Radio Frequency amplitude	1 mT	$\omega_1 = 2.5 \cdot 10^{-4} \cdot \omega_0$ rad/s
Radio Frequency	200 MHz	$\omega = \omega_0$ rad/s
Interval of integration	20	$t_\theta = 20 \cdot 10^{-6}$ s

The solution of equation (3.2.18) with the parameters of table 3.1 allows us to calculate the magnetization of one spin according to equation (3.2.47). The result is given in figure 3.5 where we retrieve the main effect of an NMR experiment, which is the induced magnetization after an appropriate RF pulse at the Larmor frequency. The maximum magnetization corresponds to the so called $\frac{\pi}{2}$ -pulse and for a double duration the π -pulse with no magnetization. The solutions $x(t)$ and $y(t)$ of equation (3.2.18) have a real and imaginary part oscillating around the Larmor frequency as shown in figure 3.6 for the real part of $x(t)$ during the RF pulse.

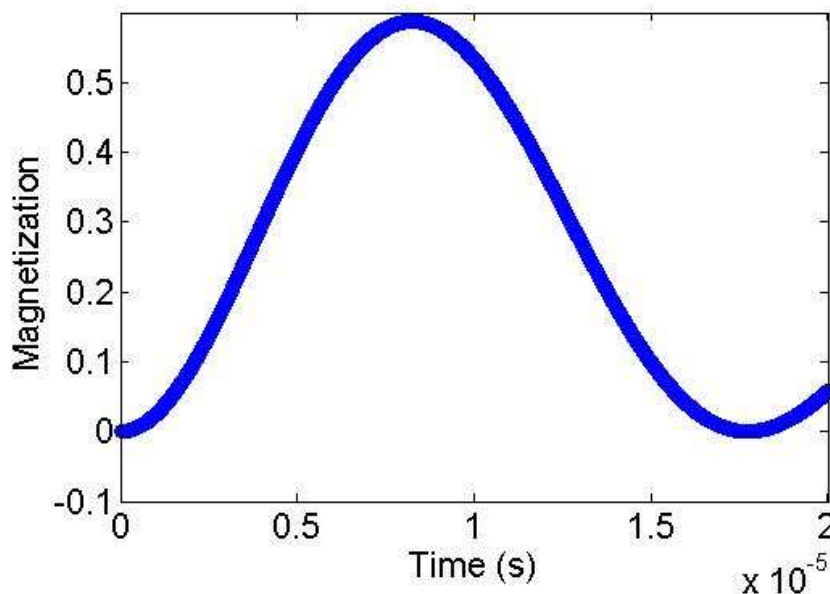


Figure 3.5: Magnetization of one spin $\frac{1}{2}$ versus the duration of the RF pulse.

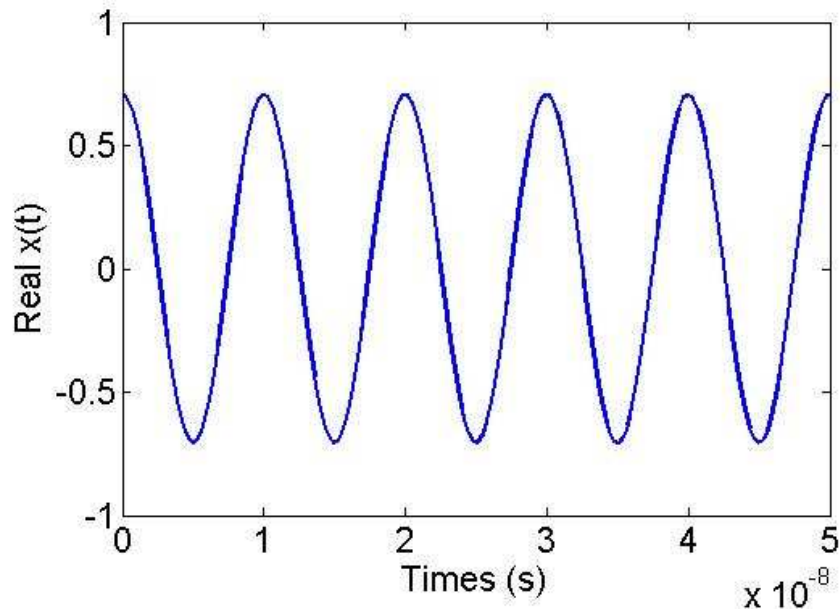


Figure 3.6: The real part of $x(t)$.

Effect of the frequency of the RF pulse

Figure 3.7 shows the effect of the frequency of the RF on the spin magnetization.

We find a well-known fact in NMR which is the inversion of the magnetization when going through the resonance frequency.

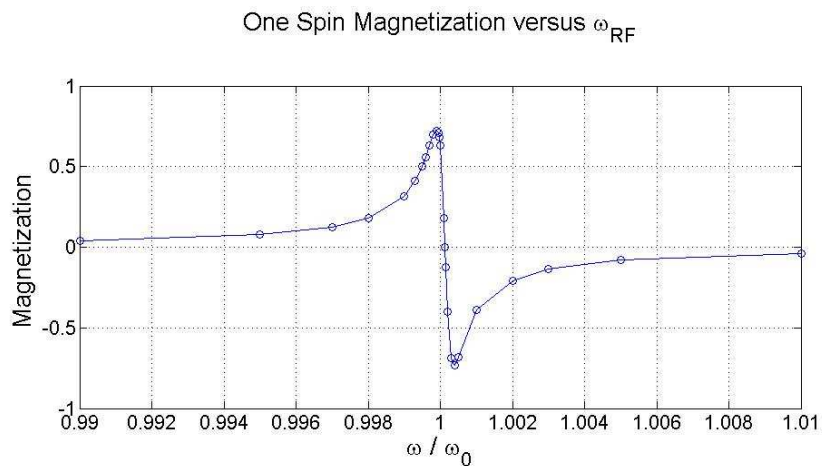


Figure 3.7: Magnetization versus the frequency ω of the RF pulse, with the RF amplitude $\omega_1 = 2 \cdot 10^{-4} \omega_0$.

Initializing the qubit

To initialize the qubit in either the state $|0\rangle = |+\rangle$ or in the state $|1\rangle = |-\rangle$ we need either maximum positive or maximum negative magnetization. For simplicity we will assume $\omega = \omega_0$. A qubit in the state $|+\rangle$ should give a magnetization of $\frac{\gamma\hbar}{2}$, while a qubit in the state $|-\rangle$ should give a magnetization of $-\frac{\gamma\hbar}{2}$. These conditions are always verified with some period T . The idea is to first observe the system and find out what this period is in order to know when these maximal magnetizations occur. At these moments, the evolving logical qubit is in the state $|0\rangle$. When we want to perform single qubit operations on one logical qubit, then we will wait to perform such an operation until the magnetization is exactly maximal. In stead of measuring, as we did for initializing the qubit, we proceed to perform a single qubit operation and measure the result only afterwards.

3.5 Conclusion and Perspective

We have exhibited a framework in which NMR quantum computing on pure states can be realized. From an experimental point of view we have shown how to initialize a qubit into the basic states $|0\rangle$ or $|1\rangle$. This result achieves the first of the five necessary steps. At the moment we are adjusting the experimental parameter settings in order to obtain a generating set of elementary one qubit gates, which will result in obtaining the second step. For the third step we will use magnetic field gradients to distinguish different qubits. This step as well as the fourth step is work in progress. For the fifth step we need to establish an interaction between different qubits. For the moment we have not yet achieved a theoretical description of this interaction Hamiltonian. Without such a Hamiltonian we cannot hope to find the correct parameter settings to achieve a CNOT operator. Therefore the crucial point in our approach is to achieve such a theoretical description for the interaction Hamiltonian. We are trying to achieve such an interaction by using long-range dipole-dipole interaction. This interaction is not averaged away to zero by random thermal motion in liquids, because at a long distance the geometrical constraints of the sample prevent a completely random thermal motion. Whether this long-range dipole-dipole interaction is large enough to serve as interaction between qubits is still work in progress.

Part II

Solving Simultaneous Pell Equations using Quantum Computation

Chapter 4

Science never solves a problem
without creating ten more.

GEORGE BERNARD SHAW

Pell equations

4.1 Introduction

Let d be a positive integer and consider the equation

$$x^2 - dy^2 = 1, \tag{4.1.1}$$

where x, y are positive integers. This equation is called the Pell equation, after the English mathematician John Pell, to whom Leonhard Euler mistakenly attributed a method of solving this type of equations. A first trivial observation shows that $(x_0, y_0) = (1, 0)$ is always a solution of equation (4.1.1) and that if d is a square there cannot be another solution in positive integers, as for $d = q^2$ we have:

$$\begin{aligned} x^2 - q^2y^2 &= x^2 - (qy)^2 \\ &= (x + qy)(x - qy). \end{aligned} \tag{4.1.2}$$

So we have

$$(x + qy)(x - qy) = 1, \tag{4.1.3}$$

which implies

$$x + qy = 1, \tag{4.1.4a}$$

$$x - qy = 1, \tag{4.1.4b}$$

with x, q, y all positive integers. This in turn implies that $(x, y) = (1, 0)$. So we can assume that d is not a square. If we find a non-trivial solution (x_1, y_1) to equation (4.1.1), then the fraction $\frac{x_1}{y_1}$ is a good approximation for the square root of d :

$$\begin{aligned} \frac{x}{y} &= \sqrt{\frac{1+dy^2}{y^2}} \\ &= \sqrt{d + \frac{1}{y^2}}. \end{aligned} \tag{4.1.5}$$

For example, if $d = 2$, and $x = 17$, $y = 12$, we have

$$17^2 - 2 \cdot 12^2 = 1, \quad (4.1.6a)$$

$$\frac{17}{12} \approx 1,4167. \quad (4.1.6b)$$

There are several questions that can be asked about Pell equations. Are there always non-trivial solutions for any integer d that is not a square? Are there infinitely many solutions? How can we compute these solutions? Can we compute these solutions quickly for any d ? It is possible to pose more technical questions about this type of equations, but we will restrict ourselves to these simple ones. It is possible to show that for any positive integer d that is not a square, there are an infinite number of solutions for equation (4.1.1). Moreover, these solutions have a simple structure, which allows us to find all solutions starting from a fundamental solution. There are several methods to solve equation (4.1.1), but not every method has the same efficiency for all integers d . We will start by looking at some classical solving techniques. These include the Indian method and the continued fraction method, which are essentially the same technique in a different form. We proceed with a more modern approach that consists of computing the regulator of an associated number field. This approach solves Pell equations more efficiently, but does not solve it in polynomial time. A quantum approach that follows the modern approach, but which uses a quantum algorithm to compute this regulator does solve the Pell equation in polynomial time.

4.2 Classical Techniques

4.2.1 Chakravala Method

The Indian approach to solve the Pell equation is called the Chakravala or cyclic method and is based upon the Brahmagupta identity and Bhaskara's lemma :

Lemma 4.1 (Brahmagupta's identity). *Let a, b, c, d, n be real numbers, then we have the following equality:*

$$(a^2 + nb^2)(c^2 + nd^2) = (ac - nbd)^2 + n(ad + bc)^2 \quad (4.2.1a)$$

$$= (ac + nbd)^2 + n(ad - bc)^2. \quad (4.2.1b)$$

Proof. The lefthandside of equation (4.2.1a) is equal to:

$$(a^2 + nb^2)(c^2 + nd^2) = a^2c^2 + n(a^2d^2 + b^2c^2) + n^2b^2d^2. \quad (4.2.2)$$

The righthandside of equation (4.2.1a) is equal to:

$$\begin{aligned} (ac - nbd)^2 + n(ad + bc)^2 &= a^2c^2 - 2nacbd + n^2b^2d^2 + n(a^2d^2 + 2adbc + b^2c^2) \\ &= a^2c^2 + n(a^2d^2 + b^2c^2) + n^2b^2d^2. \end{aligned} \quad (4.2.3)$$

The righthandside of equation (4.2.1b) is equal to:

$$\begin{aligned}(ac+nbd)^2+n(ad-bc)^2 &= a^2c^2+2nacbd+n^2b^2d^2+n(a^2d^2-2adbc+b^2c^2) \\ &= a^2c^2+n(a^2d^2+b^2c^2)+n^2b^2d^2.\end{aligned}\quad (4.2.4)$$

Therefore we have equality in both cases. \square

Lemma 4.2 (Bhaskara). *Let a, b, c, d, e be real numbers, with d not equal to 0. If*

$$a^2 = bc^2 + d, \quad (4.2.5)$$

then we have the following identity:

$$b\left(\frac{a+ec}{d}\right)^2 + \frac{e^2-b}{d} = \left(\frac{ea+bc}{d}\right)^2. \quad (4.2.6)$$

Proof. The lefthandside of equation (4.2.6) is equal to:

$$\begin{aligned}b\left(\frac{a+ec}{d}\right)^2 + \frac{e^2-b}{d} &= \frac{b(a^2+2ace+c^2e^2)}{d^2} + \frac{de^2-bd}{d^2} \\ &= \frac{b(bc^2+d+2ace+c^2e^2)+de^2-bd}{d^2} \\ &= \frac{b^2c^2+2abce+(bc^2+d)e^2}{d^2} \\ &= \frac{b^2c^2+2abce+a^2e^2}{d^2} \\ &= \left(\frac{bc+ae}{d}\right)^2.\end{aligned}\quad (4.2.7)$$

\square

In order to solve the Pell equation

$$x^2 - dy^2 = 1, \quad (4.2.8)$$

we use Brahmagupta's identity on the triples (x_1, y_1, k_1) and (x_2, y_2, k_2) that verify the equation:

$$x_1^2 - dy_1^2 = k_1, \quad (4.2.9a)$$

$$x_2^2 - dy_2^2 = k_2. \quad (4.2.9b)$$

In this manner we obtain a new triple

$$(x_3, y_3, k_3) = (x_1x_2 + dy_1y_2, x_1y_2 + x_2y_1, k_1k_2), \quad (4.2.10)$$

by multiplication:

$$(x_1^2 - dy_1^2)(x_2^2 - dy_2^2) = (x_1x_2 + dy_1y_2)^2 - d(x_1y_2 + x_2y_1)^2. \quad (4.2.11)$$

In order to solve equation (4.2.8) we start with an arbitrary triplet (x_1, y_1, k) , such that

$$x_1^2 - dy_1^2 = k, \quad (4.2.12)$$

and $\text{GCD}(x_1, y_1) = 1$. We multiply this triplet with the trivial triplet

$$(a, 1, a^2 - d), \quad (4.2.13)$$

and we obtain a new triplet $(ax_1 + dy_1, x_1 + ay_1, k(a^2 - d))$.

We use Bhaskara's lemma to obtain the following identity:

$$\left(\frac{ax_1 + dy_1}{|k|}\right)^2 - d\left(\frac{x_1 + ay_1}{|k|}\right)^2 = \frac{a^2 - d}{k}. \quad (4.2.14)$$

We choose a , such that

$$\frac{x_1 + ay_1}{k} \quad (4.2.15)$$

is an integer and

$$\frac{a^2 - d}{k} \quad (4.2.16)$$

has the smallest possible absolute value.

