# BACHELOR'S THESIS

ALGEBRAS OF OBSERVABLES AND QUANTUM COMPUTING

Vojtěch Teska

July 4, 2016

*Název práce:* **Algebry pozorovatelných a kvantové počítání**

*Autor:* Vojtěch Teska

*Obor:* Matematické inženýrství

*Zaměření:* Matematická fyzika

*Druh práce:* Bakalářská práce

*Vedoucí práce:* Prof. Ing. Jiří Tolar, DrSc., Katedra fyziky, Fakulta jaderná a fyzikálně inženýrská, České vysoké učení technické v Praze

*Konzultanti:* Ing. Petr Novotný, PhD, Katedra fyziky, Fakulta jaderná a fyzikálně inženýrská, České vysoké učení technické v Praze
Mgr. Miroslav Korbelář, PhD, Katedra matematiky, Fakulta elektrotechnická, České vysoké učení technické v Praze

*Abstrakt:* Nejprve jsou uvedeny základní pojmy a postuláty kvantové teorie (stavy, pozorovatelné, pravděpodobnosti přechodu) a blíže prozkoumány stavy a pozorovatelné na konečněrozměrných Hilbertových prostorech. Podrobně je studována asociativní C*-algebra $M_n(\mathbb{C})$ s Hilbert-Schmidtovým skalárním součinem. Jsou zkoumány jemné gradace této algebry získané pomocí maximálních grup jejích komutujících *-automorfismů (MAD-grup). Je ukázáno, že MAD-grupy korespondují s unitárními Ad-grupami, je podána jejich klasifikace a jsou uvedeny přísušné jemné gradace v jednoduchých případech. Na závěr je ilustrována kvantová komplementarita pomocí prvků Pauliho grupy.

*Klíčová slova:* kvantová mechanika, C*-algebra, jemná gradace, MAD-grupa, Pauliho grupa, kvantová komplementarita

*Title:* **Algebras of Observables and Quantum Computing**

*Author:* Vojtěch Teska

*Abstract:* A basic overview of quantum theory terms and axioms (states, observables, probabilities of transition) is presented, then states and observables in finite dimensional Hilbert spaces are examined in more detail. Associative C*-algebra $M_n(\mathbb{C})$ with Hilbert-Schmidt inner product is closely inspected. Fine gradings of this algebra obtained with maximal groups of its commuting *-automorphisms (MAD-groups) are studied. A correspondence between MAD-groups and unitary Ad-groups is shown, a classification of unitary Ad-groups is given and corresponding fine gradings in simple cases are presented. Lastly, quantum complementarity is illustrated using elements of Pauli's group.

*Key words:* quantum mechanics, C*-algebra, fine grading, MAD-group, Pauli's group, quantum complementarity

**Prohlášení**

Prohlašuji, že jsem svou bakalářskou práci vypracoval samostatně a použil jsem pouze podklady uvedené v přiloženém seznamu.

Nemám závažný důvod proti použití tohoto školního díla ve smyslu § 60 zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

V Praze dne:

Podpis:

# Contents

# Acknowledgments

I would like to thank prof. Jiří Tolar for advice, consultations and corrections in the text. I also thank my parents for support during my studies.

# Introduction

The aim of this thesis is to describe the fundamental notions and necessary mathematical structures used to build the quantum theory with focus on operators in finite dimension. It shall be shown that these are represented by the algebra $M_n(\mathbb{C})$ of complex matrices with standard matrix multiplication.

The operation of hermitian adjoining in $M_n(\mathbb{C})$ will be defined and it will be shown that it satisfies the definition of involution given in the first chapter. We shall define Hilbert-Schmidt inner product in $M_n(\mathbb{C})$ as an analogue of standard inner product in $\mathbb{C}^n$ and show that the norm defined by this inner product is not compatible with the definition of a C*-algebra and that it is necessary to equip $M_n(\mathbb{C})$ with operator norm in order to satisfy all the conditions of this definition.

After doing so, we proceed to a closer inspection of $M_n(\mathbb{C})$, examining commutativity on its irreducible subsets. We prove that a matrix commuting with all elements of an irreducible set must be a scalar multiple of identity matrix. Furthermore, we discover that a set of matrices from $M_n(\mathbb{C})$ is simultaneously diagonalizable if and only if it is a set of mutually commuting diagonalizable matrices. We conclude the second chapter with proving that all normal matrices are diagonalizable and that all automorphisms of $M_n(\mathbb{C})$ are inner.

In the third chapter, a grading of an arbitrary *-algebra is defined and the relationship between maximal groups of commuting *-automorphisms (MAD-groups) and gradings is examined. In the case of $M_n(\mathbb{C})$, we show that there is a correspondence between unitary Ad-groups and MAD-groups. A classification theorem decomposing unitary Ad-groups into tensor products of Pauli's groups $\mathscr{P}_k$ and diagonal unitary groups $\mathscr{U}_D(k)$ is proven and respective gradings of $M_n(\mathbb{C})$ are studied in simple cases.

The thesis is concluded with a brief overview of quantum computing and an illustration of quantum complementarity using elements of Pauli's group.

# Notation

| | |
|---|---|
| $\widehat{n}$ | the set $\{1, 2, 3, ..., n\}$ |
| $(\bullet, \bullet)$ | inner product |
| $e$ | multiplicative identity in algebra $\mathbf{A}$ |
| $\theta$ | zero vector |
| $\bullet^*$ | involution |
| $\sigma(A)$ | spectrum of a linear operator $A$ |
| ${}^{\mathscr{B}}A$ | matrix of linear operator $A$ in the basis $\mathscr{B}$ |
| $\bullet^H$ | hermitian adjoining |
| $I$ | identity matrix |
| $\langle \bullet, \bullet \rangle$ | Hilbert-Schmidt inner product |
| $\Vert \bullet \Vert_{\mathcal{E}}$ | Euclidean norm in $\mathbb{C}^n$ |
| $\Vert \bullet \Vert_{op}$ | operator norm |
| $\mathbb{C}^{p,q}$ | vector space of complex $p \times q$ matrices |
| $O$ | zero matrix |
| $\mathrm{Eig}(A, \lambda)$ | eigenspace of a linear operator $A$ corresponding to eigenvalue $\lambda$ |
| $\mathbf{W} \subset\subset \mathbf{V}$ | $\mathbf{W}$ is a subspace of the vector space $\mathbf{V}$ |
| $Ad_A$ | inner automorphism of $M_n(\mathbb{C})$ generated by an invertible matrix $A$ |
| $\oplus$ | direct sum |
| $U(n)$ | the group of $n \times n$ unitary matrices |
| $\mathscr{P}_k$ | $k \times k$ Pauli's group |
| $\otimes$ | tensor product |
| $G_{Ad}$ | the set of unitary matrices generating MAD-group $G$ |
| $Ad(\mathscr{G})$ | the set of *-automorphisms generated by unitary Ad-group $\mathscr{G}$ |
| $\mathscr{U}_D(n)$ | the group of diagonal unitary $n \times n$ matrices |
| $\{\bullet, \bullet\}^{(\omega_n)}$ | $\omega_n$-commutant |
| $\{\bullet, \bullet\}'$ | commutant |
| $*$ | non-zero element in a matrix |

# Chapter 1

# Algebras of observables in quantum mechanics

## 1.1 Fundamental notions of quantum theory

All physical systems are described in terms of states and observables. In classical mechanics, both are represented by functions on a given system's phase space which, in the case of Hamiltonian mechanics, is a symplectic manifold. The fundamental role in description of quantum systems is played by Hilbert spaces and their one-dimensional subspaces [1]:

**Definition 1** (Hilbert space)**.** *Hilbert space is a vector space with inner product* $(\bullet, \bullet)$ *which is complete with respect to the metric generated by its inner product.*

**Definition 2** (Ray)**.** *Let* $\mathbf{H}$ *be a complex Hilbert space. A ray in* $\mathbf{H}$ *is any one-dimensional subspace in* $\mathbf{H}$.

To avoid pathological properties in mathematical description, it is assumed that these Hilbert spaces are complex and contain a countable dense subset, and therefore are by definition separable [1]. Thus the first fundamental axiom of quantum mechanics is postulated:

**Axiom 1.** *To each quantum system* $S$ *belongs a separable complex Hilbert space* $\mathbf{H}$ *which shall be called the state space belonging to* $S$.

A proper definition of state in quantum mechanics is more complicated than in Hamiltonian mechanics because of the probabilistic character of microscopic processes, which are greatly affected by measurement, and must therefore be executed in large quantities, resulting in the impossibility to distinguish which states were originally manipulated in the experiment. Assuming that we can perform just one experiment and determine the system's state, we can proclaim it a pure state of the examined system and postulate the second fundamental axiom of quantum mechanics [1]:

**Axiom 2.** *Each pure state of a given system* $S$ *is represented by some ray* $\Phi$ *in its state space* $\mathbf{H}$ *where the probability of transition between two states* $\Phi$ *and* $\Psi$ *shall be denoted* $P(\Phi, \Psi)$ *and is given by:*

$$P(\Phi, \Psi) = \frac{|(\varphi, \psi)|^2}{(\|\varphi\|\|\psi\|)^2} \tag{1.1}$$

*where* $\varphi \in \Phi$ *and* $\psi \in \Psi$.

Since each ray $\Psi$ in $\mathbf{H}$ is generated by a unit vector $\psi : \Psi = \{\alpha\psi : \alpha \in \mathbb{C}, \|\psi\| = 1\}$, it is easy to see that $P(\Phi, \Psi)$ does not depend on the choice of $\varphi$ and $\psi$, so both vectors may be chosen with unit norm. This fact simplifies the equation (1.1) to $P(\varphi, \psi) = |(\varphi, \psi)|^2$. In addition, it implies that we can represent pure states by unit vectors generating their corresponding rays.

To complete the mathematical description of a quantum system, a definition of observables is needed. Unlike in classical physics, observables in quantum mechanics are described by different mathematical objects than states, namely linear operators on $\mathbf{H}$ with some additional properties [1]. It is possible to show that for each linear operator $\hat{T}$ on $\mathbf{H}$ such that its domain $D_{\hat{T}}$ is dense in $\mathbf{H}$ and for each $y \in \mathbf{H}$ there exists at most one $y^* \in \mathbf{H}$ such that $(y, Tx) = (y^*, x)$ for each $x \in D_{\hat{T}}$. If such $y^*$ exists, the Hermitian adjoint $\hat{T}^*$ of $\hat{T}$ can be defined in the following way:

$$D_{\hat{T}^*} = \{y \in \mathbf{H} : (\exists y^* \in \mathbf{H})((y, Tx) = (y^*, x))(\forall x \in D_{\hat{T}})\} \quad \hat{T}^* y = y^* \quad \text{for all } \ y \in D_{\hat{T}^*}.$$

For the purposes of quantum mechanics, the following definition is useful:

**Definition 3.** *If $D_{\hat{T}}$ is dense in $\mathbf{H}$ and $\hat{T}^* = \hat{T}$, then $\hat{T}$ is called self-adjoint.*

In order to be able to predict outcomes of measurements, it is needed to somehow obtain numbers from operators which map a subset of a Hilbert space $\mathbf{H}$ into $\mathbf{H}$. For that purpose we define the spectrum of an operator [1]:

**Definition 4.** *The spectrum $\sigma(\hat{T})$ of a linear operator $\hat{T}$ on a Hilbert space $\mathbf{H}$ is the set of $\lambda \in \mathbb{C}$ such that $(\hat{T} - \lambda\hat{I})$ is not a bijection.*

At this point, we can postulate the final fundamental axiom of quantum mechanics [1]:

**Axiom 3.** *Each observable $A$ of a given system $S$ is represented by a self-adjoint linear operator $\hat{A}$ on its state space $\mathbf{H}$, and possible outcomes of measurements of any given observable $A$ are elements of the spectrum $\sigma(\hat{A})$ of its corresponding operator $\hat{A}$.*

It follows from the theory of linear operators on Hilbert spaces, that since $\hat{A}$ is self-adjoint, all possible outcomes of measurements are real numbers, also it is clear from Definition 3 that the domain of any given observable must be dense in the state space $\mathbf{H}$.

## 1.2 Normed algebras

The operations performed with linear operators, which are acting as observables in quantum mechanics, can be generalized into interesting algebraic structures, several of which will be introduced in this section. However, to start, it is needed to clarify what we mean by operation. Beginning with an arbitrary set $\mathcal{M}$, we define binary operation in $\mathcal{M}$ as a map $\phi : \mathcal{M} \times \mathcal{M} \to \mathcal{M}$. We say that $\phi$ is associative if $\phi(\phi(a, b)c) = \phi(a, \phi(b, c))$ for all $a, b, c \in \mathcal{M}$ and we say that $\phi$ is commutative if $\phi(a, b) = \phi(b, a)$ for all $a, b \in \mathcal{M}$. Let us consider a set $\mathcal{R}$ equipped with two binary operations $\phi_a, \phi_b$ which shall be called addition and multiplication and let us denote $\phi_a(a, b) = a + b$, $\phi_b(a, b) = ab$. The triplet $(\mathcal{R}, \phi_a, \phi_b)$ shall be called a ring if the following conditions are satisfied [1]:

1. $(\mathcal{R}, \phi_a)$ is a commutative group
2. $a(b + c) = ab + ac$ for all $a, b, c \in \mathcal{R}$

3. $(a + b)c = ac + bc$ for all $a, b, c \in \mathcal{R}$

If there exists $e \in \mathcal{R}$ such that $ae = ea = a$ for all $a \in \mathcal{R}$, then $e$ shall be called a multiplicative identity.

Now, let $\mathbf{A}$ be a complex vector space. If we introduce a new bilinear binary operation in $\mathbf{A}$, called multiplication, then $\mathbf{A}$ becomes a ring which shall be called a complex linear algebra. We say that $\mathbf{A}$ is associative resp. commutative if its multiplication is associative resp. commutative. From this point forward, the term complex linear algebra will be abbreviated to algebra.

To generalize the process of inverting operators, we must realize that in a non-specific algebra, the existence of a multiplicative identity is not guaranteed, therefore it must be demanded in the definition of an inverse element.

