Discrete algebra and geometry applied to the Pauli group and mutually unbiased bases in quantum information theory

Olivier Albouy

N° d'ordre 077-2009
LYCEN – T 2009-08

Thèse de l'Université de Lyon
(École doctorale de Physique et Astrophysique de Lyon)

présentée devant

l'Université Claude Bernard Lyon 1

pour l'obtention du

DIPLÔME DE DOCTORAT
Spécialité PHYSIQUE THÉORIQUE

(arrêté du 7 août 2006)

par

Olivier ALBOUY

Algèbre et géométrie discrètes appliquées au groupe de Pauli et aux bases décorrélées en théorie de l'information quantique

Soutenue publiquement le 12 juin 2009
devant la Commission d'Examen

Jury:	M.	H.	de Guise	
	M.	F.	Delduc	
	M.	F.	Gieres	Président du jury
	M.	M.	Kibler	Directeur de thèse
	M.	S.	Perrine	
	M.	M.	Planat	Rapporteur
	M.	M.	Saniga	Rapporteur

Acknowledgments

I wish to thank in the first place all the people who inspired this thesis and helped me to make it. My supervisor Maurice Kibler who provided me many incentives, comments and remarks over more than three years. Michel Planat and Metod Saniga whose papers let me know about the geometrical point of view on quantum information theory. Subhash Chaturvedi and his coworkers inspired the whole chapter on the finite phase space.

I thank Michel Planat and Metod Saniga for refereeing my thesis and for their encouragements. I also thank the whole board of examiners for their proofs of interest, in particular Hubert de Guise who also corrected a deal of English mistakes.

The Institut de Physique Nucléaire de Lyon[1] provided a pleasant work environment whose staff I thank, especially my thesis fellows for so many merry lunches and other cake times.

Last but not least, I thank the city of Lyon for the quality of life I found in it. *Adrien, quand nous reviendras-tu?*

[1]Institut de Physique Nucléaire de Lyon
Domaine scientifique de la Doua, Bât. Paul Dirac
4, rue Enrico Fermi - F69622 Villeurbanne cedex

Discrete algebra and geometry applied to the Pauli group and mutually unbiased bases in quantum information theory

Abstract

A maximal set of mutually unbiased bases (MUBs) in a d-dimensional Hilbert space is known to have cardinality $d+1$ whenever d is a power of a prime. For a rectangular dimension, this is only known to be an upper-bound, but the actual maximum is still an open issue. Pauli operators are among the many tools involved in trying to answer this question: For d a power of a prime, diagonalising particular, maximally commuting sets of them results in a complete set of MUBs. Moreover, in order to account for their commutation relations, they were involved into various, finite geometrical features. In particular, they were transcribed up to a global phase as vectors in the $\mathbb{Z}_d$-module $\mathbb{Z}_d^{2n}$. The basic blocks for their commutation relations are then projective points in $\mathbb{Z}_d^{2n}$.

In our thesis, we begin by giving a way to build MUBs by means of Gauss sums, in relation with a family of irreducible representations of the Lie algebra $\mathfrak{su}(2)$. We then study $\mathbf{P}(\mathbb{Z}_d^m)$, the projective structure derived from $\mathbb{Z}_d^m$. For $m=2n$ and endowed with a symplectic product, this symplectic, finite geometry appears as a general framework that encompasses the preceding features as its algebraic or combinatorial substructures. We also show with the help of $\mathbf{P}(\mathbb{Z}_d^2)$ that, to obtain complete sets of MUBs by means of Pauli operators, tensorial products of them are mandatory.

The maximally commuting sets of Pauli operators are accounted for by the Lagrangian submodules of $\mathbb{Z}_d^{2n}$ that we fully classify. The interest in this classification is twofold. On the one hand, it enables us to discriminate which maximally commuting sets of Pauli operators are likely to yield MUBs. These are accounted for by Lagrangian half-modules. We see them as the isotropic points of the projective line $(\mathbf{P}(\mathrm{Mat}(n,\mathbb{Z}_d)^2),\omega)$. We then establish an isomorphism between Pauli unbiased bases and distant Lagrangian half-modules, which precises by the way the correspondance between Gauss sums and MUBs. As a mathematical byproduct, we give an algorithm to perform symplectic diagonalisation. Dynamical considerations are also addressed with the Clifford group. On the other hand, the classification of Lagrangian submodules is readily applicable to the finite phase space over $\mathbb{Z}_d$, namely $\mathbb{Z}_d^2$. We thus answer a technical point in the current problem of setting-up of discrete Wigner distributions over that phase space.

Finally, we turn our attention from discrete to continuous algebra. We present various tools inspired by the previous ones and confront them with classical quantum information objects. Thus we deal with cross-ratio on the Bloch sphere and projective geometry in higher dimension, Pauli operators with continuous exponents and we compare von Neumann entropy with a determinantal measure of entanglement.

Résumé

Il est connu qu'un ensemble maximal de bases décorrélées dans un espace de Hilbert de dimension d compte $d+1$ bases lorsque d est une puissance d'un premier. Mais en dimension rectangle, $d+1$ n'est plus qu'une borne supérieure dont on ne sait pas si elle est atteinte. Parmi les nombreux outils mis en œuvre pour traiter cette question figurent les opérateurs de Pauli : pour d puissance d'un premier, on obtient des bases décorrélées en diagonalisant des ensembles maximalement commutant de ces opérateurs. Leurs relations de commutation ont dès lors donné lieu à des études variées en géométrie finie, en particulier après qu'ils ont été identifiés, à une phase globale près, à des vecteurs de $\mathbb{Z}_d^{2n}$. L'étude porte alors essentiellement sur les points projectifs de ce $\mathbb{Z}_d$-module.

Dans ce mémoire, nous commençons par donner une construction de bases décorrélées en lien avec une famille de représentations irréductibles de l'algèbre de Lie $\mathfrak{su}(2)$ et faisant appel aux sommes de Gauss. Puis nous étudions $\mathbf{P}(\mathbb{Z}_d^m)$, la structure projective déduite de $\mathbb{Z}_d^m$. Pour $m = 2n$ et en munissant la structure d'un produit symplectique, les études précédentes apparaissent comme des propriétés combinatoires dans les sous-structures de cette géométrie finie symplectique. Nous montrons aussi avec $\mathbf{P}(\mathbb{Z}_d^2)$ que, pour obtenir des ensembles complets de bases décorrélées au moyen d'opérateurs de Pauli, il est nécessaire de considérer des produits tensoriels de ces opérateurs.

Les sous-modules lagrangiens de $\mathbb{Z}_d^{2n}$, dont nous donnons une classification complète, rendent compte des ensembles maximalement commutant d'opérateurs de Pauli. Cette classification présente un double intérêt. D'une part, elle permet de savoir lesquels de ces ensembles sont susceptibles de donner des bases décorrélées : ils correspondent aux demi-modules lagrangiens, qui s'interprètent encore comme les points isotropes de la droite projective $(\mathbf{P}(\mathrm{Mat}(n, \mathbb{Z}_d)^2), \omega)$. Nous explicitons alors un isomorphisme entre les bases décorrélées ainsi obtenues et les demi-modules lagrangiens distants, ce qui précise aussi la correspondance entre sommes de Gauss et bases décorrélées. Comme complément à cette étude, nous donnons un algorithme de diagonalisation symplectique. Nous traitons également les aspects dynamiques avec le groupe de Clifford. D'autre part, la classification des sous-modules lagrangiens s'adapte aussitôt à $\mathbb{Z}_d^2$, l'espace des phases discret sur $\mathbb{Z}_d$. Nous résolvons ainsi un point technique dans l'élaboration encore inachevée de fonctions de Wigner discrètes sur cet espace.

Enfin, nous quittons l'algèbre discrète pour la continue et présentons quelques outils inspirés des précédents, avant de les confronter aux objets classiques de l'information quantique. Nous traitons ainsi du rapport anharmonique sur la sphère de Bloch, de géométrie projective en dimension supérieure, des opérateurs de Pauli continus et nous comparons l'entropie de von Neumann à une mesure de l'intrication par calcul d'un déterminant.

Contents

List of Figures

Chapter 1

Introduction

An introduction classically consists of a brief historical account of the domain followed by a presentation of the whole work. Without pretention to be exhaustive, we resolve to enlarge the historical part to give a self-contained introduction to the field. It is only aimed at giving to the nonspecialist a flavour of the background material of our thesis. The last section about theoretical tools will naturally drive us to our own matter.

1.1 Quantum physics is needed for the information culture

It has become commonplace today to talk about the *information age*. Information is spread and processed ever faster. More and more people have access to a personnal computer and to the Internet and in fact are dependent on them for their everyday work. We even pay less attention to the computer itself than to the softwares we can run on it.

It was about sixty or seventy years ago that this age arose with both theoretical and technical breakthroughs. The mathematician Alan Turing launched modern computer science in a 1936 article by defining rigorously what is meant by an algorithmic process in full generality, at a time when the few available calculating machines were still physically designed for specific tasks. This construction was called a Turing machine and Turing also showed that all these machines can be simulated by a single universal one, the Universal Turing Machine. Today, even with slight modifications, it still encompasses our conception of an algorithm.

Electrical hardwares were soon designed to embody the conceptual construction of Turing, after a model by John von Neumann. But these systems came to a satisfying practical use only as John Bardeen, Walter Brattain and Will Shockley developed the transitor, in fact the first major supply of quantum physics to the so-called information age. Since then, progress in our computer capacities has been closely related to our ability in mastering the realisation of that component, making it always smaller. In 1965, Gordon Moore encoded this progress in a celebrated law dubbed after his name. Moore's law states that computer power will double approximately every two years. This law is so intimately linked to the story of the miniaturisation of transitor that it is commonly restated as the doubling in the number of transitors on a silicon chip. So far, it has been checked in this latter form. But what as we now come to the atomic scale?

At the atomic level, the effects of quantum physics become unavoidable and maybe they ought not to be avoided. Instead, if we want to keep on progressing, we should take advantage of them. The way the computer we use deal with information is typically based on our common life intuition. While transitor is quantum, it is only used as a concrete device in order to realise a circuit conceived beforehand with a classical way of thinking. The trick is to take into account the quantum laws of nature to elaborate not only more efficient hardwares but also more efficient algorithms and softwares. Before we take deeper insights in this latter idea, we have a look at the relevant principles of quantum physics.

1.2 Quantum information fundamental features

At the beginning of the twentieth century, an increasing number of discrepancies between theory and experience arose in the physics of atoms and light. Independently of the conceptual revolution brought about by general relativity, classical physics was progressively and laboriously amended in order to solve them. This endeavour resulted in a new theory of elementary particles and of their interactions. But it soon appeared that it was much more than that. It provided a completely new framework for physics as a whole, namely quantum physics. We stress that quantum physics is not quite a theory by itself, but rather principles and guidelines for building up new theories. Thus, if at first glance the standard model of particles involved in large accelerators and the paradigms we seek after to improve computing powers seem disconnected, they obey the same fundamental rules. We are going to have a look at those rules from the point of view of quantum information.

Superposition

In our classical conception of encoding and processing information, we use bits, that is to say variables that can take either one of only two values, 0 or 1. A message is built up of a series of 0's and 1's. Let us denote these two states $|0\rangle$ and $|1\rangle$ for our purpose. Then in the quantum realm, a bit as the unusual property that it can be $|0\rangle$, $|1\rangle$ or any *superposition* of these two states:

$$|\psi\rangle = c_0 |0\rangle + c_1 |1\rangle \,. \tag{1.1}$$

In this formula, $|\psi\rangle$ is a possible state for the bit, described as a weighted sum of the two fondamental states, with the coefficients c_0 and c_1 being complex numbers not both 0. The states $|0\rangle$ and $|1\rangle$ form the computational basis and $|\psi\rangle$ is dubbed a qubit, a short for quantum bit. Then to the classical series of bits corresponds one multiple qubit. If for instance our message is coded over two bits, the four computational states are

$$|00\rangle\,, |01\rangle\,, |10\rangle\,, |11\rangle \tag{1.2}$$

and they give rise to the quantum states of the type

$$|\psi\rangle = c_{00}\,|00\rangle + c_{01}\,|01\rangle + c_{10}\,|10\rangle + c_{11}\,|11\rangle\,, \tag{1.3}$$

where the c's are again complex numbers not all of them 0. This state $|\psi\rangle$ is made up of two qubits. In each term, the first digit relates to the first qubit and the second one to the second qubit. Then, what does superposition mean experimentally?

Projective measurements and projective structure

We will concentrate only on the case of a single qubit as given in (1.1). If we want to test the value of $|\psi\rangle$, we shall get either 0 or 1 as in the classical case. But in fact the value 0 is to be obtained with probability

$$p_0 = \frac{|c_0|^2}{|c_0|^2 + |c_1|^2} \tag{1.4}$$

and similarly the value 1 is to be obtained with probability p_1 as in the preceding formula after inverting c_0 and c_1. Of course we have $p_0 + p_1 = 1$. It means that if we prepare a series of qubits all of them in the same state $|\psi\rangle$ and measure on them in exactly the same way for each one in order to get 0 or 1, these values shall be obtained in proportions given by p_0 and p_1. The measurement is in an essential way probabilistic. After measurement, the qubit is no more in a superposition of the computational states, but in either state $|0\rangle$ or $|1\rangle$, in accordance with the result we got. One says that $|\psi\rangle$ was projected onto $|0\rangle$ or $|1\rangle$ and that a measurement in quantum physics is also of projective nature.

One may notice with this simple example that if we multiply $|\psi\rangle$ by any nonzero complex number λ to get

$$\lambda\,|\psi\rangle = \lambda c_0\,|0\rangle + \lambda c_1\,|1\rangle\,, \tag{1.5}$$

the probabilities p_0 and p_1 when we perform a measurement on $\lambda\,|\psi\rangle$ would remain the same. In fact, no use was ever found since the inception of quantum physics in distinguishing $\lambda\,|\psi\rangle$ from $|\psi\rangle$ and they are thus identified:

$$\text{state } \lambda\,|\psi\rangle = \text{state } \,|\psi\rangle\,, \quad \forall\lambda \neq 0. \tag{1.6}$$

One says that the set of quantum states has a projective structure[1]. So one may conveniently suppose that $|c_0|^2 + |c_1|^2 = 1$. The state $|\psi\rangle$ is said to be normalised

[1] This notion of projectivity has historically no bearing on the previous one.

then and the probabilities reduce to

$$p_0 = |c_0|^2 \quad \text{and} \quad p_1 = |c_1|^2 \,. \tag{1.7}$$

Entanglement

An entangled state involves at least two subsystems, as for instance a couple of photons or of electrons arising from the decay of a nucleus. The subsystems are described jointly by a single state. If the property we are interested in is the spin that may be up or down with respect to some measurement apparatus, we may code arbitrarily code for down with a 0 and for up with a 1. If there are actually two subsystems, then the state is of the form $|\psi\rangle$ indicated in (1.3). One says that the subsystems are entangled or that they form an entangled state, if the dynamics underwent by one of them may have some influence as such on the dynamics of the other. In mathematical words, $|\psi\rangle$ cannot be written as a single tensor product. Let us give a telling example. The most famous examples of entangled states are the Bell states for two qubits:

$$|\beta_{00}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \tag{1.8a}$$

$$|\beta_{01}\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \tag{1.8b}$$

$$|\beta_{10}\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle), \tag{1.8c}$$

$$|\beta_{11}\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle). \tag{1.8d}$$

If one gets the output 0 after measuring on the first qubit in the state $|\beta_{00}\rangle$, then one will also get the output 0 when measuring on the second qubit. Contrary to appearances at that stage, this is not a classical correlation between the states of each of the two qubits. In fact, none of the qubits has a state defined independently of the other qubit. To see this, we take say four polarisors in order to measure the polarisation of two coupled photons in the state $|\beta_{00}\rangle$, two polarisors for one of the photons and the two other ones for the other photon. For each photon, we make use at random of one the two polarisors dedicated to it and we repeat this experience a great number of times with other couples of photons prepared in the same state. We denote A, B the two random variables corresponding to the polarisors of the first photon and C, D the other two ones. Then the function

$$F = AC + AD + BC - BD \tag{1.9}$$

has always value ± 2. So, if the superposition $|\beta_{00}\rangle$ accounted for our ignorance of the photons having well-defined, separated states, then the mean value of F,

$$\langle F\rangle = \langle AC\rangle + \langle AD\rangle + \langle BC\rangle - \langle BD\rangle \tag{1.10}$$

should lie in the range $[-2; 2]$. But in 1982, when Alain Aspect and coworkers realised the experience [1], they found that

$$\langle F \rangle = 2\sqrt{2}, \tag{1.11}$$

in agreement with the prediction of quantum physics. This proves that entangled states do account for physical states.

No-clonable information

Finally, a fourth major difference between classical and quantum informations is the ability to copy information. In our classical computers, any bit of information is stored on a magnetic cell and is readable without distroying it. On the contrary, as we saw above about measurements, any attempt to read some quantum bit of information, that is to say to measure it in order to get some output, irremediably modifies it, without letting know its initial state. So, it is impossible to copy, or clone, any qubit. This result is known as the no-cloning theorem.

This feature can be used to design secure communication protocols, such as the BB84 protocol, elaborated in 1984 by Bennett and Brassard [2], by transmitting sequences of qubits instead of bits. Without describing such a protocol in details, we give the idea and connect it with the no-cloning theorem. The two communicating parties are traditionally called Allice and Bob. Alice sends some coded information to Bob through a communication channel. If Alice sent classical bits, an eavesdropper on the communication channel could learn about the state of the bits and send them forward on the channel to Bob. A way to warrant an irremediable, detectable disturbance in the message due to the presence of the eavesdropper is to relie on the probabilistic character of any quantum measurement. Indeed, the protocol provides Alice and Bob with a way to know if, on average, the qubits used in communicating were disturbed.

1.3 Quantum algorithms and protocols

Can the main features of quantum physics, namely state superposition and entanglement be exploited in order to design specifically quantum algorithms? The point in doing so is to enable one to perform impossible tasks with a classical support of information or deal in reasonable time with classically tedious tasks. The answer happens to be yes and this was the main incentive for quantum information in the 1980's. Whereas to date, only a few quantum protocols and algorithms have been found, some of them are of major importance whenever they can be realised.

As we have already seen, secure communications can be based on quantum information. In 1985, Deutsch designed the first example of quantum algorithm [3]. It was a toy algorithm in order to know whether a function

$$f : \{0, 1\} \longrightarrow \{0, 1\} \tag{1.12}$$

is constant ($f(0) = f(1)$) or balanced ($f(0) \neq f(1)$). Classically one has to perform two calculations in order to answer the question: $f(0)$ and $f(1)$. With quantum ressources, its is possible to evaluate the following logical sum, that is to say modulo 2, at once:

$$f(0) \oplus f(1), \tag{1.13}$$

which gives the answer after a single calculation. But the next two algorithms have more serious applications. In 1994, Shor showed by means of an explicit algorithm [4], that with the help of a quantum computer with sufficiently many ressources, one can factorise large numbers into their prime factors in minute's time, or to use the language of complexity theory, in polynomial time. Thus, if it could be realised, it would break the RSA code on which most of our secure communications are grounded today, for example credit card transactions. In 1995, Grover proposed an algorithm [5] to search an item within an unsorted list of N items with only $O(\sqrt{N})$ requests instead of $O(N)$ as classically. Though the gain in time is less impressive than in Shor's algorithm, the constant call on search procedures in our computing practices shows the interest in this speed-up.

1.4 Theoretical tools

Although few quantum algorithms are available till now, they exhibit a rather constant mathematical feature. As shown for instance by the BB84 protocol and its generalisations for secure communication or on the contrary by Shor's algorithm, that would highly unsecure many transactions, they often rely on the discrete Fourier transform (DFT). This core transformation is only a particular case of another major topic in quantum theory, namely mutually unbiased bases (MUBs). The characteristic of such a set of bases is that a state picked out of one of them has equal amplitude over the states of any other one in the set. In other words, a state in one of the bases is a kind of optimal superposition or mix of the states in another one. Two orthonormalised bases of a d-dimensional Hilbert space

$$A = \{|A, \alpha\rangle\}, \qquad B = \{|B, \beta\rangle\} \tag{1.14}$$

are MUBs iff they check the equalities

$$\forall \alpha, \beta, \quad |\langle A, \alpha | B, \beta \rangle| = \frac{1}{\sqrt{d}}. \tag{1.15}$$

In the matrix representing the DFT, every entry has the same modulus. Thus the basis one gets by means of the DFT is unbiased with the computational basis.

Besides Fourier transform, the notion of MUBs is widespread both in classical and quantum information theory. Schwinger unveiled them as soon as 1960 in a paper about quantum complementarity and unitary operators, but he did not name them [6]. They appear in quantum tomography [7] and in quantum games such as the Mean King problem [8–10]. As to classical information theory, one finds them in the study of Kerdock codes [11] and spherical codes [12] or in the developpement of network communication protocols [13, 14]. Recent investigations upon Feynmann path integrals also give MUBs a central role.

Since the beginning of their study in the 80's [7] [15], we know that a set of MUBs in a d-dimensional Hilbert space contains at most $d+1$ of them and that this upper-bound can be reached if d is power of a prime. But whenever d is a composite integer and despite an extensive range of mathematics involved, no conclusive information is available about the achievement of the upper-bound. As a nonexhaustive list of the mathematical tools that have been used, let us cite Galois fields and rings in relation with Gaussian sums [7, 16, 17], combinatorics, latin squares [9], unitary operator bases [6, 18], discrete phase space [19–23] and Wigner functions [10, 24, 25], DFT [26, 27], finite ring geometry [28–30] and also $\mathrm{SU}(n)$ Lie groups and their corresponding Lie algebras [31–33] with connection to positive operator-valued measures (POVMs) [34]. Throughout their story, MUBs have also been approached by a number of numerical tests. We refer to [35, 36] and references therein. All the studies concerned with dimension 6 converge to the fact that there are no more than 3 MUBs in $\mathbb{C}^6$.

Of particular relevance to our thesis are, on the one hand, the works by Bandyopadhyay *et al.* [18] and on finite phase space cited above and, on the other hand, the strand of finite geometry over rings. For d a power of a prime, Bandyopadhyay *et al.* give a recipe to find maximally commuting sets of Pauli operators that will in turn provide MUBs, but with no exhaustive correspondance between such sets and MUBs. Then studies of finite phase space enrich this construction by putting it into a geometrical framework that extracts the essential about Pauli operators, namely their commutation relations. As to finite geometry, let us cite only the two papers by Planat and Saniga [37] and jointly with Kibler [38]. These latter papers acted on us as incentives to supply a general framework encompassing the features therein together with the previous ones. The notions of neighbourhood and distance will particularly retain our attention. An extra source of inspiration that also ended in the same framework is our joint work [33] with our thesis supervisor, about MUBs and irreducible representations of the Lie algebra $\mathfrak{su}(2)$.

Thus, we begin by presenting in Chapter 2 a simplified version of that latter work. We explain how MUBs appear as the diagonalising bases of operators pertaining to irreducible representations of the Lie algebra $\mathfrak{su}(2)$, with Gauss sums as a central calculational feature. We also say why all this will finally be reduced to the same finite geometry as Pauli operators. In fact, Pauli operators prove having a richer structure than the previous ones.

In Chapter 3, we first recall the definition and main features of the Pauli group in relation with MUBs. In particular, it is explained how the problem of MUBs within the scope of Pauli operators is completely translated in the language of projective and symplectic geometry over $\mathbb{Z}_d$. Hence, we carry on in the strand of finite geometry by studying the $\mathbb{Z}_d$-module $\mathbb{Z}_d^m$ and its derived projective structure $\mathbf{P}(\mathbb{Z}_d^m)$. We generalise the notions of neighbourhood and distance in arbitrary dimension and we introduce the wedge product in order to quantify neighbourness. Neighbourhood classes will also turn up to be group orbits. However, we concentrate on picking out the evidence to show that, to get a complete set of MUBs by means of Pauli operators, tensorial products of them are mandatory. This result derives from the structure of the projective line $\mathbf{P}(\mathbb{Z}_d^2)$, for which we provide the necessary counting properties, together with illustrating diagrams.

In Chapter 4, we apply the latter features to the complete classification of maximally commuting sets of Pauli operators. This is achieved through the classification of the Lagrangian submodules of $\mathbb{Z}_d^{2n}$. From a technical point of view, it consists in diagonalising specific matrices over $\mathbb{Z}_d$ when only symplectic changes of computational basis are allowed. In order to know which maximally commuting sets of operators are likely to yield MUBs, we need a last step. In Chapter 5, we single out Lagrangian half-modules as the only suitable candidates and we find a necessary and sufficient condition for two of them to actually yield MUBs. We also give a geometric interpretation of this criterion: We see Lagrangian half-modules as the isotropic points of the projective line over $\mathrm{Mat}(n, \mathbb{Z}_d)$, the set of $n \times n$ matrices over $\mathbb{Z}_d$, and we establish an isomorphism between the unbiased bases thus obtained and distant Lagrangian half-modules, thus completing the study initiated in Bandyopadhyay *et al.*'s paper. We go on in the same chapter with an alternative graph interpretation of the same result and we finally relate to quantum computation with the Clifford group. We show how in some sense, our discrete, symplectic geometry accounts for the symplectic part of the Clifford operators, as separated from their Pauli part.

The classification of Lagrangian submodules is readily applicable to the finite phase space over $\mathbb{Z}_d$, namely $\mathbb{Z}_d^2$. In Chapter 6, we thus answer a technical point in the still in progress setting-up of discrete Wigner distributions over that phase space. We show that the isotropic lines of $\mathbb{Z}_d^2$ are nothing but its Lagrangian submodules. This identification enables us to count them under various conditions and in particular to describe their orbits under the action of the symplectic group.

Besides the main stream we have just described, we give another mathematical byproduct. The classification of Lagrangian submodules is clearly a part of the issue of symplectic diagonalisation, without any assumption on the submodules or their representing matrix. We complete this study for its own sake in Section 4.3, which emphasizes how particular Lagrangian submodules are.

In the last chapter, Chapter 7, we turn our attention from discrete to continuous algebra. We present various tools inspired by the previous ones and confront them with classical quantum information objects. In the first three sections, we try to find some notions and objects in order to account for any set of mutually unbiased bases, not only those one can reach with the help of discrete Pauli operators as above. We first answer this question for qubits as points on the Bloch sphere, which is nothing but the projective line over $\mathbb{C}$: In Section 7.1, we put forward the cross-ratio and harmonic conjugated points, and in Section 7.2, we try Pauli operators with continuous exponents. In Section 7.3, we set out a general tool pertaining to algebraic geometry, that accounts for every single vector unbiased with respect to a given basis. However, if we are able to work out the cases $d = 2$ and $d = 3$, it is seen to be clearly more difficult to handle in general. Finally in Section 7.4, we generalise the idea behind the symplectic product as we explore a determinantal measure of entanglement for pure states. We compare it with the classical von Neumann entropy and Schmidt decomposition.

We put in five appendices all the discrete mathematics we need. Appendix A sets the elementary arithmetics in the ring $\mathbb{Z}_d$ of integers modulo d. The notions of gcd, lcm, order of an element and above all the Chinese remainder theorem are stated there. Then, vector spaces have a base *field*. Appendix B presents the analogue of finite dimensional spaces over the base *ring* $\mathbb{Z}_d$. Appendix C is about reduction of vectors and matrices over $\mathbb{Z}_d$. In particular, we give an original proof of the existence of a Smith normal form (or diagonalisation) of a matrix over $\mathbb{Z}_d$. The traditional proof, as given in [39] or [40] for instance, is grounded exclusively upon algebraic properties involving the notion of ideal. In the particular case of the ring $\mathbb{Z}_d$, we can make it rely upon counting properties in $\mathbb{Z}_d$-modules. The developments contained in that appendix are not readily used in the body of the thesis. So, with the exception of Gauss algorithm stated at the end of the appendix, we will recall any results in it when required in the main text. Appendix D is about the wedge product and its particularities when considered over a ring. Though its material is simpler than in the previous appendix, it will be treated quite in the same way. Finally, Appendix E is a brief account of projective geometry over fields and ring. It also contains a criterion to know whether a submodule of $\mathbb{Z}_d^m$ is the join of projective points.

For any details about the quantum physics or quantum communication and quantum information, the reader is refered to classical treatises, such as [41–44].

Our thesis work was partially published in three papers:

- O. Albouy and M. Kibler, SU_2 Nonstandard bases: Case of Mutually Unbiased Bases. SIGMA 3 (2007), 076
- O. Albouy and M. Kibler, A unified approach to SIC-POVMs and MUBs, *Journal of Russian Laser Research, Volume 28, Number 5, 2007*
- O. Albouy, The isotropic lines of $\mathbb{Z}_d^2$, *J. Phys. A: Math. Theor.* 42 (2009) 072001
- O. Albouy, Determinantal measure for pure states entanglement, submitted to *J. Phys. A: Math. Theor.*

Chapter 2

Operator algebras and Gauss sums

Gauss sums have been among the first tools for building MUBs as they appeared in 1989 in a paper by Wootters and Fields [7]. After several investigations, they were recently connected for this purpose with irreducible representations of the Lie algebra $\mathfrak{su}(2)$ of the group SU(2), a scheme that was partly brought into form in a joint article with our thesis supervisor [33]. Though it is not our intention to work with $\mathfrak{su}(2)$ representation in this thesis, we will give a very brief idea of the background that led us to Gauss sums and MUBs. For further details, the reader is refered to the latter paper. The arithmetical properties of Gauss sums we will come across in this chapter are a particuliar incarnation of some commutation relations. Those relations and their consequences will be studied through Pauli operators in the following chapters.

2.1 Out of two quon algebras

Let us define two quon algebras $A_i = \{a_{i-}, a_{i+}, N_i\}$, $i = 1, 2$, by

$$a_{i-}a_{i+} - qa_{i+}a_{i-} = 1, \qquad [N_i, a_{i\pm}] = \pm a_{i\pm}, \qquad N_i^\dagger = N_i, \qquad (a_{i\pm})^k = 0, \tag{2.1a}$$

$$\forall x_1 \in A_1,\ \forall x_2 \in A_2, \quad [x_1, x_2] = 0, \tag{2.1b}$$

where[1]

$$q = \exp\left(\frac{2\pi i}{k}\right), \quad k \in \mathbb{N} \setminus \{0, 1\}. \tag{2.2}$$

The generators $a_{i\pm}$ and N_i of A_i are linear operators. As in the classical case where $q = 1$, we say that a_{i+} is a creation operator, a_{i-} an annihilation operator and N_i a number operator. The case $k = 2$ corresponds to fermion operators and the case $k \to \infty$ to boson operators. In other words, each of the algebras A_i describes fermions for $q = -1$ and bosons for $q = 1$. The nilpotency conditions $(a_{i\pm})^k = 0$

[1] In any expression of the form q^x, q has to be formally replaced by its expression, so as to read $\exp(2i\pi x/k)$. This is essential as we are to consider fractional values for x.

can be understood as describing a generalised exclusion principle for particles of fractional spin $1/k$ (the Pauli exclusion principle corresponds to $k = 2$).

Let $\mathcal{F}(i)$, $i = 1, 2$, be two truncated Fock–Hilbert spaces of dimension k corresponding to two truncated harmonic oscillators. We endow each space $\mathcal{F}(i)$ with an orthonormalised basis $\{|n_i);\ n_i = 0, 1, \ldots, k-1\}$. With the notation

$$[x]_q = \frac{1-q^x}{1-q}, \qquad x \in \mathbb{R}, \tag{2.3}$$

we can give the following representations among many others of the algebras A_1 and A_2 over the $\mathcal{F}(i)$'s:

$$a_{1+}|n_1) = [n_1+1]_q\,|n_1+1), \qquad a_{1+}|k-1) = 0, \tag{2.4a}$$
$$a_{1-}|n_1) = [n_1]_q\,|n_1-1), \qquad a_{1-}|0) = 0, \tag{2.4b}$$
$$a_{2+}|n_2) = [n_2+1]_q\,|n_2+1), \qquad a_{2+}|k-1) = 0, \tag{2.4c}$$
$$a_{2-}|n_2) = [n_2]_q\,|n_2-1), \qquad a_{2-}|0) = 0, \tag{2.4d}$$

and

$$N_1|n_1) = n_1|n_1), \qquad N_2|n_2) = n_2|n_2). \tag{2.4e}$$

In that framework, we define two linear operators h and v_a, $a \in \mathbb{R}$, acting on $\mathcal{F}_k = \mathcal{F}(1) \otimes \mathcal{F}(2)$:

$$h|n_1, n_2) = \sqrt{n_1(n_2+1)}|n_1, n_2), \qquad n_i = 0, 1, 2, \ldots, k-1, \quad i = 1, 2 \tag{2.5a}$$

and

$$v_a|n_1, n_2) = q^{an_2}|n_1+1, n_2-1), \qquad n_1 \neq k-1, \quad n_2 \neq 0, \tag{2.5b}$$
$$v_a|k-1, n_2) = q^{-a(k-1-n_2)/2}|0, n_2-1), \qquad n_2 \neq 0, \tag{2.5c}$$
$$v_a|n_1, 0) = q^{a(k+n_1)/2}|n_1+1, k-1), \qquad n_1 \neq k-1, \tag{2.5d}$$
$$v_a|k-1, 0) = |0, k-1). \tag{2.5e}$$

The operator h is Hermitian. v_a is unitary and has a cyclic action since $(v_a)^k = I$, with I is the identity operator. The link with ordinary operators on harmonic oscillators or angular momentum can be made after Schwinger's work on angular momentum [45]. Let us put

$$J = \frac{1}{2}\left(N_1 + N_2\right), \quad M = \frac{1}{2}\left(N_1 - N_2\right), \tag{2.6a}$$
$$|j, m\rangle = |j+m, j-m) = |n_1, n_2). \tag{2.6b}$$

For a given, admissible value of j, let $\varepsilon(j)$ be the subspace of $\mathcal{F}_k$ spanned by the corresponding vectors $|j, m\rangle$, m ranging. MUBs will appear in the $\varepsilon(j)$ of maximal dimension, that is for

$$j = \frac{k-1}{2}. \tag{2.7}$$

From now on, we shall assume that j is fixed to the latter particular value. The label m can thus take $k = 2j+1$ values, namely $m = -j, -j+1, \ldots, j$. Let us denote S_j the computational basis of $\varepsilon(j)$ defined by the $|j,m\rangle$'s:

$$S_j = \{|j,m\rangle;\ m = -j, -j+1, \ldots, j\}. \tag{2.8}$$

With this condition, $\varepsilon(j)$ is stable under the action of h and v_a and we may restrict these operators to that space:

$$\begin{aligned} h|j,m\rangle &= \sqrt{(j+m)(j-m+1)}|j,m\rangle, && (2.9)\\ v_a|j,m\rangle &= (1-\delta_{m,j})\, q^{(j-m)a}|j,m+1\rangle + \delta_{m,j}|j,-j\rangle, && (2.10)\end{aligned}$$

with $\delta_{m,j}$ the Kronecker symbol. Once restricted to $\varepsilon(j)$, h is still Hermitian and v_a unitary and cyclic, with $(v_a)^{2j+1} = I$. We can now connect the operators h and v_a with $\mathfrak{su}(2)$.

Proposition 1 *An irreducible representation of the Lie algebra* $\mathfrak{su}(2)$ *can be built out of* h *and* v_a *if one puts*

$$j_+ = hv_a, \qquad j_- = v_a^\dagger h, \qquad j_z = \frac{1}{2}(h^2 - v_a^\dagger h^2 v_a). \tag{2.11}$$

Then one has

$$\begin{aligned} j_+|j,m\rangle &= q^{(j-m)a}\sqrt{(j-m)(j+m+1)}|j,m+1\rangle, && (2.12)\\ j_-|j,m\rangle &= q^{-(j-m+1)a}\sqrt{(j+m)(j-m+1)}|j,m-1\rangle, && (2.13)\\ j_z|j,m\rangle &= m|j,m\rangle. && (2.14)\end{aligned}$$

Indeed, these definitions are in agreement with the commutation relations

$$[j_z, j_+] = j_+, \qquad [j_z, j_-] = -j_-, \qquad [j_+, j_-] = 2j_z. \tag{2.15}$$

The operators j_+ and j_- thus appear in their polar decompositions and the expression of j_z is also tailored so that

$$j^2 = j_z^2 + \frac{1}{2}(j_+j_- + j_-j_+) = \frac{1}{4}(N_1+N_2)(N_1+N_2+2). \tag{2.16}$$

With this expression of j^2, one may check that

$$\begin{aligned} j^2 &= h^2 + j_z^2 - j_z \ = \ v_a^\dagger h^2 v_a + j_z^2 + j_z && (2.17)\\ j^2|j,m\rangle &= j(j+1)|j,m\rangle. && (2.18)\end{aligned}$$

The last ingredient in order to understand why the deformed cyclic operator v_a

has bearing on MUBs is the operator z defined on $\varepsilon(j)$ by

$$z|j,m\rangle = q^{j-m}|j,m\rangle. \tag{2.19}$$

If we now consider the operators v_a for $a = 0, 1, \ldots, 2j$ only, we can express them as

$$v_a = v_0 z^a, \qquad a = 0, 1, \ldots, 2j, \tag{2.20}$$

and with the group theoretical commutator defined by

$$[x, y]_{\mathrm{g}} = xyx^{-1}y^{-1}, \tag{2.21}$$

they satisfy the commutation relations

$$[v_a, z]_{\mathrm{g}} = q, \tag{2.22}$$
$$[v_a, v_b]_{\mathrm{g}} = q^{b-a}. \tag{2.23}$$

These relations should be compared with the commutation relations of Pauli operators. Indeed, the results we will derive about Pauli operators in the following chapters are essentially based on their commutation relations and thus will adapt to the v_a's and z. In particular, it will be proved that if $2j+1$ is not prime, then the v_a's and z cannot yield a maximal set of MUBs. But for now, we are going to see how these results can be partially derived from Gauss sums.

2.2 Gauss sums and MUBs

In this part, we work with a fixed value of j. The eigenvalues and the common eigenvectors of the complete set of commuting operators $\{j^2, v_a\}$ can be easily found by using standard techniques. This leads to the following result.

Proposition 2 *The eigenvalues and the eigenvectors of the operators j^2 and v_a are given by*

$$j^2|j\alpha; a\rangle = j(j+1)|j\alpha; a\rangle, \qquad v_a|j\alpha; a\rangle = q^{ja-\alpha}|j\alpha; a\rangle, \tag{2.24}$$

where

$$|j\alpha; a\rangle = \frac{1}{\sqrt{2j+1}} \sum_{m=-j}^{j} q^{\rho(j,m,a,\alpha)}|j,m\rangle, \qquad \alpha = 0, 1, \ldots, 2j, \tag{2.25}$$

and

$$\rho(j, m, a, \alpha) = \frac{1}{2}(j+m)(j-m+1)a + (j+m)\alpha. \tag{2.26}$$

The spectrum of v_a is nondegenerate. For fixed j and a, the $2j+1$ eigenvectors $|j\alpha; a\rangle$, with $\alpha = 0, 1, \ldots, 2j$, of the operator v_a generate an orthonormalised basis

$$B_a = \{|j\alpha; a\rangle;\ \alpha = 0, 1, \ldots, 2j\} \tag{2.27}$$

of the space $\varepsilon(j)$. In addition, we have

$$|\langle j,m|j\alpha;a\rangle| = \frac{1}{\sqrt{2j+1}}, \qquad m=-j,-j+1,\dots,j, \quad \alpha=0,1,\dots,2j, \tag{2.28}$$

so that the bases B_a and S_j are mutually unbiased.

Let us now consider the overlap between two bases B_a and B_b corresponding to the schemes $\{j^2, v_a\}$ and $\{j^2, v_b\}$, respectively. We have

$$\langle j\alpha;a|j\beta;b\rangle = \frac{1}{2j+1}\sum_{m=-j}^{j} q^{\rho(j,m,b-a,\beta-\alpha)}. \tag{2.29}$$

With the help of a generalised quadratic Gauss sum $S(u,v,w)$ defined by

$$S(u,v,w) = \sum_{k=0}^{|w|-1} e^{i\pi(uk^2+vk)/w}, \tag{2.30}$$

where u, v, and w are integers such that $uw \neq 0$ and $uw+v$ is an even integer [46], we have the following result.

Proposition 3 *For $b \neq a$, the overlap $\langle j\alpha;a|j\beta;b\rangle$ can be written as*

$$\langle j\alpha;a|j\beta;b\rangle = \frac{1}{w}S(u,v,w), \tag{2.31}$$

where

$$u=a-b, \qquad v=(2j+1)(b-a)+2(\beta-\alpha), \qquad w=2j+1, \tag{2.32}$$

with $a-b=\pm1,\pm2,\dots,\pm2j$ and $\alpha,\beta=0,1,\dots,2j$. Furthermore, for $2j+1$ prime we have

$$|\langle j\alpha;a|j\beta;b\rangle| = \frac{1}{\sqrt{2j+1}}, \tag{2.33}$$

with $a-b=\pm1,\pm2,\dots,\pm2j$ and $\alpha,\beta=0,1,\dots,2j$.

We give a first proof involving only arithmetics in the ring $\mathbb{Z}_{2j+1}$ of residual integers modulo $2j+1$. It is an adaptation, in the framework of angular momentum, of the method developed in [18] in order to construct a complete set of MUBs in $\mathbb{C}^d$ with d prime. Afterwards we will give a second proof involving the same arithmetics, but more in relation with Gauss sums.

Proof. The proof of (2.31) is straightforward with a translation by j of the index of the sum in (2.29).

As to (2.33), we start from

$$v_a z^n = v_b, \qquad n=b-a \in \mathbb{Z}, \tag{2.34}$$

which can be derived from (2.20). In view of Proposition 2, the action of the operator $v_a z^n$ on some vector $|j\beta_0; b\rangle$ leads to

$$v_a z^n |j\beta_0; b\rangle = q^{j(a+n)-\beta_0} |j\beta_0; b\rangle. \tag{2.35}$$

Furthermore, if we put

$$\beta_i = \beta_0 - in, \qquad i \in \mathbb{Z}, \tag{2.36}$$

then formulae (2.19) and (2.25) give

$$z^n |j\beta_0; b\rangle = q^{2jn} |j\beta_1; b\rangle. \tag{2.37}$$

Let us consider the scalar product $\langle j\alpha; a|v_a z^n|j\beta_0; b\rangle$. This product can be calculated in two different ways owing to (2.35) and (2.37). We thus obtain

$$|\langle j\alpha; a|v_a|j\beta_1; b\rangle| = |\langle j\alpha; a|j\beta_0; b\rangle| \,. \tag{2.38}$$

Since v_a is unitary and $(v_a)^{2j+1} = I$, we can write

$$v_a = (v_a^\dagger)^{2j} \, (v_a)^{2j} \, v_a = (v_a^\dagger)^{2j} \, (v_a)^{2j+1} = (v_a^\dagger)^{2j}. \tag{2.39}$$

Finally, the introduction of (2.39) into (2.38) produces

$$|\langle j\alpha; a|j\beta_1; b\rangle| = |\langle j\alpha; a|j\beta_0; b\rangle| \,. \tag{2.40}$$

The number of different β_i modulo $2j+1$ that can be reached by repeated translations of β_0 is $(2j+1)/\gcd(2j+1, n)$. The conclusion follows. ∎

The second proof uses arithmetics in Gauss sums. It is divided into two parts, whether the first argument of the Gauss sum under consideration is even or odd. But we shall see that in fact, the method for the odd case is general enough to encompass the even case. The reader may also have a look at the following section about arithmetics applied to Gauss sums.

Proof. We have

$$(2j+1)\langle j\alpha; a|j\beta; b\rangle = S(u, v, 2j+1) = \sum_{k=0}^{2j} q^{(uk^2+vk)/2}, \tag{2.41}$$

where $u = a - b$, $v = (2j+1)(b-a) + 2(\beta - \alpha)$, and $q = e^{2\pi i/(2j+1)}$.

For $j = 1/2$, the generalised quadratic Gauss sum $S(u, v, 2)$ can be easily calculated and we then check that (2.33) is satisfied for $2j + 1 = 2$.

We carry on with $2j + 1$ equal to an odd prime number. In $S(u, v, 2j + 1)$, the integer u is such that $-2j \leq u \leq 2j$ and, for $2j + 1$ prime with $j \neq 1/2$, the integer v has the same parity as u. We shall thus consider in turn u even and u odd.

In the case where u is even, we introduce the two integers $\xi = u/2$ and $\eta = v/2$. Then we have

$$S(u, v, 2j+1) = \sum_{k=0}^{2j} q^{\xi k^2 + \eta k}, \tag{2.42}$$

where the exponent of q may be taken modulo $2j+1$. A translation of the index k gives

$$S(u, v, 2j+1) = \sum_{k=0}^{2j} q^{\xi (k+t)^2 + \eta (k+t)}. \tag{2.43}$$

Since the elements appearing in the exponent of q may be considered to be the elements of a field, there exists t such that $2\xi t + \eta = 0 \pmod{2j+1}$. With this value of t, we get

$$|S(u, v, 2j+1)| = \left| \sum_{k=0}^{2j} q^{\xi k^2} \right|. \tag{2.44}$$

The value of the rhs of (2.44) is $\sqrt{2j+1}$ (see [46]). Therefore, (2.33) is proved for $2j+1$ odd prime and u even.

In the case where u is odd, let us introduce the canonical additive character of $\mathbb{Z}/(2(2j+1))\mathbb{Z}$

$$\psi : (\mathbb{Z}/(2(2j+1))\mathbb{Z}, +) \longrightarrow (\mathbb{C}, \times), \quad x \longmapsto q^{y/2}, \tag{2.45}$$

with $y \in \mathbb{Z}$ a representative of x modulo $2(2j+1)$. Consequently, we have

$$S(u, v, 2j+1) = \sum_{k=0}^{2j} \psi(uk^2 + vk), \tag{2.46}$$

where the argument of ψ stands for a residue modulo $2(2j+1)$. In order to apply the translation trick and to get rid of the linear term, as in the even case, k has to range over a complete set of residues modulo $2(2j+1)$. For this purpose, we may for instance consider the extra sum

$$\sum_{\ell=2j+1}^{2(2j+1)-1} \psi(u\ell^2 + v\ell) = \sum_{k=0}^{2j} \psi(uk^2 + 2(2j+1)uk + u(2j+1)^2 + vk + v(2j+1)). \tag{2.47}$$

The second term of the argument of ψ in the rhs of (2.47) vanishes under ψ. Moreover,

$$u(2j+1)^2 + v(2j+1) = 2(2j+1)uj + (u+v)(2j+1) \equiv 0 \pmod{2(2j+1)} \tag{2.48}$$

since $u+v$ is even. Hence, the extra sum is equal to $S(u, v, 2j+1)$ so that

$$S(u, v, 2j+1) = \frac{1}{2} \sum_{k=0}^{2(2j+1)-1} \psi(uk^2 + vk). \tag{2.49}$$

Now let us carry out the translation:

$$u(k+t)^2 + v(k+t) = uk^2 + (2ut+v)k + ut^2 + vt. \tag{2.50}$$

Since u is odd and between $-2j$ and $2j$, it is invertible modulo $2(2j+1)$. Choosing $t \equiv u^{-1} \pmod{2(2j+1)}$, we see that

$$|S(u, v, 2j+1)| = |S(u, v+2, 2j+1)|\,, \tag{2.51}$$

where an increase of v by 2 amounts for an increase of $\beta - \alpha$ by 1. Therefore, the moduli in the lhs of (2.33) do not depend on $\beta - \alpha$. To show that they are independent of $a - b$, we need only remember that the overlaps $\langle j\alpha; a | j\beta; b\rangle$ are coefficients connecting two orthonormalised bases. Consequently

$$\sum_{\alpha=0}^{2j} |\langle j\alpha; a | j\beta; b\rangle|^2 = 1 \tag{2.52}$$

and

$$\forall \alpha \in \{0, 1, \ldots, 2j\}, \quad (2j+1)\,|\langle j\alpha; a | j\beta; b\rangle|^2 = 1, \tag{2.53}$$

so that (2.33) is proved for $2j+1$ prime and u odd. ■

At this point, it is interesting to emphasize that the method we have developed to handle the odd case works in the even case too. Suppose $u = 2^n u'$, with u' not divisible by 2. In the translation Relation (2.50), the term $2ut$ should be replaced by $2^{n+1}u't$, where u' is invertible modulo $2(2j+1)$. Thus $v+2$ in (2.51) is replaced by $v + 2^{n+1}$ and an increase of v by 2^{n+1} amounts for an increase of $\beta - \alpha$ by 2^n. Since 2^n is coprime with $2j+1$, all values of $\beta - \alpha$ will be swept over modulo $2j+1$ and the result follows.