For this value a we replace the triplet (x_1, y_1, k) by

$$(x_2, y_2, k_2) = \left(\frac{ax_1 + dy_1}{|k|}, \frac{x_1 + ay_1}{|k|}, \frac{a^2 - d}{k}\right), \quad (4.2.17)$$

and we repeat the procedure. Lagrange proved that this process always terminates with a solution. We give an example with $d = 113$. Let $x_1 = 11$, $y_1 = 1$ and $k = 8$, we have the obvious identity:

$$11^2 - 113 \times 1 = 8. \quad (4.2.18)$$

So we want to find an integer a , such that

$$\frac{11 + a}{8} \quad (4.2.19)$$

is an integer and

$$\left|\frac{a^2 - 113}{8}\right| \quad (4.2.20)$$

is minimal. In this case, $a = 13$, so we obtain the new triplet:

$$32^2 - 113 \times 3^2 = 7. \quad (4.2.21)$$

Repeating this process we find the following triplets:

$$85^2 - 113 \times 8^2 = -7, \quad (4.2.22a)$$

$$287^2 - 113 \times 27^2 = -8, \quad (4.2.22b)$$

$$776^2 - 113 \times 73^2 = -1. \quad (4.2.22c)$$

At this point we could continue the process, but instead we take the square of the last solution, using Brahmagupta's identity to obtain the solution:

$$1204353^2 - 113 \times 113296^2 = 1. \quad (4.2.23)$$

4.2.2 Continued fraction method

The other classical approach to solve Pell equations is by using the continued fraction development of \sqrt{d} . Let $a_0 = \lfloor \sqrt{d} \rfloor$, then:

$$\begin{aligned} \sqrt{d} &= a_0 + \frac{1}{a_1 + \frac{1}{\ddots + \frac{1}{a_n + \frac{1}{2a_0}}}} \\ &= [a_0, a_1, \dots, a_n, 2a_0]. \end{aligned} \quad (4.2.24)$$

When we consider the periodic part of the continued fraction development

$$\frac{x}{y} = [a_0, a_1, \dots, a_n], \quad (4.2.25)$$

then we have that

$$\left| \frac{x}{y} - \sqrt{d} \right| < \frac{1}{2a_0 y^2}. \quad (4.2.26)$$

From this we can easily derive

$$\left| \frac{x^2}{y^2} - \sqrt{d} \right| < \frac{2}{y^2}, \quad (4.2.27)$$

which leads to

$$|x^2 - dy^2| < 2. \quad (4.2.28)$$

As the lefthand side of the above inequality is an integer and since d is not a square, we immediately deduce that

$$x^2 - dy^2 = \pm 1. \quad (4.2.29)$$

Therefore there are two cases to consider. If $x^2 - dy^2 = 1$, we have a solution to the Pell equation. If $x^2 - dy^2 = -1$, then

$$(x^2 - dy^2)^2 = (x^2 + dy^2)^2 - d(2xy)^2 = 1. \quad (4.2.30)$$

In that case $x' = x^2 + dy^2$ and $y' = 2xy$ are a solution of the Pell equation. We try to compute such a solution again for the case $d = 113$. We find that

$$\sqrt{113} = [10, 1, 1, 1, 2, 2, 1, 1, 1, 20]. \quad (4.2.31)$$

This leads to the fraction

$$\frac{x}{y} = \frac{73}{776}, \quad (4.2.32)$$

which gives

$$776^2 - 113 \times 73^2 = -1. \quad (4.2.33)$$

Taking squares at both sides leads to the same solution as the Indian method.

4.3 Modern Techniques

As Lenstra remarks in his article on the Pell equation [Len02], the efficiency of the continued fraction method is conjectured to be exponentially slow for most values of d and that any method that spells out the smallest solution (x_0, y_0) of the Pell equation in full is exponentially slow for infinitely many values of d . One method to improve the algorithm would be to consider only the square-free part of each integer d , but this only helps a little bit. In order to build a faster algorithm we need to use the structure of the ring $\mathbb{Z}[\sqrt{d}]$.

Let d be a square-free integer and consider the equation

$$x^2 - dy^2 = 1. \quad (4.3.1)$$

If $\sqrt{d} \notin \mathbb{Q}$, then for rational numbers a, b, x, y we have that

$$a + b\sqrt{d} = x + y\sqrt{d}, \quad (4.3.2)$$

if and only if $a = x$ and $b = y$. It is therefore possible to uniquely encode the solution of (4.3.1) as

$$x + y\sqrt{d} \in \mathbb{R}. \quad (4.3.3)$$

Conversely we say that $\sigma \in \mathbb{R}$ is a solution of (4.3.1), if

$$\sigma = s + t\sqrt{d}, \quad (4.3.4)$$

for integers s, t , such that

$$s^2 - dt^2 = 1. \quad (4.3.5)$$

To solve the Pell equation it suffices to calculate the regulator

$$R = \log(x_1 + y_1\sqrt{d}), \quad (4.3.6)$$

for which $x_1^2 - dy_1^2 = 1$ is the smallest solution. For this it suffices to calculate the regulator of $\mathbb{Z}[\sqrt{d}]$. Let

$$\begin{aligned} K &= \mathbb{Q}[\sqrt{d}] \\ &= \{u + v\sqrt{d} \mid u, v \in \mathbb{Q}\} \end{aligned} \quad (4.3.7)$$

be a real quadratic number field. The order O of discriminant d is the subring

$$\begin{aligned} O &= \mathbb{Z}\left[\frac{d+\sqrt{d}}{2}\right] \\ &= \left\{a + b\frac{d+\sqrt{d}}{2} \mid a, b \in \mathbb{Z}\right\} \subseteq K. \end{aligned} \quad (4.3.8)$$

The units of O are of the form $\pm e^k$, with $k \in \mathbb{Z}$. The regulator of O is defined as

$$R = \log \epsilon, \quad (4.3.9)$$

with $\epsilon > 1$. The regulator satisfies the following inequalities:

$$\log(2\sqrt{d}) < R < \sqrt{d}(\log(4d) + 2). \quad (4.3.10)$$

The modern method to solve the Pell equation uses the above ingredients in combination with the notion of power products. If (x_0, y_0) is the fundamental solution of the Pell equation $x^2 - dy^2 = 1$, then a power product notation of the solution is a product of the following form:

$$x_0 + y_0\sqrt{d} = \prod_{i=1}^k (a_i + b_i\sqrt{d})^{n_i}. \quad (4.3.11)$$

We have the following theorem on the relevance of the regulator approach to solve Pell equations:

Theorem 4.1. *There are positive constants C_1, C_2 , such that*

1. *For each positive non square integer d , there exists a power product representation of the fundamental solution of its associated Pell equation with length at most $C_1(\log d)^2$.*
2. *The problem of computing such a power product representation is polynomial time equivalent to the problem of computing an integer \tilde{R} , such that $|\tilde{R} - R| < 1$, where R is the regulator of the number field $\mathbb{Z}[\sqrt{d}]$.*
3. *There exists an algorithm that for given d computes a power product representation of the fundamental solution of its associated Pell equation in time at most $\sqrt{\tilde{R}}(1 + \log d)^{C_2}$.*

The theorem above gives an algorithm for solving Pell equations that still has exponential run time. A more refined approach which uses smooth numbers over the number field $\mathbb{Z}[\sqrt{d}]$ can compute an integer approximation to a multiple of the regulator. This leads to a probabilistic algorithm that runs in time $O(e^{C\sqrt{\log d \log \log d}})$ under the assumption of the generalized Riemann hypothesis. This approach resembles the quadratic sieve for factoring integers and has the same run time halfway between exponential and polynomial time.

4.4 Quantum Computational Techniques

The quantum computational approach to solve the Pell equation is to construct a periodic function h which has the regulator R as period and to apply

an extended version of the QFT on this function to retrieve the period. The product of two subsets $I, J \subseteq K$ is the additive subgroup of K generated by the set

$$\{xy \mid x \in I, y \in J\}. \quad (4.4.1)$$

An invertible O -ideal is a subset $I \subseteq K$, with $OI = I$, for which there exists a subset $J \subseteq K$ with $IJ = O$. The set of invertible ideals of O form an Abelian group under multiplication and will be denoted \mathcal{I} . The set of principal ideals will be denoted

$$\mathcal{P} = \{O\alpha \mid \alpha \in K\}. \quad (4.4.2)$$

This is a subgroup of \mathcal{I} . An invertible ideal has the form

$$\left\{ q \left(\mathbb{Z} + \frac{-b+\sqrt{d}}{2a}\mathbb{Z} \right) \mid a, b \in \mathbb{Z}, q \in \mathbb{Q}, c = \frac{b^2-d}{4a} \in \mathbb{Z}, \text{GCD}(a, b, c) = 1 \right\}. \quad (4.4.3)$$

An ideal is reduced if

$$\left| \sqrt{d} - 2|a| \right| < b < \sqrt{d}. \quad (4.4.4)$$

The set of all reduced ideals is denoted \mathcal{R} . This is a finite set with a group-like structure under multiplication. We define the distance function as:

$$\begin{aligned} \delta: \mathcal{P} &\longrightarrow \mathbb{R}/R\mathbb{Z} \\ (a + b\sqrt{d})O &\longmapsto \frac{1}{2} \log \left| \frac{a+b\sqrt{d}}{a-b\sqrt{d}} \right| \pmod{R}. \end{aligned} \quad (4.4.5)$$

The unit ideal has distance zero. The composition of two ideals $I, J \in \mathcal{I}$ is the product $I \cdot J \in \mathcal{I}$. We have

$$\delta(IJ) = \delta(I) + \delta(J). \quad (4.4.6)$$

Reduction is a map

$$\rho: \mathcal{I} \longrightarrow \mathcal{I}, \quad (4.4.7)$$

such that after a polynomial number of steps k an ideal $\rho^k(I)$ will be in \mathcal{R} . For the exact formula for the reduction we refer to the appendices. We can give the following bounds

$$\delta(I) + \frac{1}{\sqrt{d}} \leq \delta(\rho(I)) \leq \delta(I) + \log \sqrt{d}, \quad (4.4.8a)$$

$$\delta(\rho^2(I)) > \delta(I) + \log 2. \quad (4.4.8b)$$

Multiplication is a map from the reduced ideals to itself, taking as input two reduced ideals I, J , applying the reduction ρ repeatedly on IJ , until $\rho^k(IJ)$ is a reduced ideal.

Given a rational distance x , it is possible to calculate the ideal with distance

closest to the left of x . We define this ideal I_x . We can now define the coset separating function

$$\begin{aligned} f: \mathbb{R} &\longrightarrow \mathcal{I} \times \mathbb{R} \\ x &\longmapsto (I_x, x - \delta(I_x)), \end{aligned} \quad (4.4.9)$$

which is a periodic function with as period the regulator R . Using this function f we define a function \hat{f} which is suitable for Fourier sampling.

Choose an integer $N \geq 2\sqrt{d}$, then we define the function \hat{f} as

$$\begin{aligned} \hat{f}: \mathbb{Z} &\longrightarrow \mathcal{I} \times \mathbb{Z} \\ i &\longmapsto \left(I_{\frac{i}{N}}, \left\lfloor N \left(\frac{i}{N} - \delta(I_{\frac{i}{N}}) \right) \right\rfloor \right). \end{aligned} \quad (4.4.10)$$

This function \hat{f} is periodic with period NR .

The finite set of all principal fractional reduced ideals is $\mathcal{PT}_{\text{red}}$. This set is called the principal cycle. $\delta(I)$ is the distance between an ideal I of the principal cycle and the ring of integers \mathcal{O} . We define the map h as follows

$$\begin{aligned} h: \mathbb{R} &\longrightarrow \mathcal{PT}_{\text{red}} \times \mathbb{R} \\ x &\longmapsto (I_x, \tilde{x} - \delta(I_x)), \end{aligned} \quad (4.4.11)$$

with $\tilde{x} \equiv x \pmod{R}$ and $I_x \in \mathcal{PT}_{\text{red}}$ the largest ideal in the principal cycle that verifies $\delta(I_x) < \tilde{x}$. We have the following theorem:

Theorem 4.2. *The function h is computable in polynomial time: if x is a multiple of 10^{-n} , then we can compute I_x and an approximation of $\tilde{x} - \delta(I_x)$ with precision 10^{-n} in time $\text{poly}(\log D, \log x, n)$. Moreover, h is a periodic function with period R and is one-to-one on every interval smaller than the period R .*

If we know the value of the integer closest to the regulator R we can turn this into an algorithm to approximate R with arbitrary precision:

Proposition 4.1. *If we know the value of*

$$\lceil R \rceil = \lfloor R + \frac{1}{2} \rfloor, \quad (4.4.12)$$

then there exists an algorithm that computes R with precision 10^{-n} in time $\text{poly}(n, \log D)$.

Suppose we have a function $f: \mathbb{R} \longrightarrow X$, with $f(x+R) = f(x)$. In order to be able to apply the quantum period finding algorithm, we need to discretize f by taking multiples of $\frac{1}{N}$, with N big enough. If X is continuous, it needs to be discretized as well.

Definition 4.1. For $f: \mathbb{R} \rightarrow \mathbb{R}$ we define the map \tilde{f}_N as:

$$\begin{aligned} \tilde{f}_N: \mathbb{Z} &\longrightarrow \frac{1}{N}\mathbb{Z} \\ k &\longmapsto \left\lfloor f\left(\frac{k}{N}\right) \right\rfloor_N, \end{aligned} \quad (4.4.13)$$

where $\lfloor x \rfloor_N$ is $\frac{\lfloor Nx \rfloor}{N}$, and $\lceil x \rceil_N$ is defined likewise.

We would like that \tilde{f} contains approximative information about the period R of f , however if f has a big variation in an interval of $\frac{1}{N}$ around $x = \frac{k}{N}$, then \tilde{f} can take arbitrary values. We need a notion of weak periodicity.

Definition 4.2. A function $f: \mathbb{Z} \rightarrow X$ is weakly periodic with period $S \in \mathbb{R}$, if for all $0 \leq k \leq \lfloor S \rfloor$ and for all $l \in \mathbb{Z}$, either $f(k + \lceil lS \rceil)$ or $f(k + \lfloor lS \rfloor)$ is equal to $f(k)$. We write $f(k) = f(k + \lfloor lS \rfloor)$.

We are discretizing the function h :

Definition 4.3. The discretized function of h is defined as

$$\begin{aligned} \tilde{h}_N: \mathbb{Z} &\longrightarrow \mathcal{PI}_{\text{red}} \times \frac{1}{N}\mathbb{Z} \\ k &\longmapsto \left(I_{\frac{k}{N}}, \left\lfloor \frac{k}{N} - \delta\left(I_{\frac{k}{N}}\right) \right\rfloor_N \right). \end{aligned} \quad (4.4.14)$$

The following proposition gives a further characterisation of \tilde{h}_N :

Proposition 4.2. The function \tilde{h}_N has the following properties:

1. \tilde{h}_N is one-to-one on $[0, \lfloor NR \rfloor]$.
2. $\tilde{h}_N(k)$ is computable in time $O(k^{c_1}, N^{c_2}, D^{c_3})$, so if $N, k = O(D^{c_4})$, then $\tilde{h}_N(k)$ is computable in $O(D^{c_5})$, where c_i are positive constants.
3. Let $d_{\min} = \frac{3}{32D}$ be a lower bound on the distances between reduced ideals and $\sigma = \log d$. If $\frac{1}{N} < \frac{d_{\min}}{\log d}$, then \tilde{h}_N is weakly periodic with period NR . The condition $\tilde{h}_N(k) = \tilde{h}_N(k + \lfloor lS \rfloor)$ is verified for all $0 \leq k \leq \lfloor NR \rfloor$, except possibly for a small fraction of size $\frac{1}{\log d}$.

To build a quantum algorithm that approximates the period of a weakly periodic function in polynomial time, we need the following conditions:

Theorem 4.3. Suppose that $f: \mathbb{Z} \rightarrow X$ is weakly periodic with period S and

1. $f(k)$ is computable in $O((\log k)^{c_1}, (\log S)^{c_2})$,
2. f is one-to-one on $[0, \lfloor S \rfloor]$,

3. for $m \in \mathbb{Z}$, there exists an algorithm in $O((\log S)^{c_3})$ that tests whether m is close to a multiple of S : $|jS - m| < 1$, for an integer j .

Then there exists a quantum algorithm in $O((\log S)^{c_4})$ that produces an integer a , such that $|S - a| < 1$ with probability larger than $O((\log S)^{-c_5})$, where c_i are positive constants.