**Definition 5.** *Let $\mathbf{A}$ an algebra with multiplicative identity and let $a \in \mathbf{A}$. Then $a$ is called invertible if there exists $a^{-1} \in \mathbf{A}$, called the inverse element of $a$, such that $a^{-1}a = aa^{-1} = e$.*

By generalizing the operation of hermitian adjoining as a map on the state space to algebras, the following definition is obtained [1], [2]:

**Definition 6.** *Let $\mathbf{A}$ be an algebra. Involution in $\mathbf{A}$ is a map $* : \mathbf{A} \to \mathbf{A}$ satisfying:*

1. $(\xi a + b) = \overline{\xi} a + b$ *for all $\xi \in \mathbb{C}$ and for all $a, b \in \mathbf{A}$*
2. $(a^*)^* = a$ *for all $a \in \mathbf{A}$*
3. $(ab)^* = b^* a^*$ *for all $a, b \in \mathbf{A}$*

*The element $a^*$ will be called the adjoint element of $a$ and the pair $(\mathbf{A}, *)$ will be called an involutive algebra or *-algebra.*

Note that $(a^*)^* = a$ implies that $*$ is its own inverse and therefore it must be a bijection. Having defined the adjoint element, we are able to proceed analogously as with operators, resulting in the following definition [1]:

**Definition 7.** *Let $\mathbf{A}$ be a *-algebra and let $a \in \mathbf{A}$. Then $a$ is called:*

1. *normal iff* $aa^* = a^* a$
2. *hermitian iff* $a^* = a$
3. *a projector iff* $a^* = a = a^2$.
4. *Furthermore if $\mathbf{A}$ is an algebra with multiplicative identity and $a^* = a^{-1}$, then $a$ is called unitary.*

As in the case of studying the relations between two given vector spaces, it is needed to define a map which preserves all algebraic operations (which in the case of *-algebras includes involution) [1].

**Definition 8.** *Let $\mathbf{A}$ and $\mathbf{B}$ be algebras. A map $\psi : \mathbf{A} \to \mathbf{B}$ is called a morphism iff*

1. $\psi(\xi a + b) = \xi \psi(a) + \psi(b)$ *for all $\xi \in \mathbb{C}$ and for all $a, b \in \mathbf{A}$*
2. $\psi(ab) = \psi(a)\psi(b)$ *for all $a, b \in \mathbf{A}$*

*If $\psi$ is a bijection, then it is called an isomorphism, furthermore if $\mathbf{A} = \mathbf{B}$, then $\psi$ is called an automorphism.*

**Remark 1.** *Let $\mathbf{A}, \mathbf{B}$ be algebras with multiplicative identity and let $\psi : \mathbf{A} \to \mathbf{B}$ be a morphism. Since $a = ae = ea$ for all $a \in \mathbf{V}$, we obtain that $\psi(a) = \psi(e)\psi(e) = \psi(e)\psi(a)$, thus $\psi(e) = e$. Similarly, $\psi(a) = \psi(a+\theta) = \psi(a)+\psi(\theta)$ for all $a \in \mathbf{V}$, proving that $\psi(\theta) = \theta$, where $\theta$ denotes the zero vector.*

**Definition 9.** *Let $\mathbf{A}$ and $\mathbf{B}$ be *-algebras. A morphism $\psi : \mathbf{A} \to \mathbf{B}$ is called a *-morphism iff $\psi(a^*) = (\psi(a))^*$ for all $a \in \mathbf{A}$. If $\psi$ is a bijection, then it is called a *-isomorphism, furthermore if $\mathbf{A} = \mathbf{B}$, then $\psi$ is called a *-automorphism.*

In a normed vector space, two well known relations, namely $\|a + b\| \leq \|a\| + \|b\|$, between the norm of a sum of two vectors and the norms of its summands, and $\|\theta\| = 0$ for the norm of the zero vector, are postulated. Similar relations are given in the definition of a normed algebra with regard to multiplication [1]:

**Definition 10.** *A normed algebra is an algebra $\mathbf{A}$ satisfying:*

1. *$\mathbf{A}$ is a normed vector space with the norm $\|\bullet\|$*
2. *$\|ab\| \leq \|a\|\|b\| \quad$ for all $a, b \in \mathbf{A}$*
3. *if $\mathbf{A}$ contains a multiplicative identity $e$, then $\|e\| = 1$.*

*If $\mathbf{A}$ is complete with respect to its norm, then it is called a Banach algebra.*

To finish, the relation between the norm and involution must be discussed, defining another two important types of algebras [1].

**Definition 11.** *A normed involutive algebra $\mathbf{A}$ is called a normed *-algebra iff $\|a^*\| = \|a\|$ for all $a \in \mathbf{A}$. A normed *-algebra complete with respect to its norm is called a Banach *-algebra.*

**Definition 12.** *A Banach *-algebra $\mathbf{A}$ shall be called a C*-algebra iff $\|a^*a\| = \|a\|^2$ for all $a \in \mathbf{A}$.*

Note that $\|a\|^2 = \|a^*a\|$ implies $\|a^*\| = \|a\|$. Since $\|a\|^2 = \|a^*a\| \leq \|a^*\|\|a\|$ i.e. $\|a\| \leq \|a^*\|$, analogously $\|a^*\| \leq \|a\|$ and therefore $\|a^*\| = \|a\|$.

**Remark 2.** *In quantum mechanics with a separable Hilbert space $\mathbf{H}$, the mathematical framework can be generalized in the form of C*-algebra postulate [3]:*
*A quantum system is characterized by a triplet $(S^*, \mathcal{A}, \langle \bullet, \bullet \rangle)$ where*

1. *$\mathcal{A}$, the set of its observables, is the collection of all the hermitian elements $h$ of a C*-algebra $\mathbf{A}$;*
2. *$S^*$, the set of its states, is the collection of all real-valued, positive linear functionals $\phi$ on $\mathcal{A}$, normalized by the condition $\langle \phi, h \rangle = 1$;*
3. *$\langle \bullet, \bullet \rangle$ is the prediction rule, a map $\langle \bullet, \bullet \rangle : S^* \times \mathcal{A} \to \mathbb{R}$ which attributes to every pair $(\phi, h)$, the value $\langle \phi, h \rangle = \phi(h)$, interpreted as the expectation of the observable $h$ when the system is in a state $\phi$.*

## 1.3 Quantum mechanics in finite dimension

Quantum systems with finite dimensional state space (such as when describing the spin of a particle) can be described by assigning vector space $\mathbb{C}^n$ of ordered n-tuples of complex numbers

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

with standard inner product $(x, y) = \sum_{i=1}^{n} \overline{x_i} y_i$.

Observable $A$, which is a hermitian linear operator on $\mathbb{C}^n$, is represented by its matrix ${}^{\mathscr{B}}A$ in a chosen basis $\mathscr{B} = (e_k)_{k=1}^n$ of $\mathbb{C}^n$, defined as $A_{ij} = e_i^{\#}(Ae_j)$, where $e_k^{\#}$ denotes vectors of the dual basis of $\mathscr{B}$. Using this definition, we can see that composing two observables $A, B$ is equivalent to matrix multiplication [4]:

$$
\begin{aligned}
{}^{\mathscr{B}}(AB)_{ij} &= e_i^{\#}(ABe_j) = e_i^{\#}(A(Be_j)) = e_i^{\#}(A(\sum_{k=1}^{n} e_k^{\#}(Be_j)e_k)) = e_i^{\#}(A(\sum_{k=1}^{n}({}^{\mathscr{B}}B)_{kj}e_k)) = \\
&= \sum_{k=1}^{n} {}^{\mathscr{B}}B_{kj} e_i^{\#}(Ae_k) = \sum_{k=1}^{n} {}^{\mathscr{B}}B_{kj} e_i^{\#}(\sum_{l=1}^{n} e_l^{\#}(Ae_k)e_l) = \sum_{k=1}^{n} {}^{\mathscr{B}}B_{kj} e_i^{\#}(\sum_{l=1}^{n} {}^{\mathscr{B}}A_{lk} e_l) = \\
&= \sum_{k,l=1}^{n} {}^{\mathscr{B}}A_{lk} {}^{\mathscr{B}}B_{kj} e_i^{\#}(e_l) = \sum_{k=1}^{n} {}^{\mathscr{B}}A_{ik} {}^{\mathscr{B}}B_{kj}.
\end{aligned}
$$

Furthermore, by Riesz's lemma [1], [4] for any continuous linear functional $\varphi$ on a Hilbert space $\mathbf{H}$ there exists a vector $x \in \mathbf{H}$ such that $\varphi$ can be written uniquely in the form $\varphi(y) = (x, y)$ for all $y \in \mathbf{H}$. Assuming orthonormality of $\mathscr{B}$, this statement allows to write the matrix ${}^{\mathscr{B}}A$ in the form

$$
{}^{\mathscr{B}}A = \begin{pmatrix}
(e_1, Ae_1) & (e_1, Ae_2) & ... & (e_1, Ae_n) \\
(e_2, Ae_1) & (e_2, Ae_2) & ... & (e_2, Ae_n) \\
\vdots & \vdots & \ddots & \vdots \\
(e_n, Ae_1) & (e_n, Ae_2) & ... & (e_n, Ae_n)
\end{pmatrix}.
$$

Lastly, matrix multiplication is associative [4], i.e. for any three $n \times n$ matrices $A, B, C$:

$$(A(BC))_{ij} = \sum_{k=1}^{n} A_{ik}(BC)_{kj} = \sum_{k,l=1}^{n} A_{ik}B_{kl}C_{lj} = \sum_{l=1}^{n}(AB)_{il}C_{lj} = (AB)C,$$

Thus the vector space containing all observables of a given finite dimensional system can be represented by the vector space $\mathbb{C}^{n,n}$ of $n \times n$ matrices which together with the operations of matrix multiplication forms an associative algebra, which shall be denoted $M_n(\mathbb{C})$. Properties of this algebra are studied in the following chapters.

# Chapter 2

# Associative algebras $M_n(\mathbb{C})$ of complex $n \times n$ matrices

## 2.1 Involution, inner product and norms

In this section we define involution on $M_n(\mathbb{C})$ in such a way so that it forms a C*-algebra and provide a definition of an inner product on $M_n(\mathbb{C})$. We begin by observing that the operation of Hermitian adjoining of a given matrix $A \mapsto A^H$, defined by $(A^H)_{ij} = \overline{A_{ji}}$ satisfies the definition of an involution:

1. $(\xi A + B)^H_{ij} = \overline{(\xi A + B)_{ji}} = \overline{\xi A_{ji}} + \overline{B_{ji}} = \overline{\xi}(A^H)_{ij} + (B^H)_{ij}$
2. $((A^H)^H)_{ij} = \overline{(A^H)_{ji}} = \overline{\overline{A_{ij}}} = A_{ij}$
3. $(AB)^H_{ij} = \overline{(AB)_{ji}} = \overline{\sum_{k=1}^n A_{jk}B_{ki}} = \sum_{k=1}^n \overline{A_{ik}B_{kj}} = \sum_{k=1}^n (B^H)_{jk}(A^H)_{ki} = (B^H A^H)_{ij}$

for all $i, j \in \widehat{n}$ and for all $\xi \in \mathbb{C}$, therefore the pair $(M_n(\mathbb{C}), \bullet^H)$ constitutes a *-algebra where multiplicative identity is the identity matrix $I$ (the zero vector is the zero matrix $O$, $O_{ij} = 0$ for all $i, j \in \widehat{n}$). Note that the Hermitian adjoining in general maps the vector space $\mathbb{C}^{p,q}$ of $p \times q$ matrices bijectively onto $\mathbb{C}^{q,p}$. In the following, the term involution will refer to the operation of Hermitian adjoining.

Note that if $A \in M_n(\mathbb{C})$ is invertible, then inversion commutes with involution:

$$AA^{-1} = A^{-1}A = I \Rightarrow (A^{-1})^H A^H = A^H (A^{-1})^H = I^H = I \Rightarrow (A^{-1})^H = (A^H)^{-1},$$

justifying the use of abbreviated notation $(A^{-1})^H = (A^H)^{-1} = A^{-H}$.

Using involution, it is possible to define an analogue of the standard inner product on $\mathbb{C}^n$. Define the trace of a matrix $A$ as the sum of its diagonal elements, i.e. $\mathrm{Tr}(A) = \sum_{i=1}^n A_{ii}$ for all $A \in M_n(\mathbb{C})$.

**Definition 13** (Hilbert-Schmidt inner product)**.** *Let $A, B \in M_n(\mathbb{C})$. The Hilbert-Schmidt inner product of A and B is defined as $\langle A, B\rangle = \mathrm{Tr}(A^H B)$.*

It is easy to see that $\langle A, B\rangle = \sum_{i,j=1}^n \overline{A}_{ij} B_{ij}$, hence $\langle \bullet, \bullet \rangle$ is linear in the second argument and that $\langle A, B\rangle = \overline{\langle B, A\rangle}$, moreover $\langle A, A\rangle = \sum_{i,j=1}^n |A_{ij}|^2 \geq 0$ and $\langle A, A\rangle = 0 \Leftrightarrow A = O$ and therefore it is a strictly positive sesquilinear form, inducing the norm $\|A\| = \sqrt{\langle A, A\rangle}$ for all

$A \in M_n(\mathbb{C})$. Furthermore, $M_n(\mathbb{C})$ is complete with respect to this norm, but $\|I\| = \sqrt{n}$ and so $M_n(\mathbb{C})$ paired with $\|\bullet\|$ cannot constitute a normed algebra. To satisfy all three conditions of a normed algebra (Definition 10), another norm is needed [1]:

**Definition 14** (Operator norm)**.** *Let $A \in M_n(\mathbb{C})$ and let $x \in \mathbb{C}^n$. Operator norm of $A$ is defined:*

$$\|A\|_{op} = \sup\{\|Ax\|_{\mathcal{E}} : \|x\|_{\mathcal{E}} = 1\},$$

*where $\|x\|_{\mathcal{E}}$ denotes the norm induced by the standard inner product in $\mathbb{C}^n$.*

Since all norms on finite dimensional vector spaces are equivalent [1], $M_n(\mathbb{C})$ is also complete with respect to the operator norm. Its properties are summarized in the following corollary [1], [5]:

**Corollary 1.** *Let $A, B \in M_n(\mathbb{C})$. Then:*

1. $\|AB\|_{op} \leq \|A\|_{op}\|B\|_{op}$
2. $\|A^H\|_{op} = \|A\|_{op}$
3. $\|A^H A\|_{op} = \|A\|_{op}^2$
4. $\|I\|_{op} = 1$

*Proof.* First, let $a, b \in \mathbb{C}^n = \mathbb{C}^{n,1}$ with standard inner product. Then $(a, b) = a^H b$ and so $(Aa, b) = (Aa)^H b = (a^H A^H)b = a^H(A^H b) = (a, A^H b)$, for $A^H$ we obtain $(a, Ab) = (A^H a, b)$.