We now gather Propositions 2 and 3. It is known that for a d-dimensional Hilbert space, with d prime ($d = p$) or a power of a prime ($d = p^s$, with p prime and s positive integer greater than 1), there exists a complete set of $d+1$ MUBs. In our particular context, for $d = p = 2j+1$ prime, the orthonormal bases

$$B_a = \{|j\alpha; a\rangle;\ \alpha = 0, 1, \ldots, p-1\}, \qquad a = 0, 1, \ldots, p-1$$

satisfy (2.33), so that they constitute an incomplete set of p MUBs. But according to Proposition 2, the bases S_j and B_a, with fixed a, are also unbiased. Therefore

Proposition 4 *For $p = 2j+1$ prime, the computational basis S_j given by (2.8) and the bases B_a with $a = 0, 1, \ldots, p-1$, given by (2.25), constitute a complete set of $p+1$ MUBs in $\mathbb{C}^p$.*

We close this section with a few remarks concerning the number of bases which are unbiased with a given basis. In the second proof of Proposition 3, one of the key arguments is that u or u' must be invertible modulo $2(2j+1)$, which was immediately checked since $2j+1$ was prime. This argument cannot be used when the dimension $d = 2j+1$ is a power of a prime, $d = p^s$ (p prime and s integer greater than 1). However, taking $p \neq 2$, let us consider the bases B_a ($a = 0, 1, \ldots, d-1$) whose vectors are given by (2.25), with $j = (d-1)/2$. We remark that the number of bases B_a (a ranging) that are unbiased with one of them is at least $\varphi(p^s) = p^s - p^{s-1}$, a remark that is also valid for arbitrary dimension. If $d = p_1^{s_1} p_2^{s_2} \cdots p_n^{s_n}$, with $p_i \neq 2$ for $i = 1, 2, \ldots, n$, then the number of bases B_a (a ranging) that are unbiased with one of them is at least

$$\varphi(d) = \prod_{i=1}^{n} p_i^{s_i} - p_i^{s_i - 1}.$$

These considerations can be expressed in a geometrical way in the case of a prime power dimension $d = p^s$, with $p \neq 2$. Any integer a between 0 and $p^s - 1$ can be written in the form

$$a = a_0 + a_1 p + \cdots + a_{s-1} p^{s-1},$$

with $0 \leq a_i \leq p-1$ for $i = 0, 1, \ldots, s-1$. Thus, any basis B_a corresponds to the point of coordinates $(a_0, a_1, \ldots, a_{s-1})$ in an affine space of dimension s over the Galois field $\mathbb{Z}_p$. Moreover, we see that two bases B_a and B_b are mutually unbiased if $a_0 - b_0 \neq 0$, which excludes a hyperplane of the affine space. In fact, we will see in Chapter 5, Theorem 21 p.71 that this is also a necessary condition. Whenever d is a product of prime powers, all of the primes being different from 2, a generalisation is straightforward by the use of the Chinese remainder theorem. We shall find again such a geometrical pattern in Chapter 3.

2.3 Arithmetics and Gauss sums

As a byproduct of our study, it is worthwhile to mention that arithmetical methods can be used to derive relations between generalised quadratic Gauss sums. We are going to give an example in order to know whether some Gauss sums are 0 or not without calculating them. By the way, we find and use a result (2.60) which is usually demonstrated with the help of the reciprocity theorem.

The Gauss sum $S(u, v, w)$ as defined in (2.30), with u, v, and w integers such that $w \neq 0$ and $uw + v$ even, can be rewritten as

$$S(u, v, w) = q^{(ut^2 + vt)/2} \sum_{k=-t}^{|w|-1-t} q^{(uk^2 + (v+2ut)k)/2}, \qquad t \in \mathbb{Z}, \tag{2.54}$$

with $q = e^{2\pi i/w}$. This is again a simple translation in the index k. Moreover, as a more general version of (2.48), we have

$$uw^2 + vw = (uw + v)w \equiv 0 \pmod{2w}, \tag{2.55}$$

which, as we have already learnt, shows that, in spite of the factor $1/2$ in the exponent of q, a translation by w of any of the indices k does not modify the sum in (2.54). Hence,

$$S(u,v,w) = q^{(ut^2+vt)/2} S(u, v+2ut, w). \tag{2.56}$$

For t ranging and fixed u, v, and w, the number of different values of $v+2ut$ modulo $2(2j+1)$ is $|w|/\gcd(u,w)$; the corresponding Gauss sums are equal up to a phase factor.

If there exists t such that $ut+v \equiv 0$ modulo w, then (2.56) yields

$$S(u,v,w) = \pm S(u,-v,w). \tag{2.57}$$

In details, whenever the equation

$$tu + kw = v \tag{2.58}$$

has an integer solution in (t,k), (2.56) simplifies to

$$S(u,v,w) = e^{-i\pi kt} S(u,-v,w). \tag{2.59}$$

Besides, by using again the translation method, we see that

$$S(u,v,w) = \sum_{k=-|w|+1}^{0} q^{(uk^2+vk)/2} = S(u,-v,w), \tag{2.60}$$

a result that also follows by applying twice the reciprocity theorem [46] for generalised quadratic Gauss sums whenever $u \neq 0$. So if kt is odd, $S(u,v,w) = 0$.

Let suppose kt is odd, so that k and t are odd, and examine the system

$$\begin{cases} uw + v \text{ even}, \\ tu + kw = v. \end{cases} \tag{2.61}$$

If v is odd, then u and w are odd as well from the first condition. Then $tu+kw$ is even whereas v is odd. Thus v is even. If u is odd, then w is even from the first condition. Then again $tu+kw$ is odd whereas v is even. Thus u is even. u and w having symmetric roles, w is even as well. In conclusion, we have

Proposition 5 $S(u,v,w) = 0$ *whenever the following conditions are satisfied:*

$$u, v, w \ \textit{even}, \tag{2.62a}$$

$$v_p(v) \geq \min(v_p(u), v_p(w)) \quad \textit{for all prime } p > 2, \tag{2.62b}$$

and

$$v_2(u) \neq v_2(w) \quad \Longrightarrow \quad v_2(v) = \min(v_2(u), v_2(w)) \tag{2.62c}$$

$$v_2(u) = v_2(w) \quad \Longrightarrow \quad v_2(v) \geq \min(v_2(u), v_2(w)) \tag{2.62d}$$

For instance, one may check that they are satified by $(u,v,w) = (2,6,8)$. But not by $(4,10,12)$, as both Conditions (2.62b) and (2.62d) are unchecked in that case.

Chapter 3

The projective structure $\mathbf{P}(\mathbb{Z}_d^m)$

In this chapter, we begin by presenting the Pauli group, a subgroup of $\mathrm{U}(d)$. Section 3.1 recalls all the known features about it that we shall need. The diagonalising bases of Pauli operators are well-known to provide MUBs, which we will call Pauli MUBs. Bandyopadhyay *et al.* gave a recipe for that in [18], resting on Galois fields calculations in [7]. However, they did not give a complete isomorphism between the unbiasedness relation among bases and the various maximally commuting sets of Pauli operators to diagonalise. In order to establish a suitable relation among maximally commuting sets of Pauli operators and then the desired isomorphism, we are first led to study the projective structures $\mathbf{P}(\mathbb{Z}_d^m)$. In Section 3.2, we introduce the wedge product to quantify neighbourness between projective points. We also derive there the counting properties we shall need to relate to MUBs in Section 3.3. Section 3.4 is a mathematical complement relating neighbourhood to group theory.

3.1 The Pauli group

Let d be any integer greater than or equal to 2 and put

$$q = \exp\left(\frac{2\pi i}{d}\right), \tag{3.1}$$

the canonical primitive root of unity of order d. The Hilbert space $\mathbb{C}^d$ is endowed with a computational basis which we denote with the classical convention in quantum information

$$\{|0\rangle, |1\rangle, \ldots, |d-1\rangle\} \tag{3.2}$$

where the indices are taken in $\mathbb{Z}_d$, the ring of integers modulo d. We define on $\mathbb{C}^d$ two fundamental unitary operators: the shift operator X by

$$\forall i \in \mathbb{Z}_d, \quad X|i\rangle = |i+1\rangle \tag{3.3}$$

and the clock operator Z by

$$\forall i \in \mathbb{Z}_d, \quad Z|i\rangle = q^i|i\rangle. \tag{3.4}$$

For any particular value of d, they may be written in matrix form. For $d = 2$, we get the usual Pauli matrices σ_x and σ_z of spin 1/2 physics:

$$X = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad Z = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.5}$$

For $d = 4$, one gets

$$X = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad Z = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -i \end{pmatrix}. \tag{3.6}$$

Thus Z is a diagonal operator with eigenvectors the vectors of the computational basis and X operates a circular permutation among the eigenvectors of Z. Both operators are idempotent:

$$X^d = Z^d = I_d, \tag{3.7}$$

with I_d the identity operator on $\mathbb{C}^d$. So, in the expressions X^a and Z^b with $a, b \in \mathbb{Z}_d$, the exponents may be counted modulo d, which will be the case from now on. In the same way, for any integer k, we will consider that $k \in \mathbb{Z}_d$ in the expression q^k.

We will call the operators of the form $q^c X^a Z^b$, with $a, b, c \in \mathbb{Z}_d$, the elementary Pauli operators and denote their set

$$\mathcal{P}(d) = \{q^c X^a Z^b;\ a, b, c \in \mathbb{Z}_d\}. \tag{3.8}$$

This is indeed a subgroup of the group of unitary operators on $\mathbb{C}^d$. The fundamental ingredient in the study of these operators is the following commutation relation that can be easily checked from the definitions:

$$XZ = q^{-1} ZX. \tag{3.9}$$

With the following notation for the group theoretic commutator,

$$[A, B]_{\mathrm{g}} = ABA^{-1}B^{-1}, \tag{3.10}$$

Relation (3.9) also reads

$$[X, Z]_{\mathrm{g}} = q^{-1} I_d. \tag{3.11}$$

From (3.9), one derives:

$$\forall a, b, u, v \in \mathbb{Z}_d, \quad (X^a Z^b)(X^u Z^v) = q^{bu-av} (X^u Z^v)(X^a Z^b), \tag{3.12}$$

or equivalently

$$\forall a, b, u, v \in \mathbb{Z}_d, \quad [X^a Z^b, X^u Z^v]_{\mathrm{g}} = q^{bu-av} I_d. \tag{3.13}$$

We dropped the global phase factor q^c that appears in the definition of Pauli operators as it is irrelevant for commutation relations. The exponent of q in (3.13) is nothing but the opposite of the symplectic product of the vectors (a, b) and (u, v) in $\mathbb{Z}_d^2$. The symplectic inner product is defined by

$$\omega((a,b),(u,v)) = \begin{vmatrix} a & u \\ b & v \end{vmatrix} = av - bu. \tag{3.14}$$

It is a bilinear form whose representative matrix in the canonical basis of $\mathbb{Z}_d^2$ is

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{3.15}$$

Therefore two elementary Pauli operators $X^a Z^b$ and $X^u Z^v$ commute iff

$$\omega((a,b),(u,v)) = 0. \tag{3.16}$$

As a general rule from group theory, the set of commutators

$$D(\mathcal{P}(d)) = \{[X^a Z^b, X^u Z^v]_{\mathrm{g}};\ a, b, u, v \in \mathbb{Z}_d\}, \tag{3.17}$$

called the derived group of $\mathcal{P}(d)$, may be used to factorise $\mathcal{P}(d)$ into a commutative group. Namely, $D(\mathcal{P}(d))$ is a normal subgroup of $\mathcal{P}(d)$ and $\mathcal{P}(d)/D(\mathcal{P}(d))$ is commutative. In the particular case of elementary Pauli operators, the derived group happens to be the center of $\mathcal{P}(d)$, that is to say the subset of operators in $\mathcal{P}(d)$ that commute with any other one in $\mathcal{P}(d)$:

$$D(\mathcal{P}(d)) = \{q^c I_d;\ c \in \mathbb{Z}_d\}, \tag{3.18}$$

Thus the members of the quotient group $\mathcal{P}(d)/D(\mathcal{P}(d))$ depend only on the two parameters $a, b \in \mathbb{Z}_d$ and may be represented by the two component vector $(a, b) \in \mathbb{Z}_d^2$. In fact, we get the isomorphism

$$(\mathcal{P}(d)/D(\mathcal{P}(d)), \times) \ \simeq \ (\mathbb{Z}_d^2, +). \tag{3.19}$$

Since commutation relations in $\mathcal{P}(d)$ depend only the exponents of X and Z, all the information about commutation relations is contained the quotient group. So, the isomorphism is a kind of logarithm that can be used to investigate these relations on the basis of algebra in the $\mathbb{Z}_d$-module $\mathbb{Z}_d^2$, endowed with the symplectic inner product ω.

Now we define the Pauli group $\mathcal{P}(d, n)$ to be the set of the n-th tensorial products of elementary Pauli operators in $\mathcal{P}(d)$:

$$\mathcal{P}(d,n) = \bigotimes_{i=1}^{n} \mathcal{P}(d). \tag{3.20}$$

A quotient group of $\mathcal{P}(d,n)$ is still obtained by getting rid of the global phase factor and we get the isomorphism

$$\text{class}(X^{a_1}Z^{b_1} \otimes \cdots \otimes X^{a_n}Z^{b_n}) \simeq (a_1, b_1, \ldots, a_n, b_n) \in \mathbb{Z}_d^{2n}. \tag{3.21}$$

The $\mathbb{Z}_d$-module $\mathbb{Z}_d^{2n}$ is endowed with its canonical symplectic structure so as to account for the commutation relations of the Pauli operators:

$$\begin{aligned}&[X^{a_1}Z^{b_1} \otimes \cdots \otimes X^{a_n}Z^{b_n}, X^{c_1}Z^{d_1} \otimes \cdots \otimes X^{c_n}Z^{d_n}]_{\text{g}} = I\\ &\qquad\qquad \iff \omega((a_1, b_1, \ldots, a_n, b_n), (c_1, d_1, \ldots, c_n, d_n)) = 0,\end{aligned} \tag{3.22}$$

where the bracket is the group theoretical commutator and ω is now the symplectic product over $\mathbb{Z}_d^{2n}$. In matrix form, ω is defined by the $2n \times 2n$ block diagonal matrix J_n with n blocks equal to J:

$$J_n = \text{diag}(J, \ldots, J), \qquad n \text{ blocks}, \tag{3.23}$$

so that

$$\omega((a_1, b_1, \ldots, a_n, b_n), (c_1, d_1, \ldots, c_n, d_n)) = \sum_{i=1}^{n} a_i d_i - b_i c_i. \tag{3.24}$$

In this thesis, we are to make an extensive use of the latter $\mathbb{Z}_d$-module transcription of the quotiented Pauli group. In brief, a Pauli operator on $\mathbb{C}^{d^n}$ up to a complex multiplier can be represented by a vector in $\mathbb{Z}_d^{2n}$. The commutation relations among these operators are transcribed in this algebraic framework into computing a symplectic product. In particular, two operators commute iff the corresponding symplectic product is 0.

A basic fact is that two colinear vectors have a zero symplectic product. Indeed, for $n = 1$ for example, one has

$$\omega((ka, kb), (la, lb)) = 0. \tag{3.25}$$

Since we will be interested in maximally commuting sets of Pauli operators, we are thus led to consider the projective structure of $\mathbb{Z}_d^{2n}$, namely $\mathbf{P}(\mathbb{Z}_d^{2n})$.

In the forthcoming sections of this chapter, we largely put aside the role of the symplectic product and get more familiar with the projective structures in finitely generated $\mathbb{Z}_d$-modules, by comparing with the field case. Moreover, in order to be as general as possible, we will consider $\mathbb{Z}_d^m$ instead of $\mathbb{Z}_d^{2n}$, with m either even or odd. In fact, another idea is behind the features we are going to investigate here, as we shall see in Section 7.4.

3.2 Neighbourhood and distance

In a vector space, that is to say one uses a base field as $\mathbb{R}$ or $\mathbb{C}$, any two linearly independent vectors generate the same structure, namely a 2-dimensional subspace.

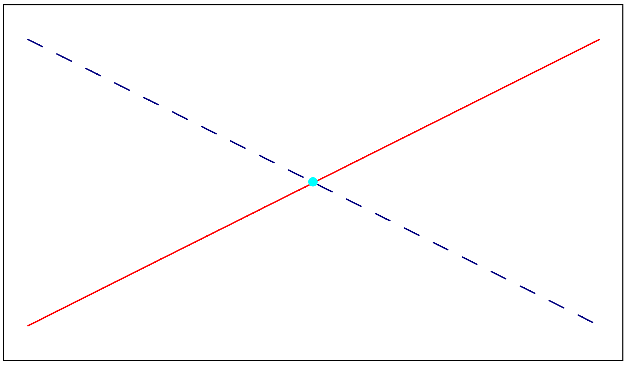

Figure 3.1: Two lines over a base field, as $\mathbb{R}$ or $\mathbb{C}$

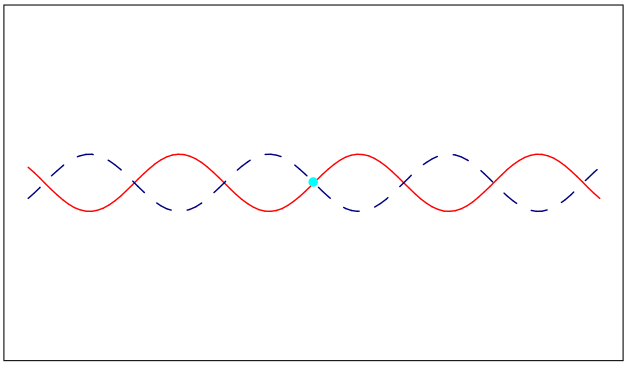

Figure 3.2: Two lines over a base ring, as $\mathbb{Z}_d$

Two vectorial lines over a field intersect only at the origin (see Figure 3.1). In the projective language, two distinct points of a vector space generate a projective line and all the lines thus obtained are isomorphic. This is no more the case in general over a ring, for instance $\mathbb{Z}_d$ whenever d is not a prime. There two lines may intersect at points other than the origin (see Figure 3.2).

We are specifically interested in the case of the ring $\mathbb{Z}_d$. The results we are going to present were known in the particular case of qubits ($d = 2$) and qutrits ($d = 3$), as well as for some other particular dimensions. These results were derived in [37, 47–49], in the framework of extensive computations with the group theoretical software GAP and comparing with previous studies in projective geometry over various finite rings [50, 51]. In this thesis, we concentrate on the overarching projective structure and prove these results mathematically so that they are valid in any dimension. We remark that our results are stated in terms of $\mathbb{Z}_d$ only. We also connect with the wedge product, as a new tool for characterising neighbour or distant points.

Let d be any integer greater than or equal to 2 and m be a positive integer. We first specialise in the case d a power of prime, say $d = p^s$, with p a prime integer and s a positive integer. The general case will be deduced from that latter one with the use of the Chinese remainder theorem. Moreover, $\nu(x)$ will denote the order of any element x in a group G and $v_p(k)$ the p-valuation of any $k \in \mathbb{Z}_d$.

3.2.1 Special case: d a power of a prime

We consider $\mathbb{Z}_d^m$, $d = p^s$, endowed with its canonical structure of $\mathbb{Z}_d$-module. A vector $x \in \mathbb{Z}_d^m$ is said to be free iff $\nu(x) = d$, or in other words iff the canonical map $f : \mathbb{Z}_d \to \mathbb{Z}_d^m, k \mapsto k \cdot x$ is injective. If x is free, we say that $\langle x \rangle = \operatorname{Im} f$, the submodule generated by x, is a projective point in $\mathbb{Z}_d^m$. The set of all projective points in $\mathbb{Z}_d^m$ will be called the projective net built over $\mathbb{Z}_d^m$ and denoted $\mathbf{P}(\mathbb{Z}_d^m)$. The cardinality of $\mathbf{P}(\mathbb{Z}_d^m)$ is the number of free vectors in $\mathbb{Z}_d^m$ divided by the number of invertible elements in $\mathbb{Z}_d$:

$$\begin{aligned}
\operatorname{Card}(\mathbf{P}(\mathbb{Z}_d^m)) &= \frac{p^{sm} - p^{(s-1)m}}{p^s - p^{s-1}} && (3.26a)\\
&= \sum_{k=0}^{m-1} p^{sk} p^{(s-1)(m-1-k)} = \sum_{k=0}^{m-1} p^{(s-1)(m-1)+k} && (3.26b)\\
&= p^{(s-1)(m-1)} \frac{p^m - 1}{p - 1} && (3.26c)
\end{aligned}$$

What we said above in comparing "spaces" over fields and rings show the interest of the following lemma.

Lemma 6 *Let $x_1, x_2 \in \mathbb{Z}_d^m$ be two free vectors. Then the following three statements are equivalent:*

1. *The following map is injective:*

$$\begin{array}{rccc} f_2 : & \mathbb{Z}_d^2 & \longrightarrow & \mathbb{Z}_d^m \\ & (k_1, k_2) & \longmapsto & k_1 x_1 + k_2 x_2. \end{array} \qquad (3.27)$$

2. $\langle x_1 \rangle \cap \langle x_2 \rangle = \{0\}$

3. *$x_1 \wedge x_2$ is a free vector of $\bigwedge^2 \mathbb{Z}_d^m$.*

Proof. (1) $\Rightarrow$ (2) is obvious. If $\langle x_1 \rangle \cap \langle x_2 \rangle = \{0\}$, then, according to Lemma 44 p.143 with $x = x_1$ and $y \in \langle x_2 \rangle$, the map

$$k \longmapsto x_1 \wedge k x_2 \qquad (3.28)$$

is injective and so $x_1 \wedge x_2$ is free. Finally, let us suppose that (3) is true and let $k_1, k_2 \in \mathbb{Z}_d$ such that $k_1 x_1 + k_2 x_2 = 0$. Then $k_1 x_1 \wedge x_2 = 0$ and so $k_1 = 0$. In the same way $k_2 = 0$. Thus f_2 is injective. ■

With the notations of the lemma, there is in fact a one-to-one correspondance between $\operatorname{Card}(\operatorname{Im} f_2)$ and $\nu(x_1 \wedge x_2)$. To see this, one may choose a computational basis such that x_1 has all its coefficients but the first one equal to 0 and x_2 has all its coefficients but the first two ones equal to 0 (see Lemma 38 p.132). Then it is obvious that the order of $x_1 \wedge x_2$ in $\bigwedge^2 \mathbb{Z}_d^m$ is the order of the second coefficient of x_2 in $\mathbb{Z}_d^m$ and so

$$\operatorname{Card}(\operatorname{Im} f_2) = \nu(x_1 \wedge x_2) d. \qquad (3.29)$$

We can now restate the definition of neighbourhood and distance (see Appendix E).

Definition 7 *Let* $\langle x_1\rangle, \langle x_2\rangle \in \mathbf{P}(\mathbb{Z}_d^m)$. *If one of the three properties of Lemma 6 is satisfied by the vectors* x_1, x_2, *then they are said to be distant and so are the corresponding projective points. Otherwise the two vectors are said to be neighbour and so are the two projective points, which is denoted* $\mathcal{V}(x_1, x_2)$ *or* $\mathcal{V}(\langle x_1\rangle, \langle x_2\rangle)$.

We have to check that the notion of distance and therefore that of neighbourhood among projective points are well-defined. It has to be independent of the choice of the free vectors x_1, x_2 to generate the projective points under consideration. In fact, Properties 2 and 3 of Lemma 6 are obviously independent of that choice. If one wished to generalise the notion of distance into a quantitative way, then $\nu(x_1 \wedge x_2)$ or $\nu(x_1 \wedge x_2) \cdot d$ would be possible candidates.

Proposition 8 *Whenever* d *is a power of a prime, the neighbourhood relation* $\mathcal{V}$ *is an equivalence relation both for vectors and projective points.*

Proof. It is obviously reflexive and symmetric. To show that it is also transitive, let $x, y, z \in \mathbb{Z}_d^m$ be three free vectors such that $\mathcal{V}(x, y)$ and $\mathcal{V}(y, z)$ and let us suppose that x and z are distant. We may suppose that x has all its coefficients but the first one equal to 0 and z has all its coefficients but the first two ones equal to 0. By this assumption and Property 3 of the lemma, x_1 and z_2 are units of $\mathbb{Z}_d$. Hence, according to $\mathcal{V}(x, y)$ and Property 3, the y_i's, $i \in \{2, \dots, m\}$, have to be noninvertible. Then in the same way, y_1 has to be noninvertible according to $\mathcal{V}(y, z)$. But this contradicts the fact that y is a free vector. The transitivity of $\mathcal{V}$ for projective points follows immediately. ∎

From now on, we will be interested in the neighbourhood relation in $\mathbf{P}(\mathbb{Z}_d^m)$. Let $\langle x\rangle \in \mathbf{P}(\mathbb{Z}_d^m)$. In order to find the cardinality of the class of $\langle x\rangle$, we may first choose a basis of $\mathbb{Z}_d^m$ such that x reads

$$x = (1, 0, \dots, 0) \tag{3.30}$$

as in the previous proof. Hence the points neighbour to x read $\langle y\rangle$ with some y of the form

$$y = (y_1, py_2', \dots, py_m'), \qquad y_1 \text{ a unit.} \tag{3.31}$$

Therefore the cardinality of any class of $\mathcal{V}$ is

$$\mathrm{Card}(\mathrm{class}\,\langle x\rangle) = \frac{(p^s - p^{s-1})p^{(s-1)(m-1)}}{p^s - p^{s-1}} = p^{(s-1)(m-1)} \tag{3.32}$$

and according to (3.26c) the number of classes is

$$\mathrm{Card}(\mathbf{P}(\mathbb{Z}_d^m)/\mathcal{V}) = \frac{p^m - 1}{p - 1}. \tag{3.33}$$

Examples: According to (3.26c), $\mathbf{P}(\mathbb{Z}_2^2)$ has 3 points (Figure 3.3) and $\mathbf{P}(\mathbb{Z}_3^2)$ 4 points (Figure 3.4). Relations (3.32) and (3.33) both shows that in those two projective structures, no two points are neighbour. More generally, the projective net $\mathbf{P}(\mathbb{Z}_d^m)$ has as many classes as points, with one point per class. But $\mathbf{P}(\mathbb{Z}_4^2)$ has 6 points, 3 neighbourhood classes and 2 points per class (Figure 3.5). The projective net $\mathbf{P}(\mathbb{Z}_4^2)$ is in fact a doubling ($p^{(s-1)(m-1)} = 2$) of $\mathbf{P}(\mathbb{Z}_2^2)$. More generally, $\mathbf{P}(\mathbb{Z}_{p^s}^m)$ can be arranged into a set of parallel lines where two points are neighbour iff they belong to one and the same line. □

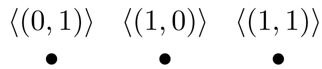

Figure 3.3: Neighbourhood relations in $\mathbf{P}(\mathbb{Z}_2^2)$

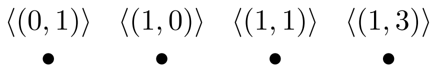

Figure 3.4: Neighbourhood relations in $\mathbf{P}(\mathbb{Z}_3^2)$

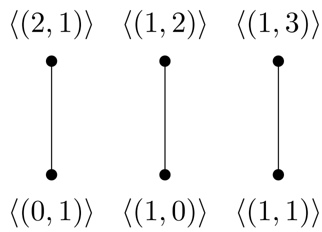

Figure 3.5: Neighbourhood relations in $\mathbf{P}(\mathbb{Z}_4^2)$

3.2.2 General case: d any integer ≥ 2

Now let us deal with the case of a general $d = \prod_i p_i^{s_i}$. With the use of the Chinese remainder theorem, we first see that the points of $\mathbf{P}(\mathbb{Z}_d^m)$ can be arranged into a multidimensional grid:

$$\mathbf{P}(\mathbb{Z}_d^m) \simeq \prod_i \mathbf{P}\left(\mathbb{Z}_{p_i^{s_i}}^m\right), \tag{3.34}$$

and so

$$\mathrm{Card}(\mathbf{P}(\mathbb{Z}_d^m)) = \prod_i \mathrm{Card}\left(\mathbf{P}\left(\mathbb{Z}_{p_i^{s_i}}^m\right)\right). \tag{3.35}$$

Also Lemma 6 is still valid and so we take up the same definition of neigbourhood and distance for a general d. With π_i the canonical projections associated to decomposition (3.34), we have for all $\langle x_1\rangle, \langle x_2\rangle \in \mathbf{P}(\mathbb{Z}_d^m)$

$$\mathcal{V}(\langle x_1\rangle, \langle x_2\rangle) \Longleftrightarrow \exists i, \mathcal{V}(\langle \pi_i(x_1)\rangle, \langle \pi_i(x_2)\rangle). \tag{3.36}$$

The neighbourhood relation is no more transitive and thus is no more an equivalence relation. But the Chinese remainder theorem and what we know about the special case of d a power of a prime are enough to give the structure of the projective net $\mathbf{P}(\mathbb{Z}_d^m)$ for a general d. If we rule out the case $m = 1$, each prime factor p_i of d contributes one or two dimensions to the grid whether $s_i = 1$ or $s_i > 1$, respectively. Whenever C_i is a class in $\mathbf{P}\left(\mathbb{Z}^m_{p_i^{s_i}}\right)$, then the hyperplane $\pi_i^{-1}(C_i)$ is a set of pairwise neighbour points in $\mathbf{P}(\mathbb{Z}_d^m)$. More precisely, two points in $\mathbf{P}(\mathbb{Z}_d^m)$ are neighbour iff they belong to one and the same of those hyperplanes.

Examples: Starting from the structures of $\mathbf{P}(\mathbb{Z}_2^2)$ and $\mathbf{P}(\mathbb{Z}_3^2)$ (Figures 3.3 and 3.4), one obtains the structure of the projective net $\mathbf{P}(\mathbb{Z}_6^2)$ (Figure 3.6). Since 6 has no square factor, this is a grid. More generally, if $d = \prod_{i=1}^k p_i$ has no square factor, then the projective net $\mathbf{P}(\mathbb{Z}_d^m)$ may be represented by a k-dimensional grid where two points are neighbour iff they belong to one and the same hyperplane. On the contrary, since 12 has a square factor, the projective structure of $\mathbf{P}(\mathbb{Z}_{12}^2)$ is more complicated (Figure 3.7). It may be represented by a $(3 \times 2) \times 4$ grid with only two families of parallel hyperplanes accounting for neighbour points, namely the hyperplanes orthogonal to the first and third canonical directions. $\square$

$$
\begin{tikzcd}
\langle(1,0)\rangle \arrow[r, dash] \arrow[d, dash] & \langle(1,2)\rangle \arrow[r, dash] \arrow[d, dash] & \langle(3,2)\rangle \arrow[r, dash] \arrow[d, dash] & \langle(5,2)\rangle \arrow[d, dash] \\
\langle(1,3)\rangle \arrow[r, dash] \arrow[d, dash] & \langle(5,1)\rangle \arrow[r, dash] \arrow[d, dash] & \langle(3,1)\rangle \arrow[r, dash] \arrow[d, dash] & \langle(1,1)\rangle \arrow[d, dash] \\
\langle(2,3)\rangle \arrow[r, dash] & \langle(2,1)\rangle \arrow[r, dash] & \langle(0,1)\rangle \arrow[r, dash] & \langle(4,1)\rangle
\end{tikzcd}
$$

Figure 3.6: Neighbourhood relations in $\mathbf{P}(\mathbb{Z}_6^2)$

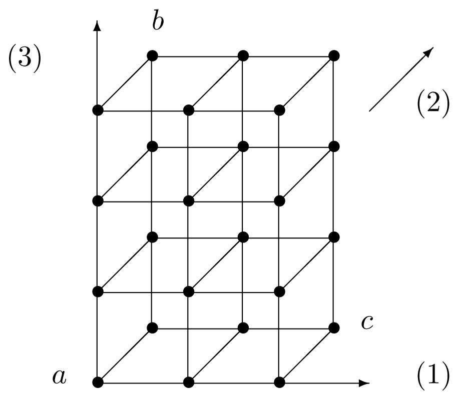

Figure 3.7: Neighbourhood relations in $\mathbf{P}(\mathbb{Z}_{12}^2)$: a is neighbour with b and c, but b and c are distant.

3.3 Neighbour points and MUBs

We here relate the features we have just developed to the study of MUBs. We have to say in anticipation of the results presented in Chapter 5 that if one works with Pauli operators that are tensorial products of k elementary Pauli operators, then the maximally commuting sets of interest in order to get MUBs are generated by k operators. Since in the present chapter, we look only at the projective points and so at the monogenic sets of operators, we must have $m = 2$ in order to relate to MUBs. That is to say we have to rule out any tensorial product. Moreover, Theorem 21 establishes that the key feature in the search for MUBs out of Pauli operators is a symplectic product. But this notion coincide with the wedge product only for $m = 2$.

In particular, it will follow from Theorem 21 that two Pauli operators $X^a Z^b$ and $X^c Z^d$ define two unbiased bases by diagonalisation iff $ad - bc$ is a unit, that is to say in the projective point language iff (a, b) and (c, d) are distant. Thus, the structure of a projective line over $\mathbb{Z}_d$ we explicited above shows that if $d = \prod_i p_i^{s_i}$ and p_j is the minimum of the prime factors appearing in the decomposition of d, then the number of MUBs one can get with such nontensorial Pauli operators is at most $p_j + 1$. This is in agreement with the more general, limiting results obtained by Archer in [17]: We have no grasp on composite dimensions. In addition, we know how to build at least $\min_i\{p_i^{s_i}\} + 1$ MUBs in $\mathbb{C}^d$, which shows that if one wants to build this number of MUBs by diagonalising Pauli operators, then tensorial products are mandatory. At this stage, we do not know how many factors must be involved in the tensor product. So, we shall retain

Proposition 9 *Let $d = p^s$ be a power of a prime, with $s > 1$. With the help of $\mathcal{P}(d)$, one is able to get only $p + 1$ MUBs, whereas a maximal set of MUBs contains $p^s + 1$ of them. Thus tensorial products of Pauli operators are mandatory in order to reach this maximum.*

The way to take this into account in a projective, geometrical framework will be developed in Chapters 4 and 5. We already know after the work by Bandyopadhyay *et al.* [18] that $\mathcal{P}(p, s)$ is a convenient Pauli group for the purpose of finding $p^s + 1$ MUBs in $\mathbb{C}^{p^s}$. We shall prove that it is indeed the only one.

A final remark is in order. In case $m = 2n$, we may retain from our first naive projective framework a very basic property. There exists a linear form $\varphi : \bigwedge^2 \mathbb{Z}_d^{2n} \to \mathbb{Z}_d$ such that for all $x_1, x_2 \in \mathbb{Z}_d^{2n}$, their symplectic product is

$$\omega(x_1, x_2) = \varphi(x_1 \wedge x_2). \tag{3.37}$$

If $\omega(x_1, x_2)$ is a unit, then $x_1 \wedge x_2$ is free. Thus Theorem 21 will show that in order to build MUBs, one needs a precise number of pairs of distant vectors or points.

3.4 Neighbourhood classes as orbits

This last section is a mathematical complement relating neighbourhood in a projective net to group theory. However, its content will not be used in the forthcoming chapters.

In Section 3.2, we gave a way to know whether two given vectors or projective points are neighbour or distant. If only one vector or point is given, we are going to show that its neighbourhood can be characterised in a geometrical way. Namely, the equivalence classes of $\mathcal{V}$ can be obtained as the orbits of a convenient group action. For simplicity, we take $d = p^s$.

Let $G = \mathrm{GL}(m, \mathbb{Z}_d)$ act naturally on $\mathbb{Z}_d^m$ and also on $\mathbf{P}(\mathbb{Z}_d^m)$ by

$$\forall g \in G,\ \forall \langle x \rangle \in \mathbf{P}(\mathbb{Z}_d^m), \quad g \cdot \langle x \rangle = \langle gx \rangle . \tag{3.38}$$

This action is compatible with $\mathcal{V}$ (see [52]):

$$\forall g \in G,\ \forall \langle x_1 \rangle , \langle x_2 \rangle \in \mathbf{P}(\mathbb{Z}_d^m), \quad \mathcal{V}(\langle x_1 \rangle , \langle x_2 \rangle) \Longleftrightarrow \mathcal{V}(\langle gx_1 \rangle , \langle gx_2 \rangle). \tag{3.39}$$

Therefore we build an action of G on the neighbourhood classes and we search for the kernel K of that latter action. For any $\Phi \in K$ and any $x \in \mathbb{Z}_d^m$, we must have $\mathcal{V}(x, \Phi x)$. In particular with x the canonical basis vectors of $\mathbb{Z}_d^m$ or the sum of two of them, we see that

1. $\forall i \in \{1, \ldots , m\}, \quad \phi_{ii}$ is a unit;
2. $\forall j \neq i,\ \exists \mu_{ij} \in \mathbb{Z}_d, \quad \phi_{ij} = p\mu_{ij}$;
3. $\forall i, j \in \{1, \ldots , m\},\ \exists \nu_{ij} \in \mathbb{Z}_d, \quad \phi_{ii} - \phi_{jj} = p\nu_{ij}$.

Those three conditions obviously define a subgroup K_0 of G that contains K. Conversely, let $\Phi \in K_0$, $x \in \mathbb{Z}_d^m$ and let us denote $\propto p$ any multiple of p. We may suppose without loss of generality that x is of the form

$$x = (u_1, \ldots , u_k, \propto p, \ldots , \propto p), \tag{3.40}$$

where $u_1, \ldots , u_k$ are units with $k \geq 1$. Then

$$\Phi x = (\phi_{11}u_1 + \propto p, \ldots , \phi_{kk}u_k + \propto p, \propto p, \ldots , \propto p) \tag{3.41}$$

and with Condition 3 we have, for all $i, j \in 1, \ldots , k$

$$\begin{vmatrix} u_i & \phi_{ii}u_i + \propto p \\ u_j & \phi_{jj}u_j + \propto p \end{vmatrix} = (\phi_{jj} - \phi_{ii})u_iu_j + \propto p = \propto p. \tag{3.42}$$

Therefore $K = K_0$.

In order to illustrate the use of Condition 3 once again, we can show directly that K is normal in G. Indeed, let $U \in G$ and $\Phi \in K$. We have

$$(\Phi U)_{kj} = \phi_{kk} u_{kj} + \propto p \tag{3.43}$$

and then, with u'_{ij} the coefficients of U^{-1},

$$\begin{aligned}
(U^{-1}\Phi U)_{ij} &= \sum_{k=1}^{m} u'_{ik}(\Phi U)_{kj} = \sum_{k=1}^{m} \phi_{kk} u'_{ik} u_{kj} + \propto p && \text{(3.44a)}\\
&= \phi_{11} \sum_{k=1}^{m} u'_{ik} u_{kj} + \sum_{k=1}^{m} (\phi_{kk} - \phi_{11}) u'_{ik} u_{kj} + \propto p && \text{(3.44b)}\\
&= \phi_{11} \delta_{ij} + \propto p. && \text{(3.44c)}
\end{aligned}$$

Hence $U^{-1}\Phi U$ satisfies the three conditions above.

In general G/K is a proper subgroup of the permutation group of the orbits. If some $g \in G$ stabilises the $2m-1$ orbits determined by the canonical basis vectors e_i and the sums $e_1 + e_j$, $j \neq 1$, for instance, then we saw that $g \in K$. Let us see when the following inequality is checked:

$$\frac{p^m - 1}{p - 1} - (2m - 1) \geq 2. \tag{3.45}$$

It also reads

$$(p^{m-1} - 2) + (p^{m-2} - 2) + \cdots + (p^2 - 2) + (p - 2) \geq 2. \tag{3.46}$$

This inequality is false only for:

- $m = 1$,
- $m = 2$ and $p = 2$,
- $m = 2$ and $p = 3$.

Chapter 4

Lagrangian submodules

In the previous chapter, we saw that we have an isomorphism

$$(\mathcal{P}(d,n)/D(\mathcal{P}(d,n)),\times) \;\simeq\; (\mathbb{Z}_d^{2n},+) \tag{4.1}$$

and that

$$[X^{a_1}Z^{b_1}\otimes\cdots\otimes X^{a_n}Z^{b_n}, X^{c_1}Z^{d_1}\otimes\cdots\otimes X^{c_n}Z^{d_n}]_{\mathrm{g}} = I \\ \Longleftrightarrow \omega((a_1,b_1,\ldots,a_n,b_n),(c_1,d_1,\ldots,c_n,d_n)) = 0, \tag{4.2}$$

where the bracket is the group theoretic commutator and ω is the symplectic product over $\mathbb{Z}_d^{2n}$. Therefore, on the one hand, maximally commuting sets of Pauli operators are accounted for by special submodules in $\mathbb{Z}_d^{2n}$. On the other hand, for any submodule M of $\mathbb{Z}_d^{2n}$, the notion of interest is the symplectic orthogonal of M defined by

$$M^\omega = \{x \in \mathbb{Z}_d^{2n};\ \forall y \in M,\ \omega(x,y) = 0\}. \tag{4.3}$$

Then a submodule M is called

- isotropic if $M \subset M^\omega$,
- coisotropic if $M^\omega \subset M$,
- Lagrangian if $M = M^\omega$.
- symplectic if $M \cap M^\omega = \{0\}$,

Let M be a Lagrangian submodule. M is isotropic. Let us suppose that there exists an isotropic submodule N such that $M \subsetneq N$. Then $M \subsetneq N \subset N^\omega \subset M^\omega$ and hence M is not Lagrangian. Thus, a Lagrangian submodule is isotropic and maximal for inclusion restricted to isotropic submodules. Theorem 13 below will show that the converse is also true. We thus see that maximally commuting sets of Pauli operators are transcribed in the module language into Lagrangian submodules, since the corresponding modules are the maximally isotropic ones. Our goal will then be to characterise Lagrangian submodules and express them in as simple a way as possible.

The general, arithmetical and algebraic tools we need are set out in the appendices in details. We extract here only the essential features adapted to our particular context.

Lemma 10 *Let $d \geq 2$, $a_1, a_2, \ldots, a_l \in \mathbb{Z}_d$ and δ be one of their* gcd*'s. For any $i \in \{1, \ldots, l\}$, one can find $k_1, k_2, \ldots, k_l \in \mathbb{Z}_d$ with $k_i \in U(\mathbb{Z}_d)$ such that*

$$\delta = \sum_{j=1}^{l} k_j a_j. \tag{4.4}$$

In the following lemma, ν_1, ν_2 are the orders of the vectors $a_1, a_2 \in \mathbb{Z}_d^{2n}$, respectively. Then $\nu_1 \vee \nu_2$ denotes the lcm of ν_1 and ν_2, and $\langle x, y\rangle$ stands for the submodule generated by x and y. By the way, $x \wedge y$ will denote the gcd of $x, y \in \mathbb{Z}_d$.

Lemma 11 *Let $a_1, a_2 \in \mathbb{Z}_d^{2n}$ of order ν_1, ν_2 respectively. There exists a linear combination a of a_1, a_2 of order $\nu = \nu_1 \vee \nu_2$ and a linear combination b of a_1, a_2 such that*

$$\langle a, b\rangle = \langle a_1, a_2\rangle . \tag{4.5}$$

Proof. Refering to the Chinese remainder theorem we may assume that d is a power of prime, say $d = p^s$. In that case, with $i = 1$ or 2 such that $\nu_i = \max(\nu_1, \nu_2)$, we simply put $a = a_i$ and b equal to the other one of the a_i's. ∎

Note that for any linear combination $x = x_1 a_1 + x_2 a_2$ of a_1 and a_2,

$$\nu x = x_1(\nu a_1) + x_2(\nu a_2) = 0. \tag{4.6}$$

Thus for all $x \in \langle a_1, a_2\rangle$, the order of x divides ν.

Now suppose that we are given a minimal basis $b = (b_1, \ldots, b_r)$ of a submodule M of $\mathbb{Z}_d^{2n}$. That is to say that one cannot find a basis of M with less than r vectors and we shall say that M is a rank-r submodule[1]. Let B be the matrix of size $2n \times r$ whose i-th column is b_i. The matrix B is called a basis matrix for M and necessarily $r \leq 2n$. Owing to Lemma 11 and associativity of lcm, we may suppose that

$$\nu(b_1) = \bigvee_{i=1}^{r} \nu(b_i), \tag{4.7a}$$

$$\forall m \in M,\ \nu(m) | \nu(b_1). \tag{4.7b}$$

An algorithm which set any matrix M that way will be called $\mathscr{A}$. It consists of an appropriate right-multiplication by an invertible matrix $R(M)$.

[1] We use the notion of rank in an extended fashion as here it applies to any submodules, including those that are not free.

4.1 A first symplectic reduction algorithm

If we put

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{4.8}$$

then in matrix form, ω is defined with respect to the canonical basis by the $2n \times 2n$ block-diagonal matrix

$$J_n = \mathrm{diag}(J, \dots, J). \tag{4.9}$$

A basis $(b_1, \dots, b_{2n})$ represented by a matrix B with respect to the canonical basis is called a symplectic basis whenever $B^T J_n B = J_n$, where B^T is the transpose of B. For instance, the canonical basis is symplectic. When one looks after the elementary divisors of a submodule, the final computational basis may be any free basis of $\mathbb{Z}_d^{2n}$. We are now interested in reduction of matrices where changes of computational basis can only be symplectic, so that in the new computational basis, the inner product ω is still to be represented by J_n. Matrices L used for left-mutiplication thus have to belong to the set of $2n \times 2n$ symplectic matrices defined as:

$$\mathrm{Sp}(n, \mathbb{Z}_d) = \{L;\ L^T J_n L = J_n\}, \tag{4.10}$$

The identity matrix is symplectic. A matrix that represents a symplectic basis with respect to another symplectic basis is symplectic. Note that in $\mathbb{Z}$, a symplectic matrix has determinant ± 1. (In fact, one can prove that the determinant is equal to 1.) The same is thus true for a symplectic matrix over $\mathbb{Z}_d$. This proves that all symplectic matrices are invertible. Moreover, the inverse of a symplectic matrix is symplectic.

Our task in this chapter is then the following. Given an $r \times r$ basis matrix B for a submodule of $\mathbb{Z}_d^{2n}$, can we find a symplectic matrix S and an $r \times r$ invertible matrix such that SBR is as simple as possible, that is to say diagonal. If yes, how can we find S and R?

Example: If we take the following basis matrix of a submodule in $\mathbb{Z}_9^4$:

$$B = \begin{pmatrix} 3 & 3 & 6 & 3 \\ 0 & 3 & 0 & 6 \\ 2 & 8 & 1 & 0 \\ 8 & 5 & 4 & 0 \end{pmatrix}, \tag{4.11}$$

then with

$$S = \begin{pmatrix} 5 & 0 & 6 & 3 \\ 7 & 2 & 7 & 5 \\ 4 & 3 & 8 & 0 \\ 0 & 0 & 4 & 8 \end{pmatrix}, \qquad R = \begin{pmatrix} 0 & 3 & 5 & 1 \\ 1 & 2 & 8 & 6 \\ 7 & 5 & 6 & 4 \\ 2 & 3 & 2 & 6 \end{pmatrix}, \tag{4.12}$$

where S is symplectic and R invertible, one obtains

$$SBR = \mathrm{diag}(3, 3, 1, 0). \tag{4.13}$$

□

We first address reduction of a single vector and afterwards that of a matrix. The case $n = 1$ is trivial and is nothing but substep 2 below. Reduction of a single vector when $n \geq 2$ relies itself on the fondamental case $n = 2$. The following substeps are elementary operations that we shall use later on in the various steps of our symplectic reduction algorithm for matrices. They form a sequence in order to reduce a vector with four components $(x, y, z, t)^T$ using only symplectic changes of basis.

► **Substep 1**: Let $x, y, z, t \in \mathbb{Z}_d$ and $\delta = x \wedge y \wedge z \wedge t$. According to corollary 10, there exist $k_1, k_2, k_3 \in \mathbb{Z}_d$ and $u \in U(\mathbb{Z}_d)$ such that

$$\underbrace{\begin{pmatrix} u & 0 & 0 & 0 \\ k_1 & u^{-1} & k_2 & k_3 \\ -k_3 u & 0 & 1 & 0 \\ k_2 u & 0 & 0 & 1 \end{pmatrix}}_{S_1} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} x_1 \\ \delta \\ z_1 \\ t_1 \end{pmatrix} \tag{4.14}$$

where S_1 is symplectic and x_1, z_1, t_1 are byproducts of the choice of k_1, k_2, k_3 and u.