In order to prove the main theorem that states that there exists a polynomial time quantum algorithm that solves the Pell equation, we need the following two technical lemmata:

Lemma 4.3. *Let S be a real number and let q be the number of qubits in the QFT register and let q be a power of 2. Let $0 \leq k \leq \lfloor S \rfloor$ and $0 \leq l < \frac{q}{S}$. If $q > 3S^2$, then*

$$\left| \frac{c}{d} - \frac{k}{l} \right| < \frac{1}{2l^2}, \quad (4.4.15)$$

where

$$c = \left\lfloor \frac{kq}{S} \right\rfloor, \quad d = \left\lfloor \frac{lq}{S} \right\rfloor. \quad (4.4.16ab)$$

Lemma 4.4. *Let $|A| \leq \frac{1}{2}$, $\xi(l)$ be an arbitrary number, such that $|\xi(l)| < \frac{1}{n}$, where $n = O(\log p)$. Then there exists a constant C , such that for all p sufficiently large we have*

$$X = \left| \sum_{l=0}^{p-1} e^{2\pi i \left(\frac{Al}{p} + \xi(l) \right)} \right|^2 \geq Cp^2. \quad (4.4.17)$$

With these lemmata, we can prove the following theorem:

Theorem 4.4. *Let d be a square-free positive integer. There exists a quantum algorithm that computes the regulator R of $\mathbb{Q}[\sqrt{d}]$ with precision 10^{-n} in time $O((\log d)^{c_1}, n^{c_2})$ with probability $O((\log d)^{-c_3}, n^{-c_4})$, if $10^{-n} < \frac{d_{\min}}{\log d}$, where c_i are positive constants.*

Chapter 5

I just invent, then wait until
man comes around to needing
what I've invented.

R. BUCKMINSTER FULLER

Simultaneous Pell equations

5.1 Introduction

Simultaneous Pell equations are equations of the form:

$$x^2 - ay^2 = 1, \quad (5.1.1a)$$

$$z^2 - by^2 = 1, \quad (5.1.1b)$$

where a, b are positive non-square integers, such that their product is not a square either. These equations are a specialized case of the more general simultaneous Fermat equations:

$$x^2 - ay^2 = c, \quad (5.1.2a)$$

$$z^2 - by^2 = d. \quad (5.1.2b)$$

Several natural questions can be posed about these type of equations. First of all, where do equations of this type occur? Are there any solutions in positive integers (x, y, z) ? Are there a finite number of solutions and if so, how many solutions can there be? Given an explicit case, can we find the solutions? Are the solutions bounded in any natural way?

We will only deal with some of these questions. We will start with an old conjecture on integer sequences. Simultaneous Pell equations occur in a simplified version of this conjecture. We will give an upper bound of the smallest solution of equations (5.1.1), if any exists, following an approach by Anglin [Ang95]. We will also reproduce a result by Cipu and Mignotte [CM] that proves that there are at most two solutions in positive integers for any pair of simultaneous Pell equations. We will combine these results with the polynomial quantum algorithm of Hallgren for a single Pell equation to produce a polynomial quantum algorithm that solves simultaneous Pell equations.

5.2 A conjecture on 5 integers

One of the problems of the Greek mathematician Diophantus was to find sets of unequal fractions, such that the product of any two of its elements is one less than a square. He found the following set of four fractions:

$$\left\{ \frac{1}{16}, \frac{33}{16}, \frac{17}{4}, \frac{105}{16} \right\}, \quad (5.2.1)$$

for which we can indeed verify:

$$\frac{1}{16} \times \frac{33}{16} = \left(\frac{17}{16} \right)^2 - 1, \quad \frac{1}{16} \times \frac{17}{4} = \left(\frac{9}{8} \right)^2 - 1, \quad (5.2.2ab)$$

$$\frac{1}{16} \times \frac{105}{16} = \left(\frac{19}{16} \right)^2 - 1, \quad \frac{33}{16} \times \frac{17}{4} = \left(\frac{25}{8} \right)^2 - 1, \quad (5.2.2cd)$$

$$\frac{33}{16} \times \frac{105}{16} = \left(\frac{61}{16} \right)^2 - 1, \quad \frac{17}{4} \times \frac{105}{16} = \left(\frac{43}{8} \right)^2 - 1. \quad (5.2.2ef)$$

In the seventeenth century, Pierre de Fermat looked for integer solutions to this type of equations. He found the set

$$\{1, 3, 8, 120\}, \quad (5.2.3)$$

for which we have:

$$1 \times 3 = 2^2 - 1, \quad 1 \times 8 = 3^2 - 1, \quad (5.2.4ab)$$

$$1 \times 120 = 11^2 - 1, \quad 3 \times 8 = 5^2 - 1, \quad (5.2.4cd)$$

$$3 \times 120 = 19^2 - 1, \quad 8 \times 120 = 31^2 - 1. \quad (5.2.4ef)$$

He tried to extend this set with a fifth integer but failed. Euler extended his set with a rational number:

$$\left\{ 1, 3, 8, 120, \frac{777480}{28792} \right\}, \quad (5.2.5)$$

but could not find a fifth integer either.

In 1969, Baker and Davenport proved [BD69] that this set cannot be extended to a fifth integer and that the only possible integer extension of the triplet 1, 3, 8 is the integer 120. It is this second part that leads to generalized simultaneous Pell equations. Suppose that we have an integer k , such that $\{1, 3, 8, k\}$ is a set with products one less than a square. In that case k has to verify the following equations:

$$1 \times k = x^2 - 1, \quad (5.2.6a)$$

$$3 \times k = y^2 - 1, \quad (5.2.6b)$$

$$8 \times k = z^2 - 1. \quad (5.2.6c)$$

By substituting k in the last two equations we obtain

$$3(x^2 - 1) = y^2 - 1, \quad (5.2.7a)$$

$$8(x^2 - 1) = z^2 - 1, \quad (5.2.7b)$$

which can be written as

$$3x^2 - y^2 = 2, \quad (5.2.8a)$$

$$8x^2 - z^2 = 7. \quad (5.2.8b)$$

The question whether there are five integers, such that the product of any two of them is one less than a square remains unanswered for the moment:

Conjecture 5.1. *There are no integers a_1, \dots, a_5 , such that for $i \neq j$ we have*

$$a_i a_j = k_{ij}^2 - 1, \quad (5.2.9)$$

where k_{ij} are positive integers.

For rational numbers a little more is known. Euler already found a set of five rational numbers. It is even possible to find six rational numbers, such that the product of any two of them is one less than a square of a rational number:

$$\left\{ \frac{11}{192}, \frac{35}{192}, \frac{155}{27}, \frac{512}{27}, \frac{1235}{48}, \frac{180873}{16} \right\}. \quad (5.2.10)$$

5.3 An upper bound

5.3.1 Diophantine Approximation

We describe the general strategy of diophantine approximation techniques. A linear form in logarithms is a form of the type:

$$\Lambda = \beta_0 + \beta_1 \log \alpha_1 + \dots + \beta_n \log \alpha_n, \quad (5.3.1)$$

where $\alpha_1, \dots, \alpha_n$ are algebraic numbers. Alan Baker obtained a lower bound for the linear form $|\Lambda|$ [Bak67], which Feldman improved [Fel71] with the following theorem:

Theorem 5.1 (Feldman). *The logarithmic form Λ verifies the inequality:*

$$|\Lambda| \geq B^{-C}, \quad (5.3.2)$$

for all algebraic numbers β_0, \dots, β_n with height at most $B > 1$, where C is effectively computable in terms of the α_i and the degree of the β_i .

For the case that the numbers β_i are integers, such that for all i we have that $|\beta_i| \leq B$ and the height $H(\alpha_i)$ of every algebraic number α_i is bounded by A_i , Baker and Wüstholz [BW93] improve this lower bound with the following theorem:

Theorem 5.2 (Baker, Wüstholz). *If the linear form $\Lambda \neq 0$, then we have the following lower bound:*

$$\log |\Lambda| > -(16nd)^{2n+4} \log B \prod_{i=1}^n \log A_i, \quad (5.3.3)$$

where d is the degree of the field $\mathbb{Q}(\alpha_1, \dots, \alpha_n)$.

Suppose we have a system of equations for which we want to obtain an upper bound for the solutions. This can be achieved by the previous theorems and the following strategy.

1. Reduce the equations if necessary to such equations for which Baker's theory can be applied.
2. Reduce these new equations to inequalities of the form

$$0 < \left| \alpha_1^{b_1} \cdots \alpha_n^{b_n} - \alpha_{n+1} \right| < c_1 e^{-c_2 B}, \quad (5.3.4)$$

where $\alpha_1, \dots, \alpha_{n+1}$ are algebraic numbers, b_1, \dots, b_n are unknown rational integers, $B = \max(|b_i|)$ and c_1, c_2 are positive constants that are independent of the integers b_i and can be effectively computed. If the bound B is large, then the inequalities (5.3.4) imply that

$$|\Lambda| \leq c_3 e^{-c_2 B}, \quad (5.3.5)$$

where the linear form

$$\Lambda = b_1 \log \alpha_1 + \cdots + b_n \log \alpha_n - \log \alpha_{n+1}, \quad (5.3.6)$$

and c_3 is a positive constant that can be effectively computed.

3. The crucial step in the general strategy is to apply Baker's theorem which gives an inequality

$$|\Lambda| \geq e^{-c_4 B}, \quad (5.3.7)$$

where c_4 is a positive constant that can be effectively computed. When we put these two inequalities together we obtain:

$$e^{-c_4 B} \leq |\Lambda| \leq c_3 e^{-c_2 B}, \quad (5.3.8)$$

which in turn leads to an explicit upper bound B_0 for B .

4. We reduce this upper bound B_0 , which is usually very big, to a much smaller upper bound B_1 , by using continued fraction techniques, Davenport's lemma or the LLL-reduction algorithm, depending on whether the linear form is either in two, three or more than three logarithms.
5. From this upper bound B_1 we deduce an upper bound for the unknowns in the original equations.
6. Using search techniques and properties of the initial equations we determine all possible solutions.

Davenport's Lemma and LLL-reduction

The lemma of Davenport is a result proved by Baker and Davenport [BD69], that can be applied to linear forms in three logarithms to show that a certain gap must exist between solutions of certain equations. In its original form it is given as follows:

Lemma 5.1 (Baker-Davenport). *Let $K, M > 6$, p, q be positive integers satisfying the following inequalities:*

$$1 \leq q \leq KM, \quad (5.3.9a)$$

$$|\theta q - p| \leq \frac{2}{KM}, \quad (5.3.9b)$$

$$\|q\beta\| \geq \frac{3}{K}, \quad (5.3.9c)$$

where θ, β are irrational numbers and $\|z\|$ is the distance of a real number z to its nearest integer, that is $\|z\| = |z - \lfloor z + \frac{1}{2} \rfloor|$. Then the inequality

$$|m\theta + n - \beta| \leq c^{-m}, \quad (5.3.10)$$

has no solution in integers (m, n) in the range

$$\frac{\log K^2 M}{\log c} < m < M. \quad (5.3.11)$$

For LLL-reduction techniques we follow Cohen's description on lattices and reduction [Coh96].

Definition 5.1. *Let K be a field of characteristic different from 2 and let V be a K -vector space. A map q from V to K is a quadratic form if the following conditions are satisfied:*

1. For every $\lambda \in K$ and $x \in V$ we have:

$$q(\lambda \cdot x) = \lambda^2 q(x). \quad (5.3.12)$$

2. Let the function $b(x, y)$ be defined by

$$b(x, y) = \frac{1}{2}(q(x + y) - q(x) - q(y)). \quad (5.3.13)$$

Then b is a symmetric bilinear form.

We have the obvious identity

$$b(x, x) = q(x). \quad (5.3.14)$$

If $K = \mathbb{R}$, and if for all $x \in V$ we have $q(x) > 0$, we say that q is positive definite.

Definition 5.2. A lattice \mathcal{L} is a free \mathbb{Z} -module of finite rank together with a positive definite quadratic form q on $\mathcal{L} \otimes \mathbb{R}$.

Let $(b_i)_{1 \leq i \leq n}$ be a \mathbb{Z} -basis for \mathcal{L} . If

$$x = \sum_{1 \leq i \leq n} x_i b_i \in \mathcal{L}, \quad (5.3.15)$$

with $x_i \in \mathbb{Z}$, then we have that

$$q(x) = \sum_{1 \leq i, j \leq n} q_{i,j} x_i x_j, \quad (5.3.16)$$

where $q_{i,j} = b(b_i, b_j)$. The matrix $Q = (q_{i,j})_{1 \leq i, j \leq n}$ is a positive definite symmetric matrix that verifies

$$b(x, y) = Y^T Q X, \quad (5.3.17)$$

where X, Y are the column vectors of the coordinates of x and y . As Q is positive definite, we have that the determinant $\det Q > 0$. The determinant $d(\mathcal{L})$ of the lattice \mathcal{L} is defined as

$$d(\mathcal{L}) = \sqrt{\det Q}. \quad (5.3.18)$$

A lattice \mathcal{L} can also be considered as a discrete subgroup of rank n of the Euclidean vector space $\mathcal{L} \otimes \mathbb{R}$. If $(b_i)_{1 \leq i \leq n}$ is a \mathbb{Z} -basis for \mathcal{L} , then the matrix of scalar products

$$Q = (b_i \cdot b_j)_{1 \leq i \leq n} \quad (5.3.19)$$

is called the Gram matrix of the vectors b_i . We have the following theorem:

Theorem 5.3. If Q is the matrix of a positive definite quadratic form, then Q is the Gram matrix of some lattice basis. Moreover, the Gram matrix of a lattice basis $(b_i)_{1 \leq i \leq n}$ determines that basis uniquely up to isometry.

The existence of an orthonormal basis in a Euclidean vector space is proved by the Gram-Schmidt orthonormalization procedure. For the normalization part of this procedure, square roots need to be taken, but the orthogonalization procedure works just as well without:

Theorem 5.4. *Let $(b_i)_{1 \leq i \leq n}$ be a basis of a Euclidean vector space E . Define*

$$b_i^* = b_i - \sum_{j=1}^{i-1} \mu_{i,j} b_j^*, \quad (5.3.20)$$

where

$$\mu_{i,j} = \frac{b_i \cdot b_j^*}{b_j^* \cdot b_j^*}. \quad (5.3.21)$$

Then the $(b_i^*)_{1 \leq i \leq n}$ form an orthogonal but not necessarily orthonormal basis of E . We have the following equality for the determinant of the lattice:

$$d(\mathcal{L}) = \prod_{1 \leq i \leq n} \|b_i^*\|^2. \quad (5.3.22)$$

The following inequality is a corollary of this theorem:

Corollary 5.1 (Hadamard's inequality). *Let (\mathcal{L}, q) be a lattice of determinant $d(\mathcal{L})$, let $(b_i)_{1 \leq i \leq n}$ be a \mathbb{Z} -basis for \mathcal{L} , then*

$$d(\mathcal{L}) \leq \prod_{i=1}^n \sqrt{q(b_i, b_i)}. \quad (5.3.23)$$

Amongst all the \mathbb{Z} -bases of a lattice \mathcal{L} , some are better than others. The bases whose elements are the shortest are called reduced bases. We can think of a reduced basis as of a basis that is almost orthogonal. A basis is called LLL-reduced (for A. K. Lenstra, H. W. Lenstra and L. Lovász) [LLL82] if the following conditions are satisfied:

1. The real numbers $\mu_{i,j}$ all verify the inequality $|\mu_{i,j}| \leq \frac{1}{2}$.
2. For all $1 \leq i \leq n$ we have the following inequality:

$$|b_i^* + \mu_{i,i-1} b_{i-1}^*|^2 \geq \frac{3}{4} |b_{i-1}^*|^2, \quad (5.3.24)$$

where the norm of a vector is defined as $|b_i| = \sqrt{q(b_i, b_i)}$.