1. Let $x \neq \theta$ then $\|ABx\|_{\mathcal{E}} = \frac{\|ABx\|_{\mathcal{E}}}{\|Bx\|_{\mathcal{E}}}\|Bx\|_{\mathcal{E}} \leq \|Bx\|_{\mathcal{E}} \sup\left\{\frac{\|ABx\|_{\mathcal{E}}}{\|Bx\|_{\mathcal{E}}} : Bx \in \mathbb{C}^n \setminus \{\theta\}\right\}$. Now substituting $z = Bx$, it follows that $\frac{\|Az\|_{\mathcal{E}}}{\|z\|_{\mathcal{E}}} = \|A(\frac{z}{\|z\|_{\mathcal{E}}})\|$ and therefore
$$\sup\left\{\frac{\|ABx\|_{\mathcal{E}}}{\|Bx\|_{\mathcal{E}}} : Bx \in \mathbb{C}^n \setminus \{\theta\}\right\} = \|A\|_{op}$$
This fact implies that $\|ABx\|_{\mathcal{E}} \leq \|A\|_{op}\|Bx\|_{\mathcal{E}}$, and the same is true for suprema, thus proving the first part of the corollary.

2. The Cauchy-Schwarz inequality $|(a, b)| \leq \|a\|\|b\|$ implies $\sup\{|(a, b)| : \|b\| = 1\} = \|a\|$. It follows that
$$\|A\|_{op}^2 = \sup\{(Ax, Ax) : \|x\| = 1\} = \sup\{(A^H Ax, x) : \|x\| = 1\} \leq$$
$$\leq \sup\{\|A^H Ax\| : \|x\| = 1\}$$
and therefore
$$\|A\|_{op}^2 \leq \|A^H A\|_{op} \leq \|A^H\|_{op}\|A\|_{op}. \tag{2.1}$$
By dividing both sides by $\|A\|_{op}$, we obtain $\|A\|_{op} \leq \|A^H\|_{op}$ and by doing the same for $A^H$, the inequality $\|A^H\|_{op} \leq \|A\|_{op}$ is obtained, giving the desired result.

3. Applying the previous result to inequality (2.1)
$$\|A\|_{op}^2 \leq \|A^H A\|_{op} \leq \|A^H\|_{op}\|A\|_{op} = \|A\|_{op}^2$$
proves the assertion.

4. By definition: $\|I\|_{op} = \sup\{\|Ix\|_{\mathcal{E}} : \|x\|_{\mathcal{E}} = 1\} = \sup\{\|x\|_{\mathcal{E}} : \|x\|_{\mathcal{E}} = 1\} = 1.$

□

Completeness of $M_n(\mathbb{C})$ with respect to the operator norm and points 1, 3, and 4 of Corollary 1 imply that $M_n(\mathbb{C})$ paired with operator norm forms a C*-algebra.

## 2.2 Irreducibility, diagonalizability and commutativity

In this section, the relationship between multiplicative commutativity and other properties of commuting matrices is investigated, beginning with commutation on irreducible sets [6].

**Definition 15.** *Let $\mathcal{U}$ be an arbitrary subset of $M_n(\mathbb{C})$. The set $\mathcal{U}$ is said to be reducible if fixed positive integers $p, q$ and a fixed invertible matrix $S$ exist such that for each $A \in \mathcal{U}$,*

$$S^{-1}AS = \begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} \end{pmatrix}$$

*where $A_{11} \in \mathbb{C}^{p,p}$ , $A_{12} \in \mathbb{C}^{p,q}$, $A_{22} \in \mathbb{C}^{q,q}$ and $O$ is zero $q \times p$ matrix. Otherwise $\mathcal{U}$ is said to be irreducible.*

**Lemma 1** (Schur's lemma)**.** *Let $\mathcal{U}$ be an irreducible subset of $M_n(\mathbb{C})$ and let $M \in M_n(\mathbb{C})$ be a fixed matrix such that for each $A \in \mathcal{U}$ there exists $\tilde{A} \in M_n(\mathbb{C})$ satisfying $AM = M\tilde{A}$. Then either $M = O$ or $M$ is invertible. Furthermore if $\tilde{A} = A$ (so that $M$ commutes with each element of $\mathcal{U}$), then there exists $\lambda \in \mathbb{C}$ such that $M = \lambda I$.*

*Proof.* Suppose that $\operatorname{rank}(M) = r < n$, and write

$$M = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$$

where $P, Q$ are invertible an $I_r$ is $r \times r$ identity matrix. Then for each $A \in \mathcal{U}$

$$AM = M\tilde{A} \Rightarrow (P^{-1}AP) \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} (Q^{-1}\tilde{A}Q) \tag{2.2}$$

Put

$$(P^{-1}AP) = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

$$(Q^{-1}\tilde{A}Q) = \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{pmatrix}$$

where $A_{11}, \tilde{A}_{11}$ are $r \times r$ matrices, $A_{12}, \tilde{A}_{12}$ are $r \times (n-r)$ matrices, $A_{21}, \tilde{A}_{21}$ are $(n-r) \times r$ matrices and $A_{22}, \tilde{A}_{22}$ are $(n-r) \times (n-r)$ matrices. Then (2.2) implies that

$$\begin{pmatrix} A_{11} & O \\ A_{21} & O \end{pmatrix} = \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ O & O \end{pmatrix}. \tag{2.3}$$

Thus $A_{21} = O$, which contradicts irreducibility of $\mathcal{U}$. Hence $r$ must be $0$ or $n$, and so either $M = O$ or $M$ is invertible. This proves the first part of the lemma.

Now suppose $AM = MA$ for each $A \in \mathcal{U}$, and choose $\lambda$ as any eigenvalue of $M$ (which must exist due to the fundamental theorem of algebra) and let $x$ be its corresponding eigenvector. Then

$$(M - \lambda I)x = Mx - \lambda x = \lambda x - \lambda x = \theta = 0x$$

hence $0 \in \sigma(M - \lambda I)$ i.e. $M - \lambda I$ is singular. In addition

$$A(M - \lambda I) = AM - A(\lambda I) = MA - (\lambda I)A = (M - \lambda I)A \Rightarrow M - \lambda I = O \Rightarrow M = \lambda I.$$

Thus proving the second part of the lemma. $\square$

An important example of an irreducible set is given in the following corollary:

**Corollary 2.** *$M_n(\mathbb{C})$ is an irreducible set.*

*Proof.* Let $S \in M_n(\mathbb{C})$ be an arbitrary invertible matrix, $p, q \in \hat{n}$ arbitrary numbers and $R \in \mathbb{C}^{q,p}, R \neq O$. Then the matrix $S \begin{pmatrix} O & O \\ R & O \end{pmatrix} S^{-1}$ satisfies

$$S^{-1}S \begin{pmatrix} O & O \\ R & O \end{pmatrix} S^{-1}S = \begin{pmatrix} O & O \\ R & O \end{pmatrix} \neq \begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} \end{pmatrix}$$

where $A_{11} \in \mathbb{C}^{p,p}$, $A_{12} \in \mathbb{C}^{p,q}$ and $A_{22} \in \mathbb{C}^{q,q}$. Thus $M_n(\mathbb{C})$ is an irreducible set. $\square$

The following theorem describes the set of $n \times n$ matrices commuting with all elements of $M_n(\mathbb{C})$, which will be needed in the subsequent chapters.

**Theorem 1.** *Let $M \in M_n(\mathbb{C})$. Then $M$ commutes with all elements of $M_n(\mathbb{C})$ if and only if $M = \lambda I$ for some $\lambda \in \mathbb{C}$.*

*Proof.* It is trivial that $M = \lambda I$ commutes with all elements of $M_n(\mathbb{C})$. To prove the converse we apply Lemma 1 on the set $M_n(\mathbb{C})$. $\square$

A simple way of describing a linear operator $A$ on a vector space $\mathbf{V}_n$ of finite dimension $n$ is by giving the image of vectors $e_1, e_2, e_3, ..., e_n$ composing a basis of $\mathbf{V}_n$. $A$ is called a diagonalizable operator iff there exists a basis $\mathscr{B} = (e_i)_{i=1}^n$ and $\lambda_i \in \mathbb{C}$ such that $Ae_i = \lambda_i e_i$ for all $i \in \widehat{n}$. The basis $\mathscr{B}$ is called a diagonal basis of $A$. The set of vectors in $\mathbf{V}_n$ satisfying $Ax = \lambda x$ for a given linear operator $A$ is called an eigenspace of $A$ corresponding to $\lambda \in \mathbb{C}$ [4]. It shall be denoted $\mathrm{Eig}(A, \lambda)$.

It follows that any matrix $X \in M_n(\mathbb{C})$ can be proclaimed a matrix of some linear operator on $\mathbb{C}^n$ in the same manner as in Chapter 1, so this definition can be carried over to $M_n(\mathbb{C})$, i.e. $X$ is diagonalizable iff there exists a diagonal basis of $\mathbb{C}^n$ for the operator defined by $X$. The following definitions and lemma are needed to study the relationship between commutativity and diagonalizability of operators.

**Definition 16.** *Let $\mathbf{V}_n$ be a vector space of finite dimension $n$ and let $\mathcal{M}$ be a set of linear operators on $\mathbf{V}_n$. Then $\mathcal{M}$ is called simultaneously diagonalizable iff there exists a basis $\mathscr{B}$ of $\mathbf{V}_n$ such that $\mathscr{B}$ is a diagonal basis of all $A \in \mathcal{M}$.*

**Definition 17.** *Let $A$ be a linear operator on a vector space $\mathbf{V}_n$ of finite dimension $n$ and let $\mathbf{W}$ be a subspace of $\mathbf{V}_n$. We say that $\mathbf{W}$ is $A$-invariant iff $A(\mathbf{W}) \subseteq \mathbf{W}$ [4].*

In the following, the relation $\mathbf{W}$ is a subspace of $\mathbf{V}$ will be denoted $\mathbf{W} \subset\subset \mathbf{V}$.

**Lemma 2.** *Let $\mathbf{V}_n$ be a vector space of finite dimension $n$ such that $\mathbf{V}_n = \bigoplus_{i=1}^{k} \mathbf{W}_k$, where $\mathbf{W}_i \subset\subset \mathbf{V}$ for all $i \in \widehat{k}$ and let $B$ be a diagonalizable linear operator on $\mathbf{V}_n$ such that $\mathbf{W}_i$ is $B$-invariant for all $i \in \hat{k}$. Then B is diagonalizable on all $\mathbf{W}_i$, where $i \in \widehat{k}$.*

*Proof.* First, we prove that the above statement is true for $k = 2$. We shall denote $\mathbf{W}_1 = \mathbf{U}$ and $\mathbf{W}_2 = \mathbf{W}$. Let $\mathbf{V}_n = \operatorname{span}((f_i)_{i=1}^n)$, where $(f_i)_{i=1}^n$ is the diagonal basis of $B$. Let $\lambda_i$ be eigenvalue of $B$ corresponding to $f_i$ for every $i \in \hat{n}$. Since it is possible to uniquely decompose each $f_i$ in the following form: $f_i = u_i + w_i$, where $u_i \in U$ and $w_i \in W$, we obtain

$$Bf_i = \lambda_i f_i = \lambda_i(u_i + w_i) = \lambda_i u_i + \lambda_i w_i$$

$$Bf_i = B(u_i + w_i) = Bu_i + Bw_i.$$

Since $\lambda_i u_i, Bu_i \in \mathbf{U}$ and $\lambda_i w_i, Bw_i \in \mathbf{W}$, we obtain $Bu_i = \lambda_i u_i$ and $Bw_i = \lambda_i w_i$.

We now prove that $\operatorname{span}((u_i)_{i=1}^n) = \mathbf{U}$ and $\operatorname{span}((w_i)_{i=1}^n) = \mathbf{W}$. Since $\operatorname{span}((u_i)_{i=1}^n) \subset\subset \mathbf{U}$ and $\operatorname{span}((w_i)_{i=1}^n) \subset\subset \mathbf{W}$,

$$\dim(\operatorname{span}((u_i)_{i=1}^n)) \leq \dim \mathbf{U} \qquad \dim(\operatorname{span}((w_i)_{i=1}^n)) \leq \dim \mathbf{W}$$

Assume $\dim(\operatorname{span}((u_i)_{i=1}^n)) + r = \dim \mathbf{U}$, where $r \in \widehat{\dim U}$. Now

$$\operatorname{span}((u_i)_{i=1}^n) \oplus \operatorname{span}((w_i)_{i=1}^n) = \mathbf{V}_n \Rightarrow \dim \mathbf{U} - r + \dim(\operatorname{span}((w_i)_{i=1}^n)) = n$$

which is equivalent to

$$\dim(\operatorname{span}((w_i)_{i=1}^n) = n - \dim \mathbf{U} + r = \dim \mathbf{W} + r > \dim \mathbf{W},$$

a contradiction.

Therefore by choosing $\dim \mathbf{U}$ linearly independent vectors from $(u_i)_{i=1}^n$ and $\dim \mathbf{W}$ linearly independent vectors from $(w_i)_{i=1}^n$ we obtain diagonal bases of $B$ in $\mathbf{U}$ and $\mathbf{W}$ respectively.

In the general case $\mathbf{V}_n = \bigoplus_{i=1}^{k} \mathbf{W}_k$ we apply this result $(n-1)$-times on vector spaces

$$\mathbf{V}_n = (\bigoplus_{i=1}^{k-1} \mathbf{W}_k) \oplus \mathbf{W}_k \quad \mathbf{V}^{(2)} = (\bigoplus_{i=1}^{k-2} \mathbf{W}_k) \oplus \mathbf{W}_{k-1} \quad \mathbf{V}^{(3)} = (\bigoplus_{i=1}^{k-3} \mathbf{W}_k) \oplus \mathbf{W}_{k-2} \quad \text{etc.}$$

thus proving the lemma. □

The following theorem describes how diagonalizability and commutativity are related. Also note that an analogous theorem holds true for matrices from $M_n(\mathbb{C})$.

**Theorem 2.** *Let $\mathbf{V}_n$ be a vector space of finite dimension $n$ and let $\mathcal{M}$ be an arbitrary set of linear operators on $\mathbf{V}_n$. Then $\mathcal{M}$ is a set of simultaneously diagonalizable operators if and only if $\mathcal{M}$ is a set of mutually commuting diagonalizable operators.*

*Proof.* Let $A, B \in \mathcal{M}$. If there exists a basis $\mathscr{D}$ in which both matrices ${}^{\mathscr{D}}A$ and ${}^{\mathscr{D}}B$ are diagonal and if $\mathscr{D} = (x_i)_{i=1}^n$ and if $\lambda_i^{(A)}$ and $\lambda_i^{(B)}$ are the eigenvalues of $A$ and $B$ corresponding to $x_i$ respectively, then we obtain:

$$ABx_i = A(\lambda_i^{(B)} x_i) = \lambda_i^{(B)}(Ax_i) = \lambda_i^{(B)}\lambda_i^{(A)} x_i = \lambda_i^{(A)}\lambda_i^{(B)} x_i = \lambda_i^{A} Bx_i = B(\lambda_i^{(A)} x_i) = BAx_i.$$

Therefore for any $y \in V_n, y = \sum_{i=1}^{n} \alpha_i x_i$, where $\alpha_i \in \mathbb{C}$ for all $i \in \hat{n}$:

$$ABy = AB\sum_{i=1}^{n} \alpha_i x_i = \sum_{i=1}^{n} \alpha_i ABx_i = \sum_{i=1}^{n} \alpha_i BAx_i = BA\sum_{i=1}^{n} \alpha_i x_i = BAy$$

We have proved that simultaneously diagonalizable operators commute.