► **Substep 2**: Then, with Euclid's algorithm to calculate a gcd, we find $v, w, k_4, k_5 \in \mathbb{Z}_d$ such that

$$vz_1 + wt_1 = z_1 \wedge t_1 = z_2, \quad -k_5 z_1 + k_4 t_1 = 0, \quad vk_4 + wk_5 = 1 \tag{4.15}$$

and we perform a second left-multiplication

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & v & w \\ 0 & 0 & -k_5 & k_4 \end{pmatrix}}_{S_2} \begin{pmatrix} x_1 \\ \delta \\ z_1 \\ t_1 \end{pmatrix} = \begin{pmatrix} x_1 \\ \delta \\ z_2 \\ 0 \end{pmatrix}. \tag{4.16}$$

where S_2 is symplectic.

► **Substep 3**: Since

$$\delta = x \wedge y \wedge z \wedge t = x_1 \wedge \delta \wedge z_1 \wedge t_1 = x_1 \wedge \delta \wedge z_2, \tag{4.17}$$

we also have

$$\delta \wedge z_2 = (x_1 \wedge \delta \wedge z_2) \wedge z_2 = \delta. \tag{4.18}$$

Thus we can find k_6 such that $k_6\delta + z_2 = 0$ and we perform a third left-mutiplication

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 & k_6 \\ 0 & 1 & 0 & 0 \\ 0 & k_6 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{S_3} \begin{pmatrix} x_1 \\ \delta \\ z_2 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 \\ \delta \\ 0 \\ 0 \end{pmatrix}. \tag{4.19}$$

where S_3 is symplectic.

If $n > 2$, we apply the process defined by this sequence of substeps $n-1$ times in order to end with a vector whose components are null except maybe the first two ones. At Step i, we set the $(2n+2-2i)$-th and the $(2n+1-2i)$-th components to 0. For a single vector, we can go further and set the second component to 0 as in the second substep above. We shall soon define a substep 4 to complete this list of elementary operations.

It is in general not possible to diagonalise nor to trigonalise a matrix using only a left-multiplication by a symplectic matrix. For instance, let us try to do even weaker a job with the matrix B in the following equality over $\mathbb{Z}_{p^s}$, $s \geq 1$:

$$\underbrace{\begin{pmatrix} \alpha & * & \gamma & k_1 \\ \beta & * & \delta & k_2 \\ 0 & * & l_1 p & * \\ 0 & * & l_2 p & * \end{pmatrix}}_{L} \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & p \\ 0 & 1 \\ 0 & 0 \end{pmatrix}}_{B} = \begin{pmatrix} \alpha & * \\ \beta & * \\ 0 & 0 \\ 0 & 0 \end{pmatrix}. \tag{4.20}$$

Our aim is to find a symplectic matrix L so as to get rid of any nonzero term in the last two rows. The first, third and fourth column vectors of L, let us call them C_1, C_3 and C_4, must be as shown in (4.20). But as L is supposed to be sympletic, C_3 must be free and $\omega(C_1, C_3) = 0$. So there exist $k_3, k_4 \in \mathbb{Z}_d$ such that $k_3\gamma + k_4\delta = 1$ and $\alpha\delta = \beta\gamma$. Hence (α, β) is a multiple of (γ, δ):

$$\alpha = (k_3\gamma + k_4\delta)\alpha = (k_3\alpha + k_4\beta)\gamma, \tag{4.21a}$$

$$\beta = (k_3\gamma + k_4\delta)\beta = (k_3\alpha + k_4\beta)\delta. \tag{4.21b}$$

Since C_1 has to be free, $(k_3\alpha + k_4\beta)$ has to be a unit. Then there exists $l \in \mathbb{Z}_d$ such that

$$\omega(C_1, C_4) = k_2\alpha - k_1\beta = (k_3\alpha + k_4\beta)(k_2\gamma - k_1\delta) = (k_3\alpha + k_4\beta)(\omega(C_3, C_4) - lp). \tag{4.22}$$

That quantity should be both 0 and invertible and L cannot be symplectic. Then we shall make use of right-multiplications to complete the reduction. The matrices used for right-mutiplication need not be symplectic. Still, it is only possible to lower-trigonalise that way. Despite that restrictive result, we are to find another way of reducing that will prove sufficient to classify Lagrangian submodules in Section 4.2. We shall also need the

Criterion 12 *Let $a, x, y, z \in \mathbb{Z}_d$, $a \neq 0$, x a multiple of a and*

$$m = \begin{pmatrix} a & x \\ 0 & y \\ 0 & z \\ 0 & 0 \end{pmatrix}. \tag{4.23}$$

There exists a symplectic matrix S such that Sm is upper-triangular iff z is a multiple of y.

Proof. If z is a multiple of y, we can trigonalise m by applying substep 3.

Given $a, x, y, z \in \mathbb{Z}_d$ as specified in the criterion, $\delta = y \wedge z$ on the one hand and $k \in \mathbb{Z}_d, v \in U(\mathbb{Z}_d)$ on the other hand such that $\delta = ky + vz$, we perform the following left-mutiplication of m by a symplectic matrix S_4:

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 & kv^{-1} \\ 0 & 1 & 0 & 0 \\ 0 & k & v & 0 \\ 0 & 0 & 0 & v^{-1} \end{pmatrix}}_{S_4} \underbrace{\begin{pmatrix} a & x \\ 0 & y \\ 0 & z \\ 0 & 0 \end{pmatrix}}_{m} = \underbrace{\begin{pmatrix} a & x \\ 0 & y \\ 0 & \delta \\ 0 & 0 \end{pmatrix}}_{m'}, \tag{4.24}$$

There exists $k' \in \mathbb{Z}_d$ such that $y = k'\delta$ and let ν be the order of δ in $\mathbb{Z}_d$. In order not to burden the argument with unessential details, we refer to the Chinese remainder theorem to suppose that d is a power of a prime, say p^s. Let $t = v_p(a) < s$. If m' is symplectically trigonalisable as set out in the criterion, the symplectic matrix to use must be as shown in the following equation:

$$\begin{pmatrix} w + k_{11}p^{s-t} & * & * & * \\ k_{21}p^{s-t} & w^{-1} + k_{22}p^{s-t} & k_{23}p^{s-t} & k_{24}p^{s-t} \\ k_{31}p^{s-t} & \alpha_1 & -\alpha_1 k' + l_1\nu & \beta_1 \\ k_{41}p^{s-t} & \alpha_2 & -\alpha_2 k' + l_2\nu & \beta_2 \end{pmatrix} \begin{pmatrix} a & x \\ 0 & y \\ 0 & \delta \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} wa & * \\ 0 & * \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \tag{4.25}$$

with $w \in U(\mathbb{Z}_d)$. We leave the checking of that form to the reader. But the symplectic inner product of the third and fourth columns of that matrix has to be 1, which proves with Bézout's theorem that k' and ν are coprime. Let $\alpha, \beta \in \mathbb{Z}_d$ be such that $\alpha k' + \beta\nu = 1$. Then $\alpha y = \alpha k'\delta = (1 - \beta\nu)\delta = \delta$. ■

We can now state our

▶ **Substep 4**: Let $x, y, z \in \mathbb{Z}_d$, $\delta = y \wedge z$ and $X = (x, y, z, 0)^T$ with respect to some symplectic basis. One can find a new symplectic basis in which X is written $(x, y, \delta, 0)^T$. The way to do so is given in (4.24).

In what follows, we shall need a refined version of the algorithm $\mathscr{A}$. Recall that for any $2n \times k$ matrix m, $k \geq 1$, there exists an $k \times k$ invertible matrix $R(m)$ such that $\mathscr{A}(m) = mR(m)$. For any $2n \times k$ matrix m, $i \in \{1, \dots, 2n\}$, $j \in \{1, \dots, k-1\}$, and $m_{[i,j]}$ the $(i, \dots, 2n; j, \dots, k)$ submatrix of m, $\mathscr{A}_{i,j}$ will be the algorithm defined by

$$\mathscr{A}_{i,j}(m) = m \begin{pmatrix} I_{j-1} & 0_{j-1,k-j+1} \\ 0_{k-j+1,j-1} & R(m_{[i,j]}) \end{pmatrix}. \tag{4.26}$$

$\mathscr{A}_{i,j}$ does essentially the same job as $\mathscr{A}$ on columns j to k of m, but it takes into account only the last $2n - i + 1$ rows to maximise the order and combines those columns on the other lines accordingly.

We now go on with the symplectic reduction algorithm for a single Chinese factor. We suppose that $d = p^s$.

Symplectic reduction algorithm $\mathscr{S}$: Suppose we are given a basis $b = (b_1, \dots, b_k)$ of a submodule M of $\mathbb{Z}_d^{2n}$ and B is the matrix of size $2n \times k$ whose i-th column is b_i. To reduce B in a symplectic way, the starting point is $i = j = 1$ and $B' = B$, where i and j are some counters. Then while $i \leq 2n - 3$ and $j \leq k - 1$, that is to say while there remain at least four lines and two columns to deal with, do:

1. Apply $\mathscr{A}_{i,j}$ to B' and perform a first left-multiplication by a symplectic matrix in order to set to 0 all the coefficients in the j-th column sarting from the $(i+1)$-th line. We obtain a matrix $B^{(1)}$.

2. Apply $\mathscr{A}_{i+1,j+1}$ to $B^{(1)}$ and perform a second left-multiplication by a symplectic matrix to set to 0 all the coefficients in the $(j+1)$-th column sarting from the $(i+4)$-th line. Indeed, as we see with the example above (Equation 4.20), a step further as we planned to make it in the substeps could affect the j-th column in a wrong way. We obtain a matrix $B^{(2)}$ whose $(i, \dots, i+3; j, j+1)$ submatrix is
$$\begin{pmatrix} b_{i,j}^{(1)} & b_{i,j+1}^{(1)} \\ 0 & b_{i+1,j+1}^{(1)} \\ 0 & b_{i+2,j+1}^{(2)} \\ 0 & b_{i+3,j+1}^{(2)} \end{pmatrix}. \tag{4.27}$$

3. Performing substeps 2 and 4, we get a matrix $B^{(3)}$ whose $(i, \dots, i+3; j, j+1)$ submatrix is of the form
$$\begin{pmatrix} b_{i,j}^{(1)} & b_{i,j+1}^{(1)} \\ 0 & xb_{i+2,j+1}^{(3)} \\ 0 & b_{i+2,j+1}^{(3)} \\ 0 & 0 \end{pmatrix}, \tag{4.28}$$
with $x \in \mathbb{Z}_d$. Notice the line index on the second line.

4. If x is a unit, apply substep 3 to get a matrix $B^{(4)}$. If x is not a unit, just take $B^{(4)} = B^{(3)}$.

5. Since every coefficient in the $(i, \dots, 2n; j+1, \dots, k)$ submatrix of $B^{(4)}$ is a multiple of $b^{(4)}_{i,j} = b^{(1)}_{i,j}$, right-multiply $B^{(4)}$ by an appropriate, invertible matrix R_1 to set to 0 the coefficients on the i-th row starting from the $(j+1)$-th column. We obtain a matrix $B^{(5)}$.

6. If x is a unit, since every coefficient in the $(i+1, \dots, 2n; j+2, \dots, k)$ submatrix of $B^{(5)}$ is a multiple of $b^{(5)}_{i+1,j+1} = b^{(4)}_{i+1,j+1}$, right-multiply $B^{(5)}$ by an appropriate, invertible matrix R_2 to set to 0 the coefficients on the $(i+1)$-th row starting from the $(j+2)$-th column. We obtain a new matrix B'. If x is not a unit, just take $B' = B^{(5)}$.

7. If x is a unit, increase i and j by 2. If not, increase i by 2 and j by 1 only. In this latter case, we need not perform Step 1 at the next pass: The new $B^{(1)}$ is just the new B'.

Once this repeating process is finished, if $i = 2n-1$, we may reduce the last two rows by means of substep 2 and a right-mutiplication, so as to have at most two nonzero coefficients on them. If $i < 2n-1$ and $j = k$, apply a last left-multilplication by a symplectic matrix to B' so as to reduce the last column as far as possible without modifying the others. As an example, if we started with a $2n \times 2n$ matrix B with $n = 8$, we may end up with a matrix of the form

$$\mathscr{S}(B) = \left(\begin{array}{cccccccccccccccc}
* &&&&&&&&&&&&&&& \\
& * &&&&&&&&&&&&&& \\
\hline
&& * &&&&&&&&&&&&& \\
&&& * &&&&&&&&&&&& \\
\hline
&&&& * &&&&&&&&&&& \\
&&&&& R & * & * & * & * & * & * & * & * & * & * \\
\hline
&&&&& * &&&&&&&&&& \\
&&&&&& * &&&&&&&&& \\
\hline
&&&&&&& * &&&&&&&& \\
&&&&&&&& R & * & * & * & * & * & * & * \\
\hline
&&&&&&&& * &&&&&&& \\
&&&&&&&&& * &&&&&& \\
\hline
&&&&&&&&&& * &&&&& \\
&&&&&&&&&&& * &&&& \\
\hline
&&&&&&&&&&&& * &&& \\
&&&&&&&&&&&&& * &&
\end{array}\right), \tag{4.29}$$

where the meaning of the letter R is explained below. ♦

Horizontal lines of stars in (4.29) beginning with an R will be called rent lines and places marked with an R rent points. A rent line can occur only on an even row. Suppose (i, j) is a rent point in the reduced matrix. Every coefficient in the $(i, \dots, 2n; j, \dots, k)$ submatrix is a multiple of the coefficient underneath the rent point, at position $(i+1, j)$. So, if this coefficient is 0, we may stop the algorithm. Last but not least about rents, it was necessary to perform the algorithm in a single Chinese factor, since a rent may occur at some position in some Chinese factor while not in another one. This reduction procedure is thus linked in an essential way to the Chinese remainder theorem.

Example: We take again the basis matrix B of a submodule in $\mathbb{Z}_9^4$ as defined in (4.11). Since the first column vector is free, we may suppose that $\mathscr{A}_{1,1}$ leaves B unchanged and so Step 1 gives

$$B^{(1)} = B. \tag{4.30}$$

Then in Step 2, we reduce the first column by applying substeps 1 to 3 plus Euclid's algorithm. We have:

$$S_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 8 & 0 \\ 0 & 0 & 1 & 0 \\ 8 & 0 & 0 & 1 \end{pmatrix}, \qquad S_1 B = \begin{pmatrix} 3 & 3 & 6 & 3 \\ 1 & 7 & 5 & 0 \\ 2 & 8 & 1 & 0 \\ 5 & 2 & 7 & 6 \end{pmatrix}, \tag{4.31}$$

$$S_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 7 & 8 \end{pmatrix}, \qquad S_2 S_1 B = \begin{pmatrix} 3 & 3 & 6 & 3 \\ 1 & 7 & 5 & 0 \\ 1 & 4 & 5 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}, \tag{4.32}$$

$$S_3 = \begin{pmatrix} 1 & 0 & 0 & 8 \\ 0 & 1 & 0 & 0 \\ 0 & 8 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad S_3 S_2 S_1 B = \begin{pmatrix} 3 & 3 & 6 & 0 \\ 1 & 7 & 5 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}, \tag{4.33}$$

$$S_2' = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 6 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad B^{(2)} = S_2' S_3 S_2 S_1 B = \begin{pmatrix} 1 & 7 & 5 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}. \tag{4.34}$$

We get a matrix $B^{(2)}$ which is already under the form required at the end of Step 3. Moreover $x = 0$, so that after Step 4 we get

$$B^{(4)} = B^{(2)}. \tag{4.35}$$

We then perform Step 5 to get $B^{(5)}$:

$$R_1 = \begin{pmatrix} 1 & 2 & 4 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad B^{(5)} = B^{(4)} R_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}. \tag{4.36}$$

In Step 6, we get

$$B' = B^{(5)} \tag{4.37}$$

since x is not a unit. Finally in Step 7, we put

$$i = 3, \qquad j = 2. \tag{4.38}$$

Since $i = 2n - 1$, the process go out of the loop and we just have to right-multiply

B' to reorder its $2n-j+1=3$ last column vectors. We thus get

$$\mathscr{S}(B) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix}. \tag{4.39}$$

There is a single rent with all coefficients on the rent line equal to 0. □

We see with that simple example that we could have reordered the 3 last columns otherwise so as to get a diagonal matrix. But in case there are several rents or even a single one with nonzero coefficients, it may be more difficult to see. So we want to improve our algorithm in order to automatically achieve diagonalisation whenever possible. Indeed, the algorithm $\mathscr{S}$ consists in choosing basis vectors $f_1, \ldots, f_{2n}$ one after the other so as to obtain a basis matrix of a particular form with respect to the free basis f thus constituted. But can we avoid rents by a judicious choice of the f_i's so as to get a diagonal basis matrix? Is it a good strategy to choose a vector of the greatest possible order as we did? If the issue of order has actually to be addressed, is it of some use to discriminate between the vectors of a given order? We answer these questions in Section 4.3, but we are now ready to classify Lagrangian submodules.

4.2 Classification of Lagrangian submodules

We can now use the symplectic reduction algorithm to find a very simple form for a minimal basis matrix of a Lagragian submodule M. As we saw, we are to suppose that $d = p^s$ is a power of a prime. Let B_0 be a basis matrix for M. The symplectic reduction $B = \mathscr{S}(B_0)$ is still a basis matrix for M. Suppose some coefficient appears on an even row, say at position $(2i, j)$, without a rent. Since M is isotropic, the symplectic product of the $(2i-1)$-th and the $(2i)$-th column vectors of $\mathscr{S}(B)$ must be zero, which can be written

$$v_p(\mathscr{S}(B)_{2i-1,j-1}) + v_p(\mathscr{S}(B)_{2i,j}) \geq s. \tag{4.40}$$

The maximality of M implies that this is in fact an equality. On the contrary, if there is a rent point at position $(2i, j)$ and if the coefficient of $\mathscr{S}(B)$ at position $(2i-1, j-1)$ has p-valuation t, then, by maximality of M, the vector

$$C = (0, \ldots, 0, p^{s-t}, 0, \ldots, 0)^T \tag{4.41}$$

with p^{s-t} at the $(2i)$-th position, is in M. We insert this column at position $2i$, that is to say between the $(2i-1)$-th and the $(2i)$-th columns of $\mathscr{S}(B)$. Since M is isotropic, every coefficient on the $(2i)$-th line is a multiple of p^{s-t} and we may set to 0 every coefficient on this line at right of the new column. We apply this trick to each rent and obtain a diagonal matrix. So there exist $k \in \{1, \ldots, n\}$ and $s_1, \ldots, s_k \in \{0, \ldots, s\}$ so that the diagonal matrix

$$D = \mathrm{diag}(p^{s_1}, p^{s-s_1}, p^{s_2}, p^{s-s_2}, \dots, p^{s_k}, p^{s-s_k}) \tag{4.42}$$

is a basis matrix for M. If $k < n$, then M would not be maximal. One could add for instance the vector

$$(0, \dots, 0, 1, 0, \dots, 0)^T \tag{4.43}$$

with 1 at the $(2k+1)$-th position and get a greater isotropic submodule. So $k = n$. By construction of $\mathscr{S}(B)$, $s_i \leq s_j$ whenever $i < j$. Also note that C, as a vector of M, has to be a linear combination of the column vectors of $\mathscr{S}(B)$. Since our trick to make good a rent always yields a new basis matrix for M, the same is true for every additional column. So, whether a diagonal coefficient of D on an even row appeared while dealing with a rent or not, our use of the algorithm $\mathscr{A}$ warrants that for all $i \in \{1, \dots, n\}$, $s_i \leq s - s_i$.

Before we conclude, a remark is in order. Suppose the $(2i)$-th diagonal coefficient of D, $i \in \{1, \dots, n-1\}$, appeared while applying the algorithm $\mathscr{S}$ to B_0, that is to say there was no rent on the $(2i)$-th line. Then $s/2 \geq s_{i+1} \geq s - s_i \geq s/2$ and so, for $j \geq i$, $s_j = s/2$. If s is odd, there is necessarily a rent on every even row of $\mathscr{S}(B_0)$ except the last one.

Now, since these results do not depend on the Chinese factor we chose, we have proved the

Theorem 13 *Let M be a submodule of $\mathbb{Z}_d^{2n}$ and $d = \prod_{i \in I} p_i^{s_i}$ be the prime factor decomposition of d. Then M is Lagrangian iff the following two conditions are satisfied. There exists a unique family*

$$(d_1, \dots, d_n) \in \left\{1, \dots, \prod_{i \in I} p_i^{\lfloor s_i/2 \rfloor}\right\}^n \tag{4.44}$$

such that $d_1 | d_2 | \dots | d_n | d$ and there exists a $2n \times 2n$ symplectic matrix S such that

$$S \times \mathrm{diag}(d_1, d/d_1, d_2, d/d_2, \dots, d_n, d/d_n) \tag{4.45}$$

be a basis matrix for M.

Example: The reader may check that in the example set out in Relation (4.11) p.47, the column vectors of B are pairwise orthogonal for the symplectic product. Thus the submodule M whose basis matrix is B, or equivalently BR, is isotropic. It is in fact Lagrangian. Indeed, if we take as a new computational basis the column vectors of S^{-1}, a new basis matrix for M is SBR. Then a vector $x = (x_1, x_2, x_3, x_4) \in M^\omega$ has to verify

$$3x_1 = 0, \qquad 3x_2 = 0, \qquad x_4 = 0, \tag{4.46}$$

so that it is a member of M. Finally,

$$BR = S^{-1} \times \mathrm{diag}(3, 3, 1, 0), \tag{4.47}$$

in agreement with Theorem 13. This could also be seen in the same way from $\mathscr{S}(B)$ in (4.39). □

As we have already pointed out, Theorem 13 classifies the maximaly commuting sets of Pauli operators. Moreover, we shall see in Chapter 5 that it enables to know which of these sets are likely to provide one with MUBs. It also finds a direct application in the field of discrete Wigner distributions over a $\mathbb{Z}_d$-phase space as we develop in Chapter 6.

Note that the proof we gave of Theorem 13 is algorithmic as it is an adaptation of the algorithm $\mathscr{S}$ with the extra information that the input submodule is Lagrangian. But this is unsatisfactory if we do not know *a priori* whether a given submodule is Lagrangian. According to (4.45), the objects of interest are the elementary divisors of the submodule together with a symplectic matrix S. Since we already know how to compute the elementary divisors with the help of the algorithm $\mathscr{G}$ (see Appendix C p.140), we need another algorithm to find S that involves at most the knowledge of the elementary divisors.

So, as a mathematical complement to the present study, we expose in the next section how to diagonalise a given $2n \times r$ matrix by means of a symplectic change of computational basis. The results we get on symplectic diagonalisation will not be used in the remaining chapters and may be skipped by the reader interested only in applications.

4.3 Symplectic diagonalisation

Lagrangian submodules are quite a particular case. In this section, we first prove with an example that it is not always possible, for some submodule M, to find a symplectic basis f and a $2n \times 2n$ diagonal matrix D such that fD be a basis of M. The diagonal entries of the D need not be arranged by increasing valuations. If such a pair (f, D) exists, we shall say that M is nearly symplectic. Our aim will then be to provide a criterion to know if a given M is nearly symplectic. That will be done with the algorithm $\mathscr{D}_\omega$ that also yields the symplectic basis f if any. We shall eventually see that as Lagrangian submodules, symplectic ones form a particular kind of nearly symplectic submodules. For the sake of simplicity, we take again in this section $d = p^s$.

4.3.1 Preliminaries

We first recall the

Theorem 14 *For any rank-r submodule M of $\mathbb{Z}_d^{2n}$, there exist a free basis f of $\mathbb{Z}_d^{2n}$ and a minimal basis b of M such that:*

1. *b is represented by a diagonal $2n \times r$ matrix B with respect to f;*
2. *for all $i, j \in \{1, \ldots, r\}$, $i < j$, we have $b_{ii} | b_{jj}$.*

Moreover, for any pair (f, b) as above, the sequence $(d/\nu(b_{ii}))_{i\in\{1,\ldots,r\}}$ of the diagonal entries of B "without unit factors" is the same and therefore is a property of M. That sequence is called the sequence of the elementary divisors of M.

The pair (f, b) in Theorem 14 is not unique. For any submodule M of $\mathbb{Z}_d^{2n}$, we denote F_M the set of all free bases f of $\mathbb{Z}_d^{2n}$ such that M has a diagonal basis matrix with respect to f as in the theorem. We also put $\Sigma_{\mathscr{D}}(M)$ the subgroup of $\mathrm{GL}(n, \mathbb{Z}_d)$ that consists of all the change of basis matrices P between the bases in F_M.

Let us study the relation between the various bases $f \in F_M$. With the notations of Theorem 14, we will denote σ_i, $i \in \{0, \ldots, s-1\}$, the number of diagonal entries of B of the form up^j, $u \in U(\mathbb{Z}_{p^s})$, $j \leq i$, and as intervals in $\mathbb{N}$

$$\forall i \in \{0, \ldots, s-1\}, \quad K_i = \{\sigma_{i-1}+1, \ldots, \sigma_i\}, \tag{4.48a}$$

$$K_s = \{r+1, \ldots, n\}. \tag{4.48b}$$

Some of these intervals may be empty. Let $(f^{(1)}, b^{(1)})$ and $(f^{(2)}, b^{(2)})$ be two convenient pairs and P the $n \times n$ change of basis matrix defined by $f^{(1)}P = f^{(2)}$. For any $k \in \{0, \ldots, 2n\}$, there exists some $i_k \in \{0, \ldots, s\}$ so that $k \in K_{i_k}$. So $p^{i_k} f_k^{(2)} \in M$ and hence

$$\forall i \in \{i_k+1, \ldots, s\},\ \forall j \in K_i, \quad p^{i-i_k} | P_{jk}. \tag{4.49}$$

Since P is invertible, we also deduce from that latter result that for any $i \in \{0, \ldots, s\}$, the $(K_i; K_i)$ diagonal block of P is an invertible matrix.

As a converse, for any convenient pair (f, b) and any invertible matrix P satisfying Relation (4.49), let b' be the family represented by the matrix $fP\mathscr{D}(b)$ and N be the submodule of M generated by b'. Since P is invertible, fP is a free family and (fP, b') is a convenient pair for N. Hence M and N have the same sequence of elementary divisors and with the help of corollary 42 of Appendix C, p.140, we see that they have the same cardinality. So $N = M$ and (fP, b') is a convenient pair for M.

Finally, the Gram matrix will be the crux of our symplectic reduction algorithm. Let $c \in \{1, \ldots, 2n\}$ and $x = (x_1, \ldots, x_c)$ a family of vectors in $\mathbb{Z}_d^{2n}$. The Gram matrix of x, $G = \mathrm{Gram}(x)$, is the $c \times c$ matrix given by

$$\forall i, j \in \{1, \ldots, c\}, \quad g_{ij} = \omega(x_i, x_j). \tag{4.50}$$

With matrices, if B is the representative matrix of x with respect to the computational basis e, then $G = B^T J_n B$ and thus G is antisymmetric, but not necessarily invertible, even if x is free. Yet, if $c = 2n$ and x is a free basis of $\mathbb{Z}_d^{2n}$, then $B, G \in \mathrm{GL}(2n, \mathbb{Z}_d)$. The discriminant of x is the determinant of its Gram matrix:

$$\Delta(x) = \det(\mathrm{Gram}(x)). \tag{4.51}$$

By restriction, the K_i's defined above in (4.48a, 4.48b) determine a partition K' of $\{1, \ldots, c\}$:

$$\forall i \in \{0, \ldots, s\}, \quad K'_i = K_i \cap \{1, \ldots, c\}. \tag{4.52}$$

For every $(i,j) \in \{0,\dots,s\}^2$, G_{ij} will be the $(K'_i;K'_j)$ block of G. We also put $\widehat{G}_{ij}$ to be a matrix so that if G_{ij} is not the empty matrix and if $s_{ij} = v_p(G_{ij})$, then $v_p\big(\widehat{G}_{ij}\big) = 0$ and $p^{s_{ij}}\widehat{G}_{ij} = G_{ij}$. The matrix $\widehat{G}_{ij}$ thus pointed out is not unique if $s_{ij} > 0$. If G_{ij} is the empty matrix, then so is $\widehat{G}_{ij}$.

4.3.2 A counter-example

For now we take $c = 2n$ only. A simplified study upon Gram matrices enables us to give the simplest example of a non-nearly-symplectic submodule. The pattern we catch a glimpse of here about those matrices will be seen in its plain form afterwards. The reader who is interested only in the general case may skip to the next part.

Let $f \in F_M$. For all $i \in \{1,\dots,2n\}$, we define

$$\alpha_M(f_i) = \min(v_p(\omega(f_i,x));\ x \in M), \tag{4.53}$$

$$\beta(f,i) = \min(j \in \{0,\dots,s\};\ \exists k \in K_j,\ g_{ik} \in U(\mathbb{Z}_d)). \tag{4.54}$$

Since f is a free basis of $\mathbb{Z}_d^{2n}$, the matrix G is invertible and hence β is well-defined. The graph on Figure 4.1 illustrates the meaning of $\alpha_M(f_i)$ and $\beta(f,i)$. For any k and v, a plain bullet at position (k,v) indicates that $v_p(g_{ik}) = v$.

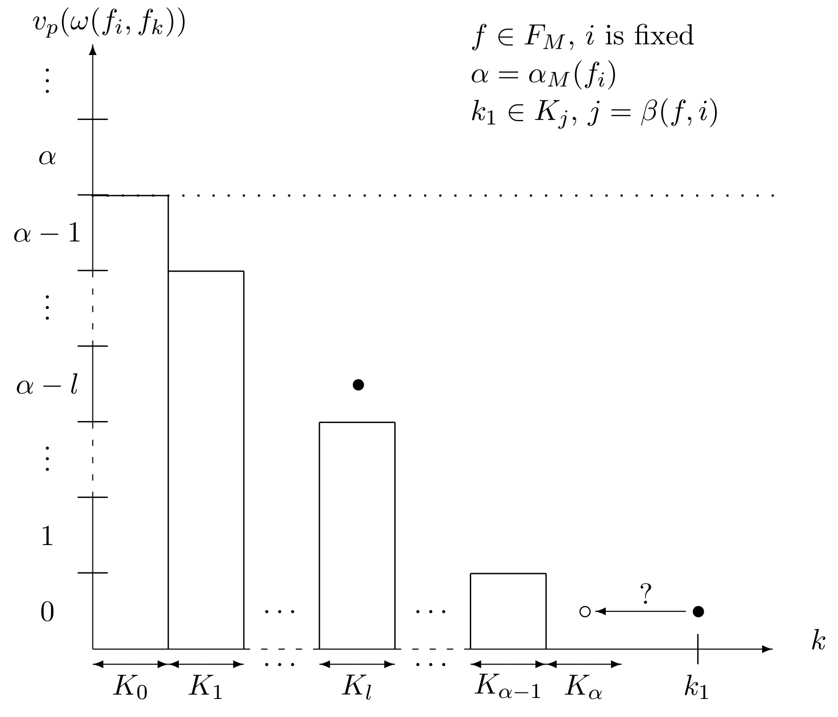

Figure 4.1: The functions α_M and β

So there must exist $l \in \{0,\dots,\alpha\}$ and $k_0 \in K_l$ so that $v_p(g_{ik_0}) = \alpha - l$. Let $i \in \{1,\dots,2n\}$, $j = \beta(f,i)$ and $k_1 \in K_j$ so that $g_{ik_1} \in U(\mathbb{Z}_d)$. Then $\alpha_M(f_i) \le v_p(\omega(f_i, p^j f_{k_1})) = j = \beta(f,i)$. This inequality is illustrated by the second plain bullet at position $(k_1, 0)$.

We then consider a nearly symplectic submodule M with a convenient pair (f, D). If $(v_p(d_{ii}))_{i=1,\dots,2n}$ is not an increasing sequence, we use a $2n \times 2n$ permutation matrix Q so that the diagonal coefficients of $Q^T DQ$ are arranged by increasing valuation. Let $f' = fQ \in F_M$. On each line of $\mathrm{Gram}(f') = Q^T J_n Q$, there is only one nonzero coefficient which is necessarily invertible, in fact 1 or -1, and it is clear that for all $i \in \{1, \dots, 2n\}$, $\alpha_M(f'_i) = \beta(f', i)$. On Figure 4.1, the equality $\alpha_M(f_i) = \beta(f, i)$ is checked iff $k_1 \in K_\alpha$.

We are now ready to find the announced non-nearly-symplectic submodule. Let $s > 1$ and M be the submodule generated by the column vectors of the matrix B in the following equation, with respect to e:

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1-p & 0 \\ 0 & -1 & p & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{L} \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}}_{B} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.55}$$

We left-multiply B by an invertible matrix L so as to obtain a diagonal matrix. Here, the diagonal coefficients of that latter matrix are already arranged by increasing valuation. $K_0 = \{1, 2\}$ and $K_s = \{3, 4\}$ are the only nonempty intervals K_i. The new computational basis is $f = eL^{-1} \in F_M$ and the Gram matrix of f is

$$G = L^{-T} J_n L^{-1} = \left(\begin{array}{cc|cc} 0 & p & -1+p & 0 \\ -p & 0 & 0 & 1 \\ \hline 1-p & 0 & 0 & 1 \\ 0 & -1 & -1 & 0 \end{array}\right), \tag{4.56}$$

with $L^{-T} = (L^{-1})^T$. Here, any matrix $P \in \Sigma_{\mathscr{D}}(M)$ (see definition p.4.3.1) is of the form

$$P = \begin{pmatrix} A_1 & A_2 \\ 0_{2,2} & A_3 \end{pmatrix}, \tag{4.57}$$

with $A_1, A_3 \in \mathrm{GL}(2, \mathbb{Z}_d)$. Then the Gram matrix of $f' = fP$ is of the form

$$P^T GP = \begin{pmatrix} pA_1^T \widehat{G}_{00} A_1 & A_4 \\ -A_4 & A_5 \end{pmatrix}, \text{ with } \widehat{G}_{00} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{4.58}$$

and $A_1^T \widehat{G}_{00} A_1, A_4 \in \mathrm{GL}(2, \mathbb{Z}_d)$. But we see that for any $i \in K_0 = \{1, 2\}$, $\alpha_M(f'_i) = 1 < \beta(f', i) = s$. Comparing to the result of the previous paragraph, this proves our claim that M is not nearly symplectic.

What if $s = 1$? In that case, the matrix obtained by swapping the second and third columns of B, namely $\mathrm{diag}(1, 0, 1, 0)$, is a convenient diagonal basis matrix for M with respect to the symplectic basis e.

4.3.3 General case

Let us now tell how to know in the general case whether a given submodule M is nearly symplectic or not and how to find a convenient pair (f, D) if any. We shall need a little more vocabulary.

Let b be a free basis of $\mathbb{Z}_d^{2n}$, $\sigma \in \mathfrak{S}_{2n}$ a permutation of $\{1, \ldots, 2n\}$ and Q the representative matrix of σ, that is to say the only nonzero coefficients of Q are equal to 1 and are located at the positions $(i, \sigma(i))_{i=1,\ldots,2n}$. We denote b_σ the free basis $(b_{\sigma(1)}, \ldots, b_{\sigma(2n)})$ of $\mathbb{Z}_d^{2n}$ and say that b is σ-symplectic if $b_\sigma = bQ^T$ is symplectic. In that case, the representative matrix of ω in basis b is $Q^T J_n Q$. A $2n \times 2n$ matrix L is said σ-symplectic if QLQ^T is symplectic or equivalently if

$$L^T(Q^T J_n Q)L = Q^T J_n Q. \tag{4.59}$$

Thus the conjugation by L preserves the matrix representative of ω in b, L is invertible and L^{-1} is still σ-symplectic. If b and L are σ-symplectic, $bL = bQ^T QL$ is still a σ-symplectic basis and if B is the representative matrix of b with respect to a σ-symplectic basis f, then B is a σ-symplectic matrix. Indeed, fQ^T and $bQ^T = (fQ^T)(QBQ^T)$ are symplectic bases and hence QBQ^T is a symplectic matrix.

The notions of scalar and set fringe we are going to define involve the K_i's and thus are meanigless unless a reference submodule or a suitable partition of $\{1, \ldots, 2n\}$ is specified. Let M be a submodule of $\mathbb{Z}_d^{2n}$. Define the K_i's accordingly and let κ be the map

$$\begin{array}{rcll} \kappa : \{1, \ldots, 2n\} & \longrightarrow & \{0, \ldots, s\}, & \text{such that } i \in K_{\kappa(i)}. \\ i & \longmapsto & \kappa(i) & \end{array} \tag{4.60}$$

Then for any Gram matrix G of size $\leq 2n$ and containing at least one unit, we define the scalar (M-)fringe of G by

$$\mathrm{fr}_M(G) = \min(\kappa(i) + \kappa(j);\ g_{ij} \in U(\mathbb{Z}_d)) \tag{4.61}$$

or equivalently

$$\mathrm{fr}_M(G) = \min(i + j;\ v_p(G_{ij}) = 0). \tag{4.62}$$

The (M-)fringe of G is the set of all coefficients g_{ij} such that $\kappa(i) + \kappa(j) \leq \mathrm{fr}_M(G)$. A block G_{ij} is said to be in the fringe of G if $\gamma_{ij} = \mathrm{fr}_M(G) - i - j \geq 0$. Whenever all the blocks G_{ij} in the fringe of G verify $v_p(G_{ij}) \geq \gamma_{ij}$, we shall say that the (M-)fringe of G is good. If there exists $i, j \in \{1, \ldots, 2n\}$ such that

$$g_{ij} \in U(\mathbb{Z}_d) \text{ while } \gamma_{\kappa(i)\kappa(j)} = 0 \tag{4.63a}$$

$$\forall k \leq i, \quad v_p(g_{kj}) \geq \gamma_{\kappa(k)\kappa(j)}, \tag{4.63b}$$

$$\forall l \leq j, \quad v_p(g_{il}) \geq \gamma_{\kappa(i)\kappa(l)}, \tag{4.63c}$$

we shall say that the (M-)fringe of G is nice. Of course a good M-fringe is a nice M-fringe. Let us give an example. If a block G_{ij} with $i+j=3$ contains a unit, then the following (Gram) matrix has a good fringe and scalar fringe 3:

$$G = \begin{pmatrix} \begin{array}{|c|c|c|c|} \hline p^3\widehat{G}_{00} & p^3\widehat{G}_{01} & p\widehat{G}_{02} & G_{03} \\ \hline p^2\widehat{G}_{10} & p\widehat{G}_{11} & pG_{12} & \multicolumn{1}{c}{} \\ \cline{1-3} p^3\widehat{G}_{20} & G_{21} & \multicolumn{2}{c}{} \\ \cline{1-2} G_{30} & \multicolumn{3}{c}{} \\ \cline{1-1} \end{array} & \cdots \\ \vdots & \end{pmatrix}. \tag{4.64}$$

We shall need the following lemma and corollary.

Lemma 15 *Let M be a submodule in $\mathbb{Z}_d^{2n}$. Let b be a free basis of $\mathbb{Z}_d^{2n}$ with Gram matrix G and assume that G has a good M-fringe. Then for any $P \in \Sigma_{\mathscr{D}}(M)$, P^TGP has a good M-fringe with the same scalar M-fringe as G.*

That is to say the form (4.64), with the particular scalar M-fringe required, is preserved under conjugation by a matrix in $\Sigma_{\mathscr{D}}(M)$.

Proof. The reference submodule is M. Let $H = GP$. For every block H_{ij} of H, we have

$$H_{ij} = \sum_{k=0}^{j-1} G_{ik}P_{kj} + G_{ij}P_{jj} + \sum_{k=j+1}^{s} G_{ik}P_{kj}. \tag{4.65}$$

As to the first sum, for every $k \in \{0, \ldots, j-1\}$, we have

$$v_p(G_{ik}) + v_p(P_{kj}) \geq v_p(G_{ik}) \geq \gamma_{ik} \geq \gamma_{ij} + 1 \tag{4.66}$$

and we refer to Relations (A.29a) and (A.29b) of Appendix A.2 to see that

$$v_p(G_{ik}P_{kj}) \geq \min(v_p(G_{ik}) + v_p(P_{kj}), s) \geq \gamma_{ij} + 1. \tag{4.67}$$

Since P_{jj} is invertible, the lines of $G_{ij}P_{jj}$ are of the same order as the lines of G_{ij} respectively and then

$$v_p(G_{ij}P_{jj}) = v_p(G_{ij}) \geq \gamma_{ij}. \tag{4.68}$$

As to the second sum, for every $k \in \{j+1, \ldots, s\}$, the inequality

$$v_p(G_{ik}) + v_p(P_{kj}) \geq (\mathrm{fr}(G) - i - k) + (k - j) = \gamma_{ij} \tag{4.69}$$

implies that

$$v_p(G_{ik}P_{kj}) \geq \gamma_{ij}. \tag{4.70}$$

So $v_p(H_{ij}) \geq \gamma_{ij}$. Let (i,j) be such that $\gamma_{ij} = 0$ and $v_p(G_{ij}) = 0$. Then the inequality in (4.69) may be modified as

$$\forall k \in \{j+1, \ldots, s\}, \quad v_p(G_{ik}) + v_p(P_{kj}) \geq 0 + (k-j) \geq 1, \tag{4.71}$$

and we see that $v_p(H_{ij}) = 0$. So H has a good fringe with scalar fringe $\mathrm{fr}(G)$. In the same manner, $P^T G P = P^T H$ has a good fringe with scalar fringe $\mathrm{fr}(G)$. ∎

Corollary 16 *Let M be a nearly symplectic submodule of $\mathbb{Z}_d^{2n}$ and $f \in F_M$. Then the matrix* $\mathrm{Gram}(f)$ *has a good M-fringe.*

Proof. By assumption, there exists $\sigma \in \mathfrak{S}_{2n}$ and f' a σ-symplectic basis in F_M. We have already seen that $G' = \mathrm{Gram}(f')$ has a good fringe (with respect to any submodule). Besides, there exists $P \in \Sigma_{\mathscr{D}}(M)$ so that $f = f'P$. So, $\mathrm{Gram}(f) = P^T G' P$ has a good M-fringe. ∎

We can now give the algorithm for symplectic diagonalisation whenever possible:

Algorithm $\mathscr{D}_\omega$: Let M be a submodule of $\mathbb{Z}_d^{2n}$, b a basis of M and B its representative basis matrix with respect to any computational basis e'.

Let $f = e' L(B)^{-1} \in F_M$, where $L(B)$ is defined within the algorithm $\mathscr{D}_0$ (see page 135), $\widetilde{M} = M$ and b' be the empty sequence with values in $\mathbb{Z}_d^{2n}$. Let also c be a counter with initial value 0.

While $c < n$ and $G = \mathrm{Gram}(f)$ has a nice $\widetilde{M}$-fringe, do

1. Choose a pair $(i, j) \in \{1, \dots, 2n - 2c\}^2$ that verifies Conditions (4.63) and perform the partial Gram-Schmidt orthogonalisation process:

$$f'_i = f_i, \quad f'_j = f_j, \tag{4.72a}$$
$$\forall k \in \{1, \dots, 2n-2c\} \setminus \{i, j\}, \quad f'_k = f_k - g_{ij}^{-1} g_{ik} f_j + g_{ij}^{-1} g_{jk} f_i. \tag{4.72b}$$

   Owing to the nice fringe condition, the corresponding change of basis matrix R is in $\Sigma_{\mathscr{D}}(M)$. With $i \leq j$ and $g_{ij} = 1$, it reads

$$R = \left(\begin{array}{ccccccccccc}
1 & & & & & & & & & & \\
 & \ddots & & & & & & & & & \\
 & & 1 & & & & & & & & \\
\hline
g_{j1} & \cdots & g_{j,i-1} & 1 & g_{j,i+1} & \cdots & g_{j,j-1} & 0 & g_{j,j+1} & \cdots & g_{j,2n} \\
\hline
 & & & & 1 & & & & & & \\
 & & & & & \ddots & & & & & \\
 & & & & & & 1 & & & & \\
\hline
-g_{i1} & \cdots & -g_{i,i-1} & 0 & -g_{i,i+1} & \cdots & -g_{i,j-1} & 1 & -g_{i,j+1} & \cdots & -g_{i,2n} \\
\hline
 & & & & & & & & 1 & & \\
 & & & & & & & & & \ddots & \\
 & & & & & & & & & & 1
\end{array}\right), \tag{4.73}$$

   where the two special rows are the i-th one and the j-th one respectively. For any $k \in \{1, \dots, 2n-2c\} \setminus \{i, j\}$, $f'_k \in \left\langle f'_i, f'_j \right\rangle^\omega$ and since $R \in \Sigma_{\mathscr{D}}(M)$, $f' = fR \in F_M$.

2. Let b' be the concatenation of b' and $(g_{ij}^{-1} f'_i, f'_j)$.

3. Rename $\widetilde{M} \cap \left\langle f'_i, f'_j \right\rangle^{\omega}$ as $\widetilde{M}$.

4. Rename $f' \setminus \{f'_i, f'_j\}$ as f.

5. Increase c by 1.

Whenever $\mathscr{D}_\omega(b) = \mathscr{D}_\omega(e', B) = b'$ has cardinality $2n$, then it is a symplectic basis of $\mathbb{Z}_d^{2n}$, M is nearly symplectic and there exists $\sigma \in \mathfrak{S}_{2n}$ so that $b'_\sigma \in F_M$. ♦

Example: Before we give the proof of the algorithm, let us process through an example. We take over, in $\mathbb{Z}_9^4$, the submodule M defined in Relation (4.11) p.47 by its basis matrix B with respect to the canonical computational basis, $e' = e$. We already know that

$$K_0 = \{1\}, \quad K_1 = \{2, 3\}, \quad K_2 = \{4\}. \tag{4.74}$$

Knowing the K_i's is a requisite in order to apply the algorithm $\mathscr{D}_\omega$. The column vectors of the following matrix F form a basis f in F_M:

$$F = \begin{pmatrix} 6 & 0 & 4 & 3 \\ 6 & 5 & 0 & 7 \\ 5 & 6 & 7 & 6 \\ 2 & 6 & 1 & 5 \end{pmatrix}, \qquad f = eF \in F_M. \tag{4.75}$$

Indeed, $F^{-1}B$ is clearly amenable to $\operatorname{diag}(1, 3, 3, 0)$ by means of a right-multiplication by an invertible matrix:

$$F^{-1} = \begin{pmatrix} 4 & 3 & 8 & 0 \\ 4 & 2 & 1 & 5 \\ 8 & 0 & 3 & 3 \\ 0 & 0 & 4 & 8 \end{pmatrix}, \qquad F^{-1}B = \begin{pmatrix} 1 & 4 & 5 & 3 \\ 0 & 6 & 0 & 6 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{4.76}$$

Note that it is not necessary to get P with the help of the algorithm $\mathscr{D}_0$. Any basis $f \in F_M$ is convenient. Then $\widetilde{M}$, b' and c are initialised to M, the empty sequence with values in $\mathbb{Z}_d^{2n}$ and 0, respectively.

The counter c is less than $n = 2$ and the Gram matrix of $f = eF$ is

$$G = \left(\begin{array}{c:cc:c} 0 & 3 & 3 & 1 \\ \hdashline 6 & 0 & 7 & 6 \\ 6 & 2 & 0 & 3 \\ \hdashline 8 & 3 & 6 & 0 \end{array}\right). \tag{4.77}$$

It obviously has a good M-fringe.

1. In Step 1 we choose

$$(i,j) = (2,3), \tag{4.78}$$

which gives

$$g_{ij}^{-1} = g_{23}^{-1} = 4 \tag{4.79}$$

and for R and f':

$$R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 6 & 1 & 0 & 3 \\ 3 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad PR = \begin{pmatrix} 0 & 0 & 4 & 6 \\ 0 & 5 & 0 & 4 \\ 8 & 6 & 7 & 0 \\ 5 & 6 & 1 & 8 \end{pmatrix}. \tag{4.80}$$

Comparing to F, we see that the second and third columns have been left unchanged in FR.