We have the following theorem:

Theorem 5.5. *Let $(b_i)_{1 \leq i \leq n}$ be an LLL-reduced basis of a lattice \mathcal{L} , then the following inequalities are satisfied:*

$$d(\mathcal{L}) \leq \prod_{i=1}^n |b_i| \leq 2^{\frac{n(n-1)}{4}} d(\mathcal{L}), \quad (5.3.25a)$$

$$|b_j| \leq 2^{\frac{i-1}{2}} |b_i^*|, \quad \text{if } 1 \leq j \leq i \leq n, \quad (5.3.25b)$$

$$|b_1| \leq 2^{\frac{n-1}{4}} \sqrt[n]{d(\mathcal{L})}. \quad (5.3.25c)$$

We also have for any linear independent vectors $x_1, \dots, x_t \in \mathcal{L}$, that

$$|b_j| \leq 2^{\frac{n-1}{2}} \max(|x_1|, \dots, |x_t|), \quad \text{with } 1 \leq j \leq t. \quad (5.3.26)$$

We can apply LLL-reduction to reduce the upper bound B_0 for linear forms in logarithms. This is done in the following manner. Let

$$0 < |b_1\alpha_1 + \dots + b_n\alpha_n + \alpha| < c_3 e^{-c_2 B}. \quad (5.3.27)$$

Let \mathcal{L} be the lattice in \mathbb{R}^{n+1} spanned by the column vectors of

$$\begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & & & 1 & 0 \\ C\alpha_1 & C\alpha_2 & \dots & C\alpha_n & C\alpha \end{pmatrix}, \quad (5.3.28)$$

where C is a constant. Let e_1 be the first basis vector of the LLL-reduced basis of the lattice \mathcal{L} . We then have the following inequality:

$$|e_1|^2 \leq 2^n |x|^2, \quad (5.3.29)$$

for all vectors $x \in \mathcal{L}$. If we choose the constant C , such that

$$|e_1| \geq \sqrt{(n+2)2^n} B_0, \quad (5.3.30)$$

then we obtain the following inequality:

$$B \leq \frac{\log c_3 C - \log B_0}{c_2} = B_1. \quad (5.3.31)$$

This reduces the upper bound B_0 to approximately $\log B_0$.

5.3.2 Upper bound for smallest solution

We want to find an upper bound for the system (5.1.1). Let (x_0, y_0) be the smallest solution in positive integers of equation (5.1.1a) and (z'_0, y'_0) that of equation (5.1.1b). Consider the algebraic numbers

$$R = x_0 + y_0 \sqrt{a}, \quad (5.3.32a)$$

$$R' = z'_0 + y'_0 \sqrt{b}. \quad (5.3.32b)$$

which have minimal polynomials

$$X^2 - 2x_0X + 1, \quad (5.3.33a)$$

$$X^2 - 2z'_0X + 1. \quad (5.3.33b)$$

The classical height $H(\alpha)$ of an algebraic number α is defined as the maximum of the absolute values of the coefficients of its minimal polynomial in $\mathbb{Z}[X]$ with the greatest common divisor of these coefficients being 1. Therefore the height of R is $2x_0$. For practical purposes we are going to assume that $a, b < 1000$, so that $H(R) < 4 \cdot 10^{37}$. This occurs for $a = 661$. Consider the polynomial

$$p(x) = \prod_{i=1}^4 (X - E_i), \quad (5.3.34)$$

where the algebraic numbers E_i are defined as follows:

$$E_1 = \frac{(x_0 + y_0\sqrt{a})\sqrt{b}}{(z'_0 + y'_0\sqrt{b})\sqrt{a}}, \quad E_2 = -\frac{(x_0 - y_0\sqrt{a})\sqrt{b}}{(z'_0 + y'_0\sqrt{b})\sqrt{a}}, \quad (5.3.35ab)$$

$$E_3 = -\frac{(x_0 + y_0\sqrt{a})\sqrt{b}}{(z'_0 - y'_0\sqrt{b})\sqrt{a}}, \quad E_4 = \frac{(x_0 - y_0\sqrt{a})\sqrt{b}}{(z'_0 - y'_0\sqrt{b})\sqrt{a}}. \quad (5.3.35cd)$$

The polynomial $p(x)$ can be written as

$$p(x) = \frac{1}{a^2}(a^2x^4 + 4a^2by_0y'_0x^3 - 2ab(1 + 2ay_0^2 + 2by_0'^2)x^2 + 4ab^2y_0y'_0x + b^2). \quad (5.3.36)$$

None of the linear polynomial factors of $p(x)$ is in $\mathbb{Q}[x]$, therefore E_1 does not have degree 1 or 3. The height of E_1 is easily bounded by 10^{86} . We have therefore the following upperbounds:

$$(1 + \log H(R)), (1 + \log H(R')), (1 + \log H(E_1)), \log R, \log R', \log |E_1| < 200. \quad (5.3.37)$$

All solutions of equation (5.1.1a) are given by:

$$x_m = \frac{(x_0 + y_0\sqrt{a})^{m+1} + (x_0 - y_0\sqrt{a})^{m+1}}{2}, \quad (5.3.38a)$$

$$y_m = \frac{(x_0 + y_0\sqrt{a})^{m+1} - (x_0 - y_0\sqrt{a})^{m+1}}{2\sqrt{a}}. \quad (5.3.38b)$$

and likewise for equation (5.1.1b):

$$z'_n = \frac{(z'_0 + y'_0\sqrt{b})^{n+1} + (z'_0 - y'_0\sqrt{b})^{n+1}}{2}, \quad (5.3.39a)$$

$$y'_n = \frac{(z'_0 + y'_0\sqrt{b})^{n+1} - (z'_0 - y'_0\sqrt{b})^{n+1}}{2\sqrt{b}}. \quad (5.3.39b)$$

So to solve simultaneous Pell equations it is sufficient to find all (m, n) , such that

$$y_m = y'_n. \quad (5.3.40)$$

We have the recurrence relations:

$$y_{m+2} = 2x_0y_{m+1} - y_m, \quad (5.3.41a)$$

$$y'_{n+2} = 2z'_0y'_{n+1} - y'_n. \quad (5.3.41b)$$

So provided we know y_0, y'_0 it is relatively straightforward to check whether there is a solution with $m, n < 100$. An upper bound to the smallest solution will tell us upto which value we need to compute (m, n) to be sure that there are no solutions. Let

$$P = \frac{(x_0 + y_0\sqrt{a})^{m+1}}{\sqrt{a}}, \quad (5.3.42a)$$

$$Q = \frac{(z'_0 + y'_0\sqrt{b})^{n+1}}{\sqrt{b}}. \quad (5.3.42b)$$

Then we have that

$$\frac{1}{P} = (x_0 - y_0\sqrt{a})^{m+1}\sqrt{a}, \quad (5.3.43a)$$

$$\frac{1}{Q} = (z'_0 - y'_0\sqrt{b})^{n+1}\sqrt{b}. \quad (5.3.43b)$$

The smallest possible value for $(x_0 + y_0\sqrt{a})$ is $c = 2 + \sqrt{3}$. We obviously have the inequalities:

$$P > c^{m-1}, \quad (5.3.44a)$$

$$Q > c^{n-1}. \quad (5.3.44b)$$

We also have the relation

$$\frac{P}{Q} = \frac{E_1 R^m}{R'^n}. \quad (5.3.45)$$

The case of equality $y_m = y'_n$ can only happen if

$$P - \frac{1}{aP} = Q - \frac{1}{bQ}. \quad (5.3.46)$$

Suppose that $P > Q$. Then if there is a solution we have:

$$\begin{aligned} \frac{P}{Q} - 1 &= \frac{1}{aPQ} - \frac{1}{bQ^2} \\ &< \frac{1000}{PQ} \\ &< \frac{1000}{c^{m-1}c^{n-1}} \\ &< c^{-\max\{m,n\}}. \end{aligned} \quad (5.3.47)$$

It therefore follows that

$$0 < \left| \log \frac{P}{Q} \right| < c^{-\max\{m,n\}}. \quad (5.3.48)$$

A similar argument for the case $P < Q$ leads to the same inequality and therefore, assuming $m, n \geq 10$, we have

$$0 < \left| \log \frac{E_1 R^m}{R^n} \right| < c^{-\max\{m,n\}}, \quad (5.3.49)$$

which can also be written as

$$0 < |m \log R - n \log R' + \log E_1| < c^{-\max\{m,n\}}. \quad (5.3.50)$$

To this linear form Λ in three logarithms we apply Baker's theory and Davenport's lemma. The logarithm of the algebraic numbers R, R', E_1 as well as the logarithm of their heights are all bounded by 200, so we can use theorem 5.2 and we obtain the following estimate:

$$-2^{101} 200^3 (\max(\log m, \log n) + 11) 11 < \log |\Lambda| < c^{-\max\{m,n\}}, \quad (5.3.51)$$

which results in the inequality

$$\max(m, n) < 10^{41}. \quad (5.3.52)$$

To this upper bound we apply Davenport's lemma, where we set:

$$\theta = \frac{\log R}{\log R'}, \quad (5.3.53a)$$

$$\beta = -\frac{\log E_1}{\log R'}, \quad (5.3.53b)$$

$$M = 10^{41}. \quad (5.3.53c)$$

This will result in the inequality in m :

$$\max(m, n) < 83. \quad (5.3.54)$$

This is sufficient as an upper bound for the smallest solution even if a second application of Davenport's lemma would reduce this upper bound even further.

5.4 Finite Number of Solutions

5.4.1 Introduction

We will highlight some elements of the proof of Cipu and Mignotte [CM] that the system

$$x^2 - az^2 = 1, \quad (5.4.1a)$$

$$y^2 - bz^2 = 1, \quad (5.4.1b)$$

has at most two distinct solutions in positive integers x, y, z .

It is interesting to note that there exist families of integers (a, b) , such that the system (5.4.1) has two positive solutions.

Let $1 < l, m$ be positive integers and define the following quantities:

$$\alpha = m + \sqrt{m^2 - 1}, \quad (5.4.2a)$$

$$n(l, m) = \frac{\alpha^{2l} - \alpha^{-2l}}{4\sqrt{m^2 - 1}}. \quad (5.4.2b)$$

Then the simultaneous Pell equations of the family (a, b) , such that

$$a = m^2 - 1, \quad (5.4.3a)$$

$$b = n(l, m)^2 - 1, \quad (5.4.3b)$$

have the following two solutions in positive integers:

$$(x_0, y_0, z_0) = (m, n(l, m), 1), \quad (5.4.4a)$$

$$(x_1, y_1, z_1) = \left(\frac{\alpha^{2l} + \alpha^{-2l}}{2}, 2n(l, m)^2 - 1, 2n(l, m) \right). \quad (5.4.4b)$$

An earlier result by Yuan [Yua02] showed that there are at most finitely many cases of simultaneous Pell equations with three solutions:

Theorem 5.6 (Yuan). *If $\max(a, b) \geq 1.4 \cdot 10^{57}$, then the system (5.4.1) has at most two distinct solutions in positive integers.*

Cipu and Mignotte prove that the system (5.4.1) have at most two solutions in positive integers x, y, z , if $a < b$ are distinct positive integers, removing these finitely many exceptions. The basic idea of the proof is a three step approach. First it is shown that any system of simultaneous Pell equations can be transformed to another system of simultaneous Pell equations with the same number of solutions. The new system has coefficients a, b of a special type so that it is straightforward to find the smallest solution x_0, y_0, z_0 of the system. The second step creates a linear form in three logarithms from this smallest solution and two hypothetical bigger solutions x_1, y_1, z_1 and x_2, y_2, z_2 . Baker's theory on these kind of forms results in an upper bound for the biggest of these two solutions. The last step is a gap principle that shows that the distance $y_2 - y_1$ must necessarily exceed some kind of lower bound. It will turn out that this lower bound will conflict with the upper bound constraint from the second step. Hence there can be no system of simultaneous Pell equations with three distinct solutions.

5.4.2 Transforming the equations

We want to transform the system (5.4.1) to a system of equations with an obvious smallest solution. We need to show that this transformation does not reduce the number of solutions. To do so we need to prove the following lemma:

Lemma 5.2. *Suppose that the system (5.4.1) has at least one solution in positive integers. Let z_0 be the smallest positive value taken by the third component of a solution (x, y, z) . Then for any solution (x_i, y_i, z_i) of (5.4.1), z_i is a multiple of z_0 .*

This lemma implies that if (x_0, y_0, z_0) is the solution of (5.4.1) with minimal third component, then this system has as many positive integral solutions as the system

$$u^2 - (x_0^2 - 1)v^2 = 1, \quad (5.4.5a)$$

$$w^2 - (y_0^2 - 1)v^2 = 1. \quad (5.4.5b)$$

So we will consider from now on that

$$a = m^2 - 1, \quad (5.4.6a)$$

$$b = n^2 - 1, \quad (5.4.6b)$$

for integers $n > m \geq 2$. We set

$$\alpha = m + \sqrt{m^2 - 1}, \quad (5.4.7a)$$

$$\beta = n + \sqrt{n^2 - 1}. \quad (5.4.7b)$$

Let (x, y, z) be a positive integer solution of (5.4.1). Then $z = U_j = U'_k$, where

$$U_j = \frac{\alpha^j - \alpha^{-j}}{2\sqrt{a}}, \quad (5.4.8a)$$

$$U'_k = \frac{\beta^k - \beta^{-k}}{2\sqrt{b}}, \quad (5.4.8b)$$

with j, k positive integers.