To prove the converse, let $x_{i,j}^{(B)}$ denote $i$-th eigenvector corresponding to $j$-th eigenvalue of $B$, which we shall denote $\lambda_j^{(B)}$, and let $\eta_j^{(B)}$ be its geometrical multiplicity. Then

$$BAx_{i,j}^{(B)} = ABx_{i,j}^{(B)} = A(\lambda_j^{(B)} x_{i,j}^{(B)}) = \lambda_j^{(B)} Ax_{i,j}^{(B)}$$

Therefore $Ax_{i,j}^{(B)} \in \mathrm{Eig}(B, \lambda_j^{(B)})$. Furthermore let for each $\lambda_j^{(B)} \in \sigma(B)$ be $y_j \in \mathrm{Eig}(B, \lambda_j^{(B)})$, so $y_j = \sum_{k=1}^{\eta_j^{(B)}} \beta_k x_{k,j}^{(B)}$, then

$$Ay_j = \sum_{k=1}^{\eta_j^{(B)}} \beta_k x_{k,j}^{(B)} = y_j = \sum_{k=1}^{\eta_j^{(B)}} \beta_k Ax_{k,j}^{(B)} = \sum_{k=1}^{\eta_j^{(B)}} \beta_k \lambda_j x_{k,j}^{(B)} \in \mathrm{Eig}(B, \lambda_j^{(B)})$$

Which means that $\mathrm{Eig}(B, \lambda_j^{(B)})$ is $A$-invariant for each $\lambda_j^{(B)} \in \sigma(B)$. Due to the fact that $\mathbf{V}_n = \bigoplus_{j=1}^{|\sigma(B)|} \mathrm{Eig}(B, \lambda_j^{(B)})$, according to Lemma 2, there exists a basis $\mathscr{B}$ of $\mathbf{V}_n$ such that ${}^{\mathscr{B}}A$ is diagonal on each $\mathrm{Eig}(B, \lambda_j^{(B)})$ and therefore also on $\mathbf{V}_n$. It is easy to see that ${}^{\mathscr{B}}B$ is also diagonal. □

## 2.3 Schur's decomposition and diagonalizability of normal matrices

It is well-known that every diagonalizable matrix $B \in M_n(\mathbb{C})$ satisfies $B = PDP^{-1}$, where $P$ denotes the transition matrix between the original basis and the diagonal basis of $A$ and $D$ denotes a diagonal matrix. More general decomposition holds for every $A \in M_n(\mathbb{C})$, as given by the following theorem [7].

**Theorem 3** (Schur's decomposition theorem)**.** *Let $A \in M_n(\mathbb{C})$. Then there exists a unitary matrix $U \in M_n(\mathbb{C})$ and an upper triangular matrix $T \in M_n(\mathbb{C})$ such that $U^H AU = T$.*

*Proof.* This theorem is obviously true for $n = 1$. Assume that the theorem holds for $n-1$,i.e. for every $A_1 \in M_{n-1}(\mathbb{C})$ there exists a unitary matrix $U_1 \in M_{n-1}(\mathbb{C})$ and an upper triangular matrix $T_1 \in M_{n-1}(\mathbb{C})$ such that $U_1^H A_1 U_1 = T_1$. Let $x_1$ be an eigenvector of $A$ corresponding to $\lambda \in \sigma(A)$. Without loss of generality, assume that $\|x_1\|_{\mathscr{E}} = 1$. Then by applying the Gramm-Schmidt algorithm, there exists an orthonormal basis of $\mathbb{C}^n$ containing $x_1$ with respect to the standard inner product, denote its vectors by $x_1, x_2, ..., x_n$. Put $Q = (x_1, x_2, ..., x_n)$. It is easy to see that $Q \in M_n(\mathbb{C})$ is unitary. Then

$$Q^H AQ = \begin{pmatrix} x_1^H \\ x_2^H \\ x_3^H \\ \vdots \\ x_n^H \end{pmatrix} A(x_1, x_2, ..., x_n) = \begin{pmatrix} x_1^H \\ x_2^H \\ x_3^H \\ \vdots \\ x_n^H \end{pmatrix} (Ax_1, Ax_2, ..., Ax_n) =$$

$$= \begin{pmatrix} x_1^H \\ x_2^H \\ x_3^H \\ \vdots \\ x_n^H \end{pmatrix} (\lambda x_1, Ax_2, Ax_3, ..., Ax_n) = \begin{pmatrix} \lambda & q^H \\ \theta & A_1 \end{pmatrix}$$

where $q \in \mathbb{C}^{n-1}, A_1 \in M_{n-1}(\mathbb{C})$. Now let $U = Q \begin{pmatrix} 1 & \theta^H \\ \theta & U_1 \end{pmatrix}$ Since both $U_1$ and $Q$ are unitary, it follows that $U$ is unitary. Now it remains to prove that $U^H AU$ is upper triangular:

$$U^H AU = \begin{pmatrix} 1 & \theta^H \\ \theta & U_1^H \end{pmatrix} Q^H AQ \begin{pmatrix} 1 & \theta^H \\ \theta & U_1 \end{pmatrix} = \begin{pmatrix} 1 & \theta^H \\ \theta & U_1^H \end{pmatrix} \begin{pmatrix} \lambda & q^H \\ \theta & A_1 \end{pmatrix} \begin{pmatrix} 1 & \theta^H \\ \theta & U_1 \end{pmatrix} =$$

$$= \begin{pmatrix} \lambda & q^H U_1 \\ \theta & U_1^H A_1 U_1 \end{pmatrix} = \begin{pmatrix} \lambda & q^H U_1 \\ \theta & T_1 \end{pmatrix}$$

Since it is assumed that $T_1$ is upper triangular, the proof is complete. □

As a consequence of this theorem, we prove another equivalent characterization of normal matrices:

**Lemma 3.** *Let $T \in M_n(\mathbb{C})$ then $T$ be an upper triangular matrix. Then $T$ is normal iff it is diagonal.*

*Proof.* It is obvious that this statement is true for $n = 1$. We proceed by induction, assuming that $T = \begin{pmatrix} \alpha & z^H \\ O & S \end{pmatrix}$, where $\alpha \in \mathbb{C}, z \in \mathbb{C}^{n-1}$ and $S \in \mathbb{C}^{n-1,n-1}$ is diagonal. Therefore $T^H = \begin{pmatrix} \overline{\alpha} & O \\ z & S^H \end{pmatrix}$. The equation $T^H T = TT^H$ gives:

$$\begin{pmatrix} |\alpha|^2 + z^H z & z^H T^H \\ Sz & SS^H \end{pmatrix} = \begin{pmatrix} |\alpha|^2 & \overline{\alpha} z^H \\ \alpha z & SS^H \end{pmatrix}$$

which implies that $\|z\|_{\mathscr{E}} = 0$ and thus $z = \theta$. The assumption that $S$ is diagonal gives the desired result. □

**Lemma 4.** *Let $A, U \in M_n(\mathbb{C})$, furthermore let $A$ be normal and let $U$ be unitary. Then $U^H AU$ is normal iff $A$ is normal.*

*Proof.* Assuming that $A$ is normal:

$$U^H AU(U^H AU)^H = U^H AUU^H A^H U = U^H AA^H U =$$

$$= U^H A^H AU = U^H A^H UU^H AU = (U^H AU)^H U^H AU.$$

Assuming the converse, $U^H AU(U^H AU)^H = (U^H AU)^H U^H AU$ gives the result $U^H AA^H U = U^H A^H AU$ and multiplying this equation by $U$ from the left and by $U^H$ from the right gives $AA^H = A^H A$, i.e. $A$ is normal. □

**Theorem 4.** *Let $A \in M_n(\mathbb{C})$. Then $A$ is normal iff there exists a unitary matrix $U \in M_n(\mathbb{C})$ such that A is diagonalizable by $U$.*

*Proof.* Let $A \in M_n(\mathbb{C})$, then by Schur's decomposition theorem, it can be written in the form $A = U^H T U$, where $U$ is unitary and $T$ is upper triangular. By Lemma 4 $A$ is normal iff $T$ is normal. Lemma 3 states that $T$ is normal iff $T$ is diagonal, so $A$ is normal iff it is diagonalizable by $U$. □

Alternative proof of the above theorem can be found in [8].

## 2.4 Automorphisms of $M_n(\mathbb{C})$

**Definition 18.** *Let $A$ be an invertible element of $M_n(\mathbb{C})$. Inner automorphism $Ad_A$ is defined $Ad_A(X) = A^{-1}XA$ for all $X \in M_n(\mathbb{C})$ [9].*

Note that linearity of $Ad_A$ follows from the linearity of matrix multiplication and that $Ad_A(XY) = A^{-1}(XY)A = (A^{-1}XA)(A^{-1}YA) = Ad_A(X)Ad_A(Y)$, hence $Ad_A$ preserves multiplication. Since $Ad_A(X) = A^{-1}XA = O \Rightarrow X = O$, thus $\ker(Ad_A) = \{O\}$ and therefore it is a bijection, so $Ad_A$ satisfies the definition of an automorphism for every invertible $A \in M_n(\mathbb{C})$. The following theorem describes the properties of inner automorphisms in relation to their generating matrices [9], [10].

**Theorem 5.** *Let $A, B$ be invertible elements of $M_n(\mathbb{C})$. Then*

1. *$Ad_A$ and $Ad_B$ commute iff there exists $\omega_n \in \mathbb{C}$ such that*
$$AB = \omega_n BA \quad \textit{and} \quad \omega_n^n = 1.$$
2. *$Ad_A$ is a *-automorphism iff it is generated by a unitary matrix.*
3. *$Ad_A$ is diagonalizable iff $A$ is diagonalizable.*

*Proof.*

1. By definition,
$$Ad_A(Ad_B) = Ad_B(Ad_A) \Leftrightarrow A^{-1}B^{-1}XAB = B^{-1}A^{-1}XBA$$
for all $X \in M_n(\mathbb{C})$, i.e. $BAB^{-1}A^{-1}X = XBAB^{-1}A^{-1}$, so by Theorem 1,
$$BAB^{-1}A^{-1} = \frac{1}{\omega_n}I \quad \text{i.e.} \quad AB = \omega_n BA.$$
Equality of determinants implies that $\det(AB) = \omega_n^n \det(BA)$ and therefore $\omega_n^n = 1$
2. First, let $A$ be unitary. It follows that
$$(Ad_A(X))^H = (A^{-1}XA)^H = A^H X^H A^{-H} = A^{-1}X^H A = Ad_A(X).$$
Assuming the converse, i.e. $Ad_A(X^H) = (Ad_A(X))^H$ is equivalent to
$$A^{-1}X^H A = A^H X^H A^{-H} \Leftrightarrow AA^H X^H = X^H AA^H$$

and therefore by Theorem 1 there exists $\lambda \in \mathbb{C}$ such that $AA^H = \lambda I$, where

$$(\lambda I)_{ii} = \lambda = \sum_{i=1}^{n} A_{ij}A_{ji}^H = \sum_{i=1}^{n} A_{ij}\overline{A_{ij}} = \sum_{i=1}^{n}|A_{ij}|^2 > 0.$$

The fact $AA^HAA^H = \lambda^2 I = AAA^HA^H$ shows that $A$ is normal and therefore the matrix $U = \frac{1}{\sqrt{\lambda}}A$ is unitary and since $A^{-1} = \frac{1}{\lambda}A^H$, the automorphism $Ad_A$ can be written in the form $Ad_A(X) = \frac{1}{\lambda}A^HXA = \frac{1}{\sqrt{\lambda}}A^HX\frac{1}{\sqrt{\lambda}}A = Ad_U(X)$ for all $X \in M_n(\mathbb{C})$, thus $Ad_A$ is generated by a unitary matrix.

3. Since the standard basis of $M_n(\mathbb{C})$ can be written in the form $E_{ij} = e_ie_j^T$, where $e_i$ denotes the standard basis of $\mathbb{C}^n$, it follows that

$$Ad_A(E_{ij}) = (A^{-1}e_i)(A^He_j)^T = (A^{-1} \otimes A^T)(e_i \otimes e_j) \Rightarrow Ad_A = A^{-1} \otimes A^T.$$

It can be easily proven that both $A^{-1}$ and $A^T$ are diagonalizable iff A is diagonalizable, and since tensor product is diagonalizable iff both is components are diagonalizable [10], the assertion is proven.

□

It is easy to see that the standard basis $(E_{ij})_{i,j=1}^n$ of $M_n(\mathbb{C})$ satisfies the following relation:

$$E_{ij}E_{kl} = \begin{cases} O & \text{for } j \neq k \\ E_{il} & \text{for } j = k \end{cases}$$

and that an automorphism $\psi$ preserves this relation. In addition, let $\{M_1, M_2, ..., M_k\}$ be a finite set of matrices from $M_n(\mathbb{C})$ and let $\alpha_1, \alpha_2..., \alpha_k \in \mathbb{C}$. Then since $\psi$ is a bijection, we obtain

$$\psi(\sum_{i=1}^{k} \alpha_iM_i) = \psi(O) = O \Leftrightarrow \sum_{i=1}^{k} \alpha_iM_i = O,$$

hence $\psi$ preserves linear independence, therefore the image of the standard basis constitutes another basis of $M_n(\mathbb{C})$, resulting in the subsequent definition:

**Definition 19.** *Any basis $(B_{ij})_{i,j=1}^n$ of $M_n(\mathbb{C})$ satisfying*

$$B_{ij}B_{kl} = \begin{cases} O & \textit{for } j \neq k \\ B_{il} & \textit{for } j = k \end{cases}$$

*shall be called a generalized standard basis of $M_n(\mathbb{C})$.*

**Theorem 6** (Skolem-Noether theorem, specialized)**.** *All automorphisms of $M_n(\mathbb{C})$ are inner, i.e. for each automorphism $\psi$ there exists an invertible matrix G such that $Ad_G = \psi$.*

*Proof.* (As suggested in [11]) First, let $(E_{ij})_{i,j=1}^n$ be a standard basis of $M_n(\mathbb{C})$ and denote $\psi(E_{ij}) = F_{ij}$ for all $i, j \in \widehat{n}$. It is clear that $(F_{ij})_{i,j=1}^n$ is a generalized standard basis of $M_n(\mathbb{C})$. In addition, it is possible to write $F_{11}$ in the following form:

$$F_{11} = \sum_{i,j=1}^{n} \alpha_{ij}E_{ij}$$

where there exists at least one ordered pair of indices, say $p, q$ such that $\alpha_{pq} \neq 0$.