2. In Step 2, we collect $4f_2'$ and f_3' and find in matrix form:

$$b' = \begin{pmatrix} 0 & 4 \\ 2 & 0 \\ 6 & 7 \\ 6 & 1 \end{pmatrix}. \tag{4.81}$$

3. In Step 3, given the remark at the end of Step 1 in the statement of $\mathscr{D}_\omega$, the new submodule $\widetilde{M}$ is generated by $\mu_1 f_1'$ and $\mu_4 f_4'$:

$$\widetilde{M} = \langle \mu_1 f_1', \mu_4 f_4' \rangle \tag{4.82}$$

for some $\mu_1, \mu_4 \in \mathbb{Z}_9$. Since $1 \in K_0$ and $4 \in K_2$, we may choose

$$\mu_1 = 1, \quad \mu_4 = 0. \tag{4.83}$$

Thus

$$\widetilde{M} = \langle f_1' \rangle . \tag{4.84}$$

4. In Step 4, we simply put, in matrix form,

$$f = (f_1', f_4') = \begin{pmatrix} 0 & 6 \\ 0 & 4 \\ 8 & 0 \\ 5 & 8 \end{pmatrix}. \tag{4.85}$$

5. Then in Step 5, c is set to 1.

Since $c = n - 1 = 1$, the forthconing passage in the loop will be the last one after we check the fringe condition. The new K_i's are

$$K_0 = \{1\}, \quad K_1 = \emptyset, \quad K_2 = \{2,3,4\} \tag{4.86}$$

and the new Gram matrix is

$$G = \mathrm{Gram}(f_1', f_4') = \left(\begin{array}{c:c} 0 & 1 \\ \hdashline 8 & 0 \end{array}\right), \tag{4.87}$$

which again has a good $\widetilde{M}$-fringe. This time, we have nothing to do in Step 1:

$$R = I_2, \qquad f' = f. \tag{4.88}$$

Thus in Step 2, we get in matrix form

$$b' = P = \begin{pmatrix} 0 & 4 & 0 & 6 \\ 2 & 0 & 0 & 4 \\ 6 & 7 & 8 & 0 \\ 6 & 1 & 5 & 8 \end{pmatrix}. \tag{4.89}$$

Steps 3 and 4 do not matter: They simply exhaust $\widetilde{M}$ and f'. Then c is set to 2 and the algorithm stops.

We have

$$P^{-1} = \begin{pmatrix} 0 & 5 & 1 & 2 \\ 7 & 0 & 3 & 6 \\ 4 & 3 & 8 & 0 \\ 0 & 0 & 4 & 8 \end{pmatrix}. \tag{4.90}$$

So, in basis $\mathscr{D}_\omega(e, B) = b'$, we obtain the following basis matrix for M:

$$P^{-1}M = \begin{pmatrix} 0 & 6 & 0 & 3 \\ 3 & 3 & 6 & 3 \\ 1 & 4 & 5 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{4.91}$$

which is diagonalisable via a right-miltiplication by an invertible matrix:

$$P^{-1}M \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 8 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 2 & 0 \end{pmatrix} = \mathrm{diag}(3, 3, 1, 0). \tag{4.92}$$

This finish proving that the algorithm $\mathscr{D}_\omega$ works for the particular case we dealt with. □

We now prove the algorithm in full generality. Since $e'L(B)^{-1}$ is a free basis of $\mathbb{Z}_d^{2n}$, its Gram matrix is invertible and thus has a well-defined M-fringe. Then the discriminant

$$\Delta(f' \setminus \{f_i', f_j'\}) = g_{ij}^{-2}\Delta(f') = g_{ij}^{-2}\Delta(f) \tag{4.93}$$

being a unit, all the forthcoming matrices G have a well-defined $\widetilde{M}$-fringe and $\mathscr{D}_\omega$

is a valid algorithm. Now if M is nearly symplectic, does this algorithm yields the matrix b' we search for? Besides, in Step 1, the pair (i, j) is not unique. So we are to prove that if M is nearly symplectic, the algorithm $\mathscr{D}_\omega$, with any choice of the pairs, builds a symplectic basis b' that endows M with a diagonal basis matrix.

Let M be a nearly symplectic submodule of $\mathbb{Z}_d^{2n}$, $\sigma \in \mathfrak{S}_{2n}$ with representative matrix Q and h a σ-symplectic basis in F_M. Let also B be a basis matrix of M with respect to any computational basis e' and let us carry out the algorithm. Corollary 16 shows that the first Gram matrix G has a nice M-fringe. We choose a convenient pair (i, j) and find the first two vectors of b' by performing Steps 1 and 2. Then with $N = M \cap \left\langle f'_i, f'_j \right\rangle^\omega$, we want to show that the Gram matrix of $f^\flat = f' \setminus \{f'_i, f'_j\}$ has a nice N-fringe. From now on, we consider N as a submodule of $\left\langle f^\flat \right\rangle$ exclusively. With that convention, $f^\flat \in F_N$, and Corollary 16 tells us that it suffices to show that N is nearly symplectic.

There exists $P \in \Sigma_{\mathscr{D}}(M)$ such that $f' = hP$. For any $m \in \{1, \dots, 2n\}$, let $\ell(m)$ be the index defined by $\omega(h_m, h_{\ell(m)}) = \pm 1$. Since $\mathrm{fr}_M(h) = \mathrm{fr}_M(f')$ as shown by Lemma 15, we have

$$\kappa(m) + \kappa(\ell(m)) \geq \kappa(i) + \kappa(j) \tag{4.94}$$

and hence

$$\kappa(m) < \kappa(i) \implies \kappa(\ell(m)) > \kappa(j), \tag{4.95a}$$

$$\kappa(m) < \kappa(j) \implies \kappa(\ell(m)) > \kappa(i). \tag{4.95b}$$

So, and because $\omega(f'_i, f'_j) = g_{ij}$ is a unit, there exist $k \in K_{\kappa(i)}$ and $l = \ell(k) \in K_{\kappa(j)}$ so that the coefficients p_{ki} and p_{lj} in P are units. So QP has a unit at position $(\sigma^{-1}(k), i)$. Let L be a symplectic matrix so that the i-th column of LQP has all but its $\sigma^{-1}(k)$-th coefficient equal to 0. Since we suppose we know where an invertible coefficient is in the i-th column of P, the substeps of the symplectic reduction algorithm are not necessary to find L. Instead, we form a symplectic matrix inspired by the Gaussian reduction. For instance, if $\sigma^{-1}(k) = 1$, then $\sigma^{-1}(l) = 2$ and L is of the form

$$L = \begin{pmatrix} 1 & & & & & & \\ k_0 & 1 & -k_4 & k_3 & \cdots & -k_{2n} & k_{2n-1} \\ k_3 & & 1 & & & & \\ k_4 & & & 1 & & & \\ \vdots & & & & \ddots & & \\ k_{2n-1} & & & & & 1 & \\ k_{2n} & & & & & & 1 \end{pmatrix}. \tag{4.96}$$

Then the i-th column of $P' = Q^T LQP$ has all but its k-th coefficient equal to 0. The basis $h' = hQ^T L^{-1} Q$ is still σ-symplectic and $f' = hP = h'P'$. Moreover, the matrix $Z = Q^T LQ$ is in $\Sigma_{\mathscr{D}}(M)$ and thus $P' \in \Sigma_{\mathscr{D}}(M)$. Indeed, the coefficients in the k-th column of Z have the right valuations by construction. The coefficients on the l-th row not in the k-th nor in the l-th columns were determined so that Z is

σ-symplectic. In particular:

$$\forall m \in \{1, \ldots, l-1\} \setminus \{k\}, \quad \omega(Z_m, Z_k) = \pm z_{\ell(m),k} \pm z_{lm} = 0, \tag{4.97}$$

where for all i, Z_i is the i-th column vector of Z. And according to Relation (4.94),

$$\forall m \in \{1, \ldots, l-1\} \setminus \{k\}, \quad v_p(z_{\ell(m),k}) \geq \kappa(\ell(m)) - \kappa(i) \geq \kappa(l) - \kappa(m). \tag{4.98}$$

Thus $v_p(z_{lm}) \geq \kappa(l) - \kappa(m)$ and that proves that $Z \in \Sigma_{\mathscr{D}}(M)$. The coefficient p'_{lj} divides g_{ij} and hence is a unit. So we apply the same kind of reduction as before to the j-th column of P' while preserving the i-th one and find a σ-symplectic basis h'' and an invertible matrix $P'' \in \Sigma_{\mathscr{D}}(M)$ so that $f' = h''P''$. We may suppose without loss of generality that $g_{ij} = 1$. Then the vectors h''_k and h''_l may be redefined under a multiplication by a unit factor so that $p''_{ki} = p''_{lj} = 1$. If we assume that $i < j$ and $k < l$ for instance, P'' is of the form

$$P'' = \begin{array}{cc} \left(\begin{array}{c|c|c|c|c} * & & * & & * \\ \hline & 1 & & 0 & \\ \hline * & & * & & * \\ \hline & 0 & & 1 & \\ \hline * & & * & & * \end{array}\right) & \begin{array}{l} \\ \leftarrow \; k \\ \\ \leftarrow \; l \\ \\ \end{array} \\ \begin{array}{ccccc} & \uparrow & & \uparrow & \\ & i & & j & \end{array} & \end{array} \tag{4.99}$$

Let $h^\flat = h'' \setminus \{h''_i, h''_j\}$ and $P^\flat$ be the matrix obtained by deleting the k-th and l-th rows as well as the i-th and j-th columns of P''. Now $f'_i = h''_k$ and $f'_j = h''_l$ so that $f^\flat = h^\flat P^\flat$. By construction, $P^\flat \in \Sigma_{\mathscr{D}}(N)$. So $h^\flat \in F_N$. But since h'' is σ-symplectic and $\omega(h''_k, h''_l) = 1$, there exists $\rho \in \mathfrak{S}_{2n-2}$ such that $h^\flat$ is ρ-symplectic. That proves that N is nearly symplectic.

4.3.4 Symplectic submodules

We end this section with a proposition that shows the difference between symplectic and nearly symplectic submodules.

Proposition 17 *Let M be a submodule of $\mathbb{Z}_d^{2n}$. Then M is symplectic iff M is nearly symplectic and such that $M + M^\omega = \mathbb{Z}_d^{2n}$. In that case, M is free and of even rank.*

Proof. If $M = \{0\}$, both terms of the equivalence are checked and M is obviously free and of even rank. So let M be a nonzero symplectic submodule and let $f \in F_M$. Since $p^{s-1}f_1 \in M \setminus \{0\}$ and $M \cap M^\omega = \{0\}$, there exists $x = \sum_{i=2}^{2n} x_i f_i \in M$ such that $\omega(p^{s-1}f_1, x) \neq 0$. Thus x is free, $\omega(f_1, x)$ is a unit, there exists $j \in K_0 \setminus \{1\}$ so that $\omega(f_1, f_j)$ is a unit and $f_1 \in M$. That proves that Gram(f) has a good fringe.

We then perform the partial Gram-Schmidt process and find a new basis $f' \in F_M$:

$$f'_1 = f_1, \quad f'_j = f_j, \tag{4.100a}$$

$$\forall k \in \{1, \ldots, 2n\} \setminus \{1, j\}, \quad f'_k = f_k - g_{1j}^{-1} g_{1k} f_j + g_{1j}^{-1} g_{jk} f_1. \tag{4.100b}$$

Since $2, j \in K_0$, we may rename without loss of generality f'_j as f'_2 and f'_2 as f'_j. Let $N = M \cap \langle f'_1, f'_2 \rangle^\omega$ and let y be some nonzero vector in N if any:

$$y = \sum_{i=3}^{r} y_i f'_i \in N \setminus \{0\}, \tag{4.101}$$

with r the rank of M. Since M is symplectic, there exists $z \in M$ so that $\omega(y, z) \neq 0$:

$$z = \sum_{i=1}^{r} z_i f'_i \in M. \tag{4.102}$$

But with $z' = z - z_1 f'_1 - z_2 f'_2 \in N$, we also have $\omega(y, z') = \omega(y, z) \neq 0$. Hence $y \notin N^\omega$ and N is symplectic. If M is larger than $\langle f'_1, f'_2 \rangle$, then $N \neq \emptyset$. We carry out again the same reasoning until we find a free basis h of $\mathbb{Z}_d^{2n}$ the first r vectors of which form a symplectic basis of M. Moreover, the last $2n - r$ vectors of h form a free basis $h^\flat$ of M^ω. Up to now, we proved that M is free, of even rank and such that $M \oplus M^\omega = \mathbb{Z}_d^{2n}$.

Since $\Delta(h^\flat) = \Delta(h)$ is a unit, then in the same manner as we showed the validity of $\mathscr{D}_\omega$, we see that we can apply the entire Gram-Schmidt orthogonalisation process to $h^\flat$. Hence, M is nearly symplectic.

Let us show the converse. Let f be a symplectic basis of $\mathbb{Z}_d^{2n}$ and D the following $2n \times 2n$ diagonal matrix such that fD is a basis matrix for M:

$$D = \operatorname{diag}(p^{s_1}, p^{s_2}, \ldots, p^{s_{2n-1}}, p^{s_{2n}}). \tag{4.103}$$

Then this other diagonal matrix D' is such that fD' is a basis matrix for M^ω:

$$D' = \operatorname{diag}(p^{s-s_2}, p^{s-s_1}, \ldots, p^{s-s_{2n}}, p^{s-s_{2n-1}}). \tag{4.104}$$

Under the assumption that $M + M^\omega = \mathbb{Z}_d^{2n}$, we have

$$s_1 < s \Rightarrow s - s_1 \geq 1 \Rightarrow s_2 = 0 \Rightarrow s - s_2 \geq 1 \Rightarrow s_1 = 0. \tag{4.105}$$

The same reasoning is true starting with any $i \neq 1$ and thus M is free: Each of the d_{ii}'s is either 1 or 0. For any $i \in \{1, \ldots, n\}$, suppose that $f_{2i} \in M$ and let $x \in M, y \in M^\omega$ so that $f_{2i-1} = x + y$. Then

$$\omega(x, f_{2i}) = \omega(x + y, f_{2i}) = 1. \tag{4.106}$$

That proves that the component of x along f_{2i-1} is 1 and hence $f_{2i-1} \in M$. By the same token, f_{2i} is in M if f_{2i-1} is. Therefore M is symplectic and of even rank. ■

Chapter 5

Lagrangian half-modules and MUBs

In the first section of this chapter, we find a necessary and sufficient condition to know whether two maximal commuting sets of Pauli operators yields unbiased bases or not. This is achieved in Theorem 21. In the following sections, we express this result as a projective geometrical feature, we translate it into a graph interpretation and we give some basic properties of that graph. As Pauli operators and their eigenstates play quite a role in quantum information and quantum communication theory, and since the notion of a symplectic product appears to be central in their behaviour, we also show how the symplectic structure relates to the Clifford group.

5.1 A criterion to get unbiased bases

Let $d = p^s$ be a power of a prime, n a positive integer and $D = p^{sn}$. A vector $(a_1, b_1, \ldots, a_n, b_n) \in \mathbb{Z}_d^{2n}$ codes for the tensor product of the n Pauli operators $X^{a_i} Z^{b_i}$ over $\mathbb{C}^d$, i ranging from 1 to n, resulting in an operator on $\mathbb{C}^D$. We denote $q = \exp(2\pi i/d)$ the canonical root of unity of order d. Since the tensor Pauli operators are unitary and their d-th power is the identity operator, their eigenvalues are integer powers of q.

In order to determine a basis of $\mathbb{C}^D$ with the help of such tensor Pauli operators, we need a maximal set of commuting operators and diagonalise them simultaneously. In the vector language over $\mathbb{Z}_d$, we need a Lagrangian submodule of $\mathbb{Z}_d^{2n}$ and we would like to know a necessary and sufficient condition bearing on those submodules in order to check whether the corresponding operators give rise to unbiased bases of $\mathbb{C}^D$. To get such bases, the intersection of the two maximal commuting sets of Pauli operators has to be reduced to identity and so the intersection of the two corresponding Lagrangian submodules have to be reduced to $\{0\}$. This latter condition has a strong bearing on the form of the submodules we search for.

Definition 18 *A half-module of the $\mathbb{Z}_d$-module $\mathbb{Z}_d^{2n}$ is a free submodule of rank n, that is to say isomorphic to $\mathbb{Z}_d^n$.*

Proposition 19 *Let M and N be two Lagrangian submodules of $\mathbb{Z}_d^{2n}$. If $M \cap N = \{0\}$, then M and N are half-modules.*

Proof. We suppose that one of the two Lagrangian submodules is not a half-module, say N, and we are to show that we can build some nonzero vector $x \in M \cap N$. Let e be a symplectic basis of $\mathbb{Z}_d^{2n}$ such that M has a diagonal basis matrix with respect to e as in Theorem 13 p.55 and let B be a basis matrix for N. The matrix B has $r \geq n+1$ columns. Let

$$s_2 = \min(v_p(b_{2,j});\ j \in \{1, \dots, r\}) \tag{5.1}$$

and let us reorder the columns of B so that $v_p(b_{2,1}) = s_2$. We may then right-multiply B by an $r \times r$ invertible matrix so as to set to 0 all the $b_{2,j}$'s, $j \in \{2, \dots, r\}$. We call the matrix we get B again. Now we do the same job on every even line from the fourth line to the $2n$-th one. One calculates

$$s_{2i} = \min(v_p(b_{2,j});\ j \in \{i, \dots, r\}), \tag{5.2}$$

with i ranging from 2 to n, and reorders the $r-i+1$ last columns. The coefficients $b_{2i,j}$ are set to 0, with j ranging from $i+1$ to r. That way, we find in the $(n+1)$-th column of B a nonzero vector x_0 of N whose coefficients on even lines are all zero. Let $k = v_p(x_0)$ and $x = p^{s-1-k}x_0 \neq 0$. With $\widetilde{s} = \lfloor s/2 \rfloor$, the floor part of $s/2$, we know that for all $i \in \{0, \dots, n-1\}$, the vector $p^{\widetilde{s}}e_{2i+1}$ is in M. As x is a linear combination of them, the intersection of M and N is not reduced to $\{0\}$. ■

Hence, we shall be interested only in Lagrangian half-modules. Secondly, we are going to check that a Lagrangian half-module actually defines a basis of eigenvectors. Indeed, since we restrict the operators for diagonalisation to be only Pauli operators, it must be proved that a maximally commuting set of them corresponding to a Lagrangian half-module yields a well-defined basis.

Lemma 20 *Let M be a $(2n \times n)$ basis matrix of a Lagrangian half-module. Then:*

1. *M contains an $n \times n$ invertible submatrix.*
2. *There exists a $2n \times n$ matrix N such that $N^T J M = I_n$.*

Proof. (1) Since the first column vector of M is free, it contains an invertible coefficient, say on line i_1. We can right-multiply M by an invertible matrix R_1 so that this coefficient becomes 1 and any other coefficient on this line is set to 0. Then, since the column vectors of $M_1 = MR_1$ still generates the same Lagrangian half-module, the second column vector of M_1 is free and contains an invertible coefficient, say on line i_2. We can right-multiply M_1 by an invertible matrix R_2 so that this latter coefficient becomes 1 and any other coefficient on line i_2 is set to 0. The line i_1 is unchanged. We repeat this process until we get a matrix M_n whose $(i; 1, \dots, n)$ submatrix is the identity matrix. Thus the $(i; 1, \dots, n)$ submatrix of M is invertible. Let us call M' this submatrix.

(2) Let N_0 be a $2n \times n$ matrix such that $N_0^T JM = M'$ and let $R = \prod_{k=1}^n R_k$. Then $(N_0 R^T)^T JM = I_n$ and we put $N = N_0 R^T$. ∎

With the notations of the lemma , let $P_1, \ldots, P_n$ be the Pauli operators corresponding to the column vectors of M and $Q_1, \ldots, Q_n$ those corresponding to the column vectors of N. Hence

$$\forall i, j \in \{0, \ldots, n\}, \quad i \neq j \Longrightarrow Q_j^\dagger P_i Q_j = P_i, \tag{5.3a}$$

$$\forall i \in \{0, \ldots, n\}, \quad Q_i^\dagger P_i Q_i = q P_i. \tag{5.3b}$$

Let also $k \in \{0, \ldots, p^s - 1\}^n$. The operator

$$\pi_k = \frac{1}{d^n} \bigotimes_{i=1}^{n} \sum_{j=1}^{d} (q^{-k_i} P_i)^j \tag{5.4}$$

is the projector on the common eigenspace of $P_1, \ldots, P_n$ with eigenvalues q^{k_i}, respectively. Then

$$\pi_k = Q^\dagger \pi_0 Q, \qquad \text{with } Q = \bigotimes_{i=1}^{n} Q_i^{-k_i}, \tag{5.5}$$

which proves that all the π_k's have the same rank. As a consequence, each set of eigenvalues $(q^{k_i})_{i=1,\ldots,n}$ does correspond to a one-dimensional eigenspace in $\mathbb{C}^D$.

We may now state our central theorem relating Pauli operators and MUBs.

Theorem 21 *Let $d = p^s$ and $D = p^{sn}$. Then let B and C be the basis matrices of two Lagrangian submodules M_B and M_C of $\mathbb{Z}_d^{2n}$. The bases of $\mathbb{C}^D$ they encode are unbiased iff M_B and M_C are Lagrangian half-modules and $B^T JC$ is invertible.*

Proof. Let us suppose that M_B and M_C are Lagrangian half-modules and that $B^T JC$ is invertible. In particular $B^T JC$ is $n \times n$. There exist $L, R \in \mathrm{GL}(n, \mathbb{Z}_d)$ so that $(BL)^T J(CR)$ is diagonal (see Theorem 40 p.139 and proof or also algorithm $\mathscr{G}$ at the end of the same appendix). So we may assume that $B^T JC$ is diagonal with diagonal entries d_i, $i \in \{1, \ldots, n\}$. Let P_i (resp. Q_i), $i \in \{1, \ldots, n\}$, be the Pauli operators encoded in B (resp. in C) and let $|B; k\rangle$ (resp. $|C; k\rangle$), $k \in \{0, \ldots, p^s - 1\}^n$, denote the eigenvectors of the P_i's (resp. the Q_i's) with eigenvalues q^{k_i}:

$$\forall i \in \{1, \ldots, n\}, \quad P_i \,|B; k\rangle = q^{k_i} \,|B; k\rangle \quad \text{and} \quad Q_i \,|C; k\rangle = q^{k_i} \,|C; k\rangle \,. \tag{5.6}$$

Then for all $i \in \{1, \ldots, n\}$,

$$P_i Q_i = q^{-d_i} Q_i P_i \implies P_i \,|C; l\rangle = q^{-l_i} P_i Q_i \,|C; l\rangle = q^{-(l_i + d_i)} Q_i P_i \,|C; l\rangle \tag{5.7}$$

and so

$$|\langle B; k_1, \ldots, k_n | C; l_1, \ldots, l_n\rangle| = |\langle B; k_1, \ldots, k_n| \, P_i \,|C; l_1, \ldots, l_n\rangle| \tag{5.8a}$$

$$= |\langle B; k_1, \ldots, k_n | C; l_1, \ldots, l_i + d_i, \ldots, l_n\rangle| \,. \tag{5.8b}$$

Therefore, if all the d_i's are units, we get two unbiased bases.

If M_B or M_C is not a half-module, Proposition 19 tells us that we cannot obtain unbiased bases. Now suppose that they are half-modules and that B^TJC is not invertible. We may still assume that B^TJC is diagonal with diagonal entries d_i, $i \in \{1, \ldots, n\}$. At least one of the d_i's is not a unit. We have already seen that for all $i \in \{1, \ldots, n\}$,

$$P_i \,|C; l_1, \ldots, l_n\rangle = e^{i\theta} \,|C; l_1, \ldots, l_i + d_i, \ldots, l_n\rangle\,, \qquad \theta \in \mathbb{R}. \tag{5.9}$$

Thus, if we neglect the factor $e^{i\theta}$, the P_i's induce an action on the $|C; l\rangle$'s that now breaks into several orbits, at least p of them. Let O_j, j ranging, be the orbits and for each index j and let π_j be the orthogonal projection onto the space generated by the $|C; l\rangle$'s in O_j. Then for all $i \in \{1, \ldots, n\}$,

$$\sum_j \pi_j \,|B; 1, \ldots, 1\rangle = |B; 1, \ldots, 1\rangle = P_i \,|B; 1, \ldots, 1\rangle = \sum_j P_i \pi_j \,|B; 1, \ldots, 1\rangle \tag{5.10}$$

and we find that for all j,

$$\pi_j \,|B; 1, \ldots, 1\rangle = P_i \pi_j \,|B; 1, \ldots, 1\rangle\,. \tag{5.11}$$

Since the eigenspace of the P_i's with eigenvalues all equal to 1 is one-dimensional, $\pi_j \,|B; 1, \ldots, 1\rangle$ is nonzero for only one j and $|B; 1, \ldots, 1\rangle$ cannot have equal amplitude over all the $|C; l\rangle$'s. ■

5.2 Geometrical interpertation

In this part and the following, it will be convenient that the matrix representing the symplectic inner product over $\mathbb{Z}_d^{2n}$ in the canonical basis be

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}. \tag{5.12}$$

Thus the tensor product of n Pauli operators $X^{a_i} Z^{b_i}$ over $\mathbb{C}^d$, i ranging from 1 to n, is now represented by the vector

$$(a_1, \ldots, a_n, b_1, \ldots, b_n) \in \mathbb{Z}_d^{2n}. \tag{5.13}$$

Let also $\mathrm{Mat}(n, \mathbb{Z}_d)$ denote the set of all $n \times n$ matrices over $\mathbb{Z}_d$. Then our search for MUBs by diagonalising Pauli operators can be summed up into two projective geometrical features.

Theorem 21 states that in order to build MUBs in $\mathbb{C}^D$ by diagonalising tensorial products of n elementary Pauli operators over $\mathbb{C}^d$, one has to find n pairs of vectors $(x_1^i, x_2^i) \in (\mathbb{Z}_d^{2n})^2$, $i \in \{1, \ldots, n\}$, such that for all i, x_1^i and x_2^i are distant vectors as defined in Section 3.2 and for any i, j, $i \neq j$, the vectors x_1^i and x_2^j are orthogonal.

Because this rather peculiar configuration is only necessary and involves already many conditions, we would like to restate and complete it into a more compact form. This is achieved with a second geometrical interpretation:

Proposition 22 *The Lagrangian half-modules of* $\mathbb{Z}_d^{2n}$ *may be identified with the isotropic points of the projective line* $\mathbf{P}(\mathrm{Mat}(n, \mathbb{Z}_d)^2)$ *endowed with a suitable symplectic product* ω *to be defined in (5.16). Two Lagrangian half-modules yield MUBs iff their corresponding points are distant for this geometry.*

The surprise in such a result is that the symplectic product is involved only in the definition of isotropic points standing for Lagrangian half-modules, not in the MUB condition. In order to alleviate the various reasonings, the reader may assume that $d = p$ and so $\mathbb{Z}_d$ is a field. They may verify with the help of the p-adic decomposition that this assumption is harmless. Indeed, we are interested only in invertibility conditions and whenever we multiply two numbers in $\mathbb{Z}_d$, the coefficients of least degree in p in their p-adic decompositions multiply together. We shall find again such a simplification in the graph interpretation.

Lemma 23 *Let* M *be a* 2×1 *matrix with coefficients in* $\mathrm{Mat}(n, \mathbb{Z}_d)$*. The following conditions are equivalent:*

1. *There exists a* 2×1 *matrix* N *over* $\mathrm{Mat}(n, \mathbb{Z}_d)$ *such that the* 2×2 *matrix* $\begin{pmatrix} M & N \end{pmatrix}$ *(in block matrix notation) over* $\mathrm{Mat}(n, \mathbb{Z}_d)$ *is invertible.*

2. *There exists a* $2n \times n$ *matrix* N *over* $\mathbb{Z}_d$ *such that the* $2n \times 2n$ *matrix* $\begin{pmatrix} M & N \end{pmatrix}$ *over* $\mathbb{Z}_d$ *is invertible.*

3. M *contains an invertible* $n \times n$ *invertible submatrix.*

4. M *is left-unimodular in* $\mathrm{Mat}(n, \mathbb{Z}_d)$*, that is to say there exist* $A, B \in \mathrm{Mat}(n, \mathbb{Z}_d)$ *such that*
$$AM_1 + BM_2 = I_n. \tag{5.14}$$

5. *There exists a* 2×1 *matrix* N *over* $\mathrm{Mat}(n, \mathbb{Z}_d)$ *such that* $N^T JM = I_n$*.*

6. *There exists a* 2×1 *matrix* N *over* $\mathrm{Mat}(n, \mathbb{Z}_d)$ *such that* $N^T JM$ *is invertible (in* $\mathrm{Mat}(n, \mathbb{Z}_d)$*).*

Proof. Condition (1) means that $\begin{pmatrix} M & N \end{pmatrix}$ is a bijective linear map from $\mathrm{Mat}(n, \mathbb{Z}_d)^2$ onto itself. Condition (2) means that $\begin{pmatrix} M & N \end{pmatrix}$ is a bijective map from the set of $2n \times n$ matrices over $\mathbb{Z}_d$ onto itself. Since matrix products may be calculated by blocks, Conditions (1) and (2) are obviously equivalent.

(2) $\Rightarrow$ (3) is obvious by the way one can express a $2n \times 2n$ determinant as a sum of two-factor products of $n \times n$ determinants. Conversely, if M contains an invertible $n \times n$ submatrix, we can suppose without loss of generality that M_1 is invertible. Then
$$N = \begin{pmatrix} 0_n \\ I_n \end{pmatrix} \tag{5.15}$$

is a convenient choice to get (2).

If (3) is true, then we have already seen in the proof of Lemma 20 that there exists a $2n \times n$ matrix K (namely $N^T J$ with the notations of that latter lemma) such that $KM = I_n$. Writing $K = \begin{pmatrix} A & B \end{pmatrix}$, we get Condition (4). Conversely, if the rows of I_n can be obtained as linear combinations of the rows of M, then the row vectors of M generates $\mathbb{Z}_d^n$ as a whole, which is nothing but (3).

Finally, (5) is a mere rewriting of (4) and (5) $\Leftrightarrow$ (6) is obvious with the help of a left-multiplication. ■

A matrix M satisfying those conditions is an admissible pair for a projective geometry over $\mathrm{Mat}(n, \mathbb{Z}_d)$. Note that the existence of a complement and unimodularity are equivalent though $\mathrm{Mat}(n, \mathbb{Z}_d)$ is not a commutative ring. Also the map

$$(\mathrm{Mat}(n, \mathbb{Z}_d)^2)^2 \longrightarrow \mathrm{Mat}(n, \mathbb{Z}_d), \quad (M, N) \longmapsto M^T J N \tag{5.16}$$

defines a symplectic product over $\mathrm{Mat}(n, \mathbb{Z}_d)^2$. We still denote this map ω. A Lagrangian half-module is represented by a $2n \times n$ matrix M over $\mathbb{Z}_d$ up to a right-multiplication by an $n \times n$ invertible matrix. Moreover, this matrix M has to verify $M^T J M = 0$. Thus, with the help of Lemma 20, a Lagrangian submodule can be thought of as an isotropic point of the projective line over $\mathrm{Mat}(n, \mathbb{Z}_d)$ endowed with the symplectic product derived from ω and still denoted in the same way. Conversely, if we take an admissible, isotropic vector of $\mathrm{Mat}(n, \mathbb{Z}_d)^2$, its column vectors generate a submodule of $\mathbb{Z}_d^{2n}$ with rank n and maximal cardinality, that is a Lagrangian submodule.

Two matrices M and N as in Condition (1) of Lemma 23 are called distant. We would like to know under which condition the symplectic product $\omega(M, N)$ of the same two matrices is invertible as in Condition (6)[1]. A sufficient condition is provided in the following lemma.

Lemma 24 *Let $M \in \mathrm{Mat}(n, \mathbb{Z}_d)^2$ be an admissible vector. If M is isotropic with respect to the symplectic product we have just defined, then*

$$\forall N \in \mathrm{Mat}(n, \mathbb{Z}_d)^2, \quad \det(M^T J N) = \det(\begin{pmatrix} M & N \end{pmatrix}). \tag{5.17}$$

Together with Theorem 21, this lemma proves that two isotropic points in $(\mathbf{P}(\mathrm{Mat}(n, \mathbb{Z}_d)^2), \omega)$ provide one with a pair of MUBs iff they are distant. This latter condition is projectively well-defined, since it does not depend on the representatives of the two points.

Proof. Let M be an isotropic vector, so that its column vectors generate a Lagrangian half-module. Theorem 13 (p.55) shows that there exist an invertible $n \times n$

[1] Remark that $N^T J M = -(M^T J N)^T$ so that $N^T J M$ and $M^T J N$ are simultaneously invertible or not.

matrix K over $\mathbb{Z}_d$ and a $2n \times 2n$ symplectic matrix S over $\mathbb{Z}_d$ such that

$$SMK = \begin{pmatrix} I_n \\ 0_n \end{pmatrix}. \tag{5.18}$$

Then

$$\begin{aligned}
\det K \cdot \det(M^T JN) &= \det(K^T M^T JN) = \det((SMK)^T J(SN)) && \text{(5.19a)}\\
&= \det((SN)_2) && \text{(5.19b)}\\
&= \det(\begin{pmatrix} SMK & SN \end{pmatrix}) && \text{(5.19c)}\\
&= \det S \cdot \det(\begin{pmatrix} M & N \end{pmatrix}) \cdot \det K && \text{(5.19d)}
\end{aligned}$$

Taking into account that $\det S = 1$, we find that

$$\det(M^T JN) = \det(\begin{pmatrix} M & N \end{pmatrix}). \tag{5.20}$$

■

To end this section, we show that whenever $d = p$ and thus $\mathbb{Z}_d$ is a field, the converse implication in Lemma 24 is true. Let us assume that Condition (5.17) holds and M is not isotropic. Since M is admissible, we can use the symplectic reduction algorithm $\mathscr{S}$ developed in Section 4.1 p.51 in order to find a symplectic matrix S and an invertible $n \times n$ matrix K so that

$$SMK = \left(\begin{array}{c:c:c} 1 & & \\ \hdashline & & M_1' \\ \hline & 1 & \\ \hdashline & & M_2' \end{array}\right), \tag{5.21}$$

with M_1' and M_2' being $(n-1) \times (n-2)$ matrices. Let us write a $2n \times n$ matrix N under the form

$$N = S^{-1} \left(\begin{array}{c} \ell_1 \\ N_1' \\ \hline \ell_2 \\ N_2' \end{array}\right), \tag{5.22}$$

where ℓ_1 and ℓ_2 are rows and N_1' and N_2' are $(n-1) \times n$ matrices. Then we have

$$\begin{aligned}
\det((SMK)^T J(SN)) &= \det K \cdot \det(M^T JN) && \text{(5.23a)}\\
&= \det S \cdot \det(\begin{pmatrix} M & N \end{pmatrix}) \cdot \det K && \text{(5.23b)}\\
&= \det(\begin{pmatrix} SMK & SN \end{pmatrix}) && \text{(5.23c)}
\end{aligned}$$

In $\det(\begin{pmatrix} SMK & SN \end{pmatrix})$, the rows ℓ_1 and ℓ_2 do not matter and since M is admissible, we can choose N_1' and N_2' so that this latter determinant is a unit. But the first two rows of $(SMK)^T J(SN)$ are ℓ_2 and $-\ell_1$, so that we get a null determinant whenever we put $\ell_1 = \ell_2$, leading to a contradiction.

5.3 Graph interpertation

In this part, we still have

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}. \tag{5.24}$$

Let us consider the Lagrangian half-modules as the vertices of a graph. We shall say that two of them, with basis matrices B and C, are at distance r if $B^T JC$ is of rank r. Two vertices are the ends of an edge if they are at distance n in this latter sense or in other words if they are distant projective points. From this point of view, our aim is to find a maximal set of vertices pairwise connected on the graph. Such a set is called a clique and the greatest cardinality a clique can have is the clique-number of the graph:

$$\text{clique-number} = \max_{\{\text{cliques}\}} \text{Card(clique)}. \tag{5.25}$$

The clique-number of the graph we are considering is the maximal number of MUBs one can find in dimension d^n with Pauli operators. To find this number for a given graph is known to be an NP-complete problem, a kind of problem that cannot be effeciently solved on our present-day computers. We will not try to solve it here, even under the particular conditions of our problem, but we will give some other counting properties of interest. Anyway, it must be stressed that we already know what the clique-number of our graph is whenever $d = p^s$ is a power of a prime: $p^s + 1$. So in prime-power dimension, the search for the clique-number is only a incentive to further investigations with graph theoretical means. As for composite dimensions, this graph interpretation is liable to the Chinese remainder theorem and so will be of no help in overcoming the bound we saw in Section 3.3 or in [17].

If we still take $d = p^s$, then according to Theorem 21, any multiple of p is irrelevant for the question of MUBs and we are only left with a kind of a covering of the graph one gets with $d = p$. So we will assume until the end of that part that indeed $d = p$ and so the coefficient ring $\mathbb{Z}_p$ is a field. With $r \in \{0, \ldots, n\}$, let us also put:

- g_r the number of $r \times r$ invertible matrices,
- σ_r the number of symmetric, $r \times r$ matrices,
- σ'_r the number of symmetric, $r \times r$ invertible matrices,
- k_r the number of $r \times (n - r)$ matrices,
- l_r the number subspaces of dimension r in $\mathbb{Z}_p^n$.

It is easy to show that:

$$g_r = (p^r - 1)(p^r - p) \cdots (p^r - p^{r-1}) \tag{5.26a}$$

$$\sigma_r = p^{r(r+1)/2} \tag{5.26b}$$

$$k_r = p^{r(n-r)} \tag{5.26c}$$

$$l_r = \frac{1}{g_r}(p^n - 1)(p^n - p) \cdots (p^n - p^{r-1}) \tag{5.26d}$$

With the notations

$$
\begin{aligned}
[x]_p &= \frac{p^x-1}{p-1}, \quad x \in \mathbb{R}, && (5.27a)\\
[n]_p! &= [n]_p[n-1]_p \cdots [1]_p, \quad n \in \mathbb{N}^*, && (5.27b)\\
[0]_p! &= 1, && (5.27c)
\end{aligned}
$$

l_r also reads

$$l_r = \frac{[n]_p!}{[r]_p![n-r]_p!}. \tag{5.28}$$

We are going to find an expression for σ'_r by induction. Obviously we have

$$
\begin{aligned}
\sigma'_0 &= 1, && (5.29a)\\
\sigma'_1 &= p-1. && (5.29b)
\end{aligned}
$$

Then in a symmetric, $r \times r$ matrix, $r \geq 2$, the upper-left coefficient is either nonzero or zero. In the first case, one may choose this coefficient in $p-1$ ways and the other coefficients on the first row and the first column in p^{r-1} ways. In order to know how to choose the remaining coefficients so as to get a symmetric, invertible matrix, the upper-left coefficient can be used to perform a bilateral Gauss reduction. One thus obtains a matrix with nul coefficients on the first row and the first column save the upper-left one, without modifying the number of possibilities in choosing the remaining coefficients. This number appears clearly to be σ'_{r-1}. In the case where the upper-left coefficient is zero, one has $p^{r-1}-1$ ways in choosing the other coefficients on the first row and the first column. After choosing a unit among them and performing a bilateral Gauss reduction with it, it works out that the coefficients at positions $(2,i)$ and $(i,2)$, $2 \leq i \leq r$, may be chosen arbitrarily but in a symmetric fashion and the remaining coefficients form a symmetric, $(r-2)\times(r-2)$ invertible matrix. So we get the induction formula:

$$\sigma'_r = (p-1)p^{r-1}\sigma'_{r-1} + (p^{r-1}-1)p^{r-1}\sigma'_{r-2}. \tag{5.30}$$

In particular

$$
\begin{aligned}
\sigma'_1 &= (p-1)\sigma'_0, && (5.31a)\\
\sigma'_2 &= (p-1)^2p + (p-1)p \;=\; (p-1)p^2 \;=\; p^2\sigma'_1. && (5.31b)
\end{aligned}
$$

We are to prove by induction that

$$
\begin{aligned}
\sigma'_{2t} &= p^{2t}\sigma'_{2t-1}, \quad && t \geq 1, && (5.32a)\\
\sigma'_{2t+1} &= (p^{2t+1}-1)\sigma'_{2t}, \quad && t \geq 0. && (5.32b)
\end{aligned}
$$

Relation (5.32a) is true for $t=1$ and Relation (5.32b) is true for $t=0$. Now we perform a double induction. If $r = 2t$, $t \geq 2$, is even, then

$$\sigma'_{2t} = (p-1)p^{2t-1}\sigma'_{2t-1} + (p^{2t-1}-1)p^{2t-1}\sigma'_{2t-2} \tag{5.33a}$$

$$= (p-1)p^{2t-1}\sigma'_{2t-1} + (p^{2t-1}-1)p^{2t-1}\frac{\sigma'_{2t-1}}{p^{2t-1}-1} \tag{5.33b}$$

$$= p^{2t}\sigma'_{2t-1}. \tag{5.33c}$$

If $r = 2t+1$, $t \geq 1$, is odd, then

$$\sigma'_{2t+1} = (p-1)p^{2t}\sigma'_{2t} + (p^{2t}-1)p^{2t}\sigma'_{2t-1} \tag{5.33d}$$

$$= (p-1)p^{2t}\sigma'_{2t} + (p^{2t}-1)p^{2t}\frac{\sigma'_{2t}}{p^{2t}} \tag{5.33e}$$

$$= (p^{2t+1}-1)\sigma'_{2t}. \tag{5.33f}$$

Thus

$$\sigma'_r = \begin{cases} p^r(p^{r-1}-1)p^{r-2}(p^{r-3}-1) \;\cdots\; p^2(p-1) & \text{if } r \text{ is even,} \\ (p^r-1)p^{r-1}(p^{r-2}-1)p^{r-3} \;\cdots\; p^2(p-1) & \text{if } r \text{ is odd.} \end{cases} \tag{5.34}$$

All the ingredients are now gathered in order to count Lagrangian half-modules under distance conditions.

Proposition 25 *Let M be a Lagrangian half-module. The number of Lagangian half-modules at distance r from M is*

$$b_r = l_r \sigma_r k_r. \tag{5.35}$$

In particular, the number of Pauli eigenbases that are unbiased with a given one is

$$b_n = \sigma_n. \tag{5.36}$$

Proof. We know from Theorem 13 (p.55) that we can choose a symplectic computational basis of $\mathbb{Z}_d^{2n}$ so that a basis matrix for M be

$$\text{basis matrix }(M) = \begin{pmatrix} I_n \\ 0_n \end{pmatrix}. \tag{5.37}$$

Then we want to find the number of admissible and isotropic, $2n \times n$ matrices B of the form

$$B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \qquad B_1, B_2 \in \mathrm{Mat}(n, \mathbb{Z}_p), \tag{5.38}$$

up to a right-multiplication by an invertible $n \times n$ matrix so that

$$\begin{pmatrix} I_n \\ 0_n \end{pmatrix}^T \times J \times \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = B_2 \tag{5.39}$$

is of rank r. Therefore we get the factor l_r, since B_2 has to be a basis matrix for an r-dimensional subspace of $\mathbb{Z}_p^n$. With a right-multiplication by an invertible matrix and a change of symplectic, computational basis, we may suppose that B is of the form

$$B = \left(\begin{array}{cc} A_1 & A_3 \\ A_2 & A_4 \\ \hline I_r & 0 \\ 0 & 0 \end{array} \right). \tag{5.40}$$

In doing so, we replace I_n by an invertible matrix in (5.37). But we do not need this expression any more. With the isotropy condition one gets that A_1 is a symmetric, $r \times r$ invertible matrix, A_2 is any $(n-r) \times r$ matrix and $A_3 = 0$. Therefore $\sigma_r k_r$. As to A_4, one has to refer to the admissibility condition to see, in the same way as we did in the proof of Lemma 20, that $A_4 \in \mathrm{GL}(n-r, \mathbb{Z}_p)$. Hence a factor g_{n-r}. But, since we are interested only in the subspace generated by the column vectors of B, we still have to divide by g_{n-r}. Therefore (5.35). ■

Proposition 26 *For any two Lagrangian half-modules at distance r from one another, there exist*

$$\lambda_r = \sigma_r' k_r \sigma_{n-r} \geq 1 \tag{5.41}$$

other Lagrangian half-modules which are distant from both of them. In particular, if two Pauli eigenbases are unbiased, there exist

$$\lambda_n = \sigma_n' \geq 1 \tag{5.42}$$

other Pauli eigenbases that are unbiased with those two ones.

Proof. Let B and C be two basis matrices of two Lagrangian half-modules at distance r from one another. After the previous discussion about b_r, we may suppose that

$$B = \begin{pmatrix} I_n \\ 0_n \end{pmatrix} \quad \text{and} \quad C = \left(\begin{array}{cc} A_1 & 0 \\ A_2 & I_{n-r} \\ \hline I_r & 0 \\ 0 & 0 \end{array} \right), \tag{5.43}$$

with A_1 a symmetric, $r \times r$ invertible matrix and A_2 any $(n-r) \times r$ matrix. We search for a $2n \times n$ matrix M such that $B^T JM$ and $C^T JM$ are invertible, in addition of suitable conditions for M to account for a Lagrangian half-module. Let M be of the form

$$M = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix}. \tag{5.44}$$

Then M_2 has to be invertible since $B^T JM$ is invertible and we may suppose that $M_2 = I_n$. Thus, the admissibility condition is checked and the number λ_r we search for is the number of suitable matrices M_1. With

$$M_1 = \begin{pmatrix} M_{11} & M_{13} \\ M_{12} & M_{14} \end{pmatrix}, \tag{5.45}$$

we have the condition

$$C^T J M = \begin{pmatrix} A_1^T - M_{11} & A_2^T - M_{13} \\ 0 & I_{n-r} \end{pmatrix} \in \mathrm{GL}(n, \mathbb{Z}_p). \tag{5.46}$$

At last, M_{11} and M_{14} have to be symmetric and $M_{12}^T - M_{13} = 0$ in order to warrant isotropy. Hence, the number of suitable matrices M_{11} is σ'_r. There are k_r possibilities in choosing M_{12} and M_{13} and σ_{n-r} possibilities in choosing M_{14}. Therefore (5.41). ■

5.4 The Clifford group

Lagrangian half-modules and distance relations among them are also suitable to have a grasp on the Clifford group. The results presented in Proposition 27 have already been studied in the works [19–23] about the discrete phase space or more recently in [53], though with different names. However, these studies were concerned with the phase space over Galois fields. We will consider the ring space $\mathbb{Z}_d^{2n}$ with d any integer. On this basis, Proposition 28 will expose an original, alternative account for the dynamics among Pauli states.

Let us consider a Pauli group $\mathcal{P} = \mathcal{P}(d, n)$ with or without tensorial product. The corresponding Clifford group $\mathrm{Cliff}(\mathcal{P})$ is the normaliser of $\mathcal{P}$, that is to say the set of all unitary operators that stabilise the Pauli group under conjugation:

$$U \in \mathrm{Cliff}(\mathcal{P}) \quad \Longleftrightarrow \quad \forall P \in \mathcal{P},\ \exists P' \in \mathcal{P},\ UPU^\dagger = P'. \tag{5.47}$$

Since the action of some $U \in \mathrm{Cliff}(\mathcal{P})$ is one-to-one and $\mathcal{P}$ is a finite group, this latter action is also onto, $U^\dagger$ is also in $\mathrm{Cliff}(\mathcal{P})$ and hence $\mathrm{Cliff}(\mathcal{P})$ is indeed a subgroup of the unitary group $\mathrm{U}(d^n)$. Since for any n, the derived group $D(\mathcal{P})$ is isomorphic to

$$\mathbf{C}_d = \{q^c;\ c \in \mathbb{Z}_d\}, \tag{5.48}$$

we may write the sequence of group inclusions:

$$\mathbf{C}_d \lhd \mathcal{P} \lhd \mathrm{Cliff}(\mathcal{P}) < \mathrm{U}(d^n). \tag{5.49}$$

We thus have a group action of $\mathrm{Cliff}(\mathcal{P})$ on its normal subgroup $\mathcal{P}$ and we are going to show that this action enables one to fully analyse the structure of $\mathrm{Cliff}(\mathcal{P})$. Just before we do so, let us tell briefly about a physical application of the Clifford group.

Entanglement is one of the primary resources of quantum information theory. The entangled states one usually refers to are entirely described as eigenvectors of Pauli operators. That is to say a state is equivalent to a generating subset of a maximally commuting set of Pauli operators, with the convention that the state to describe has eigenvalue 1 for all Pauli operators in the subset. In the course of a computation, such a state is made to evolve by means of unitary gates into a state of the same type. Such an evolution may of course be transcribed as an evolution of the defining set of Pauli operators. If $|\psi\rangle$ is the eigenvector of some Pauli operator P with eigenvalue 1 and U is a unitary operator, then $U|\psi\rangle$ is an eigenvector of $UPU^\dagger$ also with eigenvalue 1 since

$$UPU^\dagger U|\psi\rangle = UP|\psi\rangle = U|\psi\rangle . \tag{5.50}$$

This justifies our consideration for $\mathrm{Cliff}(\mathcal{P})$ as defined in (5.47). Our aim will also be to see how the decomposition of $\mathrm{Cliff}(\mathcal{P})$ we are about to go through can account for this physical application. This will be achieved in Proposition 28.