5.4.3 Linear form in three logarithms

We will build a linear form on logarithms that depend on α, β . We first observe the following inequality:

$$\alpha^j < \beta^k < \sqrt{\frac{b}{a}} \alpha^j, \quad (5.4.9)$$

which follows from the fact that $m < n$, and that the map $x \mapsto x - \frac{1}{x}$ is increasing for positive x , and the fact that

$$\frac{\alpha^j - \alpha^{-j}}{2\sqrt{a}} = \frac{\beta^k - \beta^{-k}}{2\sqrt{b}}. \quad (5.4.10)$$

Another useful pair of inequalities is the following:

$$\left(1 + \frac{4}{5a^2}\right) \frac{\beta^2}{\alpha^2} < \frac{b}{a} < \left(1 + \frac{1}{2a}\right) \frac{\beta^2}{\alpha^2}, \quad (5.4.11)$$

which can easily be derived from the fact that $a \geq 3$, $b \geq a + 5$ and the fact that

$$2x + 1 - \frac{1}{4x} < 2\sqrt{x^2 + x} < 2x + 1, \quad (5.4.12)$$

for positive x . From the inequalities (5.4.9) and (5.4.11) we obtain the following inequality in α, β :

$$\beta^{k-1} < \left(1 + \frac{1}{4a}\right) \alpha^{j-1}. \quad (5.4.13)$$

We have the following lemma:

Lemma 5.3. *Let (x, y, z) be a solution of the system (5.4.1). If $z = U_j = U'_k$, with $j > k$, then j and k have the same parity. Moreover, if $j = k + 2$, then k is even.*

This is easily proven by using the recurrence sequences and by inspection for the special case $j = k + 2$. We also have a double bound on U_t :

Lemma 5.4. *For any $t \geq 2$, we have the inequalities:*

$$\alpha^t < U_{t+1} < (2m)^t. \quad (5.4.14)$$

As a consequence, we have

$$\left\lfloor U_{t+1}^{\frac{1}{t}} \right\rfloor = 2m - 1. \quad (5.4.15)$$

This leads to the following corollary:

Corollary 5.2. *If $U_j = U'_k$, then*

$$(j - 1) \log \alpha < (k - 1) \log 2n. \quad (5.4.16)$$

We consider the linear form in three logarithms:

$$\Lambda = \frac{1}{2} \log \frac{b}{a} + j \log \alpha - k \log \beta. \quad (5.4.17)$$

This form is bounded from above by

$$\Lambda < -\log(1 - \alpha^{-2j}) < \frac{\alpha^{2-2j}}{\alpha^2 - 1}. \quad (5.4.18)$$

From this inequality we obtain

$$\log \Lambda < -2j \log \alpha + \log \left(\frac{\alpha^{2-2j}}{\alpha^2 - 1} \right) < -2j \log \alpha + 0.075. \quad (5.4.19)$$

5.4.4 Gap principles

Suppose that the system (5.4.1) has at least three solutions (x_i, y_i, z_i) . We have

$$\begin{aligned} z_i &= \frac{\alpha^{j_i} - \alpha^{-j_i}}{2\sqrt{a}} \\ &= \frac{\beta^{k_i} - \beta^{-k_i}}{2\sqrt{b}}, \end{aligned} \quad (5.4.20)$$

for integers $1 = j_1 < j_2 < j_3$ and $1 = k_1 < k_2 < k_3$. The goal of this section is to prove that if such a solution exists, then the gap between k_2 and k_3 must be rather large. So large in fact, that it will create a contradiction with the upper bound found for k_3 in the previous section. Yuan [Yua02] proves the following lemma:

Lemma 5.5 (Yuan). *There exist integers $q_j, q_k \geq 2$ and $\sigma_j, \sigma_k \in \{-1, 0, 1\}$, such that*

$$j_3 = q_j j_2 + \sigma_j, \quad (5.4.21a)$$

$$k_3 = q_k k_2 + \sigma_k, \quad (5.4.21b)$$

$$q_j \sigma_j \equiv q_k \sigma_k \equiv 0 \pmod{2}. \quad (5.4.21c)$$

Cipu and Mignotte improve this lemma to obtain the following result:

Lemma 5.6. *With the above notations, we have the equality*

$$\sigma_j = \sigma_k, \quad (5.4.22)$$

and the inequalities

$$q_j > q_k, \quad (5.4.23a)$$

$$mq_j < nq_k. \quad (5.4.23b)$$

So from here on, we drop the indices for σ . Using these lemmata, Mignotte and Cipu prove the following proposition, which gives the desired gap principle:

Proposition 5.1 (Cipu-Mignotte). *We have the following lower bound for the integer j_3 :*

$$j_3 > \begin{cases} 1.99 j_2 \beta^{\frac{2}{3}}, & \text{if } k_2 = 2, k_3 \text{ is odd, } \beta > 8000 \text{ and } l = 2, \\ 1.99 j_2 \beta^{\frac{4}{5}}, & \text{if } k_2 = 2, k_3 \text{ is odd, } \beta > 8000 \text{ and } l \geq 3, \\ 2.81 j_2 \beta^{\frac{k_2-2}{2}}, & \text{if } k_2 > 2 \text{ is even,} \\ 3.96 j_2 \beta^{\frac{k_2-3}{2}}, & \text{if } k_2 > 2 \text{ is odd,} \end{cases} \quad (5.4.24)$$

where $n = n(l, m)$, for some integers $l, m > 1$.

Using this gap principle together with a lower bound for linear forms in logarithms of three algebraic numbers from Matveev [Mat00], Cipu and Mignotte obtain the fact that for

$$\max(a, b) \geq 2.26 \cdot 10^{49}, \quad (5.4.25)$$

the simultaneous Pell equations have at most two solutions. This upper bound is now used as input in a theorem of Mignotte [Mig04, BMS06] to obtain a tighter lower bound for linear forms in logarithms of three algebraic numbers. This reduces the above upper bound so that for:

$$\max(a, b) \geq 1.2 \cdot 10^{38}. \quad (5.4.26)$$

the simultaneous Pell equations have at most two solutions. From that point on Cipu and Mignotte distinguish two cases.

In the case that the solution (x_2, y_2, z_2) verifies

$$z_2 = 2n, \quad (5.4.27)$$

an explicit computation using techniques from computational Diophantine approximation theory allows them to verify that there are only two solutions. In the case that z_2 is a higher power of the fundamental solution, they could use much tighter bounds which allow them to eliminate this case as well. Therefore a pair of simultaneous Pell equations has at most two solutions. We end this section with an as of yet unproven conjecture of Yuan [Yua04].

Conjecture 5.2 (Yuan). *The equations:*

$$x^2 - az^2 = 1, \quad (5.4.28a)$$

$$y^2 - bz^2 = 1, \quad (5.4.28b)$$

have at most one solution in positive integers, unless

$$ac^2 = m^2 - 1, \quad (5.4.29a)$$

$$bd^2 = n(l, m)^2 - 1, \quad (5.4.29b)$$

where c, d are positive integers, in which case these equations have exactly two solutions in positive integers.

5.5 Quantum algorithm for simultaneous Pell equations

We extend Hallgren's result for single Pell equations, by giving a polynomial time quantum algorithm that solves simultaneous Pell equations. This algorithm uses Hallgren's algorithm as a subroutine.

Suppose we want to solve the pair of simultaneous Pell equations

$$x^2 - az^2 = 1, \quad (5.5.1a)$$

$$y^2 - bz^2 = 1, \quad (5.5.1b)$$

in polynomial time. That is to say, in a time $O(\max(\log a, \log b)^k)$. We have the following ingredients:

1. A polynomial time quantum algorithm that computes the regulator R of $\mathbb{Q}[\sqrt{d}]$ with precision 10^{-n} in time $O((\log d)^{c_1}, n^{c_2})$ with probability $O((\log d)^{-c_3}, n^{-c_4})$, if $10^{-n} < \frac{d_{\min}}{\log d}$, where c_i are positive constants.
2. A polynomial time classical algorithm that computes a power product representation of the fundamental solution of the Pell equation from an integer \tilde{R} , such that $|\tilde{R} - R| < 1$, where R is the regulator of the number field $\mathbb{Q}[\sqrt{d}]$.
3. An upper bound for the smallest solution of a pair of simultaneous Pell equations. This upper bound is given as $\max(m, n) \leq 83$, where m, n are the powers to which the fundamental solutions of the simultaneous Pell equations need to be raised. We have some remarks for this value 83. When we followed Anglin's approach to obtain an upper bound, we restricted ourselves to $\max(a, b) \leq 1000$. On the other hand, the upper bound is a direct result of the worst-case scenario for one specific fundamental solution, the case of $d = 661$. So this upper bound will only grow every time that we hit upon a new worst-case scenario for a certain $d' > 661$.

A second point to note is that we already indicated that this upper bound of 83 could be improved by running Davenport's lemma again with this new upper bound.

So how do we proceed? We use Hallgren's algorithm to obtain the regulators of both of the Pell equations. This can be done on a quantum computer in polynomial time. From these regulators we obtain power product representations of the fundamental solutions:

$$\begin{aligned} x_0^2 - az_0^2 &= 1, \\ x_0 + z_0\sqrt{a} &= \prod_{i=1}^t (a_i + b_i\sqrt{a})^{n_i}, \end{aligned} \tag{5.5.2a}$$

$$\begin{aligned} x_0'^2 - bz_0'^2 &= 1, \\ x_0' + z_0'\sqrt{b} &= \prod_{i=1}^{t'} (a_i' + b_i'\sqrt{b})^{n_i'}, \end{aligned} \tag{5.5.2b}$$

where a_i, a_i', b_i, b_i' are rational numbers and n_i, n_i' are integers. We do not use this power product representation directly, but we will need it in the end to compute z . We try to find powers m, n , such that $z_m = z_n'$, where

$$x_m + z_m\sqrt{a} = (x_0 + z_0\sqrt{a})^m, \tag{5.5.3a}$$

$$y_n + z_n'\sqrt{b} = (y_0 + z_0'\sqrt{b})^n. \tag{5.5.3b}$$

We can derive the following relations from the regulators:

$$\log z_m = mR_a + \log \frac{2}{\sqrt{a}}, \quad (5.5.4a)$$

$$\log z'_n = nR_b + \log \frac{2}{\sqrt{b}}. \quad (5.5.4b)$$

In case of equality $z_m = z'_n$ we have

$$nR_b - mR_a = \log \frac{\sqrt{b}}{\sqrt{a}}. \quad (5.5.5)$$

So for all positive integers $m, n < 83$, we test the above equation. This can be done by fixing m and solving the equality for n . If for a pair (m_0, n_0) we have equality, then

$$z_{m_0} = z'_{n_0}. \quad (5.5.6)$$

We now use the power product representation for the smallest solution to obtain a description of this smallest solution as a difference of power product representations:

$$z_{m_0} = \frac{1}{2\sqrt{a}} \left(\prod_{i=1}^t (a_i + b_i\sqrt{a})^{m_0n_i} - \prod_{i=1}^t (a_i - b_i\sqrt{a})^{m_0n_i} \right). \quad (5.5.7)$$

If we do not find an equality before running out of bounds, then the simultaneous Pell equations do not have a solution in positive integers. The case of two solutions follows essentially the same scheme. First we filter out the obvious cases of two solutions, that is the cases where

$$ac^2 = m^2 - 1, \quad (5.5.8a)$$

$$bd^2 = n(l, m)^2 - 1, \quad (5.5.8b)$$

with c, d positive integers. For the other cases, Yuan conjectured [Yua04] that there cannot be two different solutions in positive integers. While we cannot prove this conjecture, we can test it. The idea is to obtain a similar bound for $\max(m, n)$, but now for the second smallest solution. From that point on we repeat the above procedure, expecting to run out of bounds before finding a second solution.

5.6 Conclusion and Perspective

We have exhibited a polynomial time quantum algorithm with polynomial time classical postprocessing that finds solutions to simultaneous Pell equations. The key ingredients to this algorithm are Hallgren's algorithm that computes the regulator of a number field in polynomial time on a quantum computer and the upper bound on smallest solutions of simultaneous Pell

equations obtained by Anglin. A combination of these two results gives the above algorithm. It is natural to wonder whether this procedure can be extended to other types of equations. A natural extension would be to try to solve the pair of simultaneous Fermat equations:

$$Ax^2 - Bz^2 = C, \quad (5.6.1a)$$

$$Dy^2 - Ez^2 = F, \quad (5.6.1b)$$

with the usual conditions on A, B, C, D, E, F in order to prevent having equivalent equations. Anglin [Ang95] gives an upper bound for the smallest solution of this type of equations in the special case $B = E = 1$. To extend the above algorithm we also need to have a fast method to solve the individual equations and a fast method to test for equality of individual solutions. We also need a small bound for subsequent solutions and a proof that there are only a limited number of solutions. It seems probable that any type of pair of equations where these four conditions are met can be solved in polynomial time on a quantum computer.

Appendix A

Get your facts first, then you
can distort them as you please.

MARK TWAIN

Kronecker product and sum

Let A be an $m \times n$ -matrix and B a $p \times q$ matrix. Then the Kronecker product

$$A \otimes B \tag{A.1}$$

is an $mp \times nq$ matrix with the following coefficients:

$$A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix}. \tag{A.2}$$

The Kronecker product is a special case of the tensor product and therefore has the following properties:

$$A \otimes (B + C) = A \otimes B + A \otimes C, \tag{A.3a}$$

$$(A + B) \otimes C = A \otimes C + B \otimes C, \tag{A.3b}$$

$$(kA) \otimes B = A \otimes (kB) = k(A \otimes B), \tag{A.3c}$$

$$A \otimes (B \otimes C) = (A \otimes B) \otimes C, \tag{A.3d}$$

where A, B, C are matrices and k a scalar. The Kronecker product is not commutative in general. We have the following useful result:

Lemma A.1. *Let A, B, C, D be matrices, such that the multiplication $A \otimes B$ by $C \otimes D$ is well defined. We have the following identity:*

$$(A \otimes B) \cdot (C \otimes D) = (AC \otimes BD). \tag{A.4}$$

As a consequence we have the following corollary:

Corollary A.1. *The matrix $A \otimes B$ is invertible if and only if the matrices A and B are invertible. In that case we have*

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}. \tag{A.5}$$

For the eigenvalues and eigenvectors of the Kronecker product we have the following theorem:

Theorem A.1. *Let A be an $m \times m$ matrix with eigenvalues $\lambda_1, \dots, \lambda_m$ and let B be an $n \times n$ matrix with eigenvalues μ_1, \dots, μ_n . Then the eigenvalues of $A \otimes B$ are given by $\lambda_i \mu_j$.*

If x_1, \dots, x_m are linearly independent eigenvectors of A , where the eigenvector x_i corresponds to the eigenvalue λ_i and x'_1, \dots, x'_n linearly independent eigenvectors of B , where the eigenvector x'_j corresponds to the eigenvalue μ_j , then $x_i \otimes x'_j$ are linearly independent eigenvectors of $A \otimes B$ with corresponding eigenvalues $\lambda_i \mu_j$.

From this result we can derive the trace and determinant of the Kronecker product:

Corollary A.2. *Let A be an $m \times m$ matrix and B an $n \times n$ matrix. Then*

$$\text{Tr}(A \otimes B) = \text{Tr} A \text{Tr} B = \text{Tr}(B \otimes A), \quad (\text{A.6a})$$

$$\det(A \otimes B) = (\det A)^n (\det B)^m = \det(B \otimes A). \quad (\text{A.6b})$$

Let A be an $m \times m$ matrix and B an $n \times n$ matrix. Then the Kronecker sum of A and B is defined as follows:

$$A \oplus_K B = A \otimes I_n + I_m \otimes B. \quad (\text{A.7})$$

The Kronecker sum of matrices is non-commutative in general, that is

$$A \oplus_K B \neq B \oplus_K A \quad (\text{A.8})$$

We have the following theorem regarding the eigenvalues and eigenvectors of the Kronecker sum of matrices:

Theorem A.2. *Let A be an $m \times m$ matrix with eigenvalues $\lambda_1, \dots, \lambda_m$ and let B be an $n \times n$ matrix with eigenvalues μ_1, \dots, μ_n . Then the eigenvalues of $A \oplus_K B$ are given by $\lambda_i + \mu_j$.*

If x_1, \dots, x_m are linearly independent eigenvectors of A , where the eigenvector x_i corresponds to the eigenvalue λ_i and x'_1, \dots, x'_n linearly independent eigenvectors of B , where the eigenvector x'_j corresponds to the eigenvalue μ_j , then $x_i \otimes x'_j$ are linearly independent eigenvectors of $A \oplus_K B$ with corresponding eigenvalues $\lambda_i \mu_j$.

Let both A and B be $n \times n$ matrices. We then have the following identity:

$$e^{A \oplus_K B} = e^A \otimes e^B. \quad (\text{A.9})$$

There is a straightforward generalization to a Kronecker sum of n matrices:

$$A_1 \oplus_K \dots \oplus_K A_n = A_1 \otimes \dots \otimes I + \dots + I \otimes \dots \otimes A_n. \quad (\text{A.10})$$

We have the same relation for the exponential:

$$e^{A_1 \oplus_K \dots \oplus_K A_n} = e^{A_1} \otimes \dots \otimes e^{A_n}. \quad (\text{A.11})$$

Appendix B

Imagination will often carry us
to worlds that never were. But
without it we go nowhere.

CARL SAGAN

Continued Fractions

Continued fractions are used to give good rational approximations of irrational numbers. We first define a sequence of functions.

Let a_1, \dots, a_n be real numbers and $a_1 \geq 1$.