Second, we define:

$$A = \alpha_{pq}^{-1} E_{1p} F_{11} \qquad F = \sum_{i=1}^{n} E_{i_1} A F_{1i}$$

$$B = F_{11} E_{q1} \qquad G = \sum_{j=1}^{n} F_{j_1} B E_{1j}.$$

The fact that both $(E_{ij})_{i,j=1}^{n}$ and $(F_{ij})_{i,j=1}^{n}$ are generalized standard bases implies that $AF_{11} = A$ and $BE_{11} = B$. To proceed, the three following identities are needed:

$$AB = \alpha_{pq}^{-1} E_{1p} F_{11} F_{11} E_{q1} = \alpha_{pq}^{-1} E_{1p} F_{11} E_{q1} = \alpha_{pq}^{-1} E_{1p} (\sum_{i,j=1}^{n} \alpha_{ij} E_{ij}) E_{q1} =$$

$$= \alpha_{pq}^{-1} (\sum_{j=1}^{n} \alpha_{pj} E_{1j}) E_{q1} = \alpha_{pq}^{-1} \alpha_{pq} E_{11} = E_{11}$$

$$BA = F_{11} E_{q1} \alpha_{pq}^{-1} E_{1p} F_{11} = \alpha_{pq}^{-1} F_{11} E_{qp} F_{11} = \alpha_{pq}^{-1} (\sum_{i,j=1}^{n} \alpha_{ij} E_{ij}) E_{qp} (\sum_{k,l=1}^{n} \alpha_{kl} E_{kl}) =$$

$$= \alpha_{pq}^{-1} (\sum_{i=1}^{n} \alpha_{iq} E_{ip}) (\sum_{k,l=1}^{n} \alpha_{kl} E_{kl}) = \alpha_{pq}^{-1} \sum_{i,k,l=1}^{n} \alpha_{iq} \alpha_{kl} E_{ip} E_{kl} =$$

$$= \alpha_{pq}^{-1} \alpha_{pq} \sum_{i,l=1}^{n} \alpha_{il} E_{il} = \sum_{i,l=1}^{n} \alpha_{il} E_{il} = F_{11}$$

Furthermore, it is easily seen that $\sum_{i=1}^{n} F_{ii} = \sum_{i=1}^{n} \psi(E_{ii}) = \psi(\sum_{i=1}^{n} E_{ii}) = \psi(I) = I$.

Third, we prove that $G$ is invertible and $G^{-1} = F$:

$$GF = \sum_{i,j=1}^{n} F_{j1} B E_{1j} E_{i1} A F_{1i} = \sum_{i=1}^{n} F_{i1} B E_{11} A F_{1i} = \sum_{i=1}^{n} F_{i1} B A F_{1i} =$$

$$= \sum_{i=1}^{n} F_{i1} F_{11} F_{1i} = \sum_{i=1}^{n} F_{ii} = I$$

$$FG = \sum_{i,j=1}^{n} E_{i1} A F_{1i} F_{j1} B E_{1j} = \sum_{i=1}^{n} E_{i1} A F_{11} B E_{1i} = \sum_{i=1}^{n} E_{i1} A B E_{1i} =$$

$$= \sum_{i=1}^{n} E_{i1} E_{11} E_{1i} = \sum_{i=1}^{n} E_{ii} = I$$

Fourth, we prove that for each $i, j \in \widehat{n}$, $Ad_G(E_{ij}) = F_{ij}$:

$$GE_{ij}G^{-1} = GE_{ij}F = \sum_{k,l=1}^{n} F_{k1} B E_{1k} E_{ij} E_{11} A F_{1l} = \sum_{l=1}^{n} F_{i1} B E_{1j} E_{l1} A F_{1l} =$$

$$= F_{i1} B E_{11} A F_{1j} = F_{i1} B A F_{1j} = F_{i1} F_{11} F_{1j} = F_{ij}$$

Thus inner automorphism $Ad_G$ acts in exactly the same way on the standard basis $(E_{ij})_{i,j=1}^{n}$ as $\psi$ does, hence $Ad_G = \psi$. $\square$

# Chapter 3

# Classification of fine gradings of $M_n(\mathbb{C})$

## 3.1 Gradings and automorphisms of *-algebras

In a non-specific *-algebra $\mathbf{A}$, we define operations with its subsets in the following way: let $\mathcal{A}, \mathcal{B} \subseteq \mathbf{A}$, then

$$\alpha\mathcal{A} + \mathcal{B} = \{\alpha a + b : a \in \mathcal{A}, b \in \mathcal{B}\}, \quad \mathcal{A}\mathcal{B} = \{ab : a \in \mathcal{A}, b \in \mathcal{B}\}, \quad \mathcal{A}^* = \{a^* : a \in \mathcal{A}\},$$

where $\alpha \in \mathbb{C}$. This notation allows us to define a grading in an elegant fashion [9], [10], [12]:

**Definition 20.** *Let $\mathbf{A}$ be a *-algebra and let $\mathcal{I}$ be an index set. A grading $\Gamma$ of a *-algebra $\mathbf{A}$ is a decomposition of $\mathbf{A}$ into direct sum of subspaces*

$$\Gamma : \quad \mathbf{A} = \bigoplus_{i \in \mathcal{I}} \mathbf{A}_i$$

*such that for any pair of indices $i, j \in \mathcal{I}$ there exists an index $k \in \mathcal{I}$ with the property*

$$\mathbf{A}_i \mathbf{A}_j \subseteq \mathbf{A}_k$$

*and for any index $l \in \mathcal{I}$ there exists an index $m \in \mathcal{I}$ such that*

$$\mathbf{A}_l^* \subseteq \mathbf{A}_m.$$

**Definition 21.** *Let $\mathbf{A}$ be a *-algebra and let $\Gamma$ be a grading of $\mathbf{A}$. A grading $\tilde{\Gamma}$ is called a refinement of $\Gamma$ iff it satisfies that for each $\tilde{\mathbf{A}}_i$ constituting $\tilde{\Gamma}$ there exists $\mathbf{A}_j$ constituting $\Gamma$ such that $\tilde{\mathbf{A}}_i \subseteq \mathbf{A}_j$, where at least one inclusion is proper. A grading which cannot be refined further is called fine.*

Certain gradings of a finite dimensional *-algebra $\mathbf{A}$ can be obtained by looking at the group of all its *-automorphisms. If a *-automorphism $\psi$ is diagonalizable and $a, b$ are its eigenvectors with nonzero eigenvalues $\mu, \nu \in \mathbb{C}$ respectively, then clearly

$$\psi(ab) = \psi(a)\psi(b) = (\mu a)(\nu b) = (\mu\nu)ab \quad \text{and} \quad (\psi(a))^* = (\lambda a)^* = \overline{\lambda} a^*$$

This means that $ab$ is either an eigenvector of $\psi$ with the eigenvalue $\mu\nu$ or the zero element and that $a^*$ is an eigenvector of $\psi$ corresponding to $\overline{\lambda}$. The given automorphism $\psi$ therefore leads to a decomposition of $\mathbf{A}$ into the sum of eigenspaces of $\psi$ with corresponding eigenvalues $\lambda_i$,

$$\Gamma : \quad \mathbf{A} = \bigoplus_{\lambda_i \in \sigma(\psi)} \mathrm{Eig}(\psi, \lambda_i)$$

which satisfies the definition of a grading [9].

Refinements of a given grading can be obtained by adjoining further diagonalizable *-automorphisms commuting with $\psi$. Supppose that $\phi$ and $\psi$ are commuting diagonalizable *-automorphisms, i.e. $\psi \circ \phi = \phi \circ \psi$. It follows that for any eigenvector $a$ of $\psi$ with the eigenvalue $\lambda$

$$\lambda\phi(a) = \phi(\lambda a) = (\phi \circ \psi)(a) = (\psi \circ \phi)(a) = \psi(\phi(a)) \Rightarrow \phi(a) \in \mathrm{Eig}(\psi, \lambda)$$

and so $\phi$ is $\mathrm{Eig}(\psi, \lambda)$-invariant. Diagonalizability of $\phi$ (according to Lemma 2) implies that $\phi$ is diagonalizable on $\mathrm{Eig}(\psi, \lambda)$ for each $\lambda \in \sigma(\psi)$ and therefore defines a refinement of $\Gamma$.

Moreover, assuming that $\psi$ is invertible, i.e. $0 \notin \sigma(\psi)$, we obtain $\psi(a) = \lambda a \Rightarrow \psi^{-1}(a) = \frac{1}{\lambda}a$ thus $\psi^{-1}$ has the same eigenspaces as $\psi$ only corresponding to the inverses of their respective eigenvalues. It can be easily proven that $\psi^{-1}$ preserves involution and multiplication and therefore is a *-automorphism:

$$x^* = ((\psi \circ \psi^{-1})(x))^* = \psi(\psi^{-1}(x))^* = \psi(\psi^{-1}(x)^*) \Rightarrow \psi^{-1}(x^*) = (\psi^{-1}(x))^*$$

$$\psi(\psi^{-1}(x)\psi^{-1}(y)) = [(\psi \circ \psi^{-1})(x)][(\psi \circ \psi^{-1})(y)] = xy \Rightarrow \psi^{-1}(x)\psi^{-1}(y) = \psi^{-1}(xy)$$

for all $x, y \in \mathbf{A}$.

The above observations imply that a pair consisting of *-automorphism and its inverse defines the same grading and therefore a given grading $\Gamma$ and its refinements are induced by a group $G$ of invertible diagonalizable *-automorphisms. If $\Gamma$ is fine, then $G$ must be maximal, i.e. for all $\psi \notin G$ there exists some $\phi \in G$ such that $\psi\phi \neq \phi\psi$. Maximal groups of commuting diagonalizable invertible *-automorphisms shall be called MAD-groups of a *-algebra $\mathbf{A}$ [9].

## 3.2 Classification of MAD-groups of $M_n(\mathbb{C})$

Let us consider a *-automorphism $\psi$ of $M_n(\mathbb{C})$. According to Theorem 6, it is an inner automorphism. It is clear that if for some unitary matrices $U, V$, the following implication holds: $U = V \Rightarrow Ad_U = Ad_V$. Assuming the converse, i.e. $Ad_U = Ad_V$ gives, according to Theorem 1, $UV^{-1} = \alpha I$, where $\alpha^n = 1$. Hence an *-automorphism defines an equivalence relation $U \sim V \Leftrightarrow U = \alpha V, \alpha^n = 1$ on the group of unitary matrices, which shall be denoted $U(n)$. By defining multiplication of equivalence classes $[U][V] = [UV]$, a group isomorphic to the group of all *-automorphisms of $M_n(\mathbb{C})$ is obtained.

According to Theorem 5 and Theorem 6, all *-automorphisms of $M_n(\mathbb{C})$ are diagonalizable, invertible and inner. We show that there is a one-to-one correspondence between MAD-groups and unitary Ad-groups [10], defined below:

**Definition 22.** *A subgroup $\mathscr{G}$ of $U(n)$ shall be called a unitary Ad-group iff*

1. *For any pair $U, V \in \mathscr{G}$ there exists $\omega_n \in \mathbb{C}$ such that $UV = \omega_n VU$.*
2. *$\mathscr{G}$ is maximal, i.e. for each $M \notin \mathscr{G}$ there exists $U \in \mathscr{G}$ such that $UM \neq \omega_n MU$.*

Obviously, if any unitary Ad-group contains a matrix $U$, it also contains the whole equivalence class $[U]$. Denote for any MAD-group $G$:

$$G_{Ad} = \{U \in U(n) : Ad_U \in G\}$$

and conversely for any unitary Ad-group $\mathscr{G}$:

$$Ad(\mathscr{G}) = \{Ad_U : U \in \mathscr{G}\}.$$

According to Theorem 5, the first property of Definition 22 is satisfied for any pair $Ad_U, Ad_V \in G$. It is easy to see that $G_{Ad}$ is maximal and that $Ad(G_{Ad}) = G$. We are interested in the classes of MAD-groups, in what follows we will describe their suitable representatives.

If a unitary Ad-group $\mathscr{G}$ is commutative, i.e. $UV = VU$ for all $UV \in \mathscr{G}$, then all its elements are simultaneously diagonalizable by some unitary matrix. In addition, maximality of $\mathscr{G}$ implies that it is conjugated to the group of all $n \times n$ diagonal unitary matrices, denoted $\mathscr{U}_D(n)$. The following lemma describes the case $UV = \omega_n VU$ [10].

**Lemma 5.** *Let $A, B$ be diagonalizable invertible elements of $M_n(\mathbb{C})$ such that $AB = \omega_k BA$ where $\omega_k = \exp((2\pi i)/k)$ and $k$ divides $n$. Then there exists a invertible matrix $P$ such that*

$$PAP^{-1} = \operatorname{diag}(1, \omega_k, \omega_k^2, ..., \omega_k^{k-1}) \otimes \operatorname{diag}(d_1, d_2, ..., d_{n/k})$$

*and*

$$PBP^{-1} = \begin{pmatrix} 0 & 1 & 0 & ... & 0 & 0 \\ 0 & 0 & 1 & ... & 0 & 0 \\ 0 & 0 & 0 & ... & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & ... & 0 & 1 \\ 1 & 0 & 0 & ... & 0 & 0 \end{pmatrix} \otimes \operatorname{diag}(\delta_1, \delta_2, ..., \delta_{n/k})$$

*where* $\arg d_i, \arg \delta_i \in [0, 2\pi/k)$ *for all* $i = 1, 2, 3, ..., n/k$.

*Proof.* First, order the eigenvalues $\lambda_i$ of $A$ with respect to $\arg \lambda_i$, so that $0 \le \arg \lambda_1 \le \arg \lambda_2 \le ... \le \arg \lambda_n \le 2\pi$. Consider the subspace $F \subset\subset \mathbb{C}^n$ of eigenvectors with $\arg \lambda_i \in [0, 2\pi/k)$ and denote $\dim F = s$. As $AB^k = ABB^{k-1} = \omega BAB^{k-1} = \omega^2 B^2 AB^{k-2} = ... = \omega^k B^k A = B^k A$, by applying lemma 2, a basis $\{e_1, e_2, ..., e_s\} \subset F$ consisting of common eigenvectors of $A$ and $B^k$ can be chosen, i.e.

$$Ae_i = \lambda_i e_i \quad B^k e_i = \nu_i^k e_i,$$

where geometric multiplicity greater than 1 is allowed and where we may assume $\arg \nu_i \in [0, 2\pi/k)$.