To investigate $\mathrm{Cliff}(\mathcal{P})$, we will refer to the following quotient groups:

$$\begin{aligned} \mathcal{P}' &= \mathcal{P}/\mathbf{C}_d, && (5.51)\\ \mathrm{Cliff}'(\mathcal{P}) &= \mathrm{Cliff}(\mathcal{P})/\mathrm{U}(1), && (5.52)\\ \mathrm{Cliff}''(\mathcal{P}) &= \mathrm{Cliff}'(\mathcal{P})/\mathcal{P}'. && (5.53)\end{aligned}$$

The quotient $\mathcal{P}'$ is the group of Pauli operators up to a phase, or projective Pauli group. It is the quotient set of the equivalence relation

$$\forall c \in \mathbb{Z}_d,\ \forall P \in \mathcal{P}, \quad P \sim q^c P. \tag{5.54}$$

Then, the group U(1) being the group of unit modulus complex numbers, $\mathrm{Cliff}'(\mathcal{P})$ is the group of Clifford operators up to a (continuous) phase. Finally, $\mathrm{Cliff}''(\mathcal{P})$ is the group of Clifford operators when both phases and Pauli operators are discarded. So we may identify it with the quotient set of the equivalence relation

$$\forall \theta \in \mathbb{R},\ \forall P \in \mathcal{P},\ \forall U \in \mathrm{Cliff}(\mathcal{P}), \quad U \approx e^{i\theta} PU. \tag{5.55}$$

We now state and prove the

Proposition 27 *Let $\mathcal{P}$ be a Pauli group. Any Clifford operator $U \in \mathrm{Cliff}(\mathcal{P})$ analyses into three components: a global phase $q(U) \in \mathrm{U}(1)$ (not nessecarily in $\mathbf{C}_d$), a projective Pauli operator $P(U) \in \mathcal{P}'$ and a symplectic operator $\sigma(U) \in \mathrm{Sp}(n, \mathbb{Z}_d)$. Indeed,*

$$\mathrm{Cliff}''(\mathcal{P}) \simeq \mathrm{Sp}(n, \mathbb{Z}_d) \tag{5.56}$$

and also, the map

$$\sigma : \mathrm{Cliff}(\mathcal{P}) \longrightarrow \mathrm{Sp}(n, \mathbb{Z}_d) \tag{5.57}$$

is a group homomorphism.

Moreover, the Pauli and symplectic components, $P(U)$ and $\sigma(U)$, are determined by the action of U on X and Z, whereas $q(U)$ is irrelevant to the action of U. So we may consider that the maps P and σ act on $\mathrm{Cliff}'(\mathcal{P})$.

To simplify the following discussion, we will consider in any explicit formula Pauli operators over a Hilbert space $\mathbb{C}^d$ without tensorial product:

$$\mathcal{P} = \{q^c X^a Z^b;\ a, b, c \in \mathbb{Z}_d\}. \tag{5.58}$$

That is to say we will refer to the case $n = 1$. The generalisation to Pauli operators with tensorial products is straightforward.

Let $U \in \mathrm{Cliff}(\mathcal{P})$ and let us put

$$P_1 = q^{c_1} X^{a_1} Z^{b_1} = U X U^\dagger, \tag{5.59a}$$

$$P_2 = q^{c_2} X^{a_2} Z^{b_2} = U Z U^\dagger. \tag{5.59b}$$

Relation (5.59b) shows that U diagonalises P_2, so that the column vectors C_i of U are fixed by this relation up to a global phase, one phase φ_i for each of them. Moreover, Relation (5.59a) shows that P_1 is a circular shift operator among the C_i's:

$$\forall i, \quad C_{i+1} = P_1 C_i. \tag{5.60}$$

Therefore, once one of the φ_i's is fixed, all the other ones are. In brief, through Relations (5.59a) and (5.59b), P_1 and P_2 characterise U up to one continuous phase factor $q(U)$.

Remark 1 *In particular, if*

$$U X U^\dagger = q^{c_1} X \quad \textit{and} \quad U Z U^\dagger = q^{c_2} Z, \tag{5.61}$$

then $U = X^{-c_2} Z^{c_1}$ up to a continuous phase.

Then, by comparing the two following relations

$$[P_1, P_2]_{\mathrm{g}} = q^{b_1 a_2 - a_1 b_2} I, \tag{5.62a}$$

$$[P_1, P_2]_{\mathrm{g}} = [U X U^\dagger, U Z U^\dagger]_{\mathrm{g}} = q^{-1} I, \tag{5.62b}$$

we see that

$$\sigma(U) = \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} \in \mathrm{Sp}(n, \mathbb{Z}_d). \tag{5.63}$$

The matrix $\sigma(U)$ will be called the symplectic part of U.

Conversely, let $P_1, P_2 \in \mathcal{P}$ such that

$$a_1 b_2 - a_2 b_1 = 1. \tag{5.64}$$

It is straightforward with what we have already said that we can find $U \in \mathrm{Cliff}(\mathcal{P})$ so as it fulfils (5.59a) and (5.59b). In details, U is a change of basis matrix that diagonalises P_2: We get $U^\dagger P_2 U = Z$ and thus (5.59b) is checked. The relative phases of the column vectors of U have then to be fixed, something that can be done with the help of P_1, which we can force to be a circulation operator among those vectors as in (5.60). It amounts to setting to 1 every nonzero coefficient of $U^\dagger P_1 U$ so that it be equal to X and thus (5.59a) is checked. So U is uniquely determined up to the phase $q(U)$. But what if we are given only $\sigma(U)$ instead of P_1 and P_2?

Let c_1 and c_2 vary in the definition of P_1 and P_2. In order to show that the corresponding degree of freedom in U is nothing else than a projective Pauli operator, we establish that the symplectic structure involved in $\mathrm{Cliff}(\mathcal{P})$ can be expressed as $\mathrm{Cliff}''(\mathcal{P})$. Since c_1 and c_2 are dropped out in Relation (5.64), the issue relates to $\mathcal{P}'$ and we consider the natural, reduced action ρ of $\mathrm{Cliff}(\mathcal{P})$ on $\mathcal{P}'$:

$$\rho : U \cdot \mathrm{class}(P) = \mathrm{class}(UPU^\dagger). \tag{5.65}$$

Then on the one hand, with two explicit elements $P_3, P_4 \in \mathcal{P}$, we have

$$\begin{aligned}
& & \mathrm{class}(X^{a_4}Z^{b_4}) &= \mathrm{class}(U)\cdot\mathrm{class}(X^{a_3}Z^{b_3}) && (5.66\text{a})\\
&\iff & X^{a_4}Z^{b_4} &\sim UX^{a_3}Z^{b_3}U^\dagger && (5.66\text{b})\\
&\iff & X^{a_4}Z^{b_4} &\sim (UXU^\dagger)^{a_3}(UZU^\dagger)^{b_3} && (5.66\text{c})\\
&\iff & X^{a_4}Z^{b_4} &\sim (X^{a_1}Z^{b_1})^{a_3}(X^{a_2}Z^{b_2})^{b_3} && (5.66\text{d})\\
&\iff & \begin{pmatrix} a_4 \\ b_4 \end{pmatrix} &= \sigma(U)\begin{pmatrix} a_3 \\ b_3 \end{pmatrix}. && (5.66\text{e})
\end{aligned}$$

On the other hand, we know after Remark 1 that the kernel of the reduced action is $\mathcal{P}$ up to a continuous phase, namely

$$\ker\rho = \{e^{i\theta}P;\ \theta\in\mathbb{R},\ P\in\mathcal{P}\}. \tag{5.67}$$

Thus,

$$\mathrm{Cliff}(\mathcal{P})/\ker\rho \simeq \mathrm{Cliff}''(\mathcal{P}) \tag{5.68}$$

and two operators $U_1, U_2 \in \mathrm{Cliff}(\mathcal{P})$ act in the same way on $\mathcal{P}'$, that is to say $\sigma(U_1) = \sigma(U_2)$, iff they go down to the same class in $\mathrm{Cliff}''(\mathcal{P})$:

$$\sigma(U_1) = \sigma(U_2) \quad \iff \quad \exists\theta\in\mathbb{R},\ \exists P\in\mathcal{P},\ U_2 = e^{i\theta}PU_1. \tag{5.69}$$

Remark 2 *Let us give another point of view on the kernel. Due to the commutation relations among Pauli operators, then for any $U \in \mathrm{Cliff}(\mathcal{P})$ and any $P_3, P_4 \in \mathcal{P}$, the unitary operators U and P_3UP_4 act in the same way on the projective Pauli group. In particular, both of them diagonalise $X^{a_2}Z^{b_2}$, so that the column vectors of P_3UP_4 form a permutation of the column vectors of U up to a system of phases. Under the constraint given by (5.59a), the freedom in fixing that system of phases is reduced to one parameter. But we will not dwell on the details.*

In summary, one first has to specify once for all a section from $\mathrm{Cliff}''(\mathcal{P})$ to $\mathrm{Cliff}'(\mathcal{P})$ and another one from $\mathrm{Cliff}'(\mathcal{P})$ to $\mathrm{Cliff}(\mathcal{P})$. Then, any Clifford operator U is specified with respect to these sections by the triplet

$$(q(U), P(U), \sigma(U)) \ \in\ \mathrm{U}(1)\times\mathcal{P}'\times\mathrm{Sp}(n,\mathbb{Z}_d). \tag{5.70}$$

At this point, we proved the first part of Proposition 27 and we exhibited a bijection

$$f : \text{Cliff}''(\mathcal{P}) \longrightarrow \text{Sp}(n, \mathbb{Z}_d) \tag{5.71}$$

such that if

$$\pi : \text{Cliff}(\mathcal{P}) \longrightarrow \text{Cliff}''(\mathcal{P}) \tag{5.72}$$

is the canonical projection from $\text{Cliff}(\mathcal{P})$ onto $\text{Cliff}''(\mathcal{P})$, then

$$f \circ \pi = \sigma. \tag{5.73}$$

So we want to check that σ is a morphism, that is

$$\forall U_1, U_2 \in \text{Cliff}(\mathcal{P}), \quad \sigma(U_2U_1) = \sigma(U_2)\sigma(U_1). \tag{5.74}$$

It will show that f is also a morphism and it will be proved that $\text{Cliff}''(\mathcal{P})$ is isomorphic to $\text{Sp}(n, \mathbb{Z}_d)$. In fact, if we consider the actions of U_1 and U_2 on the $\mathcal{P}'$, refering to them by the superscripts (1) and (2), respectively, we have

$$\begin{aligned} U_2U_1XU_1^\dagger U_2^\dagger &\sim U_2X^{a_1^{(1)}}Z^{b_1^{(1)}}U_2^\dagger && \text{(5.75a)}\\ &\sim (X^{a_1^{(2)}}Z^{b_1^{(2)}})^{a_1^{(1)}}(X^{a_2^{(2)}}Z^{b_2^{(2)}})^{b_1^{(1)}} && \text{(5.75b)}\\ &\sim X^{a_1^{(2)}a_1^{(1)}+a_2^{(2)}b_1^{(1)}}Z^{b_1^{(2)}a_1^{(1)}+b_2^{(2)}b_1^{(1)}} && \text{(5.75c)} \end{aligned}$$

and similarly for the action on Z. That is nothing but the morphism relation we were looking for. Whenever we consider the derived action of $\text{Cliff}''(\mathcal{P})$ on $\mathcal{P}'$, this can also be written, for any $P \in \mathcal{P}$,

$$\sigma(U_2U_1) \cdot \text{class}(P) = \sigma(U_2) \cdot (\sigma(U_1) \cdot \text{class}(P)). \tag{5.76}$$

In Figure 5.1, we list the classical examples of Clifford operators for qubits, together with their symplectic part $\sigma(U)$ and in the last but one column their Pauli parts $P(U)$. The parametrisation of Pauli operators is the one given in (5.13), corresponding to J as given in (5.12). As to $P(U)$, the reference section from $\text{Cliff}''(\mathcal{P})$ to $\text{Cliff}'(\mathcal{P})$ is the natural one that brings for instance $\sigma(U)$ as in (5.63) to the U defined in (5.59) with $c_1 = c_2 = 0$, up to an irrelevant continuous phase.

The practical use of the Pauli part $P(U)$ is precisely to define c_1 and c_2 (see Remark 1). But the function $P : U \to P(U)$ has the drawback not to be a morphism. The trick is to replace $P(U)$ by the c_i's or better, as we shall see, by their opposites:

$$\begin{aligned} v : \quad \text{Cliff}(\mathcal{P}) &\longrightarrow \mathbb{Z}_d^{2n}\\ U &\longmapsto (-c_1, -c_2, \ldots, -c_{2n}), \end{aligned} \tag{5.77}$$

where $\mathbb{Z}_d^{2n}$ is of course endowed with its additive structure. The map v is a morphism that in addition is section-free. For instance, we put it for classical Clifford operators in the last column of Figure 5.1, in column form so as to parallel the Pauli inputs and outputs.

Operation	Matrix form U	Pauli input	Pauli output	Symplectic part $\sigma(U)$	Pauli part $P(U)$	$v(U)$
CNOT	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$	X_1 X_2 Z_1 Z_2	X_1X_2 X_2 Z_1 Z_1Z_2	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$	I_2I_2	$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$
H	$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$	X Z	Z X	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	I_2	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$
S	$\begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$	X Z	Y Z	$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$	I_2	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$
X	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	X Z	X $-Z$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	X	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$
Y	$\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$	X Z	$-X$ $-Z$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	Y	$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$
Z	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	X Z	$-X$ Z	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	Z	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$

Figure 5.1: Clifford operators with their symplectic and Pauli parts

Any pair

$$(x, S) \;\in\; \mathbb{Z}_d^{2n} \times \mathrm{Sp}(n, \mathbb{Z}_d) \tag{5.78}$$

is enough to specify the action of a Clifford operator on $\mathcal{P}$:

$$\exists!\; U \in \mathrm{Cliff}'(\mathcal{P}), \quad (v(U), \sigma(U)) = (x, S), \tag{5.79}$$

(see the end of Proposition 27; the formulation in terms of $\mathrm{Cliff}(\mathcal{P})$ would be a bit cumbersome). But if we consider the physical application we gave just after the beginning of the present section, the pair (x, S) also enables us to precise completely a state $|\psi\rangle$.

Proposition 28 *For any given pair*

$$(x, S) \;\in\; \mathbb{Z}_d^{2n} \times \mathrm{Sp}(n, \mathbb{Z}_d), \tag{5.80}$$

there exists a unique state $|\psi\rangle$ such that, for every $i \in \{n, \ldots, 2n\}$, it has eigenvalue q^{x_i} with respect to the Pauli operator without phase, defined by the i-th column of S. Such a state is called a Pauli state.

Moreover, the map

$$\begin{array}{rrcl} (v, \sigma): & \mathrm{Cliff}'(\mathcal{P}) & \longrightarrow & \mathbb{Z}_d^{2n} \times \mathrm{Sp}(n, \mathbb{Z}_d) \\ & U & \longmapsto & (v(U), \sigma(U)) \end{array} \tag{5.81}$$

is a componentwise morphism. So the product group $\mathbb{Z}_d^{2n} \times \mathrm{Sp}(n, \mathbb{Z}_d)$ is an alternative for describing the dynamics among Pauli states.

For an example, just bring together (5.63) and (5.59): Whenever $|\psi\rangle$ has eigenvalue 1 with respect to $q^{c_2}X^{a_2}Z^{b_2}$, then it has eigenvalue q^{-c_2} with respect to $X^{a_2}Z^{b_2}$. Thus $|\psi\rangle$ is determined by the pair

$$\left(\begin{pmatrix} -c_1 \\ -c_2 \end{pmatrix}, \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} \right). \tag{5.82}$$

However, it should be noticed that a_1, b_1, c_1 do not intervene in the determination of $|\psi\rangle$, but only in the dynamics from a Pauli state to another one. Thus, in a series of transformations $U_1, \ldots, U_k$, where U_k is not to be followed by another transformation, the coefficients $a_1^{(k)}, b_1^{(k)}, c_1^{(k)}$ are of no use.

In fact, in the scheme proposed in the second part of Proposition 28, one may avoid any reference to vector states, Pauli operators and Clifford operators. Both the classical amount of entanglement and the way it evolves are captured in the product group $\mathbb{Z}_d^{2n} \times \mathrm{Sp}(n, \mathbb{Z}_d)$, or in other words in the geometry of Lagrangian half-modules with an extra set of phases.

Chapter 6

The isotropic lines of a discrete phase space

In this chapter, we let aside the bearing of discrete geometry over Pauli operators to apply Theorem 13 (Section 4.2 p.55) to another major tool of quantum mechanics, namely discrete Wigner distributions. These distributions offer a useful, alternative way besides density matrices of representing pure and mixed states of a quantum system. But whereas in the continuous phase space those distributions are well-defined [54] [55], there is still a need for a sound mathematical definition over a discrete phase space. In particular, the structure of such a phase space is of some importance. In 1974, Buot introduced a Wigner distribution over an $d \times d$ phase space with d an odd integer [56]. In 1980, Hannay and Berry followed another approach to build a Wigner distribution over a $2d \times 2d$ lattice [57]. Still in another way, in 2004, Gibbons *et al.* constructed Wigner distributions over lattices parametrised by finite fields [58].

More recently, in their way to set up discrete Wigner distributions on the discrete phase space $\mathbb{Z}_d^2$, Chaturvedi *et al.* [59] encountered undetermined signs $S(q,p)$, one at each point (q,p) of the lattice. A natural question then arises: To what extent can these signs be fixed by demanding that averages of Wigner distributions over isotropic lines in the lattice yield probabilities, where an isotropic line is a set of d points on the lattice such that the symplectic product of any two of them is 0 (modulo d). In order to answer this and related questions one needs a detailed knowledge of the structure of the isotropic lines in $\mathbb{Z}_d^2$. In particular, it would be useful to know their number as a whole or with special conditions and also how they are arranged in orbits under the action of the symplectic group $\mathrm{Sp}(1,\mathbb{Z}_d)$.

We will be concerned only with the mathematical properties of the isotropic lines in $\mathbb{Z}_d^2$. In Section 6.1, we derive the number of isotropic lines in $\mathbb{Z}_d^2$ and then in Section 6.2 the number of them through a given point of the lattice. This should be compared with the results obtained by Havlicek and Saniga in [60] and [61] about the number of projective points in the lattice and the number of them under the same condition. In Section 6.3, we give a full description of the orbits of isotropic lines under the action of $\mathrm{Sp}(1,\mathbb{Z}_d)$ with the help of some parameters. In a fourth part, we develop two group actions on the group $\Sigma_{\mathscr{D}}(M)$ relevant to the understanding of that latter group, whenever M is a submodule of $\mathbb{Z}_d^2$. To end this introduction, we note that the results presented here do not depend on the parity of d, contrary to what happened in [56] and [57].

6.1 The number of isotropic lines

Let ω denote the symplectic product of two vectors of $\mathbb{Z}_d^2$. In terms of matrices, it consists in computing a determinant:

$$\omega((\alpha,\beta),(\gamma,\delta)) = \begin{vmatrix} \alpha & \gamma \\ \beta & \delta \end{vmatrix} = \alpha\delta - \beta\gamma. \tag{6.1}$$

We have already defined at the beginning of Chapter 4 what the orthogonal of a submodule and an isotropic submodule are. We recall that Lagrangian submodules are the maximal isotropic submodules for inclusion, which is equivalent to $M = M^\omega$.

As a first step, we are going to find the number of isotropic lines in $\mathbb{Z}_d^2$ for d a power of a prime, say $d = p^s$, $s \geq 1$. This leads to a hint for Section 6.4, but we shall also see that there exists a shortcut. We then address the case where d is arbitrary. From now on, we shall also make use without further ado of the equality

$$\mathrm{Sp}(1,\mathbb{Z}_d) = \mathrm{SL}(2,\mathbb{Z}_d), \tag{6.2}$$

where $\mathrm{Sp}(1,\mathbb{Z}_d)$ was defined in (4.10) and $\mathrm{SL}(2,\mathbb{Z}_d)$ is the group of 2×2 matrices with determinant 1. This equality is specific to the two-dimensional phase space $\mathbb{Z}_d^2$.

6.1.1 Special case: d a power of a prime

Let $\widetilde{s} = \lfloor s/2 \rfloor$, the floor part of $s/2$. According to Theorem 13 p.55, for any Lagrangian submodule M, there exist $S \in \mathrm{Sp}(1,\mathbb{Z}_d)$ and $k \in \{0,\dots,\widetilde{s}\}$ such that M is linearly generated by the column vectors of

$$S \times \begin{pmatrix} p^k & 0 \\ 0 & p^{s-k} \end{pmatrix}. \tag{6.3}$$

In other words, with S_1 and S_2 the two column vectors of S, (S_1,S_2) is a symplectic computational basis of $(\mathbb{Z}_{p^s})^2$ and M is the set of all linear combinations of $p^k S_1$ and $p^{s-k} S_2$ with coefficients in $\mathbb{Z}_{p^s}$. As a converse, any submodule thus generated is Lagrangian. In fact, the number k is a property of M, that is to say for any convenient pair (S,k') in order to generate M as in (6.3), we have $k' = k$. We will denote $\mathbf{O}_k(p^s)$ the set of all Lagrangian submodules thus obtained for a given k and S varying. The cardinality of any $M \in \mathbf{O}_k(p^s)$ is

$$p^{(s-1)-(k-1)} p^{(s-1)-(s-k-1)} = p^s. \tag{6.4}$$

Let ℓ be an isotropic line and $\langle \ell \rangle$ the submodule it generates, the set of all finite linear combinations of vectors of ℓ. Any two vectors in $\langle \ell \rangle$ are orthogonal and hence $\langle \ell \rangle$ is an isotropic submodule containing at least p^s vectors. Thus isotropic lines and Lagrangian submodules are the same.

The number of free vectors x in $\mathbb{Z}_{p^s}$ is $p^{2s} - p^{2(s-1)}$. The number of vectors y such that for a given free x we have $\omega(x, y) = 1$ is p^s. The number of pairs (x, y) such that $\omega(x, y) = 1$ is the product of the two previous ones:

$$n_\omega = |\mathrm{SL}(2, \mathbb{Z}_{p^s})| = p^{3s} - p^{3s-2}. \tag{6.5}$$

Several symplectic matrices S may give rise to the same submodule in $\mathbf{O}_k(p^s)$ according to the form (6.3). Let $k \in \{0, \ldots, \widetilde{s}\}$ and $M \in \mathbf{O}_k(p^s)$. Let $\Sigma_{\mathscr{D}}(M)$ be the matrix group of the changes of computational basis such that if $P \in \Sigma_{\mathscr{D}}(M)$ and if M is generated by the column vectors of the matrix given in (6.3), then M is also generated by the column vectors of the matrix

$$SP \times \begin{pmatrix} p^k & 0 \\ 0 & p^{s-k} \end{pmatrix}, \tag{6.6}$$

where SP need not be symplectic. The reader interested in details about $\Sigma_{\mathscr{D}}(M)$ may look at the beginning of Section 4.3.1. In particular, we derive from there that the group $\Sigma_{\mathscr{D}}(M)$ is completely determined by the value of k. So, the number of symplectic matrices that give rise to a given $M \in \mathbf{O}_k(p^s)$ is

$$n_{\mathscr{D}}(k) = |\Sigma_{\mathscr{D}}(M) \cap \mathrm{SL}(2, \mathbb{Z}_{p^s})| \tag{6.7}$$

and hence

$$|\mathbf{O}_k(p^s)| = \frac{n_\omega}{n_{\mathscr{D}}(k)}. \tag{6.8}$$

Let us suppose that $2k < s$. In $\Sigma_{\mathscr{D}}(M)$, the number of matrices with determinant 1 is the same as the number of matrices with any other (invertible) determinant. Indeed, if $u \in U(\mathbb{Z}_{p^s})$ and $P = (P_1|P_2) \in \Sigma_{\mathscr{D}}(M) \cap \mathrm{SL}(2, \mathbb{Z}_{p^s})$, with P_1 and P_2 the first and second columns of P respectively, then $(uP_1|P_2) \in \Sigma_{\mathscr{D}}(M)$ but with determinant u. This transformation is injective so that the number of matrices in $\Sigma_{\mathscr{D}}(M)$ with determinant u is greater than or equal to the number of matrices in $\Sigma_{\mathscr{D}}(M)$ with determinant 1. The converse inequality may be shown the same way. So we have

$$\begin{aligned} n_{\mathscr{D}}(k) &= \frac{|\Sigma_{\mathscr{D}}(M)|}{|U(\mathbb{Z}_{p^s})|} = \frac{(p^s - p^{s-1})^2 \cdot p^{(s-1)-(s-2k-1)} \cdot p^s}{p^s - p^{s-1}} = (p^s - p^{s-1})p^{s+2k} && \text{(6.9a)} \\ &= p^{2s}(p^{2k} - p^{2k-1}) && \text{(6.9b)} \end{aligned}$$

and so

$$|\mathbf{O}_k(p^s)| = \frac{p^s - p^{s-2}}{p^{2k} - p^{2k-1}} = p^{s-2k-1}(p+1). \tag{6.10}$$

If $2k = s$ (which supposes that s is even), then $\Sigma_{\mathscr{D}}(M) \cap \mathrm{SL}(2, \mathbb{Z}_{p^s}) = \mathrm{SL}(2, \mathbb{Z}_{p^s})$ and so

$$\left|\mathbf{O}_{s/2}(p^s)\right| = \frac{n_\omega}{n_\omega} = 1. \tag{6.11}$$

For s odd, then $2\widetilde{s} = s - 1$,

$$\sum_{k=0}^{\widetilde{s}} p^{-2k} = \frac{1 - p^{-2(\widetilde{s}+1)}}{1 - p^{-2}} = \frac{p^{2(\widetilde{s}+1)} - 1}{p^{2(\widetilde{s}+1)} - p^{2\widetilde{s}}} = \frac{p^{s+1} - 1}{p^{s+1} - p^{s-1}} \tag{6.12}$$

and hence the number of isotropic lines is

$$n_L(p^s) = \sum_{k=0}^{\widetilde{s}} |\mathbf{O}_k(p^s)| = p^{s-1}(p+1)\frac{p^{s+1} - 1}{p^{s+1} - p^{s-1}} = \frac{p^{s+1} - 1}{p - 1}. \tag{6.13}$$

If s is even, then $2\widetilde{s} = s$,

$$\sum_{k=0}^{\widetilde{s}-1} p^{-2k} = \frac{1 - p^{-2\widetilde{s}}}{1 - p^{-2}} = \frac{p^{2\widetilde{s}} - 1}{p^{2\widetilde{s}} - p^{2\widetilde{s}-2}} = \frac{p^s - 1}{p^s - p^{s-2}} \tag{6.14}$$

and hence the number of isotropic lines is again

$$\begin{aligned} n_L(p^s) = \sum_{k=0}^{\widetilde{s}-1} |\mathbf{O}_k(p^s)| + 1 &= p^{s-1}(p+1)\frac{p^s - 1}{p^s - p^{s-2}} + 1 \\ &= p\frac{p^s - 1}{p - 1} + 1 = \frac{p^{s+1} - p + p - 1}{p - 1} = \frac{p^{s+1} - 1}{p - 1}. \end{aligned} \tag{6.15}$$

6.1.2 General case: d any integer ≥ 2

Now let d be any integer greater than or equal to 2 and

$$d = \prod_{i \in I} p_i^{s_i} \tag{6.16}$$

be the prime factor decomposition of d. With the help of the Chinese remainder theorem, we can study the structure of an isotropic line ℓ in each of the Chinese factor $(\mathbb{Z}_{p_i^{s_i}})^2$. For every $i \in I$, let $\ell_i = \pi_{p_i}(\ell)$ be the i-th Chinese projection of ℓ. As a subgroup of $(\mathbb{Z}_{p_i^{s_i}})^2$, $\langle \ell_i \rangle$ has cardinality a power of p_i, say $p_i^{t_i}$. As an isotropic submodule of $(\mathbb{Z}_{p_i^{s_i}})^2$, $\langle \ell_i \rangle$ is included in a Lagrangian submodule and then $t_i \leq s_i$. So

$$d = |\ell| \leq \prod_{i \in I} |\ell_i| \leq \prod_{i \in I} p_i^{t_i} \leq d, \tag{6.17}$$

which proves that $t_i = s_i$. Moreover, if $\ell_i \subsetneq \langle \ell_i \rangle$ for some i, the second inequality just above would be strict, which is impossible and so $\ell_i = \langle \ell_i \rangle$ is a Lagrangian submodule of $(\mathbb{Z}_{p_i^{s_i}})^2$. As to the converse, for all $i \in I$, let ℓ'_i be a Lagrangian submodule of $(\mathbb{Z}_{p_i^{s_i}})^2$. The set ℓ' of all vectors $x \in \mathbb{Z}_d^2$ such that for all i, $\pi_{p_i}(x) \in \ell'_i$, is an isotropic set with cardinality d, namely an isotropic line. The reader may check that the maps $\ell \mapsto (\ell_i)_{i\in I}$ and $(\ell'_i)_{i\in I} \mapsto \ell'$ thus defined are reciprocal of one another.

So, isotropic lines and Lagrangian submodules are the same sets of $\mathbb{Z}_d^2$ and the number of isotropic lines of $\mathbb{Z}_d^2$ is

$$n_L(d) = \prod_{i\in I} n_L\left(p_i^{s_i}\right) = \prod_{i\in I} \frac{p_i^{s_i+1}-1}{p_i-1}. \tag{6.18}$$

Remark 1 *In (6.3), the left-hand-side factor was a symplectic matrix. But in fact, any invertible matrix would be convenient since we are to consider all the linear combinations of the columns in the product. Thus we could have calculated the cardinality of an orbit as*

$$\frac{n_\omega\,|U(\mathbb{Z}_{p^s})|}{|\Sigma_{\mathscr{D}}(M)|} \text{ instead of } \frac{n_\omega}{(|\Sigma_{\mathscr{D}}(M)|\,/\,|U(\mathbb{Z}_{p^s})|)} \tag{6.19}$$

and the argument between (6.8) and (6.9) could have been avoided.

Remark 2 *Let us assume that s is even. It should be noticed that the formula for the cardinality of $\mathbf{O}_k(p^s)$ given in (6.10) is not valid for $k = s/2$. Indeed, Equation (6.10) gives $1 + 1/p$ for that particular value of k, which is not even an integer. Equivalently, $n_{\mathscr{D}}(k)$ and $|\Sigma_{\mathscr{D}}(M)|$ have no unique expression for all values of k. This must be traced back to the behaviour of $\Sigma_{\mathscr{D}}(M)$ when k is ranging up to $s/2$.*

6.2 The number of lines through a given point

We now give the number of isotropic lines through a given point of the lattice. We suppose that $d = p^s$ is a power of a prime. Let $x \in \mathbb{Z}_d^2$ and let $t = v_p(x)$ be the p-valuation of x. Since all the vectors in an isotropic line $\ell \in \mathbf{O}_k(p^s)$ have p-valuation at least k, the vector x cannot be in ℓ unless $k \leq t$. Let us assume that k is such that $s - k \leq t$, which implies that $k \leq t$. Then for any computational basis (f_1, f_2), symplectic or not, x is a linear combination of $p^k f_1$ and $p^{s-k} f_2$. Hence

$$\forall k \in \{0, \ldots, \lfloor s/2 \rfloor\},\ \forall \ell \in \mathbf{O}_k(p^s), \quad (k \geq s - t \Longrightarrow x \in \ell). \tag{6.20}$$

That case can occur only if $t \geq \lceil s/2 \rceil$, the ceiling part of $s/2$. Now, let us assume that k is such that $k \leq t < s - k$. Thus $2k < s$ and we search for the symplectic computational bases (f_1, f_2) such that x is a linear combination of $p^k f_1$ and $p^{s-k} f_2$. Let (f_1, f_2) be a symplectic computational basis and $x = af_1 + bf_2$. Since

$$v_p(\omega(x, f_2)) = v_p(a) \geq t \geq k, \tag{6.21}$$

we have no extra conditions on the choice of f_2. But we must have

$$v_p(\omega(x, f_1)) = v_p(b) \geq s - k, \tag{6.22}$$

which shows that in a symplectic basis where $x = (p^t, 0)$, f_1 must be of the form

$$f_1 = (\alpha, \beta p^{s-k-t}), \tag{6.23}$$

with $\alpha, \beta \in \mathbb{Z}_d$. The number of suitable vectors f_1 is

$$(p^s - p^{s-1}) \cdot p^{(s-1)-(s-k-t-1)} = (p^s - p^{s-1})p^{k+t}. \tag{6.24}$$

The number of suitable vectors f_2 for a given f_1 is p^s. Then the number of suitable, symplectic computational bases (f_1, f_2) is $(p^s - p^{s-1})p^{s+k+t}$. Moreover, if f is a convenient basis and

$$\left\langle p^k f_1, p^{s-k} f_2 \right\rangle = \left\langle p^k f_1', p^{s-k} f_2' \right\rangle, \tag{6.25}$$

then f' is convenient too. With (6.9a), we deduce that the number of isotropic lines in $\mathbf{O}_k(p^s)$ containing x is

$$\frac{(p^s - p^{s-1})p^{s+k+t}}{(p^s - p^{s-1})p^{s+2k}} = p^{t-k}. \tag{6.26}$$

Thus, if $t < \lceil s/2 \rceil$, the number of isotropic lines containing x is

$$\sum_{k=0}^{t} p^{t-k} = p^t \cdot \frac{1 - p^{-(t+1)}}{1 - p^{-1}} = \frac{p^{t+1} - 1}{p - 1}. \tag{6.27}$$

If $t \geq \lceil s/2 \rceil$ and $\widetilde{s} = \lfloor s/2 \rfloor$, the number of isotropic lines containing x is

$$\sum_{k=0}^{s-t-1} p^{t-k} + \sum_{k=s-t}^{\widetilde{s}} |\mathbf{O}_k(p^s)|. \tag{6.28}$$

The first term is equal to

$$p^t \cdot \frac{1 - p^{-(s-t)}}{1 - p^{-1}} = \frac{p^{t+1} - p^{2t-s+1}}{p - 1}. \tag{6.29}$$

For s odd, then $2\widetilde{s} = s - 1$,

$$\sum_{k=s-t}^{\widetilde{s}} p^{-2k} = p^{-2(s-t)} \cdot \frac{1 - p^{-2(\widetilde{s}-s+t+1)}}{1 - p^{-2}} = \frac{p^{2t-s+1} - 1}{p^{s-1}(p^2 - 1)}, \tag{6.30}$$

and the second term in (6.28) is equal to

$$p^{s-1}(p + 1)\frac{p^{2t-s-1} - 1}{p^{s-1}(p^2 - 1)} = \frac{p^{2t-s+1} - 1}{p - 1}. \tag{6.31}$$

For s even, then $2\widetilde{s} = s$,

$$\sum_{k=s-t}^{\widetilde{s}-1} p^{-2k} = p^{-2(s-t)} \cdot \frac{1 - p^{-2(\widetilde{s}-1-s+t+1)}}{1 - p^{-2}} = \frac{p^{2t-s+1} - p}{p^{s-1}(p^2 - 1)}, \tag{6.32}$$

and the second term in (6.28) is again

$$p^{s-1}(p+1)\frac{p^{2t-s+1}-p}{p^{s-1}(p^2-1)}+1=\frac{p^{2t-s+1}-1}{p-1}. \tag{6.33}$$

Hence, in any case, the number of isotropic lines containing some given vector x with p-valuation t is

$$n_L(p^s;x)=n_L(p^s;t)=\frac{p^{t+1}-1}{p-1}. \tag{6.34}$$

In particular,

$$n_L(p^s;t=0)=1 \quad \text{and} \quad n_L(p^s;t=s)=n_L(p^s). \tag{6.35}$$

That is to say the sole isotropic line containing a free vector is the submodule it generates and every isotropic line goes through the null vector.

If d is not necessarily a power of a prime, then with (6.16) and for all i, $t_i = v_{p_i}(x)$, we obtain that the number of isotropic lines containing x is

$$n_L(d;x)=n_L(d;(t_i)_{i\in I})=\prod_{i\in I}\frac{p_i^{t_i+1}-1}{p_i-1}. \tag{6.36}$$

6.3 Orbits under the action of $\mathrm{Sp}(1, \mathbb{Z}_d)$

As in Section 6.1, we first suppose that d is a power of a prime, say $d = p^s$, $s \geq 1$. Then it is obvious from (6.3) that the orbits of the left-action of $\mathrm{Sp}(1, \mathbb{Z}_d)$ among the isotropic lines are the $\mathbf{O}_k(p^s)$. Their number is $\lfloor s/2 \rfloor + 1$ and we have already seen what their cardinalities are in (6.10) and (6.11).

Now if d is a composite integer as in (6.16), then the set of the orbits is parametrised by

$$k=(k_i)_{i\in I}\in\prod_{i\in I}\{0,\dots,\lfloor s_i/2\rfloor\} \tag{6.37}$$

and the orbit with index k is

$$\mathbf{O}_k(d)=\left\{\ell\subset\mathbb{Z}_d^2;\ |\ell|=d,\ \pi_{p_i}(\ell)\in\mathbf{O}_{k_i}(p_i^{s_i})\right\}. \tag{6.38}$$

The number of orbits is

$$\prod_{i\in I}\left(\lfloor s_i/2\rfloor+1\right), \tag{6.39}$$

and the cardinality of one of them is

$$|\mathbf{O}_k(d)|=\prod_{i\in I}|\mathbf{O}_{k_i}(p_i^{s_i})|. \tag{6.40}$$

Example: Let us suppose that d contains no square factor, that is to say in (6.16), for all $i \in I$, $s_i = 1$. According to (6.3), with k necessarily equal to 0, the isotropic lines are the submodules that can be generated by a single free vector. These submodules are called the projective points of $\mathbb{Z}_d^2$. With (6.18), we find that the number of isotropic lines is

$$n_L(d) = \prod_{i\in I}(p_i + 1). \tag{6.41}$$

They all belong to the sole orbit under the action of $\mathrm{Sp}(1, \mathbb{Z}_d)$ corresponding to $k_i = 0$ for all i. □

In conclusion, we counted the isotropic lines of $\mathbb{Z}_d^2$ as a whole or through a given point and we saw how they are arranged under the action of $\mathrm{Sp}(1, \mathbb{Z}_d) = \mathrm{SL}(2, \mathbb{Z}_d)$. We used this latter equality between the symplectic group and the special linear group as a counting argument in Section 6.1.1. But in the higher dimensional phase space $\mathbb{Z}_d^{2n}$, this equality does not hold any more. Thus the case of $\mathbb{Z}_d^2$ is quite a simple one. Whenever $\mathbb{Z}_d^{2n}$ had to be considered, the method we used would have to be modified. For the same reason, the content of the next section is also particular to $\mathbb{Z}_d^2$.

6.4 Some group actions on $\Sigma_{\mathscr{D}}(M)$

In this additional section, we highlight the structure of the group $\Sigma_{\mathscr{D}}(M)$, whenever M is a submodule of $\mathbb{Z}_d^2$, by means of three group actions upon it. We also assume that $d = p^s$ is a power of a prime.

In order to establish Equation (6.9), we showed that the number of matrices in $\Sigma_{\mathscr{D}}(M)$ with determinant 1 is the same as the number of matrices in the same set with any other (invertible) determinant. The simple reasoning we used was enough in the context of Section 6.1. But we are going to introduce here two other group actions that are linked to that point and to Remarks 1 and 2. Let ρ_0 be the action of $U(\mathbb{Z}_{p^s})$ on $\Sigma_{\mathscr{D}}(M)$ defined by

$$\forall u \in U(\mathbb{Z}_{p^s}),\ \forall P = (P_1|P_2) \in \Sigma_{\mathscr{D}}(M), \quad \rho_0(u)\cdot P = (uP_1|u^{-1}P_2) \tag{6.42}$$

and ρ_1 the action of $U(\mathbb{Z}_{p^s})^2$ on $\Sigma_{\mathscr{D}}(M)$ defined by

$$\forall (u_1, u_2) \in U(\mathbb{Z}_{p^s})^2,\ \forall P = (P_1|P_2) \in \Sigma_{\mathscr{D}}(M), \quad \rho_1(u_1, u_2)\cdot P = (u_1P_1|u_2P_2). \tag{6.43}$$

All the orbits of ρ_0 (resp. ρ_1) have the same cardinality, namely $|U(\mathbb{Z}_{p^s})| = p^s - p^{s-1}$ (resp. $|U(\mathbb{Z}_{p^s})|^2$). In a given orbit of ρ_0, every matrix has the same determinant. Since $\mathbb{Z}_{p^s}$ is a commutative ring, those two actions "commute":

$$\rho_1(u_1, u_2)\cdot(\rho_0(u)\cdot P) = \rho_0(u)\cdot(\rho_1(u_1, u_2)\cdot P). \tag{6.44}$$

Let $(u_1, u_2), (v_1, v_2) \in U(\mathbb{Z}_{p^s})^2$ such that $u_1u_2 = v_1v_2$, that is to say

$$\forall P \in \Sigma_{\mathscr{D}}(M), \quad \det(\rho_1(u_1, u_2)\cdot P) = \det(\rho_1(v_1, v_2)\cdot P). \tag{6.45}$$

With $\lambda = u_2 v_2^{-1} = u_1^{-1} v_1 \in U(\mathbb{Z}_{p^s})$, we have

$$(v_1, v_2) = (\lambda u_1, \lambda^{-1} u_2). \tag{6.46}$$

Thus we have a kind of a discrete Hopf fibration. It is given by the action h of $U(\mathbb{Z}_{p^s})$ on $U(\mathbb{Z}_{p^s})^2$ defined by

$$\forall \lambda \in U(\mathbb{Z}_{p^s}),\ \forall (u_1, u_2) \in U(\mathbb{Z}_{p^s})^2, \quad h(\lambda) \cdot (u_1, u_2) = (\lambda u_1, \lambda^{-1} u_2). \tag{6.47}$$

Moreover, the action $\rho = \rho_1/(h, \rho_0)$ of $U(\mathbb{Z}_{p^s})^2/h$ on $\Sigma_{\mathscr{D}}(M)/\rho_0$ is well-defined. For any $u \in U(\mathbb{Z}_{p^s})$, let

$$D_u = \{P \in \Sigma_{\mathscr{D}}(M);\ \det P = u\}. \tag{6.48}$$

Every orbit of ρ is transversal to D_u/ρ_0. Indeed, let P be in some orbit O of ρ_1 with some determinant v. Then $(uv^{-1}P_1|P_2)$ is in O with determinant u so that there is at least one orbit of ρ_0 in $D_u \cap O$. Then if P and $Q = (u_1 P_1 | u_2 P_2)$ are in O and have the same determinant, then $u_2 = u_1^{-1}$ and thus P and Q are in the same orbit of ρ_0.

As a conclusion, we have the

Proposition 29 *The group $\Sigma_{\mathscr{D}}(M)$ can be partitioned into the family $E = \{E_{ij}\}$, where the E_{ij}'s are the orbits of ρ_0, the index $i \in U(\mathbb{Z}_{p^s})$ is the determinant of every matrix in E_{ij} and the index j stands for an orbit of ρ_1 (or equivalently of ρ).*

Let $P \in E_{i_1 j_1}$ and $Q \in E_{i_2 j_2}$. On the one hand, $\det P = i_1$ *and* $\det Q = i_2$*. On the other hand, $j_1 = j_2$ iff Q_1 and Q_2 are proportional to P_1 and P_2 respectively.*

The number of different values that j can assume is

$$n_\rho = \frac{|\Sigma_{\mathscr{D}}(M)|}{|U(\mathbb{Z}_{p^s})|^2}. \tag{6.49}$$

If $2k < s$, then $n_\rho = p^{s+2k}$ according to (6.9a). But if $k = s/2$, then

$$n_\rho = \frac{|\mathrm{GL}(2, \mathbb{Z}_{p^s})|}{|U(\mathbb{Z}_{p^s})|^2} = \frac{(p^{2s} - p^{2(s-1)}) \cdot (p^s - p^{s-1}) \cdot p^s}{(p^s - p^{s-1})^2} = p^{2s} + p^{2s-1} > p^{2s}. \tag{6.50}$$

In passing, we find again that the number of matrices in $\Sigma_{\mathscr{D}}(M)$ with some determinant u is the same as the number of matrices in $\Sigma_{\mathscr{D}}(M)$ with any other (invertible) determinant v.

Chapter 7

Towards continuous algebra and geometry

Up to now, we have dealt with a discrete set of operators, namely Pauli operators, and thus we got information only about a discrete set of bases. With this restriction, we were constantly led to reduce the problem to the case of a power of a prime dimension. The behaviour of Pauli operators is indeed liable to the Chinese remainder theorem. Thus, as far as Pauli operators are concerned, the lowest prime power factor of d is a limiting factor in the number of MUBs. But can one find a similar, algebraic or geometrical framework which would account for any set of bases or at least for any set of MUBs? Can the limitation be thus removed. In this chapter, we present some tools that are likely to provide an answer to these questions. In order to give a flavour of their pertaining to algebraic geometry, we work them out in dimensions 2 and 3.

7.1 Qubits and the cross-ratio

In this section, we put $d = 2$. Qubits are normalised vectors in $\mathbb{C}^2$ or equivalently lines through the origin in the same space. From this latter point of view, they appear as points of the projective line over $\mathbb{C}$ which is nothing but the Bloch sphere of physicists, or the Riemann sphere of mathematicians. It is well-known that orthonormal bases of qubits are represented by pairs of opposite points on the sphere. As is usual, we shall take $|0\rangle$ and $|1\rangle$ to be the north and the south poles of the Bloch sphere, respectively. In the projective, algebraic language, we fix a system of homogeneous coordinates so that $|0\rangle$ is 0 and $|1\rangle$ is the point at infinity, denoted ∞:

$$|0\rangle \leftrightarrow (1,0)\,, \qquad |1\rangle \leftrightarrow (0,1)\,. \tag{7.1}$$

This convention corresponds to the stereographic projection from the south pole. The homogeneous coordinates of a vector are thus nothing but the coefficients of that vector in the computational basis $\{|0\rangle\,, |1\rangle\}$ up to a nonzero, multiplicative complex number:

$$a\,|0\rangle + b\,|1\rangle \;\leftrightarrow\; (a,b) \;\leftrightarrow\; (ka,kb) \text{ for any } k \in \mathbb{C}^*. \tag{7.2}$$

Then any basis unbiased with the computational basis is of the form

$$\left\{ |0\rangle' = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\varphi}|1\rangle); \quad |1\rangle' = \frac{1}{\sqrt{2}}(|0\rangle - e^{i\varphi}|1\rangle) \right\}, \tag{7.3}$$

where φ is the azimuthal coordinate. This basis is represented by two opposite points on the equator of the Bloch sphere or by the homogeneous coordinates

$$|0\rangle' \leftrightarrow \left(1, e^{i\varphi}\right), \qquad |1\rangle' \leftrightarrow \left(1, -e^{i\varphi}\right), \tag{7.4}$$

so that $\{|0\rangle, |1\rangle\}$ and $\{|0\rangle', |1\rangle'\}$ are harmonic conjugated. In fact, let us calculate their anharmonic ratio:

$$\begin{aligned}
(|0\rangle, |1\rangle, |0\rangle', |1\rangle') &= \left(\begin{vmatrix} 1 & 1 \\ e^{i\varphi} & 0 \end{vmatrix} \cdot \begin{vmatrix} 1 & 0 \\ -e^{i\varphi} & 1 \end{vmatrix}, \begin{vmatrix} 1 & 0 \\ e^{i\varphi} & 1 \end{vmatrix} \cdot \begin{vmatrix} 1 & 1 \\ -e^{i\varphi} & 0 \end{vmatrix} \right) && (7.5a) \\
&= (-e^{i\varphi}, e^{i\varphi}) = (1, -1) && (7.5b) \\
&= -1. && (7.5c)
\end{aligned}$$

Conversely, if some orthonormal basis whose vectors have homogeneous coordinates $(a, b), (c, d) \in \mathbb{C}^2$ is harmonic conjugated with $\{|0\rangle, |1\rangle\}$, we have

$$\left(\begin{vmatrix} a & 1 \\ b & 0 \end{vmatrix} \cdot \begin{vmatrix} c & 0 \\ d & 1 \end{vmatrix}, \begin{vmatrix} a & 0 \\ b & 1 \end{vmatrix} \cdot \begin{vmatrix} c & 1 \\ d & 0 \end{vmatrix} \right) = (-bc, -ad) = (1, -1). \tag{7.6}$$

So $ad + bc = 0$. Writing a, b, c, d in terms of the spherical coordinates θ and φ

$$a = \cos\frac{\theta}{2}, \quad b = e^{i\varphi}\sin\frac{\theta}{2}, \tag{7.7a}$$

$$c = \sin\frac{\theta}{2}, \quad d = -e^{i\varphi}\cos\frac{\theta}{2}, \tag{7.7b}$$

we get that $\theta = \pi/2$ with no conditions on φ. Then the basis we started from is of the form (7.3). Since our choice of a computational basis was arbitrary, we have shown the

Proposition 30 *Let (x_1, x_2) and (y_1, y_2) be two orthonormal bases of $\mathbb{C}^2$. Then the following properties are equivalent:*

1. *(x_1, x_2) and (y_1, y_2) are unbiased;*

2. *On the Bloch sphere, the axis defined by (x_1, x_2) and (y_1, y_2) are orthogonal;*

3. *$(x_1, x_2, y_1, y_2) = -1$, that is the two bases are harmonic conjugated.*

The point is that any possible set of MUBs is accounted for in the latter proposition. However, as we do not have at our disposal a suitable generalisation of the Bloch sphere or of the anharmonic ratio in any dimension, we still have to find another scheme. On the one hand, since in the pair (x_1, x_2), any one of the two vectors can be deduced from the other, there is in fact a redundancy and we may search for a tool that encode a basis in a single object. This is done in the next section where we try Pauli operators with continuous exponents. On the other hand, we may think that the redundancy is specific to low dimension cases and so find a generalisation that keeps track of each vector in a basis. We will try this way in Section 7.3.