Define the sequences $(f_i)_i, (g_i)_i$, by

$$\begin{aligned} f_{-1} &= 0, & f_0 &= 1, \\ f_{n+1} &= a_{n+1}f_n(a_1, \dots, a_n) + f_{n-1}(a_1, \dots, a_{n-1}), \end{aligned} \tag{B.1a}$$

$$\begin{aligned} g_{-1} &= 1, & g_0 &= 0, \\ g_{n+1} &= a_{n+1}g_n(a_1, \dots, a_n) + g_{n-1}(a_1, \dots, a_{n-1}). \end{aligned} \tag{B.1b}$$

By induction, we can prove the following:

Theorem B.1. *The sequence $(f_i)_i$ verifies the recurrence relation:*

$$f_n(a_1, \dots, a_{n-1}, a_n + \frac{1}{a_{n+1}}) = \frac{1}{a_{n+1}} f_{n+1}(a_1, \dots, a_n, a_{n+1}). \tag{B.2}$$

An analogue result holds for the sequence $(g_i)_i$.

The following result can be derived almost directly from the previous theorem:

Theorem B.2. *For all positive integers n , we have the recurrence relation:*

$$f_n(a_1, \dots, a_n) = a_1 f_{n-1}(a_2, +\frac{1}{a_1}, a_3, \dots, a_n). \tag{B.3}$$

There is also a direct relation between f and g :

Theorem B.3. *The functions f_n and g_n verify the equations:*

$$g_n(a_1, \dots, a_n) = f_{n-1}(a_2, \dots, a_n), \tag{B.4a}$$

$$f_n g_{n-1} - f_{n-1} g_n = (-1)^n. \tag{B.4b}$$

When n tends to infinity it is possible to define a limit for the fraction $\frac{f_n}{g_n}$:

Theorem B.4. *The sequences $\frac{f_{2n+1}}{g_{2n+1}}$ and $\frac{f_{2n}}{g_{2n}}$ are respectively strictly increasing and strictly decreasing. The limit*

$$\lim_{n \rightarrow \infty} \frac{f_n}{g_n} \quad (\text{B.5})$$

is well-defined.

Let a_i be positive integers. The finite continued fraction $[a_1, \dots, a_n]$ is defined as

$$a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \frac{1}{\ddots + \frac{1}{a_n}}}}. \quad (\text{B.6})$$

We have the following relation between the sequence of functions and continued fractions:

Theorem B.5. *The fraction $\frac{f_n}{g_n}$ can be written as continued fraction:*

$$\frac{f_n}{g_n} = [a_1, \dots, a_n]. \quad (\text{B.7})$$

It is still possible to express the continued fraction as a fraction of functions if we extend the continued fraction $[a_1, \dots, a_n]$:

Theorem B.6. *For $x \geq 1$, we have*

$$[a_1, \dots, a_n, x] = \frac{x f_n + f_{n-1}}{x g_n + g_{n-1}}. \quad (\text{B.8})$$

Let r be a real number, we define the sequence $(X_i)_i$ in the following way:

$$X_1 = r, \quad (\text{B.9a})$$

$$X_{n+1} = \frac{1}{(X_n - [X_n])}, \quad (\text{B.9b})$$

provided that X_n is not an integer. In that case X_n is the n th complete quotient of r . The simple continued fraction of r of order n is equal to:

$$r_n = [[X_1], [X_2], \dots, [X_n]] \quad (\text{B.10})$$

$$= [X_1] + \frac{1}{[X_2] + \frac{1}{[X_3] + \frac{1}{\ddots + \frac{1}{[X_n]}}}}. \quad (\text{B.11})$$

Theorem B.7. *Every real number r can be expressed uniquely as a simple continued fraction. Moreover, this simple continued fraction is finite if and only if r is a rational number.*

We have the following inequalities to indicate the quality of the approximation of an irrational number by a continued fraction.

Theorem B.8. *Let n be a positive integer, x a real number and $\frac{f_n}{g_n}$ the n th convergent of x . Then*

$$\frac{1}{g_n g_{n+2}} < \left| x - \frac{f_n}{g_n} \right| \leq \frac{1}{g_n g_{n+1}}. \quad (\text{B.12})$$

The following theorem states that all good rational approximations of an irrational number x are continued fractions of x .

Theorem B.9. *Let x be an irrational number and let $\frac{p}{q} \in \mathbb{Q}$, with $q > 0$. If*

$$\left| x - \frac{p}{q} \right| < \frac{1}{2q^2}, \quad (\text{B.13})$$

then $\frac{p}{q}$ is a convergent of x .

For quadratic relations between integers, continued fractions are particularly useful.

Theorem B.10. *Let A, B, x, y be positive integers, and let $C \neq 0$ be an integer, such that $C^2 < AB$ and AB not a square. If*

$$Ax^2 - By^2 = C, \quad (\text{B.14})$$

then $\frac{x}{y}$ is a convergent of $\sqrt{\frac{B}{A}}$.

Let P, Q, R be integers, such that R is positive and not a square and Q divides $P^2 - R$. Define the sequences $(P_i)_i, (Q_i)_i$ by

$$P_1 = P, \quad P_{n+1} = \left[\frac{P_n + \sqrt{R}}{Q_n} \right] Q_n - P_n, \quad (\text{B.15a})$$

$$Q_1 = Q, \quad Q_{n+1} = \frac{R - P_{n+1}^2}{Q_n}, \quad (\text{B.15b})$$

then we have the following theorem:

Theorem B.11. *The simple continued fraction of $\frac{(P+\sqrt{R})}{Q}$ is periodic after a certain point, and for n sufficiently large we have*

$$\sqrt{R} > P_n > 0, \quad (\text{B.16a})$$

$$2\sqrt{R} > Q_n > 0, \quad (\text{B.16b})$$

$$2\sqrt{R} > X_n > 1. \quad (\text{B.16c})$$

The theorem above states that all quadratic relations between integers have a continued fractions expansion that eventually becomes periodic. The following theorem indicates under which conditions this expansion is periodic from the beginning.

Theorem B.12. *The fraction $\frac{(P+\sqrt{R})}{Q}$ is purely periodic, that is*

$$\frac{(P + \sqrt{R})}{Q} = \left[a_1, \dots, a_k, \frac{(P + \sqrt{R})}{Q} \right], \quad (\text{B.17})$$

if and only if

$$\sqrt{R} + P > Q > \sqrt{R} - P > 0. \quad (\text{B.18})$$

Appendix C

Mathematics consists in proving
the most obvious thing in the
least obvious way.

GEORGE POLYA

Algebraic Number Theory

Let a, b be integers and let d be a positive square-free integer.
Let $\xi = a + b\sqrt{d}$, then $\bar{\xi} = a - b\sqrt{d}$ is called the conjugate of ξ .

Lemma C.1. *We have the following relations for ξ :*

$$\xi = \bar{\bar{\xi}}, \quad (\text{C.1a})$$

$$\overline{\xi + \eta} = \bar{\xi} + \bar{\eta}, \quad (\text{C.1b})$$

$$\overline{\xi\eta} = \bar{\xi} \cdot \bar{\eta}. \quad (\text{C.1c})$$

The solutions (a_i, b_i) of the Pell equation (4.3.1) can be characterized by the algebraic numbers $\xi_i = a_i + b_i\sqrt{d}$. We have the following relations between solutions of this type:

Proposition C.1. *If the algebraic numbers $\xi_i = a_i + b_i\sqrt{d}$ and $\xi_j = a_j + b_j\sqrt{d}$ are solutions of (4.3.1), then so are the numbers $\bar{\xi}_i$ and $\xi_i\xi_j$. In particular,*

$$\xi_i^n = (a_i + b_i\sqrt{d})^n \quad (\text{C.2})$$

is a solution of (4.3.1) for all integers n .

So from a given solution of (4.3.1), we can generate an infinite number of different solutions. It is natural to ask whether any solution of the Pell equation is necessarily of this form. The following theorem, first proved by Lagrange, confirms this.

Theorem C.1 (Lagrange, 1768). *Let $\xi_1 = a_1 + b_1\sqrt{d}$ be the smallest solution of (4.3.1), with $a_1, b_1 > 0$. Then for every positive solution (s, t) of (4.3.1) there exists a positive integer n , such that*

$$s + t\sqrt{d} = (a_1 + b_1\sqrt{d})^n. \quad (\text{C.3})$$

We call $\xi(d) = \xi_1 = a_1 + b_1\sqrt{d}$ the fundamental solution of (4.3.1).

Let

$$\mathbb{Q}[\sqrt{d}] = \left\{ r_1 + r_2\sqrt{d} \mid r_1, r_2 \in \mathbb{Q} \right\} \quad (\text{C.4})$$

be a quadratic number field. If

$$\alpha = a + b\sqrt{d} \in \mathbb{Q}[\sqrt{d}], \quad (\text{C.5})$$

then we have the following relations:

$$\bar{\alpha} = a - b\sqrt{d} \in \mathbb{Q}[\sqrt{d}], \quad (\text{C.6a})$$

$$\frac{1}{\alpha} = \frac{\bar{\alpha}}{\alpha \cdot \bar{\alpha}} \in \mathbb{Q}[\sqrt{d}]. \quad (\text{C.6b})$$

The real number $\xi \in \mathbb{Q}[\sqrt{d}]$ is an algebraic integer if there exists an integer n , such that

$$\xi^n + a_{n-1}\xi^{n-1} + \cdots + a_1\xi + a_0 = 0, \quad (\text{C.7})$$

where all a_i are integers. The set of all algebraic integers in $\mathbb{Q}[\sqrt{d}]$ is denoted by \mathcal{O} and is sometimes called the order of discriminant d .

Proposition C.2. *The algebraic integers of \mathbb{Q} are the integers.*

We have the following sufficient condition to verify whether a number is an algebraic integer:

Lemma C.2. *Let $\gamma_1, \dots, \gamma_l$ be complex numbers and let*

$$V = \left\{ \sum_{i=1}^l k_i \gamma_i, \quad k_i \in \mathbb{Z} \right\}. \quad (\text{C.8})$$

Suppose that $\alpha \in \mathbb{C}$ verifies $\alpha\gamma \in V$, for all elements $\gamma \in V$. Then α is an algebraic integer.

Proposition C.3. *If*

$$\alpha_1, \alpha_2 \in \mathcal{O} \cap \mathbb{Q}[\sqrt{d}], \quad (\text{C.9})$$

then

$$\alpha_1 + \alpha_2, \alpha_1\alpha_2 \in \mathbb{Q}[\sqrt{d}]. \quad (\text{C.10})$$

There is another way to verify whether an algebraic number is an algebraic integer:

Proposition C.4. *If $\xi = r + s\sqrt{d}$, then ξ is an algebraic integer if and only if $2r$ and $r^2 - s^2d$ are integers.*

Let $\alpha, \beta \in \mathbb{Q}[\sqrt{d}]$, we define the following rings:

$$\mathbb{Z}[\alpha] = \{m + n\alpha \mid m, n \in \mathbb{Z}\}, \quad (\text{C.11a})$$

$$\alpha\mathbb{Z} = \{n\alpha \mid n \in \mathbb{Z}\}, \quad (\text{C.11b})$$

$$\alpha\mathbb{Z} + \beta\mathbb{Z} = \{m\alpha + n\beta \mid m, n \in \mathbb{Z}\}. \quad (\text{C.11c})$$

The following relations follow more or less easily from these definitions:

$$\mathbb{Z}[\alpha] = \mathbb{Z} + \alpha\mathbb{Z}, \quad (\text{C.12a})$$

$$\alpha\mathbb{Z} = -\alpha\mathbb{Z}, \quad (\text{C.12b})$$

$$a\mathbb{Z} + \frac{b}{2}\mathbb{Z} = a\mathbb{Z} + \frac{b'}{2}\mathbb{Z}, \quad (\text{C.12c})$$

where a, b are integers that verify

$$b' \equiv b \pmod{2a}, \quad (\text{C.13a})$$

$$a\mathbb{Z} + b\mathbb{Z} = \text{GCD}(a, b)\mathbb{Z}. \quad (\text{C.13b})$$

Theorem C.2. *The set of algebraic integers \mathcal{O} of $\mathbb{Q}[\sqrt{d}]$ can be described as follows:*

$$\mathcal{O} = \{m + n\omega \mid m, n \in \mathbb{Z}\}, \quad (\text{C.14})$$

where

$$\omega = \begin{cases} \frac{-1+\sqrt{d}}{2}, & \text{if } d \equiv 1 \pmod{4}, \\ \sqrt{d}, & \text{if } d \equiv 2, 3 \pmod{4}. \end{cases} \quad (\text{C.15})$$

We obviously have that 1 and ω are linearly independent over \mathbb{Q} . Therefore \mathcal{O} is a two-dimensional \mathbb{Z} -module. Two algebraic integers $\alpha, \beta \in \mathcal{O}$ form an integral basis of \mathcal{O} if

$$\mathcal{O} = \{m\alpha + n\beta \mid m, n \in \mathbb{Z}\}. \quad (\text{C.16})$$

If we have an integral basis of the set of algebraic integers, then it is possible to define its discriminant.

Proposition C.5. *If the pair $\{\alpha, \beta\}$ forms an integral basis of \mathcal{O} , then*

$$D = \begin{vmatrix} \alpha & \beta \\ \bar{\alpha} & \bar{\beta} \end{vmatrix}^2 \quad (\text{C.17})$$

is a positive integer, independent of the choice of integral basis. The integer D is the discriminant of $\mathbb{Q}[\sqrt{d}]$. We have

$$D = \begin{cases} d, & \text{if } d \equiv 1 \pmod{4}, \\ 4d, & \text{if } d \equiv 2, 3 \pmod{4}. \end{cases} \quad (\text{C.18})$$

It is possible to give another description of the set of algebraic integers \mathcal{O} using this discriminant:

Proposition C.6. *Let D be the discriminant of $\mathbb{Q}[\sqrt{d}]$, then*

$$\mathcal{O} = \mathbb{Z} \left[\frac{D + \sqrt{D}}{2} \right]. \quad (\text{C.19})$$

An element $\xi \in \mathcal{O}$ is a unit if its inverse $\xi^{-1} \in \mathcal{O}$ is also an algebraic integer. Units of algebraic integers can be described in the following way:

Proposition C.7. *The number $\xi = x + y\sqrt{d} \in \mathcal{O}$ is a unit if and only if $2x \in \mathbb{Z}$ and $x^2 - dy^2 = \pm 1$.*

Proposition C.8. *If $a + b\sqrt{d} > 1$ is a unit, then $a, b > 0$.*

Amongst the units of \mathcal{O} , there is one unit that is special:

Theorem C.3. *Let ϵ_0 be the smallest unit in \mathcal{O} , such that $\epsilon_0 > 1$, then the set of units is given by*

$$\left\{ \pm \epsilon_0^k \mid k \in \mathbb{Z} \right\}, \quad (\text{C.20})$$

and ϵ_0 is the fundamental unit of \mathcal{O} .

Definition C.1. *The regulator of \mathcal{O} is $\log \epsilon_0$.*

Let A, B be subsets of \mathcal{O} of $\mathbb{Q}[\sqrt{d}]$, then

$$A \cdot B = \{a_1 b_1 + \cdots + a_n b_n \mid a_i \in A, b_i \in B, n \in \mathbb{N}\}. \quad (\text{C.21})$$

Definition C.2. *A subset I of \mathcal{O} is an integral ideal of \mathcal{O} if $I \cdot \mathcal{O} = I$, and if for $\alpha, \beta \in I$, we have*

$$m\alpha + n\beta \in I, \quad (\text{C.22})$$

for all integers m, n .

Definition C.3. *A subset I of $\mathbb{Q}[\sqrt{d}]$ is a fractional ideal of \mathcal{O} if $I \cdot \mathcal{O} = I$, and if for $\alpha, \beta \in I$ we have*

$$m\alpha + n\beta \in I, \quad (\text{C.23})$$

for all integers m, n .

Definition C.4. *If $\gamma \in \mathcal{O}$, then*

$$\gamma\mathcal{O} = \{\gamma\xi \mid \xi \in \mathcal{O}\} \quad (\text{C.24})$$

is an integral ideal. Ideals of this form are called principal ideals.