Let us define

$$f_1 = e_1 \quad f_2 = \frac{1}{\nu_1} Be_1 \quad , f_3 = \frac{1}{\nu_1^2} B^2 e_1 \quad ... \quad f_k = \frac{1}{\nu_1^{k-1}} B^{k-1} e_1$$

$$f_{k+1} = e_2 \quad f_{k+2} = \frac{1}{\nu_2} Be_2 \quad f_{k+3} = \frac{1}{\nu_2^2} B^2 e_2 \quad ... \quad f_{2k} = \frac{1}{\nu_2^{k-1}} B^{k-1} e_2$$

$$\vdots$$

$$f_{(s-1)k+1} = e_s \quad f_{(s-1)k+2} = \frac{1}{\nu_s} Be_s \quad f_{(s-1)k+3} = \frac{1}{\nu_s^2} B^2 e_s \quad ... \quad f_{sk} = \frac{1}{\nu_s^{k-1}} B^{k-1} e_s$$

Obviously, $Af_1 = Ae_1 = \lambda_1 f_1, Af_2 = A(\frac{1}{\nu_1} Be_1) = q\lambda_1 f_2, Af_3 = a(\frac{1}{\nu_1^2} B^2 e_1) = q^2 \lambda_1 f_3$, which implies that $f_1, f_2, ..., f_{sk}$ are eigenvectors of $A$, each corresponds to a different eigenvalue, therefore these vectors are linearly independent.

Suppose $sk < n$. Then there exists an eigenvector $x$ of $A$, linearly independent on $f_1, f_2, ..., f_{sk}$. Let $x$ correspond to $\lambda \in \sigma(A)$ and $\arg \lambda \in [(2\pi/k)j, 2\pi/k(j+1))$. This assumption implies that $A^{-1}B^{-j}x = \omega^j B^{-j}A^{-1}x = \omega^j \frac{1}{\lambda} B^{-j}x$, so $B^{-j}x \in F$. Thus $B^{-j}x$ can be written as a linear combination of $e_1, ..., e_s$ and by applying $B^j$ on both sides of this equation, we see that $x$ can be written as a linear combination of $B^j e_1, ..., B^j e_s$, which is a contradiction, so $sk = n$ and $(f_i)_{i=1}^n$ forms a basis of $\mathbb{C}^n$.

We have obtained for some suitable $P$:

$$P^{-1}AP = \mathrm{diag}(\lambda_1, \omega\lambda_1, ..., \omega^{k-1}\lambda_1, \lambda_2, \omega\lambda_2, ..., \omega^{k-1}\lambda_2, ..., \lambda_{n/k}, \omega\lambda_{n/k}, ..., \omega^{k-1}\lambda_{n/k}) =$$

$$= \mathrm{diag}(1, \omega, ..., \omega^{k-1}) \otimes \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_{n/k})$$

Moreover,

$$Bf_1 = Be_1 = \nu_1 f_2 \quad Bf_2 = \frac{1}{\nu_1} B^2 e_1 = \nu_1 f_3 \quad Bf_3 = \frac{1}{\nu_1^2} B^3 e_1 = \nu_1 f_4 \quad ...$$

$$... \quad Bf_k = \frac{1}{\nu_1^{k-1}} B^k e_1 = \frac{\nu_1^k}{\nu_1^{k-1}} e_1 = \nu_1 f_1$$

and similarly for other eigenvalues of $B^k$, giving the result:

$$P^{-1}BP = \begin{pmatrix} 0 & 1 & 0 & ... & 0 & 0 \\ 0 & 0 & 1 & ... & 0 & 0 \\ 0 & 0 & 0 & ... & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & ... & 0 & 1 \\ 1 & 0 & 0 & ... & 0 & 0 \end{pmatrix} \otimes \mathrm{diag}(\nu_1, \nu_2, ..., \nu_{n/k})$$

$\square$

Also note that in the statement of lemma 5, it is possible to replace the interval $[0, 2\pi/k)$ by the interval $(-\pi/k, \pi/k]$.

**Definition 23.** *The $k \times k$ matrix*

$$\begin{pmatrix} 0 & 1 & 0 & ... & 0 & 0 \\ 0 & 0 & 1 & ... & 0 & 0 \\ 0 & 0 & 0 & ... & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & ... & 0 & 1 \\ 1 & 0 & 0 & ... & 0 & 0 \end{pmatrix}$$

*will be denoted $P_k$ and the matrix $\mathrm{diag}(1, \omega, \omega^2, ..., \omega^{k-1})$, where $\omega_k = \exp(2\pi i/k)$ will be denoted $Q_k$. The group $\mathscr{P}_k = \{\omega_k^l Q_k^m P_k^n : l, m, n = 0, 1, 2, ..., k-1\}$ will be called Pauli's group.*

It can be easily checked that $\mathscr{P}_k$ is a group of unitary matrices satisfying

$$P_k^k = Q_k^k = I_k \quad P_k Q_k = \omega_k Q_k P_k.$$

Furthermore, if $k$ is even, then $(P_k Q_k)^k = -I$ and if $k$ is odd, then $(P_k Q_k)^k = I$. These matrices were first studied by Weyl in [13].

**Lemma 6.** *$\mathscr{P}_k \otimes \mathscr{P}_m$ is conjugated to $\mathscr{P}_{km}$ iff $k$ and $m$ are relatively prime [14].*

*Proof.* First, let us put $\mathbb{C}^k = \operatorname{span}((e_i^{(1)})_{i=0}^{k-1})$ and $\mathbb{C}^m = \operatorname{span}((e_j^{(2)})_{j=0}^{m-1})$, where $(e_i^{(1)})_{i=0}^{k-1}$ and $(e_j^{(2)})_{j=0}^{m-1}$ denote standard bases of their respective spaces. The elements of $\mathbb{C}^{km}$ can be written in the form

$$\sum_{i=0}^{k-1}\sum_{j=0}^{m-1} \alpha_{ij}(e_i^{(1)} \otimes e_j^{(1)}) \quad \text{where} \quad \alpha_{ij} \in \mathbb{C} \quad \text{for all} \quad i,j \in \widehat{n}.$$

Now, it can be easily seen that $f_s = P_{km}^s(e_0^{(1)} \otimes e_0^{(2)})$, $s = 0,1,2,...,km-1$ runs just once through all the vectors $e_i^{(1)} \otimes e_j^{(2)}$ iff the following sets are the same:

$$\left\{ \begin{pmatrix} s \mod k \\ s \mod m \end{pmatrix} : s = 0,1,2,...,km-1 \right\} = \left\{ \begin{pmatrix} i \\ j \end{pmatrix} : i = 0,1,2,...,k-1; j = 0,1,2,...,m-1 \right\}.$$

Obviously, the inclusion $\subseteq$ holds. Hence, it is sufficient to show that the equality

$$\begin{pmatrix} s \mod k \\ s \mod m \end{pmatrix} = \begin{pmatrix} t \mod k \\ t \mod m \end{pmatrix} \quad \text{i.e.} \quad \begin{pmatrix} s-t \mod k \\ s-t \mod m \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

implies that $s = t$. Indeed, both $k$ and $m$ are divisible by $|s-t|$ and $|s-t| \in \{0,1,2,...,km-1\}$. If $k$ and $m$ have no common divisors, then their lowest common multiple is $km$, hence $|s-t| = 0$.

Therefore there exists a unique pair $p, q$ such that $P_{km}f_s = f_{s-1 \mod km} = e^{(1)}_{p-1 \mod k} \otimes e^{(2)}_{q-1 \mod m} = (P_k e_p^{(1)}) \otimes (P_m e_q^{(2)}) = (P_k \otimes P_m)(e_p^{(1)} \otimes e_q^{(2)})$.

Similarly, $Q_{km}f_s = \omega_{km}^s f_s = \omega_k^p \omega_m^q (e_p^{(1)} \otimes e_q^{(2)}) = (\omega_k^p e_p^{(1)}) \otimes (\omega_m^q e_q^{(2)}) = (\omega_k^p (e_p^{(1)}) \otimes (\omega_m^q e_q^{(2)}) = (Q_k^p e_p^{(1)}) \otimes (Q_m^q e_q^{(2)}) = (Q_k \otimes Q_m)(e_p^{(1)} \otimes e_q^{(2)})$. Which implies that both bases contain the same vectors in different order, therefore any $A_{km} \in \mathscr{P}_{km}$ can be unitarily conjugated to the tensor product $A_k \otimes A_m$ where $A_k \in \mathscr{P}_k$ and $A_m \in \mathscr{P}_m$. $\square$

Now, let $A, B \in M_n(\mathbb{C})$ and let $\omega_n = \exp(2\pi i/n)$. The set

$$\{C \in M_n(\mathbb{C}) : (\exists s,t \in \mathbb{Z})(AC = \omega_n^s CA, BC = \omega_n^t CB)\}$$

shall be called $\omega_n$-commutant of matrices $A$ and $B$ and will be denoted $\{A,B\}^{(\omega_n)}$. Obviously, $\omega_n^n = 1$. For $\omega_n = 1$, we denote $\{A,B\}^{(1)} = \{A,B\}'$ and call the set commutant of matrices $A$ and $B$. In other words, $\{A,B\}'$ is the set of matrices commuting with both $A$ and $B$.

**Lemma 7.** *Let $D_1 = \operatorname{diag}(d_1, d_2, ..., d_k)$ and $D_2 = \operatorname{diag}(\delta_1, \delta_2, ..., \delta_k)$ where $\arg d_i, \arg \delta_i \in [0, 2\pi/k)$ for all $i \in \hat{n}$. Put $A = Q_k \otimes D_1$ and $B = P_k \otimes D_2$. Then $\{A,B\}^{(\omega_k)} = \mathscr{P}_k \otimes \{D_1, D_2\}'$ [10].*

*Proof.* Let us first consider $C \in \{A,B\}' \subset \{A,B\}^{(\omega_k)}$. $A$ is a diagonal matrix with $d_i \omega_k^s$ on the diagonal, $\omega_k = \exp(2\pi/k)$. Since $\arg d_i \in [0, 2\pi/k)$, we have $d_i\omega_k^s \neq d_i \omega_k^t$ for $s \neq t$ and so $AC = CA \Rightarrow C = \bigoplus_{j=1}^k C_j$, where $C_j \in M_n(\mathbb{C})$ is invertible and for each $j$ it holds

$$C_j D_1 = D_1 C_j \tag{3.1}$$

From the equality $BC = CB$ we obtain:

$$C_j D_2 = D_2 C_j \tag{3.2}$$

for all $i \in \hat{n}$, we put $C_{k+1} = C_k$, thus $C_1 D_2^k = D_2^k C_1$, for matrix elements of $C_1$, denoted $\gamma_{ij}$, the following equation is obtained:

$$\gamma_{ij}\delta_j^k = \delta_i^k \gamma_{ij} \tag{3.3}$$

for each $i, j$. If $\gamma_{ij} \neq 0$, then $\delta_j^k = \delta_i^k \Rightarrow \delta_j = \delta_i$, i.e.

$$C_1 D_2^k = D_2^k C_1 \Leftrightarrow C_1 D_2 = D_2 C_1 \tag{3.4}$$

Moreover, $C_1 D_2 = D_2 C_1 \Rightarrow C_1 = C_2$ and analogously $C_1 = C_2 = \dots = C_k$. Therefore $C = I_k \otimes C_1$, where $C_1 \in \{D_1, D_2\}'$.

Now, let us consider $H \in \{A, B\}^{(\omega_k)}$, i.e. $HA = \omega_k^s AH$ and $HB = \omega_k^t BH$. Put $C = A^x B^y H$, where $x, y \in \mathbb{Z}$. Then

$$CA = (A^x B^y H)A = \omega_k^s \omega_k^y A(A^x B^y H) = \omega_k^{s+y} AC$$

$$CB = (A^x B^y H)B = \omega_k^t \omega_k^x B(A^x B^y H) = \omega_k^{t+x} BC$$

For $y = -s$ and $x = -t$ is $C \in \{A, B\}'$ and therefore $C = A^{-s}B^{-t}H = I_k \otimes C_1, C_1 \in \{D_1, D_2\}'$, leading to $H = Q_k^s P_k^t \otimes D_1^s D_2^t C_1$, it is obvious that $D_1^s D_2^t C_1 \in \{D_1, D_2\}'$. $\square$

Now we show that the previous lemmas imply that noncommutative unitary Ad-groups are conjugated to other unitary Ad-groups acting on smaller dimension.

**Lemma 8.** *A noncommutative subgroup $\mathscr{G}$ of $U(n)$ is a unitary Ad-group iff it is unitarily conjugated to the tensor product $\mathscr{P}_{n/s_o} \otimes \tilde{\mathscr{G}}$ for some divisor $s_0$ of $n$ and some unitary Ad-subgroup $\tilde{\mathscr{G}} \subseteq U(s_0)$ [10].*

*Proof.* Let $\mathscr{G}$ be a noncommutative unitary Ad-group, i.e. every pair $U, V \in \mathscr{G}$ satisfies

$$UV = \omega_n^{s(U,V)} VU, \quad \text{where} \quad \omega_n = e^{\frac{2\pi i}{n}}, \quad S(U,V) = 0, 1, 2, \dots, n-1.$$

Denote $s_0 = \min\{s(U,V) > 0 : U, V \in \mathscr{G}\}$ and choose $U_0, V_0$ for which $s(U_0, V_0) = s_0$. Since $U_0^k V_0^l \in \mathscr{G}_{Ad}$ for all $k, l \in \mathbb{N}_0$, the set $\mathcal{Z}(\mathscr{G}) = \{(\omega_n^{s_0})^k I_n : k \in \mathbb{N}_0\}$ forms a group isomorphic to the subgroup of the cyclic group $\mathscr{Z}_n = \{\omega_n^k : k \in \mathbb{N}_0\}$ and therefore $s$ divides $n$.