7.2 Pauli operators with continuous exponents

This section is still devoted to qubits. Indeed, we shall see in conclusion that Pauli operators with continuous, real exponents are essentially limited to qubits.

The definition of Z^a for instance, with a in $\mathbb{R}$ instead of $\mathbb{Z}$, involves a matrix logarithm, so that we must avoid any negative or null eigenvalue. Therefore, we consider iP in replacement of any Pauli operator P over $\mathbb{C}^2$. Such a replacement does not modify the eigenbasis under consideration and we may define

$$\forall a \in \mathbb{R}, \quad (iP)^a = e^{a \ln(iP)}. \tag{7.8}$$

In particular, with

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, \tag{7.9}$$

the so-called Hadamard and phase matrices respectively, we have

$$\ln(iZ) = \ln \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = \begin{pmatrix} i\pi/2 & 0 \\ 0 & -i\pi/2 \end{pmatrix} = i\frac{\pi}{2} Z, \tag{7.10a}$$

$$\ln(iX) = H \ln(iZ) H = \begin{pmatrix} 0 & i\pi/2 \\ i\pi/2 & 0 \end{pmatrix} = i\frac{\pi}{2} X, \tag{7.10b}$$

$$\ln(iY) = S \ln(iX) S^\dagger = \begin{pmatrix} 0 & \pi/2 \\ -\pi/2 & 0 \end{pmatrix} = i\frac{\pi}{2} Y. \tag{7.10c}$$

As a general feature, one has for any Pauli operator P,

$$\ln(iP) = i\frac{\pi}{2} P. \tag{7.11}$$

With $a, b \in \mathbb{R}$, one has

$$\begin{aligned} (iZ)^b &= \exp\left(ib\frac{\pi}{2} Z\right) = \cos\left(b\frac{\pi}{2}\right) + i \sin\left(b\frac{\pi}{2}\right) Z \\ &= \begin{pmatrix} e^{ib\pi/2} & 0 \\ 0 & e^{-ib\pi/2} \end{pmatrix}, \end{aligned} \tag{7.12a}$$

$$\begin{aligned} (iX)^a &= H (iZ)^a H = \cos\left(a\frac{\pi}{2}\right) + i \sin\left(a\frac{\pi}{2}\right) X \\ &= \begin{pmatrix} \cos(a\pi/2) & i\sin(a\pi/2) \\ i\sin(a\pi/2) & \cos(a\pi/2) \end{pmatrix} \end{aligned} \tag{7.12b}$$

and hence

$$\begin{aligned}(iX)^a(iZ)^b &= \cos\left(a\frac{\pi}{2}\right)\cos\left(b\frac{\pi}{2}\right)I + i\sin\left(a\frac{\pi}{2}\right)\cos\left(b\frac{\pi}{2}\right)X \\ &\quad + i\sin\left(a\frac{\pi}{2}\right)\sin\left(b\frac{\pi}{2}\right)Y + \cos\left(a\frac{\pi}{2}\right)\sin\left(b\frac{\pi}{2}\right)Z \\ &= \begin{pmatrix} \cos(a\pi/2)e^{ib\pi/2} & i\sin(a\pi/2)e^{-ib\pi/2} \\ i\sin(a\pi/2)e^{ib\pi/2} & \cos(a\pi/2)e^{-ib\pi/2} \end{pmatrix}. \qquad (7.12c)\end{aligned}$$

So we may restrict a and b to lie in $[0,4[$. The characteristic polynomial of the latter matrix, with formal variable λ, is

$$\begin{aligned}(\lambda - \cos(a\pi/2)e^{ib\pi/2})(\lambda - \cos(a\pi/2)e^{-ib\pi/2}) + \sin^2(a\pi/2) \\ = \lambda^2 - 2\cos(a\pi/2)\cos(b\pi/2)\lambda + 1 \quad (7.13a)\end{aligned}$$

with reduced discriminant

$$\Delta' = \cos^2(a\pi/2)\cos^2(b\pi/2) - 1, \qquad (7.13b)$$

negative unless $\{a,b\} \subset \{0,2\}$, that is $(iX)^a(iZ)^b = \pm I$. If we exclude that case, we have two conjugated complex eigenvalues

$$\lambda_\pm = \cos(a\pi/2)\cos(b\pi/2) \pm i\sqrt{1 - \cos^2(a\pi/2)\cos^2(b\pi/2)} \qquad (7.13c)$$

with corresponding eigenvectors

$$|ab\pm\rangle = \begin{pmatrix} i\sin(a\pi/2)e^{-ib\pi/2} \\ -\cos(a\pi/2)e^{ib\pi/2} + \lambda_\pm \end{pmatrix}. \qquad (7.13d)$$

The basis $\{|ab\pm\rangle\}$ is unbiased with the computational basis iff

$$\begin{aligned}\sin^2(a\pi/2) &= (\cos(a\pi/2)e^{ib\pi/2} - \lambda_+)(\cos(a\pi/2)e^{-ib\pi/2} - \lambda_-), \quad (7.14a) \\ \sin^2(a\pi/2) &= (\cos(a\pi/2)e^{ib\pi/2} - \lambda_-)(\cos(a\pi/2)e^{-ib\pi/2} - \lambda_+). \quad (7.14b)\end{aligned}$$

Equating the right-hand sides of these equations, we get the following necessary condition:

$$\mathrm{Re}(\cos(a\pi/2)e^{ib\pi/2}\lambda_+) = \mathrm{Re}(\cos(a\pi/2)e^{ib\pi/2}\lambda_-). \qquad (7.15)$$

If $a \neq 1$ and 3, we may get rid of the factor $\cos(a\pi/2)$ and for $(iX)^a(iZ)^b$ nontrivial, (7.15) reduces to

$$b = 0 \text{ or } 2. \qquad (7.16)$$

One may check that Conditions (7.14a) and (7.14b) are then verified. The corresponding Pauli operator is $(iX)^a$.

Now if $a = 1$ for instance, one gets

$$|1b\pm\rangle = \begin{pmatrix} ie^{-ib\pi/2} \\ \pm i \end{pmatrix}. \qquad (7.17)$$

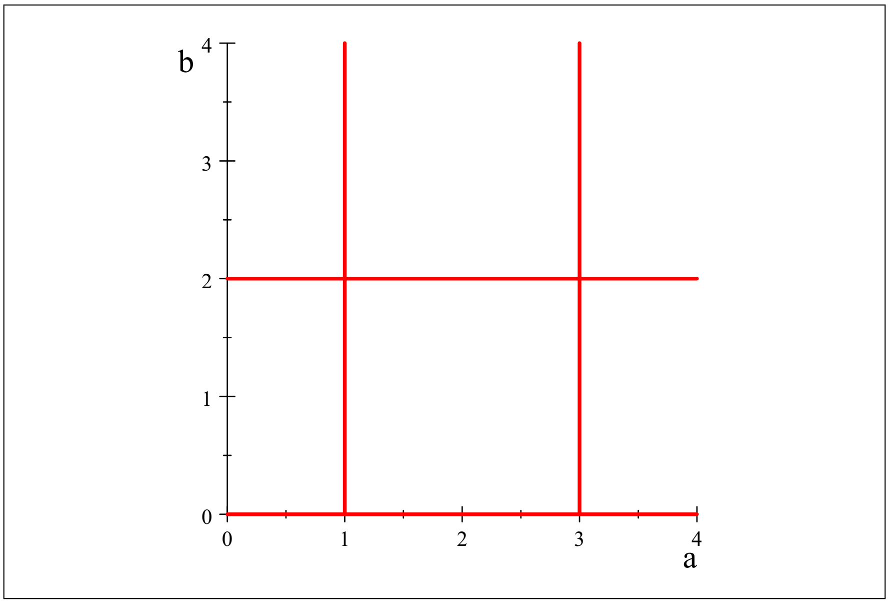

Figure 7.1: Solutions for a and b so that $(iX)^a(iZ)^b$ provide an unbiased basis with the canonical one. The points with $a = 0, 2, 4$ are excluded.

This is nothing but (7.3) up to multiplicative factor with $\varphi = b\pi/2$. The case $a = 3$ is similar: It describes the same set of bases but each of them with reverse orientation. In summary, the solutions for a and b in order that $(iX)^a(iZ)^b$ provide an unbiased basis with the canonical one are given by the four red lines in Figure 7.1.

Thus, in dimension 2, Pauli operators provide us with a complete description of the bases unbiased with a given basis. Unfortunately, when the dimension of the Hilbert space increases, the dimension of the manifold of unbiased bases with a given basis also increases, whereas the set of Pauli operators X^aZ^b still depends on only two parameters. Therefore the pattern of Pauli operators with continuous, real exponents is again specific to qubits.

7.3 MUBs as tori intersection

7.3.1 General set-up

We consider the Hilbert space $\mathbb{C}^d$, with d any integer greater than or equal to 2. A orthonormal basis of that space consists of the vectors $|i\rangle$, with i ranging from 0 to $d-1$ in $\mathbb{Z}_d$. As we saw in the historical overview as well as in Section 7.1, a quantum state can be represented by a point in the corresponding projective space $\mathbf{P}(\mathbb{C}^d)$. With this representation, the canonical computational basis is given by

$$\begin{array}{ccccc} |0\rangle\,, & |1\rangle\,, & \dots & |d-2\rangle\,, & |d-1\rangle \\ \updownarrow & \updownarrow & & \updownarrow & \updownarrow \\ 0, & \infty_1, & \dots, & \infty_{d-2}, & \infty_{d-1} \end{array} \tag{7.18}$$

A vector $x \in \mathbb{C}^d$ is said to be unbiased with the canonical basis if it has equal amplitude over the canonical basis vectors:

$$\forall i, j \in \mathbb{Z}_d, \quad |\langle i|x\rangle| = |\langle j|x\rangle|. \tag{7.19}$$

Since such an x is not 0, in particular

$$|\langle 0|x\rangle| \neq 0 \tag{7.20}$$

and x can be identified with a vector $\widetilde{x}$ in $\mathbb{C}^{d-1}$. So a necessary and sufficient condition for x to be unbiased with the canonical basis is that $\widetilde{x}$ have all its coordinates of unit modulus. That is to say $\widetilde{x}$ has to lie on the unit torus $\mathbb{T}^{d-1}$ in $\mathbb{C}^{d-1}$.

We pick up a set of d vectors on $\mathbb{T}^{d-1}$ that account for an orthonormal basis f of $\mathbb{C}^d$. As we have just seen, f is unbiased with the canonical basis. We perform a change of computational basis that takes the canonical basis to f. This induces a transformation φ on $\mathbf{P}(\mathbb{C}^d)$ that is called an orthonormal projectivity. The set of unbiased vectors with f is represented by $\varphi(\mathbb{T}^{d-1})$ and the set of unbiased vectors with both the computational basis and f is represented by $\mathbb{T}^{d-1} \cap \varphi(\mathbb{T}^{d-1})$. Let $\mathbf{T}$ be the set of all images of $\mathbb{T}^{d-1}$ under such orthonormal projectivities:

$$\mathbf{T} = \{\varphi(\mathbb{T}^{d-1});\ \varphi \text{ orthonormal projectivity in } \mathbf{P}(\mathbb{C}^d)\}. \tag{7.21}$$

Thus the problem of finding unbiased bases is translated into the study of intersections between elements of $\mathbf{T}$, which is liable to algebraic geometry. Let us say that two tori are orthogonal if one is the image of the other by an orthonormal projectivity. This is a symmetric relation and we search for maximal sets of pairwise orthogonal tori. We treat the cases $d = 2$ and $d = 3$.

7.3.2 The case $d = 2$

For $d = 2$, we have already seen in Section 7.1 that the canonical basis is represented by the north and south poles on the Bloch sphere and that, via the stereographic projection, $\mathbb{T}^1$ is the equator of the sphere. Here an orthonormal projectivity is only a rotation on the sphere. This particularity helps visualize. But we will not prove it as we are not to use it in our formal reasoning.

Let $a \in [0, 2[$ so that πa is the angular coordinate of a point on the equator. Such a point accounts for a vector unbiased with the canonical basis. We perform an orthonormal projectivity φ that brings the canonical basis, namely the eigenstates of Z, to the eigenstates of X on the equator. Then $\varphi(\mathbb{T}^1)$ is parametrised by a and it contains the points with projective coordinates

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ e^{i\pi a} \end{pmatrix} = \begin{pmatrix} 1 + e^{i\pi a} \\ 1 - e^{i\pi a} \end{pmatrix}. \tag{7.22}$$

Note that the matrix representing φ is the Fourier transform in two dimensions, namely H up to multiplicative factor. For $a = 0$, we get the north pole and for $a = 1$, we get the south pole. This could be expected from the fact that unbiasedness is a symmetric relation, so that $\varphi(\mathbb{T}^1)$ has to go through the eigenstates of Z. In fact, $\varphi(\mathbb{T}^1)$ is also a great circle and thus is orthogonal to the equator in the common sense of the word *orthogonal*. So there is only one possibilty for the third unbiased basis that will complete the two previous ones. Its vectors lie in $\mathbb{T}^1 \cap \varphi(\mathbb{T}^1)$ which contains only two points, accounting for the eigenstates of Y. To see this rigourously, we search for the points of the form given in the right-hand-side of (7.22) that also belong to the equator. So we need solve

$$\left|1 + e^{i\pi a}\right| = \left|1 - e^{i\pi a}\right|, \tag{7.23}$$

which admits only two solutions:

$$a = \pm\frac{1}{2}. \tag{7.24}$$

The corresponding states are indeed the eigenstates of Y:

$$\frac{1}{\sqrt{2}}\left((1 \pm i)\,|0\rangle + (1 \mp i)\,|1\rangle\right). \tag{7.25}$$

Since the choice of the two first bases was in fact arbitrary, we have proved the

Proposition 31 *Given an orthonormal basis b_0 in $\mathbb{C}^2$, one vector unbiased with respect to this latter basis is enough to determine the two other orthornormal bases b_1, b_2 such that (b_0, b_1, b_2) is a complete set of unbiased bases.*

7.3.3 The case $d = 3$

Here we prove, with the help of numerical computations, the following rigidity proposition which is a mere copy of the 2-dimensional case just above.

Proposition 32 *Given an orthonormal basis b_0 in $\mathbb{C}^3$, one vector unbiased with respect to this latter basis is enough to determine the three other orthornormal bases b_1, b_2, b_3 such that (b_0, b_1, b_2, b_3) is a complete set of unbiased bases.*

For $d = 3$, if we want to build a basis f which is unbiased with the canonical basis e, we may choose a first vector anywhere on $\mathbb{T}^2$. So, without loss of generality, we choose the vector whose every coefficient is equal to 1. In matrix form:

$$f_1 = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}^T. \tag{7.26}$$

Moreover, a vector of f is defined only up to a global phase. Thus we may put

$$f = \begin{pmatrix} 1 & 1 & 1 \\ 1 & f_{22} & f_{23} \\ 1 & f_{32} & f_{33} \end{pmatrix}. \tag{7.27}$$

Then, f being orthonormal,

$$1 + f_{22} + f_{32} = 1 + f_{23} + f_{33} = 0, \tag{7.28}$$

where the f_{ij}'s have unit modulus. The order of the vectors in the basis does not matter as well. So, we have no other choice than to take the Fourrier transform of the canonical basis:

$$f = \begin{pmatrix} 1 & 1 & 1 \\ 1 & j & j^2 \\ 1 & j^2 & j \end{pmatrix}, \tag{7.29}$$

with j the canonical root of unity of order 3. In other words, a basis unbiased with the canonical one is determined by one of its vectors.

Now we search for unbiased bases with respect to both e and f. We know that there are only two of them. Let φ be the orthonormal projectivity that brings e to f. In matrix form, φ is given by the latter matrix f. As in the 2-dimensional case, we parametrise $\varphi(\mathbb{T}^2)$ with parameters in $\mathbb{T}^2$. Let $a, b \in [0, 3[$ so that the generic point of $\varphi(\mathbb{T}^2)$ has projective coordinates

$$f \begin{pmatrix} 1 \\ e^{2i\pi a/3} \\ e^{2i\pi b/3} \end{pmatrix} = \begin{pmatrix} 1 + e^{2i\pi a/3} + e^{2i\pi b/3} \\ 1 + je^{2i\pi a/3} + j^2 e^{2i\pi b/3} \\ 1 + j^2 e^{2i\pi a/3} + je^{2i\pi b/3} \end{pmatrix} \tag{7.30}$$

In order to describe $\mathbb{T}^2 \cap \varphi(\mathbb{T}^2)$, we need solve the two equations

$$|1 + e^{2i\pi a/3} + e^{2i\pi b/3}| = |1 + je^{2i\pi a/3} + j^2 e^{2i\pi b/3}| \tag{7.31a}$$
$$|1 + e^{2i\pi a/3} + e^{2i\pi b/3}| = |1 + j^2 e^{2i\pi a/3} + je^{2i\pi b/3}| \tag{7.31b}$$

The solutions of (7.31a) form the set of three red, solid lines displayed in Figure 7.2. The solutions of (7.31b) form the set of three blue, dashed lines displayed in the same figure. This figure was obtained as an implicit plot by numerical computation. From a topological point of view, each set of lines is the join of a parallel, a meridian and a Hopf fiber intersecting in a common point[1]. The two sets of lines intersect in six points that account for nothing but the two expected unbiased bases. One obtains those bases b_1 and b_2 under transformation by f:

$$b_1 = f \begin{pmatrix} 1 & 1 & 1 \\ j & 1 & j^2 \\ j & j^2 & 1 \end{pmatrix} = \begin{pmatrix} i\sqrt{3} & 2+j^2 & 2+j^2 \\ 2+j^2 & i\sqrt{3} & 2+j^2 \\ 2+j^2 & 2+j^2 & i\sqrt{3} \end{pmatrix}, \tag{7.32}$$

$$b_2 = f \begin{pmatrix} 1 & 1 & 1 \\ j^2 & j & 1 \\ j^2 & 1 & j \end{pmatrix} = \begin{pmatrix} -i\sqrt{3} & 2+j & 2+j \\ 2+j & -i\sqrt{3} & 2+j \\ 2+j & 2+j & -i\sqrt{3} \end{pmatrix}. \tag{7.33}$$

[1] With perhaps a far-fetched argument, we remark that f is symmetric, so that each of the two intersecting points appears precisely as the point that gave rise to the equation with solutions passing through it.

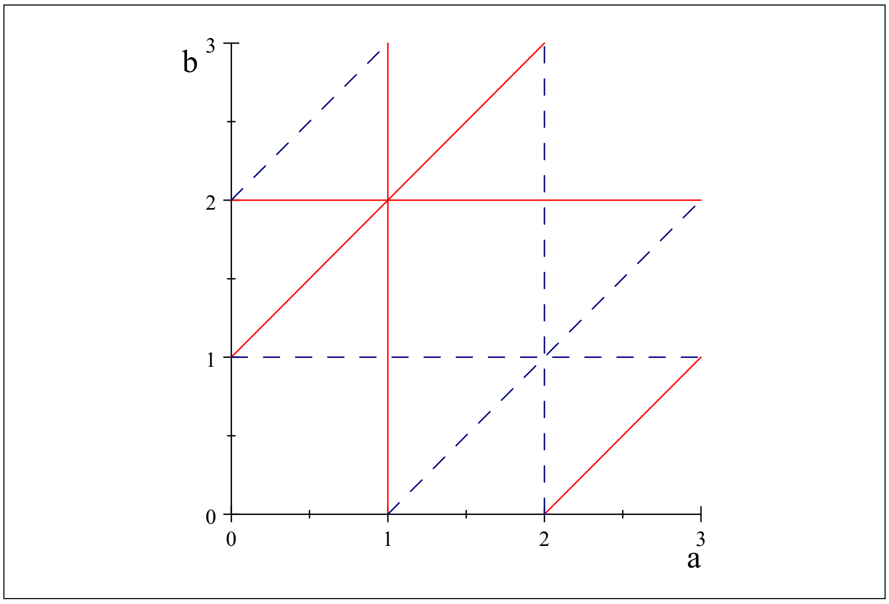

Figure 7.2: Intersection of tori in $\mathbf{P}(\mathbb{C}^3)$

To conclude the case $d = 3$, we remark that the parametrisations of b_1 and b_2 on $\mathbb{T}^2$ can be obtained as translations of f on the discrete, 3×3 net obviously appearing in Figure 7.2.

In higher dimensions, it can be seen that we have no more the same rigidity. So the method of investigation has to change or improve.

7.4 Pure states entanglement measure

In Section 3.2, we saw that two lines over a ring may intersect at some point other than the origin. If two lines intersect only at the origin, they are said distant. Otherwise they are said to be neighbour. We also remarked that the notions of distance and neighbourhood can be refined in considering the set of all linear combinations of two vectors generating the lines. The greater the cardinality of that set, the more the lines are distant or, we can say, different. What is more, the wedge product is a suitable tool in order to measure that difference.

In a vector space over $\mathbb{R}$ or $\mathbb{C}$, we may consider for any two vectors the area of the parallelogram they determine. We suppose that the angle between the vectors varies while their norms remain constant. Then the area of the parallelogram is zero when the vectors are colinear and maximal when they are orthogonal. From a quantum physical viewpoint and if the vectors are normalised, it means that this area accounts for how much the states can be distinguished from one another. If we recall that the area can be calculated as the modulus of the wedge or cross product of the vectors, this is quite similar to the discrete case.

In the present section, we explore this idea and connect it to two usual tools of quantum information, the von Neumann entropy and the Schmidt decomposition.

7.4.1 A determinantal measure

Let us consider a discrete, bipartite quantum system, the first and second subsystems having d_1 and d_2 levels, respectively:

$$|\psi\rangle = \sum_{i,j} c_{ij} |ij\rangle \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}, \tag{7.34}$$

with $|\psi\rangle$ being a normalised vector. In this section, we will restrict to

$$d_1 = d_2 = d \geq 2, \tag{7.35}$$

but it will be worth thinking of the case $d_1 \neq d_2$. Moreover, we suppose that the kets $|ij\rangle$ are nondegenerate levels with respect to well-chosen measurements on each of the subsystems. In particular, the $|ij\rangle$'s have to be orthonormal.

An intuitive way to know how entangled $|\psi\rangle$ is, is to know how much the state of one of the subsystems varies according to the result of a measurement on the other subsystem. To do so, we refer to the duality operator D that associates to any ket the corresponding bra in the dual Hilbert space:

$$D : \mathbb{C}^d \to (\mathbb{C}^d)^*, \quad |x\rangle \mapsto \langle x| \,. \tag{7.36}$$

Then the row vectors of the $d \times d$ matrix

$$\Psi = (I \otimes D) |\psi\rangle \tag{7.37}$$

stand for the various states of the second subsystem after a measurement has occured on the first one. Indeed, a row is a bra, but under the action of D one obtains one of the announced kets. Similarly, the column vectors of Ψ are the various states of the first subsystem after a measurement has occured on the second one. Since $|\psi\rangle$ is normalised, Ψ is such that

$$\|\Psi\|_2 = \sqrt{\operatorname{tr}(\Psi^\dagger \Psi)} = 1. \tag{7.38}$$

We say that Ψ is normalised for norm 2. Then our intuitive measurement of entanglement is nothing but the volume of the parallelepiped whose defining edges are the column vectors of Ψ, or equivalently its row vectors. Thus we define the entanglement of $|\psi\rangle$ by

$$E(|\psi\rangle) = d^d \left|\det \Psi\right|^2 . \tag{7.39}$$

Since

$$|\det \Psi^\dagger| = \left|\det \Psi\right| , \tag{7.40}$$

this definition is symmetric with respect to the two subsystems. The reasons for the normalisation factor d^d and for the square on the determinant will appear later on.

In the introduction of the section, we took two normalised vectors only. Here we take d vectors with a range of norms. The norm of the j-th column vector of Ψ is the probability for the j-th eigenvalue to come out when a measurement is performed on the second subsystem. Similarly, the norm of the i-th row vector of Ψ is the probability for the i-th eigenvalue to come out when a measurement is performed on the first subsystem. So our determinantal measure of entanglement takes also into account the statistics.

Example: We take a bipartite, 2-level system

$$|\psi\rangle = c_{00}\,|00\rangle + c_{01}\,|01\rangle + c_{10}\,|10\rangle + c_{11}\,|11\rangle \tag{7.41}$$

and apply $I \otimes D$ to it:

$$\begin{aligned} \Psi &= (I \otimes D)\,|\psi\rangle && (7.42a) \\ &= c_{00}\,|0\rangle\,\langle 0| + c_{01}\,|0\rangle\,\langle 1| + c_{10}\,|1\rangle\,\langle 0| + c_{11}\,|1\rangle\,\langle 1| && (7.42b) \\ &= \begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix}. && (7.42c) \end{aligned}$$

Then

$$E(|\psi\rangle) = 4\,|\det\Psi|^2 = 4\,|c_{00}c_{11} - c_{01}c_{10}|^2\,. \tag{7.43}$$

□

In what range of values does $E(|\psi\rangle)$ lies? It is of course a nonnegative quantity, so that we are to determine its upper-bound and show if it can be reached.

Theorem 33 *For any bipartite, d-level system* $|\psi\rangle$

$$0 \leq E(|\psi\rangle) \leq 1. \tag{7.44}$$

Moreover, the upper-bound is reached iff there exist two unitary operators $U_1, U_2 \in \mathrm{U}(d)$ *such that*

$$(U_1 \otimes U_2)\,|\psi\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |ii\rangle\,. \tag{7.45}$$

In that case, we say that $|\psi\rangle$ *is maximally entangled.*

This result explains the normalisation factor in the definition of E.

Proof. For a given $|\psi\rangle$, we can find, by the singular value decomposition of Ψ, two unitary matrices U_1, U_2 corresponding to two changes of basis in local measurements, such that

$$\Psi' = U_1 \Psi U_2^\dagger \tag{7.46}$$

is diagonal with nonnegative entries. Moreover

$$E(|\psi\rangle) = d^d\,\left|\det\Psi'\right|^2, \tag{7.47}$$

where the determinant is now the mere product of the diagonal entries λ_i of Ψ'. Then

$$\sum_{i=1}^{d} \lambda_i^2 = \left\|\Psi'\right\|_2^2 = \|\Psi\|_2^2 = 1, \tag{7.48}$$

which indicates that $\lambda = (\lambda_i)_i$ is a point on the unit sphere. Let us parametrise the λ_i's with the help of multidimensional spherical coordinates. We put

$$\forall i \in \{1, \ldots, d-1\}, \quad \theta_i \in [0; \pi/2] \tag{7.49}$$

and

$$\begin{aligned}
\lambda_1 &= \cos\theta_{d-1}\cos\theta_{d-2}\cdots\cos\theta_2\cos\theta_1 && (7.50a)\\
\lambda_2 &= \cos\theta_{d-1}\cos\theta_{d-2}\cdots\cos\theta_2\sin\theta_1 && (7.50b)\\
&\vdots \\
\lambda_{d-2} &= \cos\theta_{d-1}\cos\theta_{d-2}\sin\theta_{d-3} && (7.50c)\\
\lambda_{d-1} &= \cos\theta_{d-1}\sin\theta_{d-2} && (7.50d)\\
\lambda_d &= \sin\theta_{d-1}. && (7.50e)
\end{aligned}$$

With the additional notation

$$f = \left|\det \Psi'\right|, \tag{7.51}$$

we have to study

$$f(\theta_1, \ldots, \theta_{n-1}) = \prod_{i=1}^{d-1} \lambda_i = \prod_{i=1}^{d-1} \sin\theta_i \ (\cos\theta_i)^i. \tag{7.52}$$

Since f is defined on a compact set and is continuous, it reaches its upper-bound. But f is zero whenever for one i

$$\theta_i = 0 \text{ or } \pi/2 \tag{7.53}$$

and is positive othewise. So the upper-bound is positive and it is reached at some point in the open set

$$\prod_{i=1}^{d-1}]0; \pi/2[. \tag{7.54}$$

At such a point, all the partial derivatives of f annihilate. The partial derivative of f with respect to θ_i is

$$f_i' = \prod_{j \neq i} \sin\theta_j \ (\cos\theta_j)^j \cdot ((\cos\theta_i)^{i+1} - i(\sin\theta_i)^2(\cos\theta_i)^{i-1}). \tag{7.55}$$

and it annihilates for

$$\theta_i = \overline{\theta}_i = \arctan\frac{1}{\sqrt{i}}. \tag{7.56}$$

We get

$$\sin\overline{\theta}_i = \frac{\tan\overline{\theta}_i}{\sqrt{1+\tan^2\overline{\theta}_i}} = \frac{1}{\sqrt{1+i}}, \tag{7.57a}$$

$$\cos\overline{\theta}_i = \frac{1}{\sqrt{1+\tan^2\overline{\theta}_i}} = \frac{\sqrt{i}}{\sqrt{1+i}} \tag{7.57b}$$

and

$$f(\overline{\theta}_1,\ldots,\overline{\theta}_{n-1}) = \prod_{i=1}^{d-1} \frac{\sqrt{i}^i}{\sqrt{1+i}^{1+i}} = \frac{1}{\sqrt{d}^d}. \tag{7.58}$$

So, there exists a unique point on the sphere with positive coordinates such that $E(|\psi\rangle)$ is maximal. For all $i \in \{1,\ldots,d\}$,

$$\lambda_i(\overline{\theta}_1,\ldots,\overline{\theta}_{n-1}) = \frac{1}{\sqrt{i}} \prod_{j=i}^{d-1} \frac{\sqrt{j}}{\sqrt{1+j}} = \frac{1}{\sqrt{d}}. \tag{7.59}$$

The corresponding state $|\psi\rangle$ is as announced in the theorem. The calculation of the entanglement with this latter $|\psi\rangle$ is straightforward and gives

$$E(|\psi\rangle) = 1. \tag{7.60}$$

■

Remark 1 *In this proof, we used a diagonalisation of Ψ. But in order to calculate the entanglement, it is sufficient to trigonalise Ψ.*

7.4.2 Comparison with von Neumann entropy

The determinantal measurement of entanglement we have set out is basically concerned with the physical structure of a state, not with the informational using of the state from the point of view of quantum information theory. So we now compare it with the notion of entropy. However, we shall limit ourselves to the case of a bipartite, 2-level system as in the example above. It will be obvious that the method we use is not quite adapted to higher level systems.

The classical Shannon entropy $H(X)$ measures the uncertainty in the outcome of a random variable X. It depends only on the probability distribution of X. If $(p_i)_{i\in I}$ is the probability distribution, one has

$$H(X) = -\sum_{i\in I} p_i \log p_i, \tag{7.61}$$

where the logarithm is to be understood throughout this section in basis 2. In quantum physics, the variable X is replaced by a density matrix ρ describing a statistical set of states, also called a state for short. This gives rise to von Neumann's entropy defined as

$$S(\rho) = -\operatorname{tr}(\rho\log\rho). \tag{7.62}$$

If $(p_i)_{i \in I}$ are the eigenvalues of ρ, one recovers

$$S(\rho) = -\sum_{i \in I} p_i \log p_i. \tag{7.63}$$

In the same spirit as to the definition of $E(|\psi\rangle)$, we are interested in the variations of one of the subsystems when measurements are performed on the other one. So we will calculate the von Neumann entropy of reduced density matrices and compare to our measure of entanglement as defined in (7.39). Since for a bipartite system, the eigenvalues of the two reduced density matrices are the same, we may choose either of these matrices in order to calculate S. In other words, S is symmetric with respect to the two subsystems.

As a general rule, since the basis for each of the subsystems is assumed orthonormal, then the two reduced matrices of a bipartite, d-level system $|\psi\rangle$ are the Gram matrices of the row and column vectors of the matrix Ψ, respectively. If the L_i's are the row vectors of Ψ, the density matrix for the first subsystem is

$$\begin{aligned}\rho_1 = \mathrm{tr}_2(|\psi\rangle\langle\psi|) = \mathrm{tr}_2\left(\sum_{i,j,k,l} c_{ik}c^*_{jl}\,|ik\rangle\langle jl|\right)\\ = \sum_{i,j,k} c_{ik}c^*_{jk}\,|i\rangle\langle j| = \sum_{i,j} |i\rangle\langle L_j|L_i\rangle\langle j|\,,\end{aligned} \tag{7.64}$$

where tr_2 denotes the partial trace over the second subsystem. Similarly, if the C_i's are the column vectors of Ψ and tr_1 denotes the partial trace over the first subsystem, the density matrix of the second subsystem is

$$\rho_2 = \mathrm{tr}_1(|\psi\rangle\langle\psi|) = \sum_{k,l} |k\rangle\langle C_l|C_k\rangle\langle l|\,. \tag{7.65}$$

Example: Let us take for $|\psi\rangle$ a separated state

$$|\psi\rangle = |00\rangle\,. \tag{7.66}$$

The density matrix of the first subsystem is

$$\rho_1 = |0\rangle\langle 0| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{7.67}$$

so that the entanglement of $|\psi\rangle$ and the entropy of ρ_1 are both null:

$$E(|\psi\rangle) = S(\rho_1) = 0. \tag{7.68}$$

Now if we take for $|\psi\rangle$ a Bell state

$$|\beta_{00}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \tag{7.69}$$

the first reduced density matrix is

$$\rho_1 = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}, \tag{7.70}$$

so that the entropy of ρ_1 is

$$S(\rho_1) = -2\,\frac{1}{2}\log\frac{1}{2} = 1, \tag{7.71}$$

which matches the entanglement of $|\psi\rangle$:

$$E(|\psi\rangle) = 2^2 \left|\det\begin{pmatrix} 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} \end{pmatrix}\right|^2 = 1. \tag{7.72}$$

□

In the case of the general pure state set out in (7.41), the density matrix is

$$\rho = |\psi\rangle\langle\psi| \tag{7.73}$$

and we consider the density matrix ρ_1 of the first subsystem after measurements on the second one:

$$\rho_1 = \begin{pmatrix} |c_{00}|^2 + |c_{01}|^2 & c_{00}c_{10}^* + c_{01}c_{11}^* \\ c_{10}c_{00}^* + c_{11}c_{01}^* & |c_{10}|^2 + |c_{11}|^2 \end{pmatrix}. \tag{7.74}$$

Then ρ_1 has characteristic polynomial, with variable Λ,

$$\chi(\rho_1) = \Lambda^2 - \Lambda + \frac{1}{4}E(|\psi\rangle) \tag{7.75}$$

and eigenvalues

$$\Lambda_\pm = \frac{1}{2}(1 \pm \sqrt{1 - E(|\psi\rangle)}). \tag{7.76}$$

The entropy is

$$S(\rho_1) = -\Lambda_+ \log\Lambda_+ - \Lambda_- \log\Lambda_-. \tag{7.77}$$

It is interesting to compare $S(\rho_1)$ with $E(|\psi\rangle)$ in two ways. In Figure 7.3, we plot $S(\rho_1)$ and $E(|\psi\rangle)$ as functions of $|\det\Psi|$, whereas in Figure 7.5, we plot them as functions of $E(|\psi\rangle)$. In Figures 7.4 and 7.6, we plot the difference $S(\rho_1) - E(|\psi\rangle)$ as function of the same variables, $|\det\Psi|$ and $E(|\psi\rangle)$, respectively.

We notice that

$$E(|\psi\rangle) = 4\Lambda_+\Lambda_-. \tag{7.78}$$

This is not a coincidence as we are going to see in the next section.

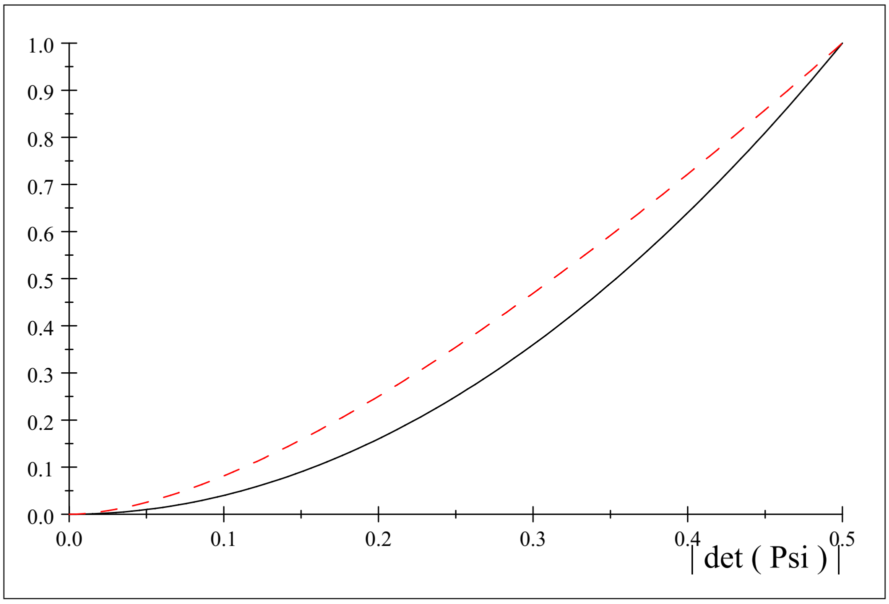

Figure 7.3: $S(\rho_1)$ (dashed, red) and $E(|\psi\rangle)$ (solid, black) as functions of $|\det \Psi|$

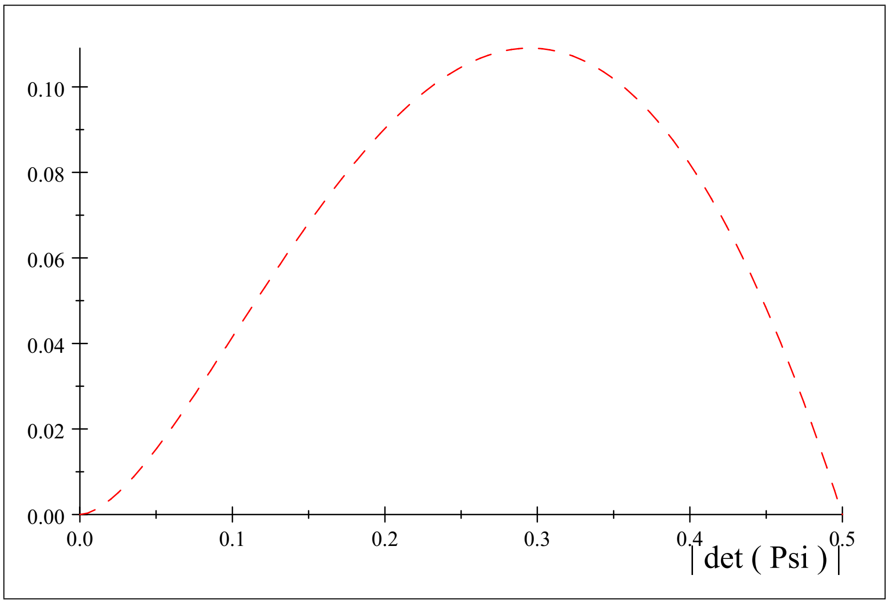

Figure 7.4: $(S(\rho_1) - E(|\psi\rangle))$ as function of $|\det \Psi|$

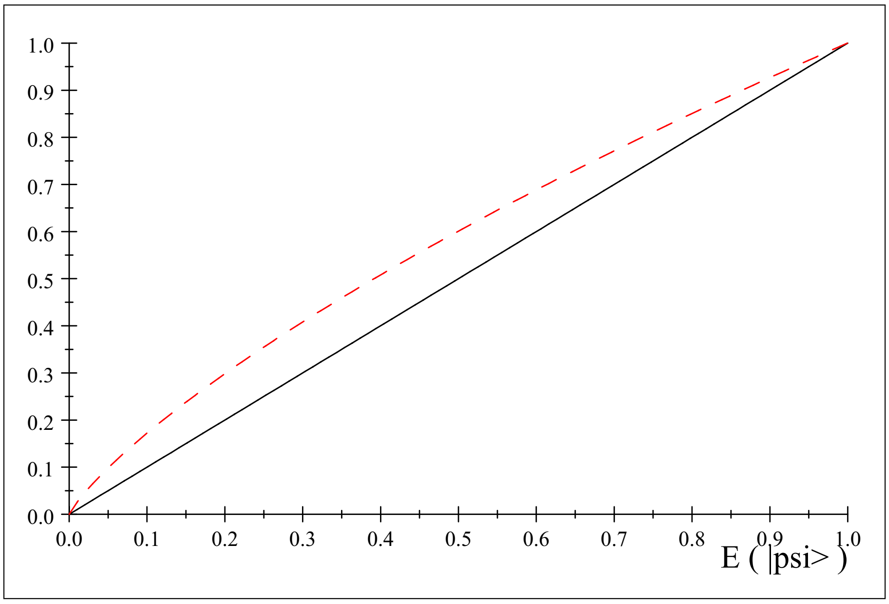

Figure 7.5: $S(\rho_1)$ (dashed, red) and $E(|\psi\rangle)$ (solid, black) as functions of $E(|\psi\rangle)$

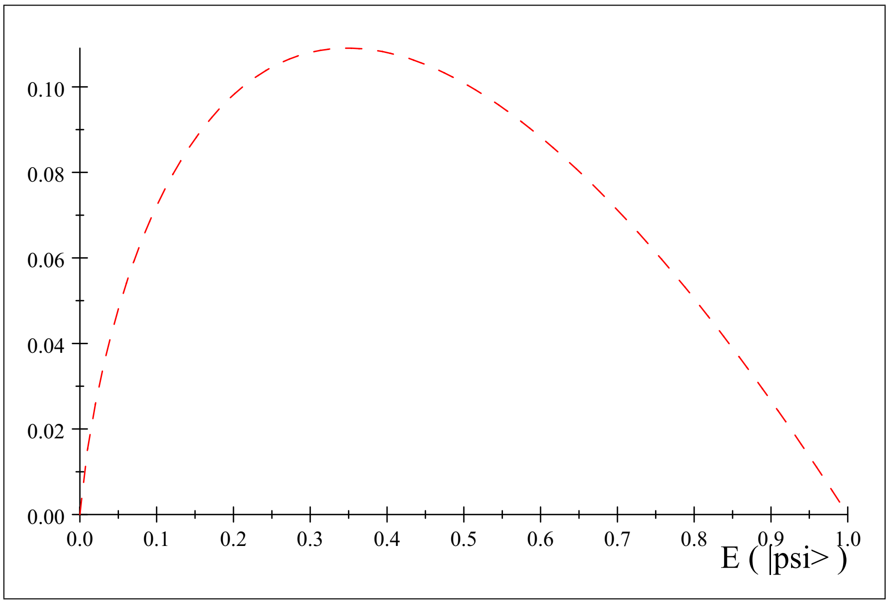

Figure 7.6: $(S(\rho_1) - E(|\psi\rangle))$ as function of $E(|\psi\rangle)$

7.4.3 Comparison with Schmidt decomposition

Let $|\psi\rangle$ be a bipartite, d-level system. According to Schmidt decomposition, there exists an orthnormal basis $(|1i\rangle)_i$ for the first subsystem, an orthonormal basis $(|2i\rangle)_i$ for the second subsystem and a set of nonnegative real numbers $(\lambda_i)_i$ such that

$$|\psi\rangle = \sum_{i=1}^{d} \lambda_i \, |1i\rangle \otimes |2i\rangle \,. \tag{7.79}$$

That is to say $|\psi\rangle$ can be written in a suitable basis as a sum of tensorial products. The λ_i's are called the Schmidt coefficients of this decomposition. This is a corollary of the singular value decomposition we have already used in the proof of Theorem 33. Then we have

$$\Psi = \sum_{i=1}^{d} \lambda_i \, |1i\rangle \, \langle 2i| \,, \qquad \rho_1 = \sum_{i=1}^{d} \lambda_i^2 \, |1i\rangle \, \langle 1i| \,, \tag{7.80}$$

and if we denote $\Lambda_i = \lambda_i^2$ the eigenvalues of ρ_1, we obtain

$$E(|\psi\rangle) = d^d \prod_{i=1}^{d} \Lambda_i, \qquad S(\rho_1) = -\sum_{i=1}^{d} \Lambda_i \log \Lambda_i. \tag{7.81}$$

If one of the λ_i's is null, then $E(|\psi\rangle) = 0$, though $|\psi\rangle$ may contain some entanglement as a lower level bipartite sytem. Thus the number of nonzero λ_i's is a first, discrete evaluation of the entanglement, which is nothing but the rank of Ψ, ranging from 1 to d. Whenever it is d, but unfortunately only in that case, the determinantal measure precises the amount of entanglement. In turn, von Neumann entropy measures the entanglement continuously but does not supply obvious information about the rank. So, in order to detect and evaluate localised entanglement in $|\psi\rangle$, we have also to consider the minors of Ψ.

There are many ways to do so and indeed this is an intricate topic of algebraic geometry. We here consider only an example as to 2×2 minors. If Δ_2 is the set of all 2×2 minors, let us define

$$E_2(|\psi\rangle) = \frac{1}{2} \sum_{\delta \in \Delta_2} |\delta|^2 \,. \tag{7.82}$$

A factor $1/4$ corrects for multiple counting of identical terms and an extra normalisation factor 2 that will be explained below is included. $E_2(|\psi\rangle)$ is invariant under local unitary transformations (see [52]), that is

$$\forall U_1, U_2 \in \mathrm{U}(d), \quad E_2((U_1 \otimes U_2) \, |\psi\rangle) = E_2(|\psi\rangle) \tag{7.83}$$

Moreover, it is zero iff the rank of Ψ is 1. Therefore, $E_2(|\psi\rangle)$ is nonzero iff $|\psi\rangle$ is entangled in the common meaning of the word, that is to say $|\psi\rangle$ cannot be written as a tensorial product. In a way, this is the most basic measurement in the determinantal family. It can be applied to any bipartite system, even if the subsystems do not have the same number of levels, i.e. $d_1 \neq d_2$. In the formal case where $d_1 = d_2 = d$, if we take $|\psi\rangle$ in its Schmidt decomposition form, $E_2(|\psi\rangle)$ reads

$$E_2(|\psi\rangle) = \left(\sum_{i=1}^{d} \Lambda_i\right)^2 - \sum_{i=1}^{d} \Lambda_i^2 = 1 - \mathrm{tr}(\rho_1^2). \tag{7.84a}$$

which is nothing but the linear entropy of $|\psi\rangle$.

If we set aside Schmidt decomposition, this can be proved in a way that relates to scalar and wedge products in a simple manner. One may prove by induction on d that, for any two $x, y \in \mathbb{C}^d$,

$$|\langle x|y\rangle|^2 + \|x \wedge y\|^2 = \|x\|^2 \|y\|^2 , \tag{7.85}$$

which is nothing but an analogue in higher dimension of

$$\cos^2\theta + \sin^2\theta = 1. \tag{7.86}$$

So, denoting by L_i the row vectors of Ψ and taking into account (7.64), we have

$$\begin{aligned} E_2(|\psi\rangle) &= \sum_{i,j=1}^{d} \|L_i \wedge L_j\|^2 = \sum_{i,j=1}^{d} \|L_i\|^2 \|L_j\|^2 - |\langle L_i|L_j\rangle|^2 && (7.87a)\\ &= \left(\sum_{i=1}^{d} \|L_i\|^2\right)^2 - \sum_{i,j=1}^{d} |\langle L_i|L_j\rangle|^2 && (7.87b)\\ &= \|\Psi\|_2^2 - \|\rho_1\|_2^2 && (7.87c)\\ &= 1 - \mathrm{tr}(\rho_1^2) && (7.87d) \end{aligned}$$

A similar sequence of equalities is obtained if we consider the column vectors of Ψ.