Proposition C.9. *We have the following equivalence on principal ideals:*

$$\begin{aligned}\alpha\mathcal{O} &= \beta\mathcal{O}, \\ &\Leftrightarrow \\ \alpha &= \beta\epsilon,\end{aligned}\tag{C.25}$$

with ϵ a unit in $\mathbb{Q}[\sqrt{d}]$.

Definition C.5. *The set of all principal fractional ideals is denoted by*

$$\mathcal{PI} = \left\{ \xi\mathcal{O} \mid \xi \in \mathbb{Q}[\sqrt{d}] \right\}.\tag{C.26}$$

Proposition C.10. *Every principal fractional ideal I is of the form*

$$\begin{aligned}I &= \alpha\mathbb{Z} + \beta\mathbb{Z} \\ &= \left\{ m_1\alpha + m_2\beta \mid m_1, m_2 \in \mathbb{Z}, \alpha, \beta \in \mathbb{Q}[\sqrt{d}] \right\},\end{aligned}\tag{C.27}$$

with α, β linearly independent over \mathbb{Q} .

Proposition C.11. *Let $\{\alpha, \beta\}$ be an integral basis of the fractional ideal I , then $\{\alpha', \beta'\}$ is another integral basis of I if and only if*

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = M \begin{pmatrix} \alpha' \\ \beta' \end{pmatrix},\tag{C.28}$$

where M is a 2×2 matrix with integer coefficients and determinant 1.

We define the norm of a fractional ideal I with integral basis $\{\alpha, \beta\}$ as

$$\mathcal{N}(I) = \frac{1}{\sqrt{D}} \left| \det \begin{pmatrix} \alpha & \beta \\ \bar{\alpha} & \bar{\beta} \end{pmatrix} \right|.\tag{C.29}$$

Proposition C.12. *The norm $\mathcal{N}(I)$ is independent of the choice of integral basis $\{\alpha, \beta\}$. If $I = \gamma\mathcal{O}$ is a principal integral ideal, then*

$$\mathcal{N}(I) = |\gamma\bar{\gamma}|.\tag{C.30}$$

Proposition C.13. *Every fractional ideal I has an integral basis $\{\alpha, \beta\}$, with $\mathbb{Q} \ni \alpha > 0$. Moreover, α is uniquely defined as the smallest positive rational number in I . If I is an integral ideal, then α is an integer.*

Proposition C.14. *An ideal $I \subseteq \mathbb{Q}[\sqrt{d}]$ is a fractional ideal if and only if there exists a positive integer m , such that mI is an integral ideal.*

Definition C.6. *For $a, b \in \mathbb{Z}$, $a \neq 0$, let $\tau(a, b)$ be the unique integer, such that*

$$\tau \equiv b \pmod{2a},\tag{C.31}$$

and with

$$-a < \tau \leq a, \quad \text{if } a > \sqrt{D},\tag{C.32a}$$

$$\sqrt{D} - 2a < \tau \leq \sqrt{D}, \quad \text{if } a < \sqrt{D}.\tag{C.32b}$$

Proposition C.15. *A subset $I \subseteq \mathbb{Q}[\sqrt{d}]$ is an integral ideal of \mathcal{O} if and only if we can write I as*

$$I = k \left(a\mathbb{Z} + \frac{b+\sqrt{D}}{2}\mathbb{Z} \right), \quad (\text{C.33})$$

where a, b, k are integers, with $a, k > 0$, and $b = \tau(a, b)$, and $4a|(b^2 - D)$. Moreover, I is uniquely represented by the triplet (a, b, k) : ak is the smallest rational number in I , $\frac{k}{2}$ is the smallest positive coefficient of \sqrt{D} of all elements of I , $b = \tau(a, b)$ is uniquely determined and $\mathcal{N}(I) = k^2 a$.

Using the previous two propositions we obtain a unique representation of a fractional ideal as

$$I = \frac{k}{l} \left(a\mathbb{Z} + \frac{b+\sqrt{D}}{2}\mathbb{Z} \right), \quad (\text{C.34})$$

with $l \in \mathbb{N}$ the smallest possible integer. This representation is called the standard form of I .

We can define a principal ideal $I = \gamma\mathcal{O}$ either by the algebraic number γ or by the parameters $a, b, k \in \mathbb{Z}$.

Proposition C.16. *Let x, y be integers and*

$$\alpha = \frac{x + y\sqrt{D}}{2} \in \mathcal{O}, \quad k = \text{GCD} \left(y, \frac{x+yD}{2} \right), \quad (\text{C.35})$$

and let u, v be integers, such that $uy + v(x + yD)/2 = k$, then

$$\alpha\mathcal{O} = k \left(a\mathbb{Z} + \frac{b+\sqrt{D}}{2}\mathbb{Z} \right), \quad (\text{C.36})$$

where

$$a = |\alpha\bar{\alpha}|, \quad b = \tau \left(a, \frac{(ux + \frac{v}{2}(x+yD))}{k} \right). \quad (\text{C.37})$$

Let $\alpha \in I$, where I is a fractional ideal. Consider the coordinates

$$\hat{\alpha} = (\alpha, \bar{\alpha}) \in \mathbb{R}^2. \quad (\text{C.38})$$

We say that α is a minimum of I , if $\alpha > 0$, and if there is no $\beta \in I$, $\beta \neq 0$, with $|\beta| < |\alpha|$ and $|\bar{\beta}| < |\bar{\alpha}|$. In other words, $\hat{\alpha}$ is in the first quadrant of \mathbb{R}^2 and the rectangle $(\pm\alpha, \pm\bar{\alpha})$ does not contain any element of I , except $(0, 0)$. A fractional ideal is called reduced, if $1 \in I$, and 1 is a minimum of I .

Proposition C.17. *If I is reduced, then it can be written in standard form as*

$$I = \mathbb{Z} + \frac{b + \sqrt{D}}{2a}\mathbb{Z}. \quad (\text{C.39})$$

Proposition C.18. *If the ideal*

$$I = \mathbb{Z} + \frac{b + \sqrt{D}}{2a} \mathbb{Z} \quad (\text{C.40})$$

is a reduced ideal in standard form, then $a, |b| < \sqrt{D}$. Therefore there are only finitely many reduced ideals.

Proposition C.19. *If a fractional ideal I can be written as*

$$I = \mathbb{Z} + \frac{b + \sqrt{D}}{2a} \mathbb{Z}, \quad (\text{C.41})$$

then I is reduced if and only if $b \geq 0$ and $b + \sqrt{D} > 2a$.

Corollary C.1. *The ideal*

$$I = \mathbb{Z} + \frac{b + \sqrt{D}}{2a} \mathbb{Z} \quad (\text{C.42})$$

is reduced if $a \leq \frac{\sqrt{D}}{2}$.

Let

$$I = \mathbb{Z} + \frac{b + \sqrt{D}}{2a} \mathbb{Z} \quad (\text{C.43})$$

be an ideal that is not necessarily reduced. Let $\gamma(I) = \frac{b + \sqrt{D}}{2a}$.

Definition C.7. *Let ρ be a mapping from principal ideals to principal ideals.*

$$\begin{aligned} \rho(I) &= \frac{1}{\gamma(I)} I \\ &= \mathbb{Z} + \frac{2a}{b + \sqrt{D}} \mathbb{Z}. \end{aligned} \quad (\text{C.44})$$

We can write this as

$$\rho(I) = \mathbb{Z} + \frac{b' + \sqrt{D}}{2a'} \mathbb{Z}, \quad (\text{C.45})$$

where $a' = \frac{|D-b^2|}{4a} = c$ and $b' = \tau(-b, c)$.

Proposition C.20. *Let*

$$I = \mathbb{Z} + \frac{b + \sqrt{D}}{2a} \mathbb{Z} \quad (\text{C.46})$$

be an ideal that is not necessarily reduced. Let $I_0 = I$ and

$$\begin{aligned} I_i &= \rho(I_{i-1}) \\ &= \mathbb{Z} + \frac{b_i + \sqrt{D}}{2a_i} \mathbb{Z}. \end{aligned} \quad (\text{C.47})$$

If I_i is not reduced, then $a_i < \frac{a_{i-1}}{2}$, and therefore there exist an

$$i \leq \left\lceil \log_2 \frac{a}{\sqrt{D}} \right\rceil + 1, \quad (\text{C.48})$$

such that I_i is reduced. Let i_{red} be the first such i . Then

$$\alpha = \prod_{j=1}^{i_{\text{red}}-1} \gamma(I_j) \quad (\text{C.49})$$

is a minimum in I and

$$I_{\text{red}} = I_{i_{\text{red}}} = \frac{1}{\alpha} I. \quad (\text{C.50})$$

Definition C.8. The right neighbour of a minimum α of the ideal I is the minimum $\beta_R \in I$, such that $\beta_R > \alpha$. The left neighbour of α is $\beta_L \in I$, such that $|\beta_L| > |\alpha|$.

Proposition C.21. Let $\alpha \in \mathbb{Q}[\sqrt{d}]$ and $\alpha > 0$. For every fractional ideal I , the map $I \mapsto \alpha I$ is a bijection that sends minima to minima and left and right neighbours to left and right neighbours.

Proposition C.22. If $I = \mathbb{Z} + \gamma(I)\mathbb{Z}$ is reduced, then we have the following properties:

- (i) $\gamma(I) > 1$ and $-1 < \overline{\gamma(I)} < 0$,
- (ii) $\gamma(I)$ is a minimum of I and $\rho(I)$ is reduced,
- (iii) $\gamma(I) \in I$ is a right neighbour of 1 in I .

We can write the set \mathcal{O} of algebraic integers of $\mathbb{Q}[\sqrt{d}]$ as

$$\begin{aligned} \mathcal{O} &= \mathbb{Z} + \frac{D + \sqrt{D}}{2} \mathbb{Z} \\ &= \mathbb{Z} + \frac{\tau(D, 2) + \sqrt{D}}{2} \mathbb{Z}, \end{aligned} \quad (\text{C.51})$$

therefore \mathcal{O} is a reduced principal ideal. Thus $\alpha_0 = 1 \in \mathcal{O}$ is a minimum. For integers i we say that α_{i-1} is the left and α_{i+1} is the right neighbour of a minimum $\alpha_i \in \mathcal{O}$.

$$\begin{aligned} J_i &= \frac{1}{\alpha_i} \mathcal{O} \\ &= \mathbb{Z} + \gamma_i \mathbb{Z}. \end{aligned} \quad (\text{C.52})$$

The real number $\frac{\alpha_{i+1}}{\alpha_i}$ is a right neighbour of 1 in J_i , and we have that:

$$\alpha_{i+1} = \gamma_i \alpha_i, \quad (\text{C.53a})$$

$$J_{i+1} = \rho(J_i). \quad (\text{C.53b})$$

Proposition C.23. For every integer i we have the following inequalities:

$$\frac{3D}{32} \leq \log\left(1 + \frac{3}{16D}\right) \leq \log \frac{\alpha_{i+1}}{\alpha_i} \leq \log \sqrt{D}, \quad (\text{C.54a})$$

$$\log 2 \leq \log \left(\frac{\alpha_{i+1}}{\alpha_{i-1}} \right). \quad (\text{C.54b})$$

Proposition C.24. The sequence $\{\alpha_i\}_i$ contains all minima of \mathcal{O} .

Theorem C.4 (Reduced Principal Ideals Cycle). We have the following properties of the reduced principal ideals cycle:

(i) The sequence $\{J_i\}_i$ is periodic with period $k_0 \in \mathbb{N}$. The repeating segment $\{J_0, \dots, J_{k_0-1}\}$ of reduced principal ideals is called the principal cycle.

(ii) Let $\epsilon = \frac{\alpha_{k_0}}{\alpha_0} = \alpha_{k_0}$, then $\epsilon = \epsilon_0$ is the fundamental unit of \mathcal{O} .

(iii) Let I be a reduced fractional principal ideal, then I is in the principal cycle.

Proposition C.25. We have the following inequalities:

$$\frac{2R}{\log D} \leq k_0 \leq \frac{2R}{\log 2}, \quad (\text{C.55})$$

where $R = \log \epsilon_0$ is the regulator of \mathcal{O} .

It is obvious that the map ρ is only invertible for the reduced ideals of the principal cycle.

Definition C.9. Let

$$\begin{aligned} I &= \mathbb{Z} + \frac{b+\sqrt{D}}{2a}\mathbb{Z} \\ &= \mathbb{Z} + \gamma\mathbb{Z} \end{aligned} \quad (\text{C.56})$$

be a reduced ideal. The conjugate ideal of I is defined as:

$$\begin{aligned} \sigma(I) &= \bar{I} \\ &= \mathbb{Z} + \frac{b-\sqrt{D}}{2a}\mathbb{Z} \\ &= \mathbb{Z} + \frac{\tau(a,-b)+\sqrt{D}}{2a}\mathbb{Z}. \end{aligned} \quad (\text{C.57})$$

Geometrically this can be seen as a reflection by the line $y = x$.

Lemma C.3. We have the following properties for conjugate ideals:

(i) I is reduced if and only if \bar{I} is reduced.

(ii) If α is a minimum of I , then $|\bar{\alpha}|$ is a minimum of \bar{I} .

(iii) If α is a right neighbour of a minimum β in I , then $|\bar{\alpha}|$ is a left neighbour of a minimum $|\bar{\beta}|$ in \bar{I} .

Proposition C.26. *The inverse of a reduced fractional principal ideal I is:*

$$\rho^{-1}(I) = \mathbb{Z} + \frac{b_* + \sqrt{D}}{2c_*} \mathbb{Z}, \quad (\text{C.58})$$

where $b_* = \tau(a, -b)$, and $c_* = \frac{D-b_*^2}{4a}$. We have

$$\rho^{-1}(I) = \sigma\rho\sigma(I). \quad (\text{C.59})$$

Definition C.10. *Let I_1, I_2 be fractional principal ideals of \mathcal{O} , such that*

$$I_1 = \gamma I_2, \quad (\text{C.60})$$

with $\gamma \in \mathbb{Q}[\sqrt{d}]$. The distance between the ideals I_1 and I_2 is defined as

$$\delta(I_1, I_2) = \log |\gamma| \pmod{R}, \quad (\text{C.61})$$

where R is the regulator. If $I_1 \neq \gamma I_2$ for some $\gamma \in \mathbb{Q}[\sqrt{d}]$, then the distance between I_1 and I_2 is undefined. We write $\delta(I)$ instead of $\delta(\mathcal{O}, I)$.

For the principal cycle we have

$$J_i = \frac{1}{\alpha_i} \mathcal{O}, \quad (\text{C.62a})$$

$$\delta(J_i) = \log \alpha_i, \quad (\text{C.62b})$$

$$\delta(J_i, J_k) = \log \frac{\alpha_k}{\alpha_i}. \quad (\text{C.62c})$$

Proposition C.27. *For every integer i we have the following inequalities:*

$$\frac{3}{32D} \leq \delta(J_i, \rho(J_i)) = \log \gamma_i \leq \log \sqrt{D}. \quad (\text{C.63})$$

Proposition C.28. *For every integer i we have the following inequality:*

$$\log 2 \leq \delta(J_i, \rho^2(J_i)). \quad (\text{C.64})$$

Proposition C.29. *Let*

$$I = \mathbb{Z} + \frac{b + \sqrt{D}}{2a} \mathbb{Z} \quad (\text{C.65})$$

be a fractional principal ideal that is not necessarily reduced. Place the ideals on the real line \mathbb{R} at positions that correspond to their distance to \mathcal{O} .