If $s_0 = 1$ then any $W \in \mathscr{G}$ lies also in $\{U_0, V_0\}^{(\omega_n)}$. We show that it is also true for $s_0 > 0$. Suppose the contrary, i.e. the exists $W \in \mathscr{G}$ such that

$$ks_0 < s(W, U_0) < (k+1)s_0 \quad \text{for some} \quad k = 1, 2, 3, \dots, n/s_0.$$

Then

$$0 < -ks_0 + s(W, U_0) = s(V_0^{n/s_0 - k} W, U_0) < s_0,$$

which is a contradiction to the minimality of $s_0$ because $V_0^{n/s_0-k} W \in \mathscr{G}$. Therefore $s(W, U_0) = ks_0$ and analogously $s(W, V_0) = ls_0$ for $k, l = 0, 1, 2, \dots, n/s_0 - 1$. Thus any $W \in \mathscr{G}$ lies in $\{U_0, V_0\}^{(\omega_n^{s_0})}$, i.e. $\mathscr{G} \subseteq \{U_0, V_0\}^{(\omega_n^{s_0})}$. Using Lemma 5 we may assume that there exists a unitary matrix $A \in U(n)$ such that

$$A^H U_0 A = Q_{n/s_0} \otimes D_1, \quad A^H V_0 A = P_{n/s_0} \otimes D_2$$

and using Lemma 7 we obtain

$$\mathscr{G} \subseteq \{U_0, V_0\}^{(\omega_n^{s_0})} = \mathscr{P}_{n/s_0} \otimes \{D_1, D_2\}',$$

where $D_1, D_2$ are unitary. By repeating the process for the group $\mathscr{P}_n$, we obtain

$$\mathscr{P}_n \subseteq \{P_n, Q_n\}' \subseteq \mathscr{P}_n \otimes \{(1)\}' = \mathscr{P}_n,$$

i.e. $\mathscr{P}_n$ is maximal. This inclusion and assumption of maximality for $\mathscr{G}$ imply that $\mathscr{P}_{n/s_0} \otimes I_{s_0} \subseteq \mathscr{G}$, i.e. there exists $\tilde{\mathscr{G}} \subseteq \{D_1, D_2\}'$ such that $\mathscr{G} = \mathscr{P}_{n/s_o} \otimes \tilde{\mathscr{G}}$, hence the group $\tilde{\mathscr{G}}$ is again maximal.

It remains to answer the question whether for a given divisor $s_0$ of $n$ there exists $\mathscr{G}$ such that $\mathcal{Z}(\mathscr{G}) = \mathscr{Z}_{n/s_0}$. An answer is affirmative because for any unitary Ad-group $\tilde{\mathscr{G}} \subseteq U(s_0)$ the group $\mathscr{P}_{n/s_0} \otimes \tilde{\mathscr{G}} \subseteq U(n)$ is maximal. $\square$

**Theorem 7.** *$\mathscr{G} \subseteq U(n)$ is a unitary Ad-group iff it is unitarily conjugated to one of the finite groups*

$$\mathscr{P}_{\pi_1} \otimes \mathscr{P}_{\pi_2} \otimes \mathscr{P}_{\pi_3} \otimes ... \otimes \mathscr{P}_{\pi_s} \otimes \mathscr{U}_D(n/\pi_1\pi_2\pi_3...\pi_s)$$

*where $\pi_1, \pi_2, \pi_3, ..., \pi_s$ are powers of primes and their product divides $n$.*

Note that this classification of unitary matrices is similar to the classification of invertible matrices given in [9] and [10].

*Proof.* Since $\mathscr{P}_1 = \{(1)\}$, it is clear that the theorem holds if $\mathscr{G}$ is a commutative unitary Ad-group.

Now assume that $\mathscr{G}$ is noncommutative and let $n = p_1^{\alpha_1} p_2^{\alpha_2} p_3^{\alpha_3} ... p_r^{\alpha_r}$, where $p_i$ are primes and $\alpha_i > 0$ for all $i \in \widehat{r}$, be unique prime decomposition of $n$. Now choose $\pi_1, \pi_2, \pi_3, ..., \pi_s$ satisfying the following conditions:

1. for each $j \in \widehat{s}$ there exists $i \in \widehat{r}$ and a natural number $\beta$ such that $\pi_j = p_i^{\beta}$
2. $\pi_1\pi_2\pi_3...\pi_s$ divides $n$

It is clear from Lemma 6 that the unitary Ad-groups of the form

$$\mathscr{P}_{\pi_1} \otimes \mathscr{P}_{\pi_2} \otimes \mathscr{P}_{\pi_3} \otimes ... \otimes \mathscr{P}_{\pi_s} \otimes \mathscr{U}_D(n/\pi_1\pi_2\pi_3...\pi_s)$$

are mutually nonconjugated and that it is sufficient to consider just powers of prime numbers as values for $\pi_j$. Lemma 8 states that $\mathscr{G}$ is a unitary Ad-group iff it is conjugated to $\mathscr{P}_{\pi_1} \otimes \tilde{\mathscr{G}}$, where $\tilde{\mathscr{G}} \subseteq U(n/\pi_1)$. By repeatedly applying the above lemma on the residual unitary Ad-group $\tilde{\mathscr{G}}$, we obtain the final result. $\square$

## 3.3 Unitary Ad-groups and corresponding fine gradings of $M_n(\mathbb{C})$

The simplest form of unitary Ad-group is $\mathscr{U}_D(n) = \{\mathrm{diag}(u_1, u_2, u_3, ..., u_n) : |u_i|^2 = 1, i \in \widehat{n}\}$. For any $U \in \mathscr{U}_D(n)$, we have

$$Ad_U(E_{ij}) = U^H E_{ij} U = U^H (e_i e_j^T) U = U^H e_i (U^T e_j)^T = (U^H e_i) \otimes (U^T e_j) = \overline{u_i} u_j E_{ij}$$

hence the standard basis of $M_n(\mathbb{C})$ is the diagonal basis for all elements of $Ad(\mathscr{U}_D(n))$. Clearly, $E_{ii} \in \mathrm{Eig}(Ad_U, 1)$ for every $U \in \mathscr{U}_D(n)$ and for every $i \in \widehat{n}$.

Now let $U$ be an element of $\mathscr{U}_D(n)$, $U = \mathrm{diag}(u_1, u_2, u_3, ..., u_n)$. Let $V = \mathrm{diag}(v_1, v_2, v_3, ..., v_n) \in \mathscr{U}_D(n)$ satisfy the relation $u_i \neq v_i$ and $u_j \neq v_j$ for some pair of indices $(i, j) \in \widehat{n} \times \widehat{n}$. We obtain

$$Ad_U(E_{ij}) = \overline{u_i} u_j E_{ij} \quad \text{and} \quad Ad_V(E_{ij}) = \overline{v_i} v_j E_{ij} \Rightarrow E_{ij} \notin \mathrm{Eig}(Ad_V, \overline{u_i} u_j),$$

i.e. for any pair of indices $(i, j), i \neq j$ and for every $U \in \mathscr{U}_D(n)$ there exists a matrix $V \in \mathscr{U}_D(n)$ which will decompose each eigenspace $\mathrm{Eig}(Ad_U, \overline{u_i} u_j)$ of $Ad_U$ into the direct sum $\mathbf{E}_{ij} \oplus \mathrm{Eig}(Ad_V, \overline{u_i} u_j)$, hence the decomposition

$$\Gamma(\mathscr{U}_D(n)): \quad M_n(\mathbb{C}) = \left( \bigoplus_{\substack{i,j=1 \\ i \neq j}}^{n} \mathbf{E}_{ij} \right) \oplus \mathbf{D}_n,$$

where $\mathbf{D}_n = \mathrm{span}(E_{11}, E_{22}, E_{33}, ..., E_{nn})$ cannot be further refined. The properties of standard basis imply that $\Gamma(\mathscr{U}_D(n))$ satisfies the definition of a grading.

Another simple example of a unitary Ad-group is $\mathscr{P}_n$. Let us denote $A_{ab} = Q_n^a P_n^b$ where $a, b = 0, 1, 2, ..., n-1$. The set of $n^2$ matrices $\{A_{ab} : a, b = 0, 1, 2, ..., n-1\}$ shall be denoted $\mathcal{A}_n$.

**Lemma 9.** *For all different pairs of indices $(a, b) \neq (c, d)$, the matrices $A_{ab}$ and $A_{cd}$ are orthogonal with respect to the Hilbert-Schmidt inner product [15].*

*Proof.* Let $(a, b) \neq (c, d)$, then

$$\langle A_{ab}, A_{cd} \rangle = \langle Q_n^a P_n^b, Q_n^c P_n^d \rangle = \mathrm{Tr}((Q_n^a P_n^b)^H Q_n^c P_n^d) = \mathrm{Tr}(P_n^{-b} Q_n^{-a} Q_n^c P_n^d)$$

and since trace is invariant under cyclic permutation of matrices,

$$\langle A_{ab}, A_{cd} \rangle = \mathrm{Tr}(P_n^d P_n^{-b} Q_n^{-a} Q_n^c).$$

Without loss of generality, we can assume that $c \geq a$ and $d \geq b$, giving the result

$$\langle A_{ab}, A_{cd} \rangle = \mathrm{Tr}(P_n^{d-b} Q_n^{c-a}).$$

If $b \neq d$, then $P_n^{d-b}$ is traceless matrix multiplied by a diagonal matrix $Q_n^{c-a}$, giving a traceless matrix. In the case $b = d$ and $c > a$, a diagonal matrix with powers of $\omega_n$ on the diagonal is obtained. It follows that

$$\mathrm{Tr}(\mathrm{diag}(1, \omega_n^{c-a}, \omega_n^{2(c-a)}, ..., \omega_n^{(n-1)(c-a)})) = \sum_{i=0}^{n-1} \omega_n^{i(c-a)} = \frac{\omega_n^{n(c-a)} - 1}{\omega_n - 1} = 0.$$

$\square$

Since orthogonal matrices are linearly independent, the above lemma implies that the set $\mathcal{A}_n$ constitutes an orthonormal basis of $M_n(\mathbb{C})$, in addition $\|A_{ab}\| = \sqrt{\mathrm{Tr}(A_{ab}^H A_{ab})} = \sqrt{\mathrm{Tr}(I)} = \sqrt{n}$, so the set $\frac{1}{\sqrt{n}} \mathcal{A}_n$ is an orthonormal basis of $M_n(\mathbb{C})$.

The relations $P_n^n = Q_n^n = I$ imply that $(P_n^k)^H = (P_n^k)^{-1} = P_n^{(-k \mod n)}$ and analogously $(Q_n^k)^H = (Q_n^k)^{-1} = Q_n^{(-k \mod n)}$ for all $k \in \mathbb{Z}$, hence

$$Ad_{A_{ab}} A_{cd} = (A_{ab})^H A_{cd} A_{ab} = P^{-b} Q^{-a} Q^c P^d Q^a P^b =$$

$$= \omega_n^{ad} P^{-b} Q^c P^{b+d} = \omega_n^{(ad+bc)} Q^c P^d = \omega_n^{(ad+bc)} A_{cd}$$

i.e. $\mathcal{A}_n$ is a diagonal basis of $M_n(\mathbb{C})$ for all elements of $Ad(\mathscr{P}_n)$. Put $\mathbf{A}_{ab} = \mathrm{span}(A_{ab})$ for all $a, b = 0, 1, 2, ..., n-1$. It follows that

$$A_{ab} A_{cd} = Q^a P^b Q^c P^d = \omega_n^{bc} Q^{a+c} P^{b+d} \quad \Rightarrow \quad \mathbf{A}_{ab} \mathbf{A}_{cd} \subseteq \mathbf{A}_{a+c, b+d}$$

$$A_{ab}^H = (Q^a P^b)^H = P^{-b} Q^{-a} = P^{n-b} Q^{n-a} = \omega_n^{(n-a)(n-b)} P^{n-a} Q^{n-b} \quad \Rightarrow \quad \mathbf{A}_{ab} \subseteq \mathbf{A}_{n-a, n-b}$$

where addition and subtraction are modulo $n$.

For any two pairs of indices $(a, b), (c, d) \in \{0, 1, 2, ..., n-1\}^2$, the following inequality holds

$$(ad + bc) \mod n \neq (ad + (b+1)c) \mod n$$

since $c \in \{0, 1, 2, ..., n-1\}$. Therefore for any $A_{cd} \in \mathscr{P}_n$ there exists $A_{a'b'} = A_{a, b+1} \in \mathscr{P}_n$ such that $A_{cd} \notin \mathrm{Eig}(Ad_{A_{a'b'}}, \omega_n^{ad+bc})$, thus the decomposition

$$\Gamma(\mathscr{P}_n): \quad M_n(\mathbb{C}) = \bigoplus_{i,j=0}^{n-1} \mathbf{A}_{ij}$$

is a fine grading.

**Definition 24.** *Let $\mathbf{A}$ be an algebra with inner product and let $\Gamma$ be a decomposition of $\mathbf{A}$ into a direct sum of subspaces. $\Gamma$ is called an orthogonal decomposition of $\mathbf{A}$ iff each subspace constituting $\Gamma$ is orthogonal to all the others.*

Lemma 9 implies that $\Gamma(\mathscr{P}_n)$ is an orthogonal decomposition of $M_n(\mathbb{C})$ with respect to the Hilbert-Schmidt inner product.

Now, we will examine the fine gradings of $M_{p_1 p_2}(\mathbb{C})$, where $p_1, p_2$ are different prime numbers. According to Theorem 7 and Lemma 6, we obtain the following nonconjugated unitary Ad-groups:

1. $\mathscr{U}_D(p_1 p_2)$
2. $\mathscr{P}_{p_1} \otimes \mathscr{U}_D(p_2)$
3. $\mathscr{P}_{p_2} \otimes \mathscr{U}_D(p_1)$
4. $\mathscr{P}_{p_1 p_2}$.

Cases 2. and 3. are clearly analogous and cases 1. and 4. were already examined. Let us consider the case $\mathscr{P}_{p_1} \otimes \mathscr{U}_D(p_2)$. It can be easily verified that $\langle A \otimes B, C \otimes D \rangle = \langle A, C \rangle \langle B, D \rangle$, thus the set $\{A_{ij} \otimes E_{kl} : i, j = 0, 1, 2, ..., p_1 - 1 : k, l = 0, 1, 2, ..., p_2 - 1\}$ constitutes an orthogonal basis of $M_{p_1 p_2}(\mathbb{C})$. Let $\mathbf{A} \subset\subset M_p(\mathbb{C})$ and $\mathbf{B} \subset\subset M_q(\mathbb{C})$ denote $\mathbf{A} \otimes \mathbf{B} = \{A \otimes B : A \in \mathbf{A}, b \in \mathbf{B}\}$. The relation $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$ implies that $\mathscr{P}_{p_1} \otimes \mathscr{U}_D(p_2)$ gives the grading

$$\Gamma(\mathscr{P}_{p_1} \otimes \mathscr{U}_D(p_2)): \quad M_{p_1 p_2}(\mathbb{C}) = \left( \bigoplus_{i,j=0}^{p_1-1} \bigoplus_{\substack{k,l=1 \\ k \neq l}}^{p_2} \mathbf{A}_{ij} \otimes \mathbf{E}_{kl} \right) \oplus \left( \bigoplus_{m,n=0}^{p_1-1} \mathbf{A}_{mn} \otimes \mathbf{D}_{p_2} \right)$$

Now let us consider the case $p_1 = p_2$, i.e. the algebra $M_{p^2}(\mathbb{C})$, where $p$ denotes a prime number. The following unitary Ad-groups are mutually nonconjugated due to Theorem 7 and Lemma 6:

1. $\mathscr{U}_D(p^2)$
2. $\mathscr{P}_p \otimes \mathscr{U}_D(p)$
3. $\mathscr{P}_p \otimes \mathscr{P}_p$
4. $\mathscr{P}_{p^2}$.