The quantity $\mathrm{tr}(\rho_1^2)$ is well-known to range in $]0, 1]$. Moreover, it is equal to 1 iff the reduced state whose density matrix is ρ_1 is pure. We thus see that the measure of localised entanglement at scale 2 (as we consider 2×2 minors) is related to the amount of mixture in each of the subsystems. The localised entanglement measure $E_2(|\psi\rangle)$ is 0 iff the subsytems of $|\psi\rangle$ appear as pure states whenever they are looked at independently of one another, and we saw that this is equivalent to $|\psi\rangle$ being a separable state.

Remark 2 *For a bipartite, two-level system, one has*

$$E(|\psi\rangle) = 2E_2(|\psi\rangle). \tag{7.88}$$

To conclude this section, a final remark is in order. There are many other ways of qualifying or quantifying entanglement, but the search for a convenient, universal tool (if any) is still an open issue. In connection with determinants, one may see [62, 63] for two other viewpoints. The relation between symplectic transformations and entanglement is approached numerically in [64].In the previous section, we evoked a Hopf fiber in the search for MUBs. Hopf fibers also intervene in the study of entanglement [65]. But to date, all these tools are quite limited. For a review of more classical tools, as entanglement witnesses, positive maps and partial transposition, one may see [66].

Chapter 8

Conclusion

Quantum information theory requires a tremendous range of mathematics. But if one looks at the educational treatises about basic quantum physics or quantum information and communication, one can see nowhere a word on MUBs. As we wrote in the introduction, only specific applications involve them. Nevertheless, as a theoretical feature, they appear to be tied up to very common objects, such as irreducible representations of the Lie algebra $\mathfrak{su}(2)$ or the quite as much wide-spread Pauli group, a discrete version of the Weyl-Heisenberg group. Any endeavour in order to understand MUBs better is not only worth being carried on for the sake of MUBs themselves, but also to enlight those objects in a new fashion. We note that the crux in those tools, whether they are considered for themselves or in connection with physical applications, is the commutation relations of operators.

So, our thesis was concerned mainly, in Chapters 3, 4 and 5, with the connection between maximally commuting sets of Pauli operators and the unbiasedness relation among their diagonalising bases. The relevant mathematical framework was known to be projective geometry over the ring $\mathbb{Z}_d$ of integers modulo d. In Chapter 3, we recalled how Pauli operators are encoded up to an irrelevant global phase as the vectors of $\mathbb{Z}_d^{2n}$. Besides, their commutation relations are transcribed as the symplectic product of the vectors. Maximally commuting sets are thus unions of projective points. In that framework, we first generalised the notions of distance and neighbourhood in arbitrary dimension and introduced the wedge product in order to quantify neighbourness between two vectors or points. We then provided counting properties in $\mathbf{P}(\mathbb{Z}_d^m)$ with respect to the neighbourhood relation. With these properties applied to the projective line $\mathbf{P}(\mathbb{Z}_d^2)$, we proved that, in order to build a complete set of MUBs in dimension $d = p^s$ a power of prime, with $s > 1$, by means of Pauli operators, tensorial products of them are mandatory.

In Chapters 4 and 5, we fully classified Lagrangian submodules of $\mathbb{Z}_d^{2n}$ and selected among them Lagrangian half-modules, as we showed that the latters are the only ones for which the corresponding Pauli operators can yield MUBs. We established an isomorphism between distant Lagrangian half-modules and Pauli unbiased bases. The isomorphism was presented in three ways: 1) with regards to symplectic algebra in $\mathbb{Z}_d^{2n}$, 2) with regards to symplectic geometry in $(\mathbf{P}(\mathrm{Mat}(n, \mathbb{Z}_d)^2), \omega)$, 3) embeded in a graph interpretation with various counting results.

This mainstream study arose with five byproducts.

1) It completed the investigation of an alternative derivation of MUBs by means of Gauss sums and in relation with a family of irreducible representations of the Lie algebra $\mathfrak{su}(2)$. In Chapter 2, we exposed the mechanism of this derivation and gave a sufficient condition to get MUBs. Since the operators of interest satisfy commutation relations analogous to those of Pauli operators, our main study enabled us to turn this condition into also a necessary one.

2) As to pure mathematics, we gave in Section 4.3 an algorithm in order to diagonalise a $2n \times r$ matrix by means of a symplectic change of computational basis. We also gave an example where such a reduction is impossible and we compared with the case of symplectic submodules.

3) As to quantum information theory, we related in Section 5.4 to the Clifford group in any dimension and exposed how it analyses into three parts: a phase, a Pauli part and a symplectic part. We called Pauli states those states that are obtained by diagonalising Pauli operators, and we showed that the product group $\mathbb{Z}_d^{2n} \times \mathrm{Sp}(n, \mathbb{Z}_d)$ accounts for any dynamics among these states, thus providing an alternative to the classical Clifford group expression of such a dynamics.

4) In Chapter 6, physical systems were addressed, as the classification of Lagrangian submodules was readily applicable to the finite phase space over $\mathbb{Z}_d$, namely $\mathbb{Z}_d^2$. We counted the isotropic lines in that space, not only as a whole but also under the condition that they go through a given point. The way they arrange under the action of $\mathrm{Sp}(1, \mathbb{Z}_d) = \mathrm{SL}(2, \mathbb{Z}_d)$ was described in full. We thus answered a technical point in the current problem of setting-up of discrete Wigner distributions over $\mathbb{Z}_d^2$.

5) Finally, as we showed in Chapter 7, the ideas we developed in the framework of discrete algebra and geometry were partially suitable for generalisation in continuous mathematics. Pauli operators with real exponents showed up to be specific to qubits. But other tools adapt to higher dimensions. On the one hand, in projective geometry, harmonic conjugation accounts for unbiasedness among bases of qubits, with an obvious geometrical translation on the Bloch sphere as orthogonal great circles. A picture that we generalised in any dimension with the notion of orthogonal tori. The cases $d = 2$ and $d = 3$ were worked out and thus we exhibited a rigidity condition in those cases: once a basis is given, one vector unbiased with respect to it is enough to determine the entire remaining bases in order to form a complete set of MUBs. On the other hand, the symplectic and wedge products were completed by the use of determinants. Measuring volumes is the counterpart of the idea of distance in discrete geometry. But quite surprisingly, whereas the notion of discrete distance related to unbiasedness, the use of determinants related to entanglement. We thus compared our determinantal measure of entanglement with the von Neumann entropy and the Schmidt decomposition. The resemblance is striking but in need for further investigations. We also saw that mixture of statistical states can this way be approached. It will be uttermost interesting to fulfil this study, for bipartite systems as well as in general for multipartite systems.

The question of the celebrated upper-bound to the number of MUBs, $d+1$ in dimension d, remains an open one. The most recent numerical tests by Brierley and Weigert [36] tend to show that there are no more than 3 MUBs in dimension 6, that is the limiting factor 2 plus 1, a fact that have theoretical counterparts. In our thesis, we refered to the perhaps best-known group theoretical tool in order to build MUBs, namely the Pauli group. Since it is liable to the Chinese remainder theorem, the building of MUBs with it is bound to prime power dimensions. This is in agreement with the expression of MUBs issued from Pauli operators [18] and the theoretical conclusion of Archer [17]. So, the problem was reduced to prime power dimensions, that is to say in our case from $\mathbb{Z}_d$, with a general d, to $\mathbb{Z}_{p^s}$, with p a prime. Moreover, all the conditions we found relate to invertibilty, so that only the terms of lowest valuation are of interest. The problem thus reduced further to $\mathbb{Z}_p$ as the base ring. A recent paper by Kibler [67] investigates the connection between the Heisenberg group and MUBs, with the same stubborn restriction on the base ring. All this should be compared fruitfully to a work by Howe [68] connecting Lagrangian submodules to the Heisenberg group, MUBs and nice error bases. He starts from a general group theoretical point of view and finally comes to the conclusion that "In some sense, the Heisenberg group over the field $\mathbf{Z}/p$, of integers mod p, is the best group for constructing mutually unbiased bases".

Appendix A

Arithmetics in $\mathbb{Z}$ and $\mathbb{Z}_d$

A.1 gcd, lcm and order

In $\mathbb{Z}$, the notion of greatest common divisor (gcd for short) has an intuitive meaning. But it is equivalent to a little bit more abstract property which will generalise to residue class rings $\mathbb{Z}_d = \mathbb{Z}/d\mathbb{Z}$, $d \geq 2$. This equivalence is called Bézout's theorem. To see how it works, note that the sets of the form $k\mathbb{Z}$, $k \in \mathbb{Z}$, are the sole subrings of $\mathbb{Z}$. Bézout's theorem states that if δ is the gcd of $a_1, \ldots, a_n \in \mathbb{Z}$:

$$\delta = \bigwedge_{i=1}^{n} a_i, \tag{A.1}$$

then δ is characterised up to its sign by the set equation

$$\delta\mathbb{Z} = \sum_{i=1}^{n} a_i\mathbb{Z}, \tag{A.2}$$

that is to say $\delta\mathbb{Z}$ is the set of all linear combinations of the a_i's over $\mathbb{Z}$. We immediately deduce from that theorem Gauss's theorem for integers: If a divides the product bc and is coprime with b then a divides c. It is also quite obvious from Bézout's theorem that the following three properties are equivalent:

1. a is coprime with d;

2. The residue class $\overline{a}$ in the quotient ring $\mathbb{Z}_d$ is invertible. In that case, we also say that a is invertible modulo d;

3. $\overline{a}$ is a generator of $\mathbb{Z}_d$:

$$\overline{a}\mathbb{Z}_d = \{\overline{a}x;\ x \in \mathbb{Z}_d\} = \mathbb{Z}_d. \tag{A.3}$$

The invertible elements of $\mathbb{Z}_d$ are also called its units and hence their set is denoted $U(\mathbb{Z}_d)$, or $\mathbb{Z}_d^*$.

In the case of $\mathbb{Z}_d$, Equation (A.2) is retained in order to define a notion of gcd. A residue $\overline{\delta} \in \mathbb{Z}_d$ is a gcd for a set of $\overline{a_i}$'s in $\mathbb{Z}_d$ if

$$\overline{\delta}\mathbb{Z}_d = \sum_{i=1}^{n} \overline{a_i}\mathbb{Z}_d. \tag{A.4}$$

So, if δ is the gcd of the a_i's, $\overline{\delta}$ is a gcd for the $\overline{a_i}$'s. As for $\mathbb{Z}$, this gcd is determined only up to an invertible mutiplier as will be shown in Appendix A.2. The computation of a gcd is still associative and commutative. As is the case for $\mathbb{Z}$, the $\overline{a_i}$'s will be said coprime if $\overline{\delta}$ is invertible. In this case, $\overline{\delta}\mathbb{Z}_d = \mathbb{Z}_d$. The interpretation in terms of linear combinations is still valid. The intuive one in terms of prime factor decomposition or division order is also still valid if one takes into account the slight modification indicated by the following property:

$$\overline{\delta} = \bigwedge_{i=1}^{n} \overline{a_i} \text{ in } \mathbb{Z}_d \quad \text{iff} \quad \delta \wedge d = \left(\bigwedge_{i=1}^{n} a_i\right) \wedge d \text{ in } \mathbb{Z}. \tag{A.5}$$

Indeed, if we come back to representatives of residue classes, definition (A.4) reads

$$\delta\mathbb{Z} + d\mathbb{Z} = \left(\sum_{i=1}^{n} a_i\mathbb{Z}\right) + d\mathbb{Z}, \tag{A.6}$$

which is nothing but the second member of equivalence (A.5). So, there is an additional d in each member of that latter expression. It means that the power k of a prime factor in δ or in any one of the a_i's must first be replaced by the minimum of k and the power of the same prime factor in d. Light is shed on that recipe in Appendix A.2 with the Chinese remainder theorem and p-adic decomposition.

If $\overline{\delta}$ is a gcd for the $\overline{a_i}$'s, we shall call $\overline{\delta \wedge d}$ *the* gcd of the $\overline{a_i}$'s. In fact, it is a gcd and if $\overline{\delta_1}$ and $\overline{\delta_2}$ are two gcd's then according to (A.5)

$$\overline{\delta_1 \wedge d} = \overline{\delta_2 \wedge d}. \tag{A.7}$$

That gcd is also the first one according to the lexicographic order from $\overline{0}$ to $\overline{d-1}$ since for any positive δ such that $\overline{\delta}$ is a gcd, $\delta \wedge d \leq \delta$.

In the same manner, one defines a lowest common multiple (lcm for short) of $a_1, \ldots, a_n \in \mathbb{Z}$ (resp. $\overline{a_1}, \ldots, \overline{a_n} \in \mathbb{Z}_d$) to be an element μ_1 (resp. $\overline{\mu_2}$) such that

$$\mu_1\mathbb{Z} = \bigcap_{i=1}^{n} a_i\mathbb{Z} \qquad \left(\text{resp. } \overline{\mu_2}\mathbb{Z}_d = \bigcap_{i=1}^{n} \overline{a_i}\mathbb{Z}_d\right). \tag{A.8}$$

The lcm operation is associative and commutative in both case and is denoted by the vee symbol $\vee$:

$$\mu_1 = \bigvee_{i=1}^{n} a_i \qquad \left(\text{resp. } \overline{\mu_2} = \bigvee_{i=1}^{n} \overline{a_i}\right). \tag{A.9}$$

Those two notions of lcm are related by

$$\overline{\mu} = \bigvee_{i=1}^{n} \overline{a_i} \text{ in } \mathbb{Z}_d \quad \text{iff} \quad \mu \wedge d = \left(\bigvee_{i=1}^{n} a_i \right) \wedge d \text{ in } \mathbb{Z}. \tag{A.10}$$

Indeed, since the map $x \mapsto \overline{x}$ is onto, the first equality means

$$\mu\mathbb{Z} + d\mathbb{Z} = \bigcap_{i=1}^{n} (a_i\mathbb{Z} + d\mathbb{Z}) \tag{A.11}$$

and the second one means

$$\mu\mathbb{Z} + d\mathbb{Z} = \left(\bigcap_{i=1}^{n} a_i\mathbb{Z} \right) + d\mathbb{Z}. \tag{A.12}$$

We are thus to prove that

$$\bigcap_{i=1}^{n} (a_i\mathbb{Z} + d\mathbb{Z}) = \left(\bigcap_{i=1}^{n} a_i\mathbb{Z} \right) + d\mathbb{Z}. \tag{A.13}$$

Since all operations involved here are associative and the intersection of two subrings is still a subring, we can prove this equality by induction. So let us suppose that $n = 2$ and let x be in the first set:

$$x = k_1 a_1 + l_1 d = k_2 a_2 + l_2 d. \tag{A.14}$$

Divide each member by $a_1 \wedge a_2 \wedge d$:

$$x' = k_1 a_1' + l_1 d' = k_2 a_2' + l_2 d'. \tag{A.15}$$

Then $a_1' \wedge a_2'$ divides $k_1 a_1' - k_2 a_2' = (l_2 - l_1) d'$ and is coprime with d'. So there exist $n_1, n_2 \in \mathbb{Z}$ such that $n_1 a_1' - n_2 a_2' = l_2 - l_1$. Let us call $y = n_1 a_1' + l_1 = n_2 a_2' + l_2$. We have

$$x' - y d' = (k_1 - n_1 d') a_1' = (k_2 - n_2 d') a_2' \tag{A.16}$$

and eventually

$$x - y d \in \bigcap_{i=1}^{n} a_i\mathbb{Z}. \tag{A.17}$$

The converse inclusion for (A.13) is trivial.

Note that (A.13) was quite obvious with the prime factor decomposition interpretation of gcd and lcm since each of those two operations in $\mathbb{Z}$ is distributive with respect to the other.

Finally, the order $\nu(a)$ of $a \in \mathbb{Z}_d$ is the cardinality of the subring $a\mathbb{Z}_d = \{ka; k \in \mathbb{Z}_d\}$. This is also the first positive natural number n such that na is a multiple of d. The only residue whose order is 1 is 0, a is invertible modulo d iff $\nu(a) = d$, and

$\nu(a)\mathbb{Z}$ is the kernel of the linear map

$$\begin{array}{ccc} \mathbb{Z} & \longrightarrow & \mathbb{Z}_d \\ k & \longmapsto & ka. \end{array} \tag{A.18}$$

It is well known from group theory that the cardinality of a subgroup H of a finite group G is a divisor of the cardinality of G. For any $a \in \mathbb{Z}$, since $\overline{a}\mathbb{Z}_d$ is a subgroup of $\mathbb{Z}_d$, $x = d/\nu(\overline{a})$ is a well-defined integer such that the order of $\overline{x}$ is $\nu(\overline{a})$. Let us carry out the Euclidean division of a by x: $a = qx + r$ with $0 \leq r < x$ and suppose that $r \neq 0$. From the definition of r and according to that latter assumption, $\nu(\overline{r}) > \nu(\overline{x}) = \nu(\overline{a})$. But $\nu(\overline{a})\overline{r} = \nu(\overline{a})\overline{a} - \overline{q}(\nu(\overline{a})\overline{x}) = 0$ so that $\nu(\overline{r}) \leq \nu(\overline{a})$ and hence there is a contradiction. Thus $\overline{a} \in \overline{x}\mathbb{Z}_d$ and $\overline{a}\mathbb{Z}_d \subset \overline{x}\mathbb{Z}_d$. Since those two sets have the same cardinality they are equal and we have just seen that no residue class $\overline{r}$ with $0 \leq r < x$ can generate this set, except for the case when $\overline{a} = \overline{x} = \overline{r} = 0$. We deduce that $\overline{x}$ is the gcd of the one-element family $(\overline{a})$. We shall say that it is the gcd of the element $\overline{a}$.

So, we can compute the order of $\overline{a}$ as

$$\nu(\overline{a}) = \frac{d}{a \wedge d}. \tag{A.19}$$

It means that if

$$d = \prod_{i=1}^{n} p_i^{s_i} \quad \text{and} \quad a = \prod_{i=1}^{n} p_i^{s'_i} \prod_{j=1}^{m} p_j^{s''_j} \tag{A.20}$$

are the prime factor decompositions of d and a, then

$$\nu(\overline{a}) = \prod_{i=1}^{n} p_i^{s_i - \min(s_i, s'_i)}. \tag{A.21}$$

Hence one can find again the equivalence we first deduced from Bézout's theorem.

A.2 The Chinese remainder theorem

In the previous section of this appendix, we saw that $\overline{a}\mathbb{Z}_d = \overline{x}\mathbb{Z}_d$ with $x = a \wedge d$. One may wonder from $\nu(\overline{a}) = \nu(\overline{x})$ and from (A.20) and (A.21) if there is no invertible factor $\lambda \in \mathbb{Z}_d$ such that $\overline{a} = \lambda\overline{x}$. Moreover, it will prove the claim after (A.4) that the gcd is determined up to an invertible factor. Since if δ_1 and δ_2 are two possible gcd's, then there shall exist two invertible λ_1 and λ_2 such that

$$\delta_k = \lambda_k \overline{\left(\frac{d}{\nu(\delta_k)}\right)} \quad \text{for } k = 1, 2, \tag{A.22}$$

and so $\delta_2 = \lambda_2 \lambda_1^{-1} \delta_1$. It will also proves that for any gcd δ of the $\overline{a_i}$'s, $\overline{d/\nu(\delta)}$ is the gcd of the $\overline{a_i}$'s.

If for any i, $s'_i \leq s_i$, the existence of λ is obvious: $\lambda = \overline{q}$ answers the question. But it is not any more so obvious when there is one i for which $s'_i > s_i$. A fundamental idea to refer to and that we use many other times in this thesis is to prove a property for d a power of prime ($d = p^s$) and then deduce that it is true for any composite d as in (A.20). This idea is carried out by the so-called Chinese remainder theorem.

Theorem 34 (Chinese remainder) *If $d = \prod_{i=1}^n p_i^{s_i}$ is the prime factor decomposition of d, then we have the following isomorphism of rings:*

$$\begin{array}{rrcl} \pi: & \mathbb{Z}_d & \xrightarrow{\sim} & \prod_{i=1}^n \mathbb{Z}_{p_i^{s_i}} \\ & \overline{a} & \longmapsto & (a_1, \ldots, a_n) \end{array} \tag{A.23}$$

where $a_i = \pi_{p_i}(\overline{a})$ is the residue class of a modulo $p_i^{s_i}$. Addition and multiplication on the right-hand side of (A.23) are componentwise:

$$(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n), \tag{A.24a}$$
$$(a_1, \ldots, a_n)(b_1, \ldots, b_n) = (a_1 b_1, \ldots, a_n b_n). \tag{A.24b}$$

The $\mathbb{Z}_{p_i^{s_i}}$ in the theorem are called the Chinese factors of $\mathbb{Z}_d$. According to (A.24b), a is invertible iff all its Chinese components a_i are. Thus, to solve our problem, we can equivalently search for a λ_i in each Chinese factor such that $a_i = \lambda_i x_i$. Moreover, we are going to give a first cumbersome proof of the existence of λ_i to show the necessity for the p-adic decomposition in each Chinese factor. Let us suppose that $d = p^s$, let $\nu = \nu(\overline{x}) = \nu(\overline{a})$ and suppose that both q and $q + \nu$ are noninvertible modulo d, that is to say p divides q and $q + \nu$. We are to prove this is impossible and thus there exists an invertible λ modulo d such that $\overline{a} = \overline{\lambda}\overline{x}$. Indeed, since $a = qx$ and $p|q$ (p divides q) the properties $p^n|a$ and $p^{n-1}|x$ are true for $n = 1$. Suppose they are true for some positive integer n. We know that $\overline{x}$ is a multiple of $\overline{a}$ in $\mathbb{Z}_d$ and thus there exist $k, l \in \mathbb{Z}$ such that $x = ka + ld = ka + l\nu x$. Since $p|(q + \nu) - q = \nu$ and $p^{n-1}|x$, $p^n|\nu x$ and then $p^n|x$ due to the induction hypothethis and to the previous expression for x. And since $p|q$ and $a = qx$, we deduce that $p^{n+1}|a$. Hence the property $p^n|a$ should be true for all positive integer n, which is clearly nonsense when $a \neq 0$. If $a = 0$, we can just replace it by d. We are now going to introduce the p-adic decomposition in $\mathbb{Z}_{p^s}$ and compare with a proof using it.

Let a be a nonnegative integer and p be prime number. Writing a in numeration basis p, we get the numbers $r \in \mathbb{N}$ and $\alpha_0, \ldots, \alpha_r \in \{0, \ldots, p\}$ such that

$$a = \alpha_0 + \alpha_1 p + \cdots + \alpha_r p^r. \tag{A.25}$$

This is the p-adic decomposition of a. The p-valuation of a is

$$v_p(a) = \begin{cases} \min(i \in \{0, \ldots, r\}; \alpha_i \neq 0) & \text{for } a \neq 0, \\ +\infty \quad \text{for } a = 0. \end{cases} \tag{A.26}$$

For instance, if $a = \prod_{i=1}^n p_i^{s_i} \neq 0$ is the prime factor decomposition of a then for any $i \in \{1, \ldots, n\}$, $v_{p_i}(a) = s_i$.

Every class $\overline{a} \in \mathbb{Z}_{p^s}$ is uniquely represented by an integer $a \in \{0, \dots, p^s - 1\}$. So there exist one single $(\alpha_0, \dots, \alpha_{s-1}) \in \{0, \dots, p\}^s$ such that

$$\overline{a} = \alpha_0 \overline{1} + \alpha_1 \overline{p} + \cdots + \alpha_{s-1} \overline{p}^{s-1}. \tag{A.27}$$

This is the p-adic decomposition of $\overline{a}$. The p-valuation of $\overline{a}$ is

$$v_p(\overline{a}) = \begin{cases} \min(i \in \{0, \dots, s-1\}; \alpha_i \neq 0) & \text{for } a \neq 0, \\ s \quad \text{for } a = 0. \end{cases} \tag{A.28}$$

The order of $\overline{a}$ is then $p^{s - v_p(\overline{a})}$ and $\overline{a}$ is invertible iff its valuation is 0. Moreover, for all $\overline{a}, \overline{b} \in \mathbb{Z}_{p^s}$,

$$v_p(\overline{a} + \overline{b}) \geq \min(v_p(\overline{a}), v_p(\overline{b})), \tag{A.29a}$$
$$v_p(\overline{a}\overline{b}) = \min(v_p(\overline{a}) + v_p(\overline{b}), s), \tag{A.29b}$$

where equality in the latter formula relies on the fact that p is prime.

To check their understanding of p-adic decomposition, the reader should be able to literally see the following equalities, for any finite set $\{a_1, \dots, a_n\} \subset \mathbb{Z}$ of divisors of some $d \geq 2$:

$$\left(\bigwedge_{i=1}^n a_i\right)\left(\bigvee_{i=1}^n d/a_i\right) = d, \tag{A.30a}$$
$$\left(\bigvee_{i=1}^n a_i\right)\left(\bigwedge_{i=1}^n d/a_i\right) = d. \tag{A.30b}$$

Now, let us hark back to our search for λ_i. Since they are of the same order, a_i and x_i are both zero or nonzero. If they are nonzero, then according to (A.29b) applied to $a_i = q_i x_i$, q_i is of p_i-valuation 0. Hence it is invertible in $\mathbb{Z}_{p_i^{s_i}}$ and we take $\lambda_i = q_i$. If they are null, then $\nu_i = \pi_{p_i}(\overline{\nu}) = \overline{1}$ and either q_i or $q_i + \nu_i$ is of p_i-valuation 0 so that we get our λ_i. That is a simple proof of the

Lemma 35 *Let $d \geq 2$ and $a, b \in \mathbb{Z}_d$. The two following assertions are equivalent:*

1. *a, b are of the same order.*
2. *There exist $\lambda \in U(\mathbb{Z}_d)$ such that $a = \lambda b$.*

If one of them is satified, a and b are said to be associated. This is the case in particular if a and b are two gcd*'s of a same set of elements in $\mathbb{Z}_d$.*

What about the computation of the gcd of given $a_1, \dots, a_m \in \mathbb{Z}_d$ using the Chinese remainder theorem. Let $a_{ij} = \pi_{p_j}(a_i)$ for any $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$. In order to lighten notations, we avoid the bar over residue classes in this paragraph. The set to which any element belongs will be known from the context. Let also $\delta = \bigwedge_{i=1}^m a_i$ in $\mathbb{Z}_d$ and $\delta_j = \pi_{p_j}(\delta)$. It is quite obvious that in the

j-th Chinese factor of $\mathbb{Z}_d$ the gcd of the a_{ij}'s is

$$\bigwedge_{i=1}^{m} a_{ij} = p_j^{k_j}, \quad \text{with } k_j = \min(v_{p_j}(a_{ij});\ i \in \{1, \ldots, m\}) \leq s_j. \tag{A.31}$$

Indeed, if i_0 is an index for which $v_{p_j}(a_{i_0 j}) = k_j$, then $a_{i_0 j} = p_j^{k_j} u$, where u is invertible. Thus $p_j^{k_j}$ may be obtained as a linear combination of the a_{ij}'s and any linear combination of them is a multiple of $p_j^{k_j}$. Moreover $p_j^{k_j} = p_j^{k_j} \wedge p_j^{s_j}$ in $\mathbb{Z}$. Since δ is a linear combination of the a_i's, δ_j is a linear combination of the a_{ij}'s and so $v_{p_j}(\delta_j) \geq k_j$. Then, a_{i_0} being a multiple of δ, $a_{i_0 j}$ is a multiple of δ_j and so $v_{p_j}(\delta_j) = k_j$. Hence $\delta = \prod_{j=1}^{m} p_j^{k_j}$. All this is nothing but the usual way to compute gcd's by means of prime factor decomposition.

Another useful lemma is the following one. It is not often found in literature maybe for the crux is easy to see.

Lemma 36 *Let $d \geq 2$ and $a, b, \delta \in \mathbb{Z}_d$ such that δ is a* gcd *for a and b. If one of the following conditions is verified:*

- *d is odd,*
- *d is even and $v_2(a) \neq v_2(b)$,*
- *d is even and $v_2(a) = v_2(b) = v_2(d)$;*

then one can choose $u, v \in U(\mathbb{Z}_d)$ such that $\delta = ua + vb$. If not, then only u or v can be chosen invertible.

Proof. In this proof, in order to distinguish classes and representatives, we shall note $\overline{a}, \overline{b}, \overline{\delta}$ instead of a, b, δ as in the terms of the lemma. Using the Chinese remainder theorem, we search for u and v in each Chinese factor separately. So suppose $d = p^s$, with p odd to begin with. Also note that owing to of Lemma 35, it suffices to prove Lemma 36 for any gcd $\overline{\delta}$ of $\overline{a}$ and $\overline{b}$. So we will choose $\delta = a \wedge b$, taking into account the remark just following (A.4). By definition, there exist $u_0, v_0 \in \mathbb{Z}$ such that $\delta = u_0 a + v_0 b$, and dividing by δ we obtain

$$1 = u_0 a' + v_0 b' \tag{A.32}$$

where $a' = a/\delta$, $b' = b/\delta$. We see that u_0 and v_0 cannot be both multiples of p. At least one of $\overline{u_0}$ and $\overline{v_0}$, say $\overline{u_0}$, is a unit. Suppose $\overline{v_0}$ is not a unit, that is to say v_0 is a multiple of p. If $v_0 + a'$ were a multiple of p, then so would a', which would contradict (A.32) once more. So $v_0 + a'$ is a unit and so is $v_0 - a'$. Besides, if $u_0 \pm b'$ were both mutiples of p, so would be $2b'$, b' and then u_0. We may now conclude that at least one of the three pairs

$$(\overline{u_0}, \overline{v_0}), \quad (\overline{u_0 + b'}, \overline{v_0 - a'}), \quad (\overline{u_0 - b'}, \overline{v_0 + a'}) \tag{A.33}$$

is in $U(\mathbb{Z}_d)^2$. That proves the lemma as to the first condition.

If $p = 2$ and $v_2(\overline{a}) \neq v_2(\overline{b})$, then in (A.32), one of a' and b' is odd, say a', and the other is even, say b'. Moreover, u_0 has to be odd too. Then one of the two pairs

$$(\overline{u_0}, \overline{v_0}), \qquad (\overline{u_0 + b'}, \overline{v_0 - a'}) \tag{A.34}$$

is in $U(\mathbb{Z}_d)^2$.

If $p = 2$ and $v_2(\overline{a}) = v_2(\overline{b}) = v_2(d)$, then $\overline{a} = \overline{b} = \overline{\delta} = 0$ and $u = v = 1$ suit the lemma.

Still with $p = 2$, if $v_2(\overline{a}) = v_2(\overline{b}) \neq v_2(d)$, we have already seen that at least one of $\overline{u_0}$ and $\overline{v_0}$, say $\overline{u_0}$, is a unit. But $\overline{v_0}$ cannot be a unit, since in that case $u_0 a' + v_0 b'$ should be even. Because we only need $u_0 a' + v_0 b'$ to be odd, we can choose which one of $\overline{u_0}$ and $\overline{v_0}$ is invertible. ■

By induction and associativity of gcd, we have the

Corollary 37 *Let $a_1, a_2, \ldots, a_n \in \mathbb{Z}_d$ and δ be one of their* gcd*'s. For any $i \in \{1, \ldots, n\}$, one can find $k_1, k_2, \ldots, k_n \in \mathbb{Z}_d$ with $k_i \in U(\mathbb{Z}_d)$ such that*

$$\delta = \sum_{j=1}^{n} k_j a_j. \tag{A.35}$$

Appendix B

Finitely generated modules over $\mathbb{Z}_d$

Let d and n be two positive integers with $d \geq 2$. The set product $\mathbb{Z}_d^n$ is endowed with its canonical structure of $\mathbb{Z}$-module and its elements will be called vectors. Addition is componentwise:

$$\begin{array}{rcl} \mathbb{Z}_d^n \times \mathbb{Z}_d^n & \longrightarrow & \mathbb{Z}_d^n \\ ((a_1, \ldots, a_n), (b_1, \ldots, b_n)) & \longmapsto & (a_1 + b_1, \ldots, a_n + b_n) \end{array} \tag{B.1}$$

and the product map is

$$\begin{array}{rcl} \mathbb{Z} \times \mathbb{Z}_d^n & \longrightarrow & \mathbb{Z}_d^n \\ (k, (a_1, \ldots, a_n)) & \longmapsto & (ka_1, \ldots, ka_n). \end{array} \tag{B.2}$$

This can also be denoted $k \cdot (a_1, \ldots, a_n)$ or even $k(a_1, \ldots, a_n)$ without a symbol. Obviously, such a product depends only on the residue class of k modulo d, so that we may consider $\mathbb{Z}_d^n$ either as a $\mathbb{Z}$-module or a $\mathbb{Z}_d$-module. So, when the context is clear or the distinction useless, one can avoid the bar to denote residue classes.

A submodule of $\mathbb{Z}_d^n$ is a module over $\mathbb{Z}_d$ included in $\mathbb{Z}_d^n$. When $n = 1$, submodules are called ideals of $\mathbb{Z}_d$. Let I be a finite index set and $x = (x_i)_{i \in I}$ be a family of vectors in $\mathbb{Z}_d^n$. The submodule those vectors generate is the set of all their linear combinations over $\mathbb{Z}_d$ and is noted $\langle x \rangle$, or $\langle x_1, \ldots, x_r \rangle$ whenever $I = \{1, \ldots, r\}$. It is the tiniest submodule that contains all the x_i's. The family x is a generating system or basis of that submodule. Moreover, any submodule of $\mathbb{Z}_d^n$ is generated by some basis, since the whole submodule itself is such a basis. The family x is free if for all family $(c_i)_{i \in I}$ of elements of $\mathbb{Z}_d$,

$$\sum_{i \in I} c_i x_i = 0 \implies \forall i \in I, c_i = 0. \tag{B.3}$$

In other words, the linear map

$$\begin{array}{rccc} f_x : & \mathbb{Z}_d^I & \longrightarrow & \mathbb{Z}_d^n \\ & (c_i)_{i\in I} & \longmapsto & \sum_{i\in I} c_i x_i \end{array} \tag{B.4}$$

has kernel 0. A basis of a submodule which is also free is called a free basis of that submodule, and a submodule for which there exists a free basis is called a free submodule. The computational basis of $\mathbb{Z}_d^n$ is of course a free basis and it will be denoted by $e = (e_i)_{i=1\ldots n}$. For any vector a, $e_i^*(a) = a_i$ is the i-th component of a with respect to e.

A vector a such that the one-element family (a) is free is called a free vector. If moreover $n = 1$, then a is just said regular.

A submodule M is said to be of rank r if the minimal number of vectors needed to generate it is r. This notion of rank should not be confused with the rank of the matrix whose columns are a set of generating vectors of M with respect to some free basis of $\mathbb{Z}_d^n$ (see [40]). Those two notions of rank for submodules and matrices do not overlap.

A minimal basis for a rank-r submodule M is a basis of M with r elements. Such a basis need not be free. For instance in $\mathbb{Z}_4^2$, $((2,0))$ is a basis for the rank-1 submodule $\{(0,0),(2,0)\}$ but is not free. But if M is free, minimal and free bases are the same ones. Indeed, let $(m_i)_{i=1,\ldots,r}$ be a minimal basis of M and $(m_i')_{i\in I}$ be a free basis of M. By minimality of m, $r \leq |I|$ and by freedom of m', $|\mathrm{Im}\, f_{m'}| = d^{|I|}$. So

$$|M| = |\mathrm{Im}\, f_m| \leq d^r \leq d^{|I|} = |\mathrm{Im}\, f_{m'}| = |M|\,. \tag{B.5}$$

Thus on the one hand $|\mathrm{Im}\, f_m| = d^r$ and f_m must be injective, so that m is free. On the other hand, $|I| = r$ implies that m' is minimal.

Let $a = (a_1,\ldots,a_n) \in \mathbb{Z}_d^n$. The order $\nu(a)$ of a is the cardinality of the set $\mathbb{Z}_d \cdot a = \{ka; k \in \mathbb{Z}_d\}$. The only vector whose order is 1 is the null vector and a is a free vector iff $\nu(a) = d$. Endly, we note that $\nu(a)\mathbb{Z}$ is the kernel of the linear map

$$\begin{array}{rccc} f : & \mathbb{Z} & \longrightarrow & \mathbb{Z}_d^n \\ & k & \longmapsto & ka. \end{array} \tag{B.6}$$

This kernel is the intersection of the $\ker(e_i^* \circ f) = \nu(a_i)\mathbb{Z}$ and thus

$$\nu(a) = \bigvee_{i=1}^{n} \nu(a_i). \tag{B.7}$$

With (A.19) and (A.30a) we also deduce that

$$\nu(a) = \frac{d}{(\bigwedge_{i=1}^{n} a_i) \wedge d}. \tag{B.8}$$

Appendix C

Matrix reduction

Let d be any integer ≥ 2. As is the case for vector space theory over a field, vectors in finitely generated modules and linear maps between such modules can be represented by matrices. The canonical computational basis for vectors will be denoted e. A $k \times l$ matrix m is upper-triangular (resp. lower-triangular) if for all $i \in \{1, \ldots, k\}$, $j \in \{1, \ldots, l\}$, $i > j$ (resp. $i < j$), we have $m_{ij} = 0$. The matrix m is diagonal if for all $i \in \{1, \ldots, k\}$, $j \in \{1, \ldots, l\}$, $i \neq j$, we have $m_{ij} = 0$. The m_{ii}'s of any matrix will be called its diagonal coefficients. We extend to matrices the factor projections π_p defined in the Chinese remainder theorem (see Appendix A.2): If m is a $k \times l$ matrix over $\mathbb{Z}_d$ and p is a prime factor of d, then $\pi_p(m)$ is the $k \times l$ matrix over $\mathbb{Z}_{p^s}$, $s = v_p(d)$, whose (i, j) coefficient is $\pi_p(m_{ij})$. Also p-valuation is extended to matrices:

$$v_p(m) = \min(v_p(m_{ij});\ i \in \{1, \ldots, k\},\ j \in \{1, \ldots, l\}). \tag{C.1}$$

We will also adopt the conventions that a $*$ in a matrix denotes an arbitrary or unknown coefficient or submatrix, and a blank denotes a null coefficient or submatrix. The $k \times k$ indentity matrix will be written I_k and the $k \times l$ null matrix $0_{k,l}$ if necessary.

In this appendix we address trigonalisation and diagonalisation of matrices whose columns are basis vectors of a submodule of $\mathbb{Z}_d^n$. A left-multiplication by an invertible matrix is to be interpreted either as an active transformation, that is to say an automorphism of $\mathbb{Z}_d^n$, or as a passive transformation, that is to say a change of computational (free) basis. A right-multiplication by an invertible matrix stands for a change of basis of the submodule under consideration. The structure of the given submodule will be much easier to study after reduction. The reader interested in a more abstract treatment of matrix reduction and in particular diagonalisation of matrices over more general rings may have a look to [39,40,69]. By the way, we shall also have an insight into generalisation over $\mathbb{Z}_d$ of the "Incomplete basis theorem". The set of invertible matrices over $\mathbb{Z}$ is denoted $\mathrm{GL}(n, \mathbb{Z})$ and the set of invertible matrices over $\mathbb{Z}_d$ is denoted $\mathrm{GL}(n, \mathbb{Z}_d)$. Note that left-multiplication by an invertible matrix does not modify the order of a column vector and hence does not modify the gcd of its coefficients. The same is true for right-multiplication and row vectors.

The only preliminary result we shall admit is that a square matrix with coefficients in a commutative ring is invertible iff its determinant is an invertible element of that ring (see [40]). In fact, the proof is a mere copy of the field case, using the comatrix of the matrix under consideration.

Before we go on, a general remark is in order about the algorithms presented here. Except the algorithm $\mathscr{G}$, they are "blind" algorithms, that is to say we do not suppose we know where invertible coefficients are located in the matrices, which would be mandatory to use the classical Gaussian reduction for instance.

Lemma 38 *Let $a \in \mathbb{Z}_d^n$ be an n-dimensional vector. Then*

$$\exists L \in \mathrm{GL}(n, \mathbb{Z}_d),\ \exists k \in \mathbb{Z}_d, \quad La = ke_1. \tag{C.2}$$

The column vectors $C_1, \ldots, C_n$ of L^{-1} form a free basis of $\mathbb{Z}_d^n$ such that $kC_1 = a$.

Proof. Our calculations to prove this lemma will be in $\mathbb{Z}$. The results will only have to be sent onto residue classes at the end. Let $a \in \mathbb{Z}^n$, $\delta_{n-1} = a_{n-1} \wedge a_n$, $a'_{n-1} = a_{n-1}/\delta$, $a'_n = a_n/\delta$. There exist $k_1, l_1 \in \mathbb{Z}$ such that $k_1 a_{n-1} + l_1 a_n = \delta_{n-1}$ so that we have the active transformation on a:

$$\underbrace{\left(\begin{array}{c|cc} I_{n-2} & & \\ \hline & k_1 & l_1 \\ & -a'_n & a'_{n-1} \end{array}\right)}_{L^{(n-1)} \in \mathrm{GL}(n,\mathbb{Z})} \underbrace{\left(\begin{array}{c} * \\ \hline a_{n-1} \\ a_n \end{array}\right)}_{a} = \underbrace{\left(\begin{array}{c} * \\ \hline \delta_{n-1} \\ 0 \end{array}\right)}_{a^{(n-1)}}. \tag{C.3}$$

Repeating this trick on $a^{(n-1)}$ with components $n-1$ and $n-2$ and so on, we bring the vector a onto a multiple of e_1. Of course, the order of k in $\mathbb{Z}_d$ is the same as the order of a in $\mathbb{Z}_d^n$. In details:

$a^{(n)} = a, \delta_n = a_n,$

$$\forall i \in \{1, \ldots, n-1\}, \left\{ \begin{array}{rcl} \delta_{n-i} & = & a_{n-i} \wedge \delta_{n-i+1} \\ a'_{n-i} & = & a_{n-i}/\delta_{n-i} \\ \delta'_{n-i+1} & = & \delta_{n-i+1}/\delta_{n-i} \\ \multicolumn{3}{l}{\exists k_i, l_i \in \mathbb{Z}_d, k_i a_{n-i} + l_i \delta_{n-i+1} = \delta_{n-i}} \end{array} \right. ,$$

$$\underbrace{\left(\begin{array}{c|cc|c} I_{n-i-1} & & & \\ \hline & k_i & l_i & \\ & -\delta'_{n-i+1} & a'_{n-i} & \\ \hline & & & I_{i-1} \end{array}\right)}_{L^{(n-i)}} \underbrace{\left(\begin{array}{c} * \\ \hline a_{n-i} \\ \delta_{n-i+1} \\ \hline 0_{i-1,1} \end{array}\right)}_{a^{(n-i+1)}} = \underbrace{\left(\begin{array}{c} * \\ \hline \delta_{n-i} \\ 0 \\ \hline 0_{i-1,1} \end{array}\right)}_{a^{(n-i)}}. \tag{C.4}$$

Each $L^{(i)}$ has determinant 1, so that the complete transformation given by the product $L = \prod_{i=1}^{n-1} L^{(i)}$ also has and therefore is an automorphism. So we have shown what we were looking for:

$$\exists L \in \mathrm{GL}(n,\mathbb{Z}),\ \exists k \in \mathbb{Z}, \quad La = ke_1. \tag{C.5}$$

■

Lemma 39 *Let $a_1, a_2 \in \mathbb{Z}_d^n$ of order ν_1, ν_2 respectively. There exists a linear combination a of a_1, a_2 of order $\nu_1 \vee \nu_2$. Moreover, if d is odd, we can build a such that*

$$\langle a, a_1 \rangle = \langle a, a_2 \rangle = \langle a_1, a_2 \rangle . \tag{C.6}$$

If d is even, then in general we can have only

$$\langle a, a_1 \rangle \text{ or } \langle a, a_2 \rangle = \langle a_1, a_2 \rangle . \tag{C.7}$$

Proof. If a_1 or a_2 is equal to 0, the lemma is obvious. We now suppose that they are not and that d is odd. Let $A = (a_1|a_2)$ be the $n \times 2$ matrix whose columns are a_1, a_2 and with the help of Lemma 38, left-multiply A by an invertible matrix L such that La_1 has all but its first coefficient equal to 0. The matrix L is to be interpreted as a change of basis. If $k_1, \ldots, k_n$ are the coefficients of the second column of LA, let $\delta = k \wedge k_1$. According to Lemma 36 of Appendix A.2, there exist $u, v \in U(\mathbb{Z}_d)$ such that

$$\delta = uk + vk_1. \tag{C.8}$$

Then we put

$$(a_1'|a) = LA \begin{pmatrix} 0 & u \\ -u^{-1} & v \end{pmatrix} \quad \text{and} \quad (a|a_2') = LA \begin{pmatrix} u & 0 \\ v & u^{-1} \end{pmatrix} \tag{C.9}$$

or

$$(a_1'|a) = LA \begin{pmatrix} v^{-1} & u \\ 0 & v \end{pmatrix} \quad \text{and} \quad (a|a_2') = LA \begin{pmatrix} u & -v^{-1} \\ v & 0 \end{pmatrix}. \tag{C.10}$$

In any case, a answers the lemma since, with Lemma 35 and Equations (A.19), (A.30a) and (B.8), the order of a is

$$\begin{aligned} \nu(a) = \frac{d}{\delta \wedge (\bigwedge_{i=2}^n k_i) \wedge d} &= \frac{d}{(k \wedge d) \wedge (\bigwedge_{i=1}^n k_i \wedge d)} \\ &= \left(\frac{d}{k \wedge d}\right) \vee \left(\frac{d}{\bigwedge_{i=1}^n k_i \wedge d}\right) = \nu(a_1) \vee \nu(a_2). \end{aligned} \tag{C.11}$$

And for $i = 1, 2$,

$$\langle a, a_i \rangle = \langle a, a_i' \rangle = \langle a_1, a_2 \rangle . \tag{C.12}$$

To complete the proof, let us deal with the case where $d = 2^s$. With $i = 1$ or 2 such that $\nu(a_i) = \max(\nu(a_1), \nu(a_2))$, we simply put $a = a_i$. ■

Note that for any linear combination $b = b_1 a_1 + b_2 a_2$ of a_1 and a_2,

$$\nu(a)b = b_1(\nu(a)a_1) + b_2(\nu(a)a_2) = 0. \tag{C.13}$$

Thus for all $b \in \langle a_1, a_2\rangle$, $\nu(b)$ divides $\nu(a)$.

Given two minimal bases $f = (f_1, \ldots, f_r)$ and $g = (g_1, \ldots, g_r)$ of a submodule M, it is in general not possible to find an automorphism of M that brings f_i onto g_i for all i, even if $\nu(f_i) = \nu(g_i)$ for all i. Indeed in $\mathbb{Z}_6$, we cannot find $a \in \mathbb{Z}_6$ and $b \in U(\mathbb{Z}_6)$ so that

$$\begin{pmatrix} 1 & a \\ 0 & b \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}. \tag{C.14}$$

We can take b to be 1 or 5. But a should be such that $1+3a = 2$, which is impossible. As to diagonalisation, left-multiplication is still not sufficient, especially because the order of respective column vectors from one basis to the other is not preserved:

$$\begin{pmatrix} 1 & a \\ 0 & b \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}. \tag{C.15}$$

We shall make use of Lemma 39 to perform diagonalisation with left- and right-multiplications. For instance, the latter inequation is solved trivially:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}. \tag{C.16}$$

Suppose that we are given a minimal basis $b = (b_1, \ldots, b_r)$ of a submodule M of $\mathbb{Z}_d^n$ and B is the matrix of size $n \times r$ whose i-th column is b_i. The matrix B is called a basis matrix for M. With the help of Lemma 38, we could easily put B in an upper-triangular form by means of left-multiplications. But we are going to transform it into a new, diagonal matrix whose column vectors still generate M. Because of Lemma 39 and associativity of lcm, we may suppose that

$$\nu(b_1) = \bigvee_{i=1}^{r} \nu(b_i), \tag{C.17a}$$

$$\forall m \in M,\ \nu(m) | \nu(b_1). \tag{C.17b}$$

An algorithm which set any matrix that way will be called $\mathscr{A}$. It consists of an appropriate right-multiplication by an invertible matrix. We left-multiply B by a matrix L_1 with determinant 1 so that $L_1 b_1$ has all but its first coefficient equal to 0. Let $\widetilde{B} = L_1 B$. If one of the coefficients of $\widetilde{B}$ but in the first column, say $\widetilde{b}_{ij}$, $j \geq 2$, were not a multiple of the upper-left coefficient $\widetilde{b}_{11}$, then $\nu\left(\widetilde{b}_{ij}\right)$ would not be a divisor of $\nu\left(\widetilde{b}_{11}\right) = \nu(b_1)$ and according to Relation (B.7) of Appendix B and to Lemma 39 again, there would exist a linear combination of b_1 and b_j of order greater than $\nu(b_1)$, which is impossible by assumption. Since we are only interested in a basis of M we can put all but the first coefficient of the first row to 0 and obtain a matrix B_1. This is equivalent to a right-multiplication by an appropriate invertible matrix. Carrying on this process, we obtain a diagonal matrix B_r whose column vectors still form a minimal basis of M. Let us describe the algorithm in details.