Let i_{red} be the smallest integer, such that

$$\begin{aligned} I_{\text{red}} &= \rho^{i_{\text{red}}}(I) \\ &= \frac{1}{\alpha} I \end{aligned} \quad (\text{C.66})$$

is reduced. Let J_k be the ideal in the principal cycle that is closest to I and that verifies $\bar{\alpha}_k \bar{\alpha} < 0$. Then I lies between J_{k-1} and J_{k+1} and I_{red} is one of the J_{k-1}, J_k, J_{k+1} , with

$$|\delta(I, I_{\text{red}})| < \log D, \quad (\text{C.67a})$$

$$\delta(I) < \delta(\rho^2(I_{\text{red}})). \quad (\text{C.67b})$$

The cardinality of the principal cycle is exponential in $\log D$ so to locate ideals by repeatedly applying ρ to \mathcal{O} can take exponentially long. Therefore a technique to jump ideals in the principal cycle is needed.

Definition C.11. Let I_1, I_2 be ideals, then $I_1 \cdot I_2$ is a vector space on \mathbb{Z} , with vectors

$$\{\alpha \cdot \beta \mid \alpha \in I_1, \beta \in I_2\}. \quad (\text{C.68})$$

The vector space $I_1 \cdot I_2$ is an ideal, and if $\{\alpha_1, \beta_1\}, \{\alpha_2, \beta_2\}$ are integral bases of I_1, I_2 , then

$$\{\alpha_1\beta_1, \alpha_1\beta_2, \alpha_2\beta_1, \alpha_2, \beta_2\} \quad (\text{C.69})$$

is an integral basis of $I_1 \cdot I_2$.

If $I_1 = \xi_1 \mathcal{O}, I_2 = \xi_2 \mathcal{O}$ are principal ideals, then

$$I_1 \cdot I_2 = \xi_1 \xi_2 \mathcal{O}. \quad (\text{C.70})$$

Proposition C.30. Let

$$I_i = a_i \mathbb{Z} + \frac{b_i + \sqrt{D}}{2} \mathbb{Z}, \quad (\text{C.71})$$

for $i \in \{1, 2\}$ be principal ideals. Let

$$k = \text{GCD}(a_1, a_2, \frac{b_1 + b_2}{2}). \quad (\text{C.72})$$

Let u, v, w be integers, such that

$$ua_1 + va_2 + w \frac{b_1 + b_2}{2} = k, \quad (\text{C.73})$$

then

$$\begin{aligned} I_3 &= I_1 \cdot I_2 \\ &= k \left(a_3 \mathbb{Z} + \frac{b_3 + \sqrt{D}}{2} \mathbb{Z} \right), \end{aligned} \quad (\text{C.74})$$

where

$$a_3 = \frac{a_1 a_2}{k^2}, \quad b_3 = \tau \left(a_3, \frac{ua_1 b_2 + va_2 b_1 + w \frac{b_1 b_2 + D}{2}}{k} \right). \quad (\text{C.75ab})$$

If we do not reduce modulo R , we have that $\delta(I_1 \cdot I_2) = \delta(I_1) + \delta(I_2)$. However it is not necessarily true that if I_1 and I_2 are reduced that $I_1 \cdot I_2$ is reduced as well. If

$$I = \frac{1}{a} \left(a\mathbb{Z} + \frac{b+\sqrt{D}}{2}\mathbb{Z} \right) \quad (\text{C.76})$$

is a reduced ideal, then

$$\begin{aligned} I_2 &= I^2 = I \cdot I \\ &= \frac{k'}{a'} \left(a'\mathbb{Z} + \frac{b'+\sqrt{D}}{2}\mathbb{Z} \right), \end{aligned} \quad (\text{C.77})$$

where

$$\begin{aligned} k' &= \text{GCD}(a, b) \\ &= ua + vb, \end{aligned} \quad (\text{C.78a})$$

$$a' = \frac{a^2}{(k')^2}, \quad b' = \tau \left(a', \frac{ua+vb}{k'} \frac{b^2+D}{2} \right). \quad (\text{C.78bc})$$

So we have that

$$I_2 = \frac{1}{k'} \left(\mathbb{Z} + \frac{b'+\sqrt{D}}{2a'}\mathbb{Z} \right), \quad (\text{C.79a})$$

$$\delta(I_2) = 2\delta(I). \quad (\text{C.79b})$$

The ideal I_2 is not necessarily reduced, but consider

$$\begin{aligned} I'_2 &= k'I_2 \\ &= \mathbb{Z} + \frac{b'+\sqrt{D}}{2a'}\mathbb{Z}. \end{aligned} \quad (\text{C.80})$$

We find that the distance between these ideals is

$$|\delta(I_2, I'_2)| = \log k' < \log \sqrt{D}. \quad (\text{C.81})$$

We can construct an ideal I''_2 from I'_2 by repeatedly applying ρ , until we have a reduced ideal. We have that

$$|\delta(I''_2, I'_2)| < \log D, \quad (\text{C.82})$$

and therefore that

$$|\delta(I''_2, I_2)| < \frac{3}{2} \log D. \quad (\text{C.83})$$

So if we apply ρ or ρ^{-1} , $2n$ times on I''_2 , where

$$n < \frac{3 \log D}{\log 4} = O(\log D), \quad (\text{C.84})$$

then we can localize the first element J_k of the principal cycle that verifies the condition $\delta(J_k) > 2\delta(I)$.

Definition C.12. For every ideal I of the principal cycle, we define the operator $*$ to be the operator that associates to I the element J_k of the principal cycle:

$$J_k = I * I. \quad (\text{C.85})$$

Proposition C.31. Let I be a reduced principal ideal. The ideal $I * I$ can be computed in $O(\text{polylog}D)$. Moreover, if we consider the sequence

$$\begin{aligned} I \longmapsto I * I = I^{(2)} \longmapsto \dots \longmapsto I^{(2^n)} \\ = I^{2^{n-1}} * I^{2^{n-1}}, \end{aligned} \quad (\text{C.86})$$

then the final ideal $I^{(2^n)}$ has distance

$$\delta(I^{(2^n)}) > 2^n \delta(I), \quad (\text{C.87})$$

and can be computed in $O(\text{polylog}D, n)$.

Definition C.13. Let I_1, I_2 be reduced ideals, with

$$I_i = \mathbb{Z} + \frac{b_i + \sqrt{D}}{a_i} \mathbb{Z}. \quad (\text{C.88})$$

The ideal $I_1 * I_2$ is the first element in the principal cycle, such that its distance exceeds $\delta(I_1) + \delta(I_2)$.

List of Symbols and Acronyms

$\vec{\mu}$	Magnetic moment
$\vec{\sigma}$	The Pauli spin operator
\vec{B}	Magnetic field
c	Speed of light in vacuum: $299\,792\,458\text{ ms}^{-1}$
\vec{J}	Angular momentum
γ	Gyromagnetic ratio, $\gamma_{\text{H}} = 267.513 \cdot 10^6\text{ rad s}^{-1}\text{ T}^{-1}$
\hbar	Reduced Planck constant: $1.055 \cdot 10^{-34}\text{ Js}$
\mathbb{C}	The complex numbers
\mathbb{N}	The positive integers
\mathbb{Q}	The rational numbers
\mathbb{R}	The real numbers
\mathbb{Z}	The integers
\mathcal{H}	The Hamiltonian of a physical system
ω	Angular frequency
τ_Q	Decoherence time for a physical realization
τ_{op}	Duration of an operation on one qubit
k_{B}	The Boltzmann constant: $1.381 \cdot 10^{-23}\text{ JK}^{-1}$
T	Temperature in Kelvin
T_1	Spin-lattice relaxation time
T_2	Transverse relaxation time
DFT	Discrete Fourier Transform

FID	Free Induction Decay
GCD	Greatest Common Divisor
LCM	Least Common Multiple
NMR	Nuclear Magnetic Resonance
QFT	Quantum Fourier Transform
RF	Radio Frequency
RMN	Résonance Magnétique Nucléaire
SMIS	Spectrométrie de Masse à Ionisation Secondaire
LLL	Short for Lenstra, Lenstra and Lovász
ode45	Matlab routine to solve ordinary differential equations

Bibliography

- [Ang95] W. S. Anglin. *The Queen of Mathematics, An Introduction to Number Theory*. Kluwer Academic Publishers, 1995.
- [Bak67] A. Baker. Linear forms in the logarithms of algebraic numbers III. *Mathematika*, 14:220–228, 1967.
- [BBG90] R. Bowtell, R. M. Bowley, and P. Glover. Multiple spin echoes in liquids in a high magnetic field. *Journal of Magnetic Resonance*, 88(3):643–651, July 1990.
- [BCJ⁺99] S. L. Braunstein, C. M. Caves, R. Jozsa, N. Linden, S. Popescu, and R. Schack. Separability of very noisy mixed states and implications for NMR quantum computing. *Phys. Rev. Lett.*, 83(5):1054–1057, 1999.
- [BD69] A. Baker and H. Davenport. The equations $3x^2 - 2 = y^2$ and $8x^2 - 7 = z^2$. *Q. J. Math. Oxford*, 20:129–137, 1969.
- [Ben73] C. H. Bennett. Logical reversibility of computation. *IBM Journal of Research and Development*, 17(6):525–532, 1973.
- [Ben82] C. H. Bennett. The Thermodynamics of Computation—A Review. *International Journal of Theoretical Physics*, 21(12):905–940, 1982.
- [BHP46] F. Bloch, W. W. Hansen, and M. Packard. Nuclear Induction. *Phys. Rev.*, 69:127, 1946.
- [BMS06] Y. Bugeaud, M. Mignotte, and S. Siksek. Classical and modular approaches to exponential Diophantine equations II: The Lebesgue-Nagell equation. *Composition Math.*, 142:31–62, 2006.
- [BW93] A. Baker and G. Wüstholz. Logarithmic forms and group varieties. *J. Reine Angew. Math.*, 442:19–62, 1993.
- [CM] Mihai Cipu and Maurice Mignotte. On the number of solutions of simultaneous Pell equations.
<http://hal.archives-ouvertes.fr/hal-00129723/fr>.

- [Coh96] Henri Cohen. *A Course in Computational Algebraic Number Theory*, volume 138 of *Graduate Text in Mathematics*. Springer-Verlag, Berlin and Heidelberg, 1996.
- [CP54] H. Y. Carr and E. M. Purcell. Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments. *Phys. Rev.*, 94(3):630–638, 1954.
- [CT65] James W. Cooley and John W. Tukey. An algorithm for the machine calculation of complex Fourier series. *Math. Comput.*, 19:297–301, 1965.
- [CTDL77] C. Cohen-Tannoudji, B. Diu, and F. Laloë. *Quantum Mechanics*. Wiley, New York, 1977.
- [CZ95] J. I. Cirac and P. Zoller. Quantum Computations with Cold Trapped Ions. *Phys. Rev. Lett.*, 74(20):4091–4094, 1995.
- [DBD79] G. Deville, M. Bernier, and J. M. Delrieux. NMR multiple echoes observed in solid ^3He . *Phys. Rev. B*, 19(11):5666–5688, June 1979.
- [Deu85] D. Deutsch. Quantum theory, the Church-Turing principle and the universal quantum computer. *Proceedings of the Royal Society of London A*, 400(1818):97–117, 1985.
- [DJ92] D. Deutsch and R. Jozsa. Rapid solutions of problems by quantum computation. *Proceedings of the Royal Society of London A*, 439(1907):553–558, 1992.
- [Fel71] N. I. Feldman. An effective refinement of the exponent in Liouville’s theorem. *Izv. Akad. Nauk*, 35:973–990, 1971.
- [Fey82] R. Feynman. Simulating physics with computers. *Int. J. of Theor. Phys.*, 21:467–488, 1982.
- [Gro97] Lov K. Grover. Quantum mechanics helps in searching a needle in a haystack. *Phys. Rev. Lett.*, 79(2):325–328, 1997.
- [Hah50] E. L. Hahn. Spin Echoes. *Phys. Rev.*, 80(4):580–594, 1950.
- [JVB95] J. Jeener, A. Vlassenbroek, and P. Broekaert. Unified derivation of the dipolar field and relaxation terms in the Bloch Redfield equations of liquid NMR. *J. Chem. Phys.*, 103(4):1309–1333, 1995.
- [Lan61] R. Landauer. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5:183–191, 1961.

- [Len02] H. W. Lenstra Jr. Solving the Pell Equation. *Notices of the AMS*, 49(2):182–192, 2002.
- [LLL82] A. K. Lenstra, H. W. Lenstra Jr., and L. Lovász. Factoring polynomials with rational coefficients. *Math. Ann.*, 261:515–534, 1982.
- [Llo93] S. Lloyd. A potentially realizable quantum computer. *Science*, 261:1569–1571, 1993.
- [Lom] Chris Lomont. The hidden subgroup problem - review and open problems. arXiv:quant-ph/0411037.
- [LP01] Noah Linden and Sandu Popescu. Good Dynamics versus Bad Kinematics: Is Entanglement Needed for Quantum Computation? *Phys. Rev. Lett.*, 87(4):047901, 2001.
- [LRVW96] S. Lee, W. Richter, S. Vathyam, and W. S. Warren. Quantum treatment of the effects of dipole-dipole interactions in liquid nuclear magnetic resonance. *J. Chem. Phys.*, 105(3):874–901, 1996.
- [Man80] Y. Manin. Computable and uncomputable. *Sovetskoye Radio*, 1980.
- [Mat00] E. M. Matveev. An explicit lower bound for a homogeneous rational linear form in logarithms of algebraic numbers II. *Izv. Ross. Akad. Nauk*, 64:1217–1269, 2000.
- [MG58] S. Meiboom and D. Gill. Modified Spin-Echo Method for Measuring Nuclear Relaxation Times. *Rev. Sci. Instrum.*, 29(8):688–691, 1958.
- [Mig04] M. Mignotte. A kit on linear forms in three logarithms, 2004. preprint, www-irma.u-strasbg.fr/~bugeaud/travaux/kit.ps.
- [NC00] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.
- [Pau24] W. Pauli. Zur Frage der theoretischen Deutung der Satelliten einiger Spektrallinien und ihrer Beeinflussung durch magnetische Felder. *Naturwiss.*, 12(37):741–743, 1924.
- [Pop75] R. P. Poplavskii. Thermodynamical models of information processing. *Uspekhi Fizicheskikh Nauk*, 115(3):465–501, 1975.
- [PTP46] E. M. Purcell, H. C. Torrey, and R. V. Pound. Resonance Absorption by Nuclear Magnetic Moments in a Solid. *Phys. Rev.*, 69:37–38, 1946.

- [Sak94] J. J. Sakurai. *Modern Quantum Mechanics*. Addison-Wesley, 1994.
- [Sha76] Derek Shaw. *Fourier Transform NMR Spectroscopy*. Elsevier, 1976.
- [Sho97] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. of Comput.*, 26(5):1474–1483, 1997.
- [Sli80] C. P. Slichter. *Principles of Magnetic Resonance*, volume 1 of *Solid-State Sciences*. Springer-Verlag, 1980.
- [Ste21] O. Stern. Ein Weg zur experimentellen Prüfung der Richtungsquantelung im Magnetfeld. *Zeitschrift für Physik A Hadrons and Nuclei*, 7:249–253, 1921.
- [vN66] John von Neumann. *Theory of Self-Reproducing Automata*. University of Illinois Press, 1966.
- [VSB⁺01] L. M. K. Vandersypen, M. Steffen, G. Breyta, C. S. Yannoni, M. H. Sherwood, and I. L. Chuang. Experimental realization of Shor’s quantum factoring algorithm using nuclear magnetic resonance. *Nature*, 414:883–887, 2001.
- [Wat95] John Watrous. On One-Dimensional Quantum Cellular Automata. In *In 36th Annual Symposium on Foundations of Computer Science*, pages 528–537. Society Press, 1995.
- [Yua02] P. Yuan. On the number of solutions of simultaneous Pell equations. *Acta Arithm.*, 101:215–221, 2002.
- [Yua04] P. Yuan. Simultaneous Pell equations. *Acta Arithm.*, 115:119–131, 2004.