The third case is the only one that was not already examined. The previous remarks imply that the group $\mathscr{P}_p \otimes \mathscr{P}_p$ gives the fine grading

$$\Gamma(\mathscr{P}_p \otimes \mathscr{P}_p): \quad \bigoplus_{i,j,k,l=0}^{p-1} \mathbf{A}_{ij} \otimes \mathbf{A}_{kl}.$$

## 3.4 Examples of gradings induced by *-automorphisms of $M_n(\mathbb{C})$

In this section, possibly non-zero matrix entries will be denoted $*$. Note that the following examples are not fine gradings.

**Example 1.** $n = 2$, $U = Q_2 = \mathrm{diag}(1, -1)$

$Ad_U E_{ij}$ gives the following decomposition into eigenspaces corresponding to the powers of $\omega_2$:

$$\omega_2^0 : \begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix} \quad \omega_2^1 : \begin{pmatrix} 0 & * \\ * & 0 \end{pmatrix}$$

**Example 2.** $n = 3$, $U = Q_3$

$$\omega_3^0 : \begin{pmatrix} * & 0 & 0 \\ 0 & * & 0 \\ 0 & 0 & * \end{pmatrix} \quad \omega_3^1 : \begin{pmatrix} 0 & * & 0 \\ 0 & 0 & * \\ * & 0 & 0 \end{pmatrix} \quad \omega_3^2 : \begin{pmatrix} 0 & 0 & * \\ * & 0 & 0 \\ 0 & * & 0 \end{pmatrix}$$

**Example 3.** $n = 3$, $U = \mathrm{diag}(1, \omega_3, \omega_3)$

$$\omega_3^0 : \begin{pmatrix} * & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{pmatrix} \quad \omega_3^1 : \begin{pmatrix} 0 & 0 & 0 \\ * & 0 & 0 \\ * & 0 & 0 \end{pmatrix} \quad \omega_3^2 : \begin{pmatrix} 0 & * & * \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$Ad_{A_{ab}}(A_{cd}) = \omega_3^{(ad+bc)} A_{cd}$ gives the following decomposition into eigenspaces corresponding to the powers of $\omega_3$:

**Example 4.** $n = 3$, $U = A_{01} = P_3$

$$M_3(\mathbb{C}) = \mathrm{span}(A_{00}, A_{01}, A_{02}) \oplus \mathrm{span}(A_{10}, A_{11}, A_{12}) \oplus \mathrm{span}(A_{20}, A_{21}, A_{22})$$

**Example 5.** $n = 3$, $U = A_{10} = Q_3$

$$M_3(\mathbb{C}) = \mathrm{span}(A_{00}, A_{10}, A_{20}) \oplus \mathrm{span}(A_{01}, A_{11}, A_{21}) \oplus \mathrm{span}(A_{02}, A_{12}, A_{22})$$

**Example 6.** $n = 3$, $U = A_{11} = Q_3 P_3$

$$M_3(\mathbb{C}) = \mathrm{span}(A_{00}, A_{12}, A_{21}) \oplus \mathrm{span}(A_{01}, A_{10}, A_{22}) \oplus \mathrm{span}(A_{02}, A_{11}, A_{20}).$$

# Chapter 4

# Orthogonal decompositions and quantum complementarity

## 4.1 Quantum bits

Classical computation and information is based on the concept of a *bit*, a physical object which can be found in two states, usually denoted $0$ and $1$. Quantum computation and information are built on a different concept, called *quantum bit* or *qubit* for short [16]. The difference between classical bits and qubits is that quantum bits can be found in any linear combination (often called superposition) of the basis states $|0\rangle$ and $|1\rangle$. Therefore, a qubit is described by a two-dimensional complex vector space where the states corresponding to those of a classical bit constitute a basis, assumed to be orthonormal. Using the bra-ket notation, we can write

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle.$$

Another difference is that we can observe whether a bit is in the state $0$ or $1$, computers do this when they retrieve the contents of their memory. However, to examine a qubit means perform a quantum measurement. Measuring a qubit yields either the result $|0\rangle$ with the probability $|\alpha|^2$ or the result $|1\rangle$ with the probability $|\beta|^2$. (Naturally, $|\alpha|^2 + |\beta|^2 = 1$, since any state is represented by a normalized vector.) Furthermore, measurement changes the state of a qubit, collapsing it from the superposition of states $|0\rangle$ and $|1\rangle$ to the specific state consistent with the measurement result, i.e. only one bit of information can be obtained from a single measurement of one qubit.

Suppose we have two qubits, thus having four possible outcomes: $|00\rangle, |01\rangle, |10\rangle$ and $|11\rangle$. A pair of qubits can exist in a superposition of these four states:

$$|\psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle,$$

where the corresponding outcomes appear with the probabilities $|\alpha_{ij}^2|$. Now suppose that only the first qubit was measured with the result $|0\rangle$ with probability $|\alpha_{00}|^2 + |\alpha_{01}|^2$, leaving the system in the normalized post-measurement state

$$|\psi'\rangle = \frac{1}{\sqrt{|\alpha_{00}|^2 + |\alpha_{01}|^2}}(\alpha_{00}|00\rangle + \alpha_{01}|01\rangle).$$

An important two qubit state is the Bell state (also called the EPR pair),

$$\frac{1}{\sqrt{2}}(|00\rangle + |01\rangle)$$

This state has the property that upon measuring the first qubit, one obtains two possible results: $|0\rangle$ with probability 1/2, leaving the post-measurement state $|\psi'\rangle = |00\rangle$, and $|1\rangle$ with probability 1/2, leaving $|\psi'\rangle = |11\rangle$; it is apparent that a measurement of the second qubit gives the same result as the measurement of the first one. These results were the first intimation that quantum mechanics allows information processing beyond what is possible in the classical world [16].

## 4.2 Complementarity structures

The definition of complementarity concerns very specific relation between quantum observables [15].

**Definition 25.** *Let $\mathbf{H}_n$ be a complex Hilbert space of finite dimension $n$. Two observables $A$ and $B$ are called complementary iff their eigenvalues are non-degenerate and any two normalized eigenvetors $u_i$ of $A$ and $v_i$ of $B$ satisfy $(u_i, v_j) = \frac{1}{\sqrt{n}}$.*

It is apparent that in an eigenstate $u_i$ of $A$ all eigenvalues of $B$ are measured with equal probabilities, and vice versa. Therefore exact knowledge of the measured value of $A$ implies maximal uncertainty to any measured value of $B$. Note that in the next definition the eigenvalues of $A$ and $B$ are in fact irrelevant, since only the corresponding orthonormal bases are involved.

**Definition 26.** *Let $\mathbf{H}_n$ be a complex Hilbert space of finite dimension $n$ and let $\mathscr{A} = (a_i)_{i=1}^n$ and $\mathscr{B} = (b_j)_{j=1}^n$ be two orthonormal bases of $\mathbf{H}_n$. The bases $\mathscr{A}$ and $\mathscr{B}$ are called mutually unbiased iff for all $i, j \in \widehat{n}$ the vectors $a_i$ and $b_j$ satisfy $|(a_i, b_j)| = \frac{1}{\sqrt{n}}$.*

Now we will show that the matrices $P_n$ and $Q_n$ satisfy the criterion of complementarity. Clearly, the vectors of standard basis of $\mathbb{C}^n$ satisfy $Q_n e_j = \omega_n^{j-1} e_j$ and therefore $Q_n$ has a non-degenerate spectrum and $(e_j)_{j=1}^n$ are its eigenvectors. Since $\det(P_n - \lambda I) = (-1)^n(\lambda^n - 1)$, it follows that $\sigma(P_n) = \{1, \omega_n, \omega_n^2, ..., \omega_n^{n-1}\}$. Denote $p_0, p_1, p_2, ..., p_{n-1}$ the eigenvectors corresponding to the respective powers of $\omega_n$. These are:

$$p_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \quad p_1 = \begin{pmatrix} 1 \\ \omega_n \\ \omega_n^2 \\ \vdots \\ \omega_n^{n-1} \end{pmatrix} \quad p_2 = \begin{pmatrix} 1 \\ \omega_n^2 \\ \omega_n^4 \\ \vdots \\ \omega_n^{2(n-1)} \end{pmatrix} \quad ... \quad p_{n-1} = \begin{pmatrix} 1 \\ \omega_n^{n-1} \\ \omega_n^{2(n-1)} \\ \vdots \\ \omega_n^{(n-1)(n-1)} \end{pmatrix}.$$

All these vectors have the same euclidean norm of $\sqrt{n}$, thus for the normalized vectors $\frac{1}{\sqrt{n}} p_k$ computing the standard inner product yields the result

$$|(e_j, \frac{1}{\sqrt{n}} p_k)| = \frac{1}{\sqrt{n}} |\omega_n^{jk}| = \frac{1}{\sqrt{n}}.$$

Thus giving the result that the grading generated by Pauli's group decomposes $M_n(\mathbb{C})$ into $n^2$ subspaces generated by its elements and thus the orthogonal decomposition of $M_n(\mathbb{C})$ into subspaces generated by the powers of complementary observables gives mutually unbiased bases of $\mathbb{C}^n$ with respect to the standard inner product.

## 4.3 Generalized quantum bits and physical interpretation

It is sometimes useful to generalize the notion of *qubit* to higher dimensions. A qudit is defined as a quantum system which has $d$ basis states. A multiple qudit physical system is described by a tensor product of single-qudit systems [15].

Theorem 7 states that any unitary Ad-group is conjugated to a tensor product of Pauli's groups acting on dimension equal to a power of a prime number and a diagonal unitary group. The spaces on which these Pauli's groups acts shall be called elementary qudits. (That is, elementary qudits have dimension equal to some power of a prime number.) The physical meaning of the matrices $P_n$ and $Q_n$, whose powers constitute Pauli's group is known; they represent the analogues of momentum and position in finite-dimensional quantum kinematics [15].

In the previous chapter, possible gradings of $M_n(\mathbb{C})$ were examined for $n = p_1 p_2$, where $p_1$ and $p_2$ are prime numbers. The results imply that if $p_1 = p_2$, it is possible to view such system also as a tensor product of two systems acting on a tensor product of two qudits, each with dimension equal to $p_1 = p_2$. Systems that can be viewed as a product of other quantum systems of smaller dimension as well as one big system shall be called composite systems.

The apparent contradiction that $M_n(\mathbb{C})$ represents linear operators for any $n$-dimensional quantum system and at the same time there may exist multiple nonconjugated unitary Ad-groups for given $n$ is resolved in the following way: from physical point of view, $M_n(\mathbb{C})$ is the operator algebra not only for a single $n$-dimensional system but also for all other members of the set of inequivalent quantum systems for this $n$ [15]. These systems correspond to different physical realizations of composite quantum systems. Of course, each such system has its preferred set of quantum operators.

# Conclusion

In this thesis, we have described the fundamental notions and necessary mathematical structures used to build the quantum theory with focus on operators in finite dimension, represented by the *-algebra $M_n(\mathbb{C})$. It was proven that all its automorphisms are inner and the relationship between fine gradings, maximal groups of commuting *-automorphisms (MAD-groups) and unitary Ad-groups was shown.

A classification of unitary Ad-groups using Pauli's groups $\mathscr{P}_k$ and diagonal unitary groups $\mathscr{U}_D(n)$ was given and the corresponding fine gradings of $M_n(\mathbb{C})$ were examined. In the final chapter, quantum complementarity was illustrated using two elements of Pauli's group.

There are several unsolved problems in regarding composite systems. For example, no physical interpretation of the matrices constituting the diagonal unitary group $\mathscr{U}_D(n)$ is known. Moreover, if the dimension of a composite system is not equal to a product of two primes, there are many ways of decomposing a unitary Ad-group into tensor products of Pauli's groups and diagonal unitary groups.

It is unclear how to determine if these decompositions are equivalent (in the sense that they describe just different realizations of one physical system) or that they depict different quantum systems.

# Bibliography


[1] J. Blank, P. Exner, M. Havlíček: *Lineární operátory v kvantové fyzice*, Karolinum, Praha 1993 (in Czech)

[2] J. Tolar: *Quantization Methods*, lecture notes, Universität Clausthal 1977

[3] G. G. Emch: *Mathematical and Conceptual Foundations of 20th-Century Physics*, North-Holland, Amsterdam 1984

[4] J. Pytlíček: *Lineární algebra a geometrie*, ČVUT, Praha 2008 (in Czech)

[5] O. Bratteli, D. W. Robinson: Operator Algebras and Quantum Statistical Mechanics 1, Springer, New York 1987

[6] M. Newman: *Two classical theorems on commuting matrices*, Journal of Research of the National Bureau of Standards, 71.B (1967), 69-71

[7] R. A. Horn, C. R. Johnson: *Matrix Analysis*, Cambridge University Press, Cambridge 1985

[8] S. Mac Lane, G. Birkhoff: *Algebra*, ALFA, Bratislava 1974 (in Slovak)

[9] M. Havlíček, J. Patera, E. Pelantová, J. Tolar: *Automorphisms of the fine gradings of* $sl(n, \mathbb{C})$ *associated with the generalized Pauli matrices*, Journal of Mathematical Physics 43 (2002), 1083-1094

[10] M. Havlíček, J. Patera, E. Pelantová: On Lie gradings II, Linear algebra and its applications 227 (1998), 97-125

[11] K. Szymiczek: *Bilinear Algebra: An Introduction to the Algebraic Theory of Quadratic Forms*, Overseas Publishers Association, Amsterdam 1997

[12] Yu. Bahturin, S. K. Sehgal, M. V. Zaicev: *Group gradings on associative algebras*, Journal of Algebra 241 (2001) 677-698

[13] H. Weyl, H.P. Robertson: *The Theory of Groups and Quantum Mechanics*, Dover Publications, New York, 1950

[14] P. Šťovíček, J. Tolar: *Quantum mechanics in a discrete space-time*, Reports on Mathematical Physics 20 (1984), 157-170

[15] J. Tolar: *A classification of finite quantum kinematics*, Journal of Physics: Conference Series 538 (2014)

[16] M. A. Nielsen, I. L. Chuang: *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge 2010