Algorithm $\mathscr{D}_0$: The starting point is the empty matrix D_0 with no lines and no columns, and as an argument a $k \times l$ matrix B. Let $B_0 = B$. Then for i from 0 to $\mu = \min(k-1, l-1)$, we go on the following steps:

1. $R^{(1)}_{i+1} = \begin{pmatrix} I_i & 0 \\ 0 & R' \end{pmatrix}$ with R' a $(l-i) \times (l-i)$ invertible matrix such that $\mathscr{A}(B_i) = B_i R'$.
2. $L_{i+1} = \begin{pmatrix} I_i & 0 \\ 0 & L' \end{pmatrix}$ with L' an $(k-i) \times (k-i)$, determinant-1 matrix given by Lemma 38 such that $B' = L'\mathscr{A}(B_i)$ has all its first column coefficients but the first one equal to 0.
3. $R^{(2)}_{i+1} = \begin{pmatrix} I_i & 0 \\ 0 & R'' \end{pmatrix}$ with R'' a $(l-i)\times(l-i)$ invertible matrix such that $B'R''$ has all its first line coefficients but the first one equal to 0.
4. $D_{i+1} = \begin{pmatrix} D_i & 0 \\ 0 & b'_{11} \end{pmatrix}$.
5. B_{i+1} is given from B' by deleting the first row and the first column of this latter one.

The results of the algorithm are the change of basis matrices $L(B) = \prod_{i=1}^{\mu+1} L_{\mu+2-i}$, $R(B) = \prod_{i=1}^{\mu+1} R^{(1)}_i R^{(2)}_i$ and the $k \times l$ diagonal matrix $\mathscr{D}_0(B)$ defined to be

$$\begin{pmatrix} D_{\mu+1} \\ 0_{k-l,l} \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} D_{\mu+1} & 0_{k,l-k} \end{pmatrix} \tag{C.18}$$

whether $k \geq l$ or $k \leq l$ respectively. For all $i, j \in \{1, \dots, r\}$, $i < j$, we have $(D_{\mu+1})_{ii} | (D_{\mu+1})_{jj}$. ♦

As to the minimal basis b, the second case for $\mathscr{D}_0(B)$ is impossible and thus $r \leq n$. The minimality of b also implies that none of the diagonal coefficients of $D_{\mu+1} = D_r$ is 0. Hence, the column vectors of $\mathscr{D}_0(B)$ still form a minimal basis of M. Additionally, note that if we replace every diagonal entry of $\mathscr{D}_0(B)$ by 1, the column vectors of the matrix we obtain form a free basis $\widehat{b}$ of a free, rank-r submodule $M_{\widehat{b}}$ containing M.

The remaining features stated in Theorem 40 below are already known from the classification of finite, commutative groups and general commutative ring theory. However, we prove them in a new way as an illustration of the material above, with counting properties in $\mathbb{Z}_d$-modules.

If we start with a nonminimal basis of M, say b' with $r+r'$ vectors, $r' \geq 1$, the algorithm $\mathscr{D}_0$ yields a matrix of the form

$$\mathscr{D}_0(B') = \begin{pmatrix} D & 0_{n,r+r'-k} \end{pmatrix}, \tag{C.19}$$

where D is a diagonal matrix with k columns, all of them nonzero. Since M is of rank r, we have $k \geq r$. Suppose $k > r$ and let $\widetilde{D}$ be the $(1,\ldots,n;1,\ldots,r+1)$ submatrix of D. There exists an $r \times (r+1)$ matrix E whose j-th column, $j \in \{1,\ldots,r+1\}$, contains the components of the j-th column vector of $\widetilde{D}$ with respect to the free basis $\widehat{b}$. A linear combination of the column vectors of $\widetilde{D}$ with some factors is null iff the linear combination of the respective column vectors of E with these same factors is null. In other words, $\widetilde{D}$ and E have the same kernel as linear maps. Applying the algorithm $\mathscr{D}_0$ to E, we construct a null linear combination of its column vectors the factors of which are located in the last column C of $R(E)$. Now, let us choose a prime factor p of d such that $t = v_p(d_{r+1,r+1}) < v_p(d)$, so that no diagonal entry of $\pi_p(\widetilde{D})$ is null. There exists such a p because $d_{r+1,r+1} \neq 0$. Since $R(E)$ is invertible, at least one of the factors contained in $\pi_p(C)$ is a unit. But in that case, $\pi_p(\widetilde{D}C)$ cannot be null as expected. So $k = r$. Thus we may add to the algorithm $\mathscr{D}_0$ a final step to get the

Simple reduction algorithm $\mathscr{D}$: Let M be a rank-r submodule of $\mathbb{Z}_d^n$, b a basis of M containing $s \geq r$ vectors and B the corresponding basis matrix. By deleting the last $s-r$ null columns of $\mathscr{D}_0(B)$, one gets a minimal basis matrix for M. The matrix $\mathscr{D}(b) = \mathscr{D}(B)$ thus obtained is called the simple reduction of the basis b or of the basis matrix B. ♦

Let $b^{(1)}$ and $b^{(2)}$ be two bases of M. In the next three paragraphs, we are going to work in a single Chinese factor, say with prime factor p, and we are to prove that for every $i \in \{1,\ldots,r\}$, the i-th diagonal entries of $\mathscr{D}(b^{(1)})$ and $\mathscr{D}(b^{(2)})$ are associated. In order to make notations lighter, we even suppose that d is a power of a prime, say p^s. There is a slight difference, since in the latter case, r may vary with the Chinese factor one chose initially. The reader may check that such a trick is allowed. Let $B^{(a)} = L(b^{(a)})^{-1}\mathscr{D}(b^{(a)})$, $a \in \{1,2\}$, $\widehat{B}$ be the representative matrix of $\widehat{b}$ with respect to the computational basis and P^{12}, P^{21} and E be three $r \times r$ matrices such that

$$B^{(1)}P^{12} = B^{(2)}, \quad B^{(2)}P^{21} = B^{(1)}, \quad \widehat{B}E = B^{(1)}. \tag{C.20}$$

So we have

$$\widehat{B}EP^{12}P^{21} = B^{(1)}P^{12}P^{21} = B^{(1)} = \widehat{B}E \tag{C.21}$$

and then

$$\mathscr{D}(E)P = \mathscr{D}(E), \quad \text{with } P = R(E)^{-1}P^{12}P^{21}R(E). \tag{C.22}$$

If some diagonal entry of $\mathscr{D}(E)$ were zero, then the column vectors of $B^{(1)}R(E) = \widehat{B}L(E)^{-1}\mathscr{D}(E)$ would form a basis of M with at most $r-1$ elements, which is impossible. So there exists an $r \times r$ matrix Q such that $P = I_r + pQ$. Hence P is invertible, and so are P^{12} and P^{21}. For $a \in \{1,2\}$, consider the maps

$$\begin{array}{rccc} f^{(a)}: & (\mathbb{Z}_{p^s})^r & \longrightarrow & M \\ & X & \longmapsto & B^{(a)}X \end{array} \tag{C.23}$$

where elements of $(\mathbb{Z}_{p^s})^r$ are presented as column vectors and let $n_i^{(a)}$, $i \in \{0,\ldots,s\}$, be the number of vectors X so that $f^{(a)}(X)$ is of order p^{s-i}. For every X such that $B^{(1)}X$ is of order p^{s-i}, the vector $Y = P^{21}X$ is such that $B^{(2)}Y$ is of order p^{s-i} as well. Since P^{21} is injective as a linear map, we have $n_i^{(2)} \geq n_i^{(1)}$. The converse inequality can be shown the same way and so $n_i^{(1)} = n_i^{(2)}$.

Now let b be any basis of M and r_i, $i \in \{0,\ldots,s-1\}$, be the number of diagonal entries of $D = \mathscr{D}(b)$ of the form up^i, $u \in U(\mathbb{Z}_{p^s})$. We also define the following two related objects

$$\forall i \in \{-1,\ldots,s-1\}, \quad \sigma_i = \sum_{j=0}^{i} r_j, \tag{C.24}$$

and as intervals in $\mathbb{N}$

$$\forall i \in \{0,\ldots,s-1\}, \quad K_i = \{\sigma_{i-1}+1,\ldots,\sigma_i\}. \tag{C.25}$$

The cardinality of a K_i is of course r_i. We are to prove by induction on i that the r_i's do not depend on the choice of b and so are properties of M. As in the previous paragraph, consider the map

$$\begin{array}{rccc} f: & (\mathbb{Z}_{p^s})^r & \longrightarrow & M \\ & X & \longmapsto & DX. \end{array} \tag{C.26}$$

The number of vectors X such that $f(X)$ is of order p^{s-i}, $i \in \{0,\ldots,s-1\}$, is

$$\begin{aligned} n_i = \sum_{j=0}^{i} \Bigg\{ \left[\prod_{k=0}^{j-1} p^{((s-1)-(i-k))r_k} \right] \times \left[p^{((s-1)-(i-j-1))r_j} - p^{((s-1)-(i-j))r_j} \right] \times \\ \times \left[\prod_{k=j+1}^{i} p^{((s-1)-(i-k-1))r_k} \right] \Bigg\} \times \prod_{l=i+1}^{s-1} p^{sr_l}. \end{aligned} \tag{C.27}$$

Indeed, one can consider the bar graph in Figure C.1 to see where that latter expression comes from.

The individual positions on the horizontal axis have not been displayed. Instead, only the relevant intervals of them have been. For any $X \in (\mathbb{Z}_{p^s})^r$ and any $a \in \{0,\ldots,r\}$, the vertical bars in plain or dashed lines and the horizontal dotted line above the a-th position show lower bounds for the p-valuation of the a-th coefficient of $f(X)$. Thus as a property of D, if $a \in K_l$, we have $v_p(f(X)_a) \geq l$ as shown by

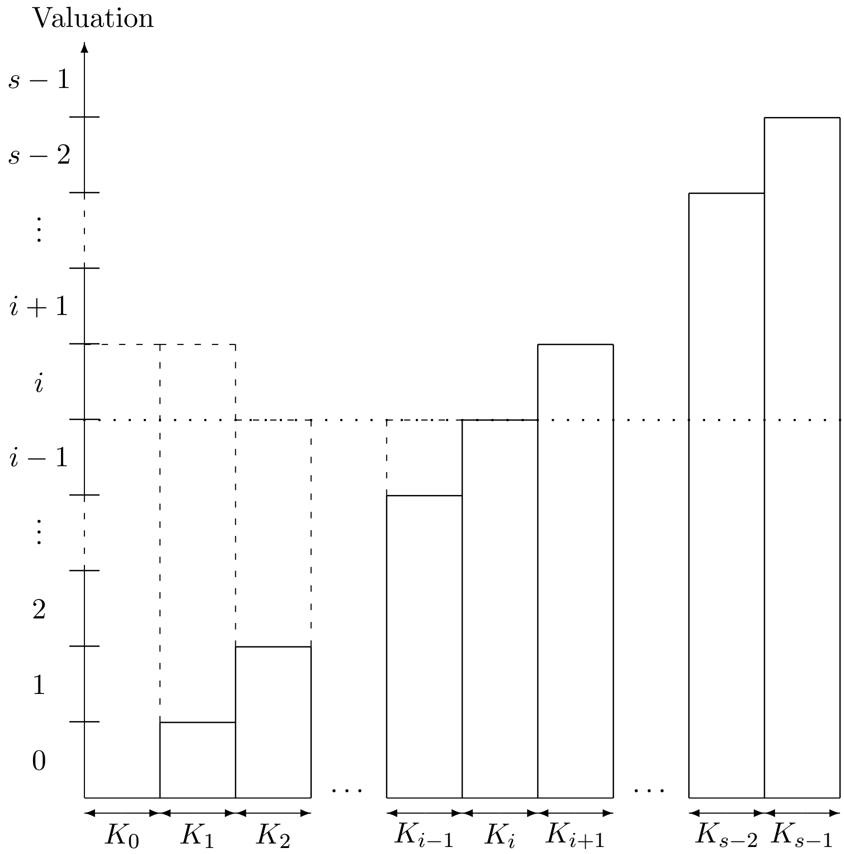

Figure C.1: How to calculate the n_i's

the plain line bars. Since the order of $f(X)$ is prescribed to be p^{s-i}, $v_p(f(X)_a) \geq i$ as shown by the dotted line. Finally, we put

$$j = \min(k \in \{0, \dots, s\};\ \exists l \in K_k,\ v_p(f(X)_l) = i). \tag{C.28}$$

There exists such a j ($j = 2$ in the example on the graph) and if $a \leq j-1$, $v_p(f(X)_a) > i$ as shown by the dashed line bars. These various lower bounds partition $\{0, \dots, r\}$ into four subintervals corresponding to the four factors in the above expression of n_i. The sum amounts for all the possibilties for j.

For $i = 0$, we have

$$n_0 = (p^{sr_0} - p^{(s-1)r_0})p^{s(r-r_0)} = p^{sr}\left(1 - \frac{1}{p^{r_0}}\right). \tag{C.29}$$

This quantity, which is a property of M, would increase strictly with r_0. Hence r_0 does not depend on the choice of b. For $i \geq 1$, we suppose that for $j \leq i-1$, the r_j's do not depend on b. Then there exist a nonnegative integer α and a positive integer β such that

$$n_i = (\alpha p^{sr_i} + \beta(p^{sr_i} - p^{(s-1)r_i}))p^{s(r-\sigma_{i-1}-r_i)} = p^{s(r-\sigma_{i-1})}\left(\alpha + \beta - \frac{\beta}{p^{r_i}}\right). \tag{C.30}$$

Again, we conclude that r_i does not depend on b. The number n'_i of vectors of order p^{s-i} in M, $i \in \{0, \ldots, s\}$, would have been much more obvious a property of M. But for $i \in \{0, \ldots, s-1\}$:

$$\begin{aligned} n'_i = \sum_{j=0}^{i} \Bigg\{ \left[\prod_{k=0}^{j-1} p^{((s-k-1)-(i-k))r_k}\right] \times \left[p^{((s-j-1)-(i-j-1))r_j} - p^{((s-j-1)-(i-j))r_j}\right] \times \\ \times \left[\prod_{k=j+1}^{i} p^{((s-k-1)-(i-k-1))r_k}\right]\Bigg\} \times \prod_{k=i+1}^{s-1} p^{((s-k-1)-(-1))r_k} \\ = \sum_{j=0}^{i} \Bigg\{ \left[\prod_{k=0}^{j-1} p^{(s-i-1)r_k}\right] \times \left[p^{(s-i)r_j} - p^{(s-i-1)r_j}\right] \times \left[\prod_{k=j+1}^{i} p^{(s-i)r_k}\right]\Bigg\} \times \prod_{k=i+1}^{s-1} p^{(s-k)r_k}, \end{aligned} \tag{C.31}$$

with a cumbersome $\sum_{k=i+1}^{s-1} kr_k$ appearing as an exponent in the last factor. Even if we looked at n'_i/n'_{i-1} to handle the induction, that exponent would stay for the initialisation at $i = 0$.

Finally, harking back to the case where d is not necessarily a power of a prime, the number $r_s = r - (r_0 + \ldots + r_{s-1})$ of diagonal entries of $\mathscr{D}(b)$ with p-valuation $v_p(d)$ is a property of M. We sum up our results about simple reduction in the

Theorem 40 *For any rank-r submodule M of $\mathbb{Z}_d^n$, there exist a free basis f of $\mathbb{Z}_d^n$ and a minimal basis b of M such that:*

1. *b is represented by a diagonal $n \times r$ matrix B with respect to f;*
2. *for all $i, j \in \{1, \ldots, r\}$, $i < j$, we have $b_{ii} | b_{jj}$.*

Such a pair of bases (f, b) can be found from any basis b_0 of M by the simple reduction algorithm $\mathscr{D}$. Moreover, for any pair (f, b) as above, the sequence $(d/\nu(b_{ii}))_{i \in \{1,\ldots,r\}}$ of the diagonal entries of B "without unit factors" is the same and therefore is a property of M. That sequence is called the sequence of the elementary divisors of M.

The pair (f, b) is not unique and with the notations of the theorem, M is free iff for all $i \in \{1, \ldots, r\}$, b_{ii} is a unit in $\mathbb{Z}_d$, or in other words iff its sequence of elementary divisors contains only 1's.

Corollary 41 *Let $\beta_0 = (b_1, \ldots, b_r)$ be a free family of $\mathbb{Z}_d^n$. Then $r \leq n$ and there exist $n - r$ vectors $b_{r+1}, \ldots, b_n \in \mathbb{Z}_d^n$ so that $\beta = (b_1, \ldots, b_r, b_{r+1}, \ldots, b_n)$ is a free basis of $\mathbb{Z}_d^n$.*

Proof. Indeed, with D the $(1, \ldots, r; 1, \ldots, r)$ submatrix of the $r \times n$ diagonal matrix $\mathscr{D}(\beta_0)$, a representative matrix for such a β with respect to the computational basis is

$$L(\beta_0)^{-1} \operatorname{diag}(DR(\beta_0)^{-1}, I_{n-r}). \tag{C.32}$$

■

Corollary 42 *For any two submodules M and N of $\mathbb{Z}_d^n$ with the same elementary divisors, which implies that they have the same rank, there exists an automorphism of $\mathbb{Z}_d^n$ that brings M onto N.*

Proof. Let (f, b) (resp. (h, c)) be a convenient pair for M (resp. N) as in theorem 40. Then the automorphism of $\mathbb{Z}_d^n$ defined by $b_i \mapsto c_i$, $i \in \{1, \ldots, n\}$, brings M onto N. ■

To close this appendix, we give a simple algorithm $\mathscr{G}$, which is in fact a adaptation of Gaussian reduction of matrices, in order to find the elementary divisors of a given submodule M together with a diagonalising basis f of $\mathbb{Z}_d^n$.

Algorithm $\mathscr{G}$: Let B be a basis matrix for M. The following steps must be performed in each Chinese factor separately.

1. Swap two rows on the one hand and two columns of the other hand so that the new upper-left coefficient has p-valuation $v_p(B)$. This consists in a left- and a right-multiplication by invertible matrices, respectively.

2. Apply the next step of classical Gaussian reduction so as to set to 0 every other coefficient on the first column and on the fisrt row. It consists again in left- and right-multiplication by invertible matrices.

3. Repeat the process with the submatrix obtained by deleting the first column and the first row.

Each left-multiplication has to be interpreted as a change of computational basis so that f can be deduced from the sequence of these multiplications. The right-mutiplications only stands for a change of basis of M. ♦

Appendix D

Submodules and wedge product

As we shall see with the lemmas and claim below, the notion of wedge product is intimately related to that of a submodule and its bases on the one hand, and on the other hand it can be a tool to show whether a vector is in a given submodule or not. For that, we shall need the notion of a $\mathbb{Z}_d$-algebra.

Definition 43 *A $\mathbb{Z}_d$-algebra A is a $\mathbb{Z}_d$-module endowed with an extra operation $*$ such that:*

1. *$*$ is distributive over addition in A;*

2. $\forall \lambda \in \mathbb{Z}_d,\ \forall x, y \in A, \quad \lambda(x * y) = (\lambda x) * y = x * (\lambda y)$.

Over a vector space E, a first way to define the wedge product is to completely antisymmetrise tensor products. But that requires to divide by some factorials in order to warrant associativity for the wedge product. A second way is to define in a first place the exterior algebra $\bigwedge E$ as a quotient of the tensor algebra $\bigotimes E$ by a well-chosen ideal of it. If e is a basis of E, this ideal is generated by

$$\{e_i \otimes e_j + e_j \otimes e_i;\ e_i, e_j \in e\}. \tag{D.1}$$

In the case of the $\mathbb{Z}_d$-module $\mathbb{Z}_d^n$, we have to prevent ourselves from refering to any definition involving divisions and thus we retain the second one. So let us build up the exterior algebra over $\mathbb{Z}_d^n$. But our construction will not use the notion of an ideal nor of a quotient ring. We opt for an equivalent logical way.

Let $F(\mathbb{Z}_d^n)$ be the free $\mathbb{Z}_d$-module a free basis of which consists of all finite sequences with values in $\mathbb{Z}_d^n$, including the empty sequence. This means that $F(\mathbb{Z}_d^n)$ is the set of all formal, finite linear combinations of those sequences and is by the same token endowed with a natural structure of $\mathbb{Z}_d$-module. Within $F(\mathbb{Z}_d^n)$, for any $k \in \mathbb{N}$, we single out the finitely generated submodule $F^k(\mathbb{Z}_d^n)$ which is generated by the sequences with values in $\mathbb{Z}_d^n$ of length k. Thus, the rank of $F^k(\mathbb{Z}_d^n)$ is d^{nk} and for instance

$$F^0(\mathbb{Z}_d^n) \simeq \mathbb{Z}_d, \quad F^1(\mathbb{Z}_d^n) \simeq \mathbb{Z}_d^n. \tag{D.2}$$

We endow $F(\mathbb{Z}_d^n)$ with a strucuture of $\mathbb{Z}_d$-algebra. Restricted to sequences, the extra operation $*$ is just concatenation. It is then extended to $F(\mathbb{Z}_d^n)$ as a whole by linearity. For instance:

$$\begin{aligned}(x_1,\dots,x_k)*(\lambda_1(y_1,\dots,y_l)+\lambda_2(z_1,\dots,z_m))\\ =\lambda_1(x_1,\dots,x_k,y_1,\dots,y_l)+\lambda_2(x_1,\dots,x_k,z_1,\dots,z_m).\end{aligned} \tag{D.3}$$

If one prescribes the following scheme of equalities among vectors of $F(\mathbb{Z}_d^n)$ (resp. $F^k(\mathbb{Z}_d^n)$),

$$\forall\lambda_1,\lambda_2\in\mathbb{Z}_d,\ \forall x_1,x_2\in\mathbb{Z}_d^n,\quad (\dots,\lambda_1x_1+\lambda_2x_2,\dots)=\lambda_1(\dots,x_1,\dots)+\lambda_2(\dots,x_2,\dots), \tag{D.4}$$

one obtains the tensor algebra $\bigotimes\mathbb{Z}_d^n$ (resp. the k-th tensor power $\bigotimes^k\mathbb{Z}_d^n$) of $\mathbb{Z}_d^n$. In the tensor algebra, the operation $*$ is called the tensor product and is noted with the tensor symbol $\otimes$. Any finite sequence $(x_1,\dots,x_k)$ with values in $\mathbb{Z}_d^n$ is then denoted

$$x_1\otimes\dots\otimes x_k. \tag{D.5}$$

The rank of $\bigotimes^k\mathbb{Z}_d^n$ is n^k and a free basis for it is the set of all

$$e_{i_1}\otimes\dots\otimes e_{i_k}, \tag{D.6}$$

with $e=(e_1,\dots,e_n)$ the computational basis of $\mathbb{Z}_d^n$ and i any sequence of length k with values in $\{1,\dots,n\}$. If in addition, one prescribes the following antisymmetry scheme of equalities,

$$\forall x_1,x_2\in\mathbb{Z}_d^n,\quad (\dots,x_1,\dots,x_2,\dots)=-(\dots,x_2,\dots,x_1,\dots), \tag{D.7}$$

then one obtains the exterior algebra $\bigwedge\mathbb{Z}_d^n$ (resp. the k-th exterior power $\bigwedge^k\mathbb{Z}_d^n$) of $\mathbb{Z}_d^n$. In the exterior algebra, the operation $*$ is called the wedge product and is noted with the wedge symbol $\wedge$. Any finite sequence $(x_1,\dots,x_k)$ with values in $\mathbb{Z}_d^n$ is then denoted

$$x_1\wedge\ldots\wedge x_k. \tag{D.8}$$

The rank of $\bigwedge^k\mathbb{Z}_d^n$ is $\binom{n}{k}$, the number of subsets of cardinality k within a set of n elements, and a free basis for it whenever $1\le k\le n$ is the set of all

$$e_{i_1}\wedge\ldots\wedge e_{i_k}, \tag{D.9}$$

with i any strictly increasing sequence of length k with values in $\{1,\dots,n\}$. Thence $\bigwedge\mathbb{Z}_d^n$ is finitely generated with rank

$$\operatorname{rank}\left(\bigwedge\mathbb{Z}_d^n\right)=\sum_{k=0}^{\infty}\binom{n}{k}=2^n. \tag{D.10}$$

Let m be an $n \times k$ matrix over $\mathbb{Z}_d$, $1 \leq k \leq n$, and m_i that of its square $k \times k$ submatrices with j-th line the (i_j)-th line of m. The reader may check that the component on $e_{i_1} \wedge \ldots \wedge e_{i_k}$ of the wedge product of the column vectors of m is $\det(m_i)$.

Lemma 44 *Let $x, y \in \mathbb{Z}_d^n$ with x a free vector. Then*

$$y \in \langle x \rangle \Longleftrightarrow x \wedge y = 0 \tag{D.11}$$

Proof. Since $x = \sum_{i=1}^n x_i e_i$ is free, there exists $(k_1, \ldots, k_n) \in \mathbb{Z}_d^n$ such that

$$\sum_{i=1}^n k_i x_i = 1. \tag{D.12}$$

Besides, with $y = \sum_{i=1}^n y_i e_i$, we suppose

$$x \wedge y = \sum_{i<j} (x_i y_j - x_j y_i) e_i \wedge e_j = 0 \tag{D.13}$$

and then

$$\forall i, j \in \{1, \ldots, n\}, \quad x_i y_j = x_j y_i. \tag{D.14}$$

So

$$\forall j \in \{1, \ldots, n\}, \quad y_j = \left(\sum_{i=1}^n k_i y_i \right) x_j. \tag{D.15}$$

The converse implication is obvious. ■

Lemma 45 *Let $x_1, x_2, y, z \in \mathbb{Z}_d^n$ so that y is a free vector of $\langle x_1, x_2 \rangle$. Then*

$$z \in \langle x_1, x_2 \rangle \quad \Longleftrightarrow \quad \exists k \in \mathbb{Z}_d,\ y \wedge z = k x_1 \wedge x_2. \tag{D.16}$$

Proof. There exists $(k_1, k_2) \in \mathbb{Z}_d^2$ such that $y = k_1 x_1 + k_2 x_2$. Then, if we suppose that the right-hand side of the latter equivalence holds,

$$y \wedge (k_1 z - k x_2) = 0 \tag{D.17}$$

and according to Lemma (44), there exists $l_1 \in \mathbb{Z}_d$ such that $k_1 z = l_1 y + k x_2 = k_1 l_1 x_1 + (k_2 l_1 + k) x_2 \in \langle x_1, x_2 \rangle$. By the same token, $k_2 z \in \langle x_1, x_2 \rangle$. Since y is free, $\gcd(k_1, k_2) = 1$ and hence $z \in \langle x_1, x_2 \rangle$. The converse is obvious. ■

Thus, Lemma (44) gives an equation for vectors in a free rank-1 submodule, whereas Lemma (45) characterises vectors in a rank-2 submodule containing a free vector.

Claim 46 *Let $r \in \{1, \ldots, n\}$ and $x_1, \ldots, x_r \in \mathbb{Z}_d^n$. Then $x = (x_1, \ldots, x_r)$ is a free basis of $M = \langle x_1, \ldots, x_r \rangle$ iff $x_1 \wedge \ldots \wedge x_r$ is a free vector of $\bigwedge \mathbb{Z}_d^n$.*

Proof. Contrary to what happens to the sympletic product, a change of computational basis never changes the way of computing a wedge product. Hence we choose a computational basis f so that the representative matrix of x be upper-triangular. x is a free basis for M iff all the diagonal coefficients of that matrix are invertible. But this is equivalent to the fact that their product k is invertible. The conclusion arises from the facts that $x_1 \wedge \ldots \wedge x_r = kf_1 \wedge \ldots \wedge f_r$ and that $f_1 \wedge \ldots \wedge f_r$ is free according (D.9). ■

These lemmas and claim are enough to see the importance of the notion of a free vector in the use of the wedge product, in contrast to the field case.

Appendix E

Projective geometry

In this thesis, we make use of the notion of a projective structure over the rings $\mathbb{Z}_d$ and $\mathrm{Mat}(n, \mathbb{Z}_d)$. In order to show the difference of projective geometry over a ring compared with the field case, we first recall briefly how a projective geometry is build over a vector space. By the way, we also recall the notion of anharmonic ratio (or cross-ratio) in projective geometry over a field.

E.1 The field case

Let $\Bbbk$ be a field and $n \in \mathbb{N}$. One defines the colinearity relation over $\Bbbk^n \setminus \{0\}$ by

$$\forall x, y \in \Bbbk^n \setminus \{0\}, \quad x \sim y \iff \exists \lambda \in \Bbbk^*,\ x = \lambda y. \tag{E.1}$$

This is an equivalence relation whose classes are the vectorial lines of $\Bbbk^n$ deprived of 0. They are called the (projective) points of the projective structure $\mathbf{P}(\Bbbk^n)$. A projective point is often identified with any of its representatives

$$\mathrm{class}(x) \longleftrightarrow x = (x^0, x^1, \ldots, x^{n-1}). \tag{E.2}$$

For convenience, the coordinates of x are indexed by intergers ranging from 0 to $n-1$. Then the following notations are classical for projective points:

$$\begin{aligned} 0 &= (1, 0, \ldots, 0), && \text{(E.3)}\\ \infty_i &= (0, \ldots, 0,\ (x^i =)\ 1, 0, \ldots, 0), \quad i \in \{1, \ldots, n-1\}. && \text{(E.4)} \end{aligned}$$

Finally, the projective points x of the form

$$(1, x^1, \ldots, x^{n-1}) \tag{E.5}$$

are in bijective correspondance with the vectors of an $(n-1)$-dimensional vector space over $\Bbbk$, namely $\Bbbk^{n-1}$, so that they are classically identified with those vectors and denoted

$$(x^1, \ldots, x^{n-1}). \tag{E.6}$$

Accordingly, the corresponding subset of $\mathbf{P}(\Bbbk^n)$ is identified with $\Bbbk^{n-1}$ and its complement $\mathbf{P}(\Bbbk^n) \setminus \Bbbk^{n-1}$ is called the hyperplane at infinity in $\mathbf{P}(\Bbbk^n)$.

Whenever $n = 2$, one calls $\mathbf{P}(\Bbbk^2)$ the projective line over $\Bbbk$. Let $x_i = (x_i^0, x_i^1) \in \Bbbk^2$, $i \in \{1, \ldots, 4\}$, be the representatives of four points on that line and let us denote

$$\omega_{ij} = \begin{vmatrix} x_i^0 & x_j^0 \\ x_i^1 & x_j^1 \end{vmatrix}, \qquad i, j \in \{1, \ldots, 4\}. \tag{E.7}$$

The anharmonic ratio of the x_i's is denoted (x_1, x_2, x_3, x_4) and is defined to be the point in $\mathbf{P}(\Bbbk^2)$ with representative

$$(\ \omega_{31}\omega_{42}\ ,\ \omega_{32}\omega_{41}\ ). \tag{E.8}$$

With the classical notations -1 for $(1, -1) \in \mathbf{P}(\Bbbk^2)$, one says that the x_i's are harmonic conjugated if

$$(x_1, x_2, x_3, x_4) = -1. \tag{E.9}$$

E.2 The ring case

Now let R be a (not necessarily commutative) ring with unity and $n \in \mathbb{N}$. Let also e_1 be the first canonical basis vector of the R-module R^n. A vector $x \in R^n$ is called an admissible vector iff there exists an invertible $n \times n$ matrix M over R such that $x = Me_1$. In other words, one can find $n-1$ vectors such that together with x they form a free basis of R^n. Then a projective point in R^n is any set of the form xR, with x an admissible vector. In this thesis, we call the set of all projective points the projective net derived form R^n, which is denoted $\mathbf{P}(R^n)$.

A vector $x = (x^0, x^1, \ldots, x^{n-1}) \in R^n$ is said unimodular if there exist $a_0, \ldots, a_{n-1} \in R$ such that

$$a_0 x^0 + \cdots + a_{n-1} x^{n-1} = 1. \tag{E.10}$$

If R is commutative, then admissibility is equivalent to unimodularity. If R is a local ring, that is to say unitary and commutative with only one maximal ideal, then the admissiblity of x is also equivalent to one of the coordinates x^i being a unit. Indeed, the sole maximal ideal of a local ring is the set of the noninvertible elements of that ring.

One the one hand, $\mathbb{Z}_d$, with $d = p^s$ a power of a prime, is a local ring. This is the simplest case one can encounter in projective geometry over a ring. On the other hand, though $\mathrm{Mat}(n, \mathbb{Z}_d)$ is not even commutative, admissibility is still equivalent to unimodularity, as we see in Section 5.2. Besides, for $k \in \mathbb{N} \setminus \{0, 1\}$, $\mathrm{Mat}(n, \mathbb{Z}_d)^k$ contains admissible vectors no coordinates of which are invertible.

If an arbitrary commutative ring R is given, two admissible vectors (a, b) and (c, d) in R^2 are said to be distant if the matrix

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \tag{E.11}$$

is invertible. In fact, this condition is equivalent to unimodularity of any of the two vectors: A vector (a, b) is admissible iff there exists another vector (c, d) which is distant from it; in that case, both (a, b) and (c, d) are admissible. Two vectors that are not distant are said to be neighbour. In Section 3.2 of the thesis, we take up these definitions of distance and neighbourhood. We generalise and extend them in a quantitative way with the help of the wedge product.

We close this section by connecting projective geometry over $\mathbb{Z}_d$ with Theorem 40 on elementary divisors. The following proposition is quite simple, which may explain that it is not found in literature. As a consequence, we coined the name of a *saturated* submodule. This latter notion and also that of an outliner the proposition refers to are not used elsewhere in the thesis.

Proposition 47 *Let M be a $\mathbb{Z}_d$-submodule of $\mathbb{Z}_d^m$. Then M is the join of the projective points it contains iff for any prime factor p of d and any elementary divisor x of M,*

$$v_p(x) = 0 \text{ or } v_p(d). \tag{E.12}$$

In that case, we will say that M is saturated. Otherwise, the points that are not parts of projective points are called the outliners of M.

Proof. Let us assume without loss of generality that $d = p^s$ is a power of a prime. According to Theorem 40, there exists a computational basis f of $\mathbb{Z}_d^m$ such that M has the following diagonal basis matrix with respect to f:

$$\operatorname{diag}(p^{s_1}, \ldots, p^{s_m}), \tag{E.13}$$

with

$$s_1 \leq s_2 \leq \cdots \leq s_m. \tag{E.14}$$

If there is an $i \in \{1, \ldots, m\}$ such that $0 < s_i < v_p(d)$, then $p^{s_i} f_i \in M$ but is colinear to no free vector in M. On the contrary, if Condition (E.12) holds, let x be a vector of M:

$$x = \sum_{i=1}^{m} p^{t_i} f_i \quad \text{with } t_i \geq s_i. \tag{E.15}$$

If we put

$$k = \max(i;\ s_i = 0), \tag{E.16a}$$

$$t = \min(t_i;\ i \in \{1, \ldots, m\}), \tag{E.16b}$$

then

$$\widetilde{x} = \sum_{i=1}^{k} p^{t_i - t} f_i \tag{E.17}$$

is a free vector in M and $x = p^t \widetilde{x} \in \langle \widetilde{x} \rangle$. ■

The reader with further interest in projective geometry over rings may consult [70] and also [71, 72].

Bibliography

[1] A. Aspect, P. Grangier, and G. Roger. Experimental test of realistic local theories via Bell's theorem. *Phys. Rev. Lett.*, 47:460, 1982.

[2] C. H. Bennett and G. Brassard. In *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing*, pages 175–179. Bangalore, India (IEEE, New York), 1984.

[3] D. Deutsch. Quantum theory, the Church-Turing principle and the universal quantum computer. *Proc. R. Soc. Lond. A*, 400:97, 1985.

[4] P. W. Shor. Algorithms for quantum computation: discrete logarithm and factoring. In S. Goldwasser, editor, *Proceedings of the 35th. Annual Symposium on the Fonudations of Computers Science*, page 124. IEEE Computer Society Press, Los Alamitos, CA, 1994.

[5] L. K. Grover. A fast quantum mechanical algorithm for database search. In *Proc. of the 28th Annual ACM Symposium on the Theory of Computing*, page 212. ACM Press, New York, 1996. also available at quant-ph/9605043.

[6] J. Schwinger. Unitary operators bases. In *Proceedings of the National Academy of Sciences*, volume 46, pages 570–579, 1960.

[7] W. K. Wootters and B. D. Fields. Optimal state-determination by mutually unbiased measurements. *Annals of Physics*, 191:363–381, 1989.

[8] L. Vaidman, Y. Aharonov, and D. Z. Albert. How to ascertain values of σ_x, σ_y, and σ_z of a spin-$\frac{1}{2}$ particle. *Phys. Rev. Lett.*, 58(14):1385–1387, 1987.

[9] A. Hayashi, M. Horibe, and T. Hashimoto. Mean king's problem with mutually unbiased bases and orthogonal latin squares. *Phys. Rev. A*, 71:052331, 2005.

[10] J. P. Paz, A. J. Roncaglia, and M. Saraceno. Qubits in phase space: Wigner function approach to quantum error correction and the mean king problem. *Phys. Rev. A*, 72:012309, 2005.

[11] A. R. Calderbank, P. J. Cameron, W. M. Kantor, and J. J. Seidel. $\mathbb{Z}_4$-Kerdock codes, orthogonal spreads, and extremal euclidean line-sets. In *Proceedings of the London Mathematical Society*, volume 75 n.3, pages 436–480, 1997.

[12] J. M. Renes. *Frames, designs, and spherical codes in quantum information theory.* PhD thesis, University of New Mexico, Albuquerque, New Mexico, 2004. Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy Physics, http://kerr.physik.uni-erlangen.de/ renes/publications/diss2side.pdf.

[13] J. M. Renes. Optimal protocols and tradeoffs in quantum key distribution. In *AIP Conference Proceedings*, 2004.

[14] R. W. Heath, T. Strohmer, and A. J. Paulraj. On quasi-orthogonal signatures for CDMA systems. *IEEE Transactions on Information Theory*, 52(3):1217–1226, 2006.

[15] I. D. Ivanović. Geometrical description of quantal state determination. *J. Phys. A: Math. Gen.*, 14:3241–3245, 1981.

[16] A. Klappenecker and M. Rötteler. Constructions of mutually unbiased bases. In *Proceedings International Conference on Finite Fields and Applications*, volume 2948, pages 137–144. LNCS, Springer, 2004.

[17] C. Archer. There is no generalization of known formulas for mutually unbiased bases. *J. Math. Phys.*, 46:022106, 2005.

[18] S. Bandyopadhyay, P. O. Boykin, V. Roychowdhury, and F. Vatan. A new proof for the existence of mutually unbiased bases. *Algorithmica*, 34:512–528, 2002.

[19] R. Balian and C. Itzykson. Observations sur la mécanique quantique finie. *C. R. Acad. Sci. Paris*, 303:Série I, n.16, 773–777, 1986.

[20] M. Havlíček, J. Patera, E. Pelantová, and J. Tolar. Automorphisms of the fine grading of $sl(n, C)$ associated with the generalised Pauli matrices. *J. Math. Phys.*, 43:1083–1094, 2002.

[21] P. Sulč and J. Tolar. Group Theoretical Construction of Mutually Unbiased Bases in Hilbert Spaces of Prime Dimension.

[22] W. K. Wootters. Quantum measurements and finite geometry. *Foundations of Physics*, 36:112, 2006.

[23] A. Vourdas. Galois quantum systems. *J. Phys. A: Math. Gen.*, 38:8453–8471, 2005.

[24] E. F. Galvão. Discrete Wigner functions and quantum computational speedup. *Phys. Rev. A*, 71:042302, 2005.

[25] A. O. Pittenger and M. H. Rubin. Wigner functions and separability for finite systems. *J. Phys. A: Math. Gen.*, 38:6005–6036, 2005.

[26] A. B. Klimov, L. L. Sánchez-Soto, and H. de Guise. A complementary-based approach to phase in finite-dimensional quantum systems. *Journal of Optics B*, 7:283–287, 2005.

[27] A. B. Klimov, L. L. Sánchez-Soto, and H. de Guise. Multicomplementary operators via finite Fourier transform. *J. Phys. A: Math. Gen.*, 38:2747–2760, 2005.

[28] M. Saniga, M. Planat, and H. Rosu. Mutually unbiased bases and finite projective planes. *Journal of Optics B*, 6:L19–20, 2004.

[29] M. Saniga and M. Planat. Sets of Mutually Unbiased Bases as Arcs in Finite Projective Planes? *Chaos, Solitons and Fractals*, 26:1267 – 1270, 2005.

[30] M. Planat, H. Rosu, S. Perrine, and M. Saniga. A Survey of Finite Algebraic Geometrical Structures Underlying Mutually Unbiased Quantum Measurements. *Foundations of Physics*, 36:1662–1680, 2006.

[31] M. Kibler. Angular momentum and mutually unbiased bases. *Int. J. Mod. Phys. B*, 20(11-13):1792–1801, 2006.

[32] M. Kibler and M. Planat. A $SU(2)$ recipe for mutually unbiased bases. *Int. J. Mod. Phys. B*, 20(11-13):1802–1807, 2006.

[33] O. Albouy and M. Kibler. SU_2 Nonstandard Bases: Case of Mutually Unbiased Bases. *SIGMA*, 3:076, 2007.

[34] O. Albouy and M. Kibler. A unified approach to SIC-POVMs and MUBs. *Journal of Russian Laser Research*, 28(5), 2007.

[35] M. Grassl. On SIC-POVMs and MUBs in Dimension 6. In *Proc. ÉRATO Conference on Quantum Information Science (EQUIS) J. Gruska (ed)*, 2004.

[36] S. Brierley and S. Weigert. Constructing Mutually Unbiased Bases in Dimension 6. 2009. available at arXiv:quant-ph/0901.4051.

[37] M. Planat and M. Saniga. On the Pauli graphs of N-qudits. *Quantum Information and Computation*, 8(1-2):127–146, 2008.

[38] M. Planat, M. Saniga, and Kibler M. Quantum Entanglement and Projective Ring Geometry. *SIGMA*, 2:066, 2006.

[39] I. Kaplansky. Elementary divisors and modules. *Transactions of the American Mathematical Society*, 66:464–491, 1949.

[40] W. C. Brown. *Matrices over commutative rings.* Pure and applied mathematics. Marcel Dekker, Inc., 1993.

[41] C. Cohen-Tannoudji, B. Diu, and F. Laloë. *Mécanique quantique.* Hermann, Paris, 1973.

[42] M. A. Nielsen and I. L. Chuang. *Quantum computation and Quantum Information.* Cambridge University Press, 2000.

[43] G. Benenti, G. Casati, and G. Strini. *Principles of Quantum computation and Information.* World Scientific, 2004.

[44] M. Le Bellac. *Physique quantique.* EDP Sciences, CNRS ÉDITIONS, 2007.

[45] J. Schwinger. *Quantum Theory of Angular Momentum*, chapter On angular momentum, pages 229–279. Editors L.C. Biedenharn and H. van Dam, Academic Press, New York, 1965.

[46] B. C. Berndt, R. J. Evans, and K. S. Williams. *Gauss and Jacobi sums.* Wiley, New York, 1998.

[47] M. Planat and M. Saniga. Pauli graph and finite projective lines/geometries. *Photon Counting Applications, Quantum Optics and Quantum Cryptography - Optics and Optoelectronics, Prague*, 2007.

[48] M. Planat, A.-C. Babouin, and M. Saniga. Multi-Line Geometry of Qubit-Qutrit and Higher-Order Pauli Operators. *Int. J. Theor. Phys.*, 47:1127–1135, 2008.

[49] M. Planat and A.-C. Babouin. Qudits of composite dimension, mutually unbiased bases and projective ring geometry. *J. Phys. A: Math. Theor.*, 40:F1005, 2007.

[50] M. Saniga, M. Planat, M. Kibler, and P. Pracna. A Classification of the Projective Lines over Small Rings. *Chaos, Solitons and Fractals*, 33(4):1095 – 1102, 2007.

[51] M. Saniga, M. Planat, M. Kibler, and P. Pracna. A Classification of the Projective Lines over Small Rings II. Non-Commutative Case. 2006. available at arXiv:math/0606500.

[52] W. v. D. Hodge and D. Pedoe. *Methods of algebraic geometry*, volume 1, chapter 7. 1947, reissued in 1994.

[53] Ahlbrecht A., L. S. Georgiev, and R. F. Werner. Implementation of Clifford gates in the Ising-anyon topological quantum computer. *Phys. Rev. A*, 79:032311, 2009.

[54] E. P. Wigner. On the Quantum Correction For Thermodynamic Equilibrium. *Physical Review*, 40:749–759, 1932.

[55] S. de Groot. *La transformation de Weyl et la fonction de Wigner: une forme alternative de la mécanique quantique*. Les Presses Universitaires de Montréal, 1975.

[56] F. A. Buot. Method for calculating TrH^n in solid-state theory. *Physical Review B*, 10:3700–3705, 1974.

[57] J. H. Hannay and M. V. Berry. On the Quantum Correction For Thermodynamic Equilibrium. *Physica D*, 1:267, 1980.

[58] K. S. Gibbons, F. J. Hoffman, and W. K. Wootters. Discrete phase space based on finite fields. *Physical Review A*, 70:062101, 2004.

[59] S. Chaturvedi, E. Ercolessi, G. Marmo, G. Morandi, N. Mukunda, and R. Simon. Wigner-Weyl correspondence in quantum mechanics for continuous and discrete systems - a Dirac-inspired view. *J. Phys. A: Math. Gen.*, 39:1405–1423, 2006.

[60] H. Havlíček and M. Saniga. Projective Ring Line of a Specific Qudit. *J. Phys. A: Math. Theor.*, 40:F943–F952, 2007.

[61] H. Havlíček and M. Saniga. Projective Ring Line of an Arbitrary Single Qudit. *J. Phys. A: Math. Theor.*, 41:015302, 2008.

[62] J.-L. Brylinski and R. Brylinski. Invariant polynomial functions on k qubits. In R. K. Brylinski and G. Chen, editors, *Mathematics of Quantum Computation*, 2002.

[63] H. Heydari. Entanglement monotones for multi-qubit states based on geometric invariant theory. *J. Math. Phys.*, 47:012103, 2006.

[64] L. Wang, H. Al Hadhrami, and A. Vourdas. Symplectic transformations and entanglement in multipartite finite systems. *Eur. Phys. J. D*, 49:265–272, 2008.

[65] R. Mosseri and R. Dandoloff. Geometry of entangled states, Bloch spheres and Hopf fibrations. *J. Phys. A*, 34:10243, 2001.

[66] J. I. Cirac. Entanglement and Distillability. In R. Arvieu and S. Weigert, editors, *Physics of Entangled States*, 2001.

[67] M. Kibler. Variations on a theme of Heisenberg, Pauli and Weyl. *J. Phys. A: Math. Theor.*, 41:375302, 2008.

[68] R. Howe. Nice error bases, mutually unbiased bases, induced representations, the Heisenberg group and finite geometries. *Indag. Mathem., N.S.*, 16(3-4):553–583, 2005.

[69] M. D. Larsen, W. J. Lewis, and T. S. Shores. Elementary divisor rings and finitely presented modules. *Transactions of the American Mathematical Society*, 187(1):231–248, 1974.

[70] A. Blunk and H. Havlíček. Projective Representations I. *Abh. Math. Sem. Univ. Hamburg*, 70:287–299, 2000.

[71] F. D. Veldkamp. Geometry over rings. In F. Buekenhout, editor, *Handbook of Incidence Geometry*, pages 1033–1084. Elsevier, Amsterdam, 1995.

[72] U. Brehm, M. Greferath, and S. E. Schmidt. Projective geometry over modular lattices. In F. Buekenhout, editor, *Handbook of Incidence Geometry*, pages 1115–1142. Elsevier, Amsterdam, 1995.