CS 766/QIC 820 Theory of Quantum Information (Fall 2011)

Lecture 1: Mathematical preliminaries (part 1)

Welcome to CS 766/QIC 820 Theory of Quantum Information. The goal of this lecture, as well as the next, is to present a brief overview of some of the basic mathematical concepts and tools that will be important in subsequent lectures of the course. In this lecture we will discuss various facts about linear algebra and analysis in finite-dimensional vector spaces.

1.1 Complex Euclidean spaces

We begin with the simple notion of a *complex Euclidean space*. As will be discussed later (in Lecture 3), we associate a complex Euclidean space with every discrete and finite physical system; and fundamental notions such as states and measurements of systems are represented in linear-algebraic terms that refer to these spaces.

1.1.1 Definition of complex Euclidean spaces

For any finite, nonempty set Σ, we denote by $\mathbb{C}^\Sigma$ the set of all functions from Σ to the complex numbers $\mathbb{C}$. The collection $\mathbb{C}^\Sigma$ forms a vector space of dimension $|\Sigma|$ over the complex numbers when addition and scalar multiplication are defined in the following standard way:

1. Addition: given $u, v \in \mathbb{C}^\Sigma$, the vector $u + v \in \mathbb{C}^\Sigma$ is defined by the equation $(u + v)(a) = u(a) + v(a)$ for all $a \in \Sigma$.
2. Scalar multiplication: given $u \in \mathbb{C}^\Sigma$ and $\alpha \in \mathbb{C}$, the vector $\alpha u \in \mathbb{C}^\Sigma$ is defined by the equation $(\alpha u)(a) = \alpha u(a)$ for all $a \in \Sigma$.

Any vector space defined in this way for some choice of a finite, nonempty set Σ will be called a *complex Euclidean space*.

Complex Euclidean spaces will generally be denoted by scripted capital letters near the end of the alphabet, such as $\mathcal{W}$, $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$, when it is necessary or helpful to assign specific names to them. Subsets of these spaces will also be denoted by scripted letters, and when possible our convention will be to use letters near the beginning of the alphabet, such as $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$, when these subsets are not themselves necessarily vector spaces. Vectors will typically be denoted by lowercase Roman letters, again near the end of the alphabet, such as u, v, w, x, y, and z.

In the case where $\Sigma = \{1, \ldots, n\}$ for some positive integer n, one typically writes $\mathbb{C}^n$ rather than $\mathbb{C}^{\{1,\ldots,n\}}$. For a given positive integer n, it is typical to view a vector $u \in \mathbb{C}^n$ as an n-tuple $u = (u_1, \ldots, u_n)$, or as a column vector of the form

$$u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}.$$

The convention to write u_i rather than $u(i)$ in such expressions is simply a matter of typographic appeal, and is avoided when it is not helpful or would lead to confusion, such as when vectors are subscripted for another purpose.

It is, of course, the case that one could simply identify $\mathbb{C}^\Sigma$ with $\mathbb{C}^n$, for $n = |\Sigma|$, with respect to any fixed choice of a bijection between Σ and $\{1, \ldots, n\}$. If it is convenient to make this simplifying assumption when proving facts about complex Euclidean spaces, we will do that; but there is also a significant convenience to be found in allowing for arbitrary (finite and nonempty) index sets, which is why we define complex Euclidean spaces in the way that we have.

1.1.2 Inner product and norms of vectors

The *inner product* $\langle u, v \rangle$ of vectors $u, v \in \mathbb{C}^\Sigma$ is defined as

$$\langle u, v \rangle = \sum_{a \in \Sigma} \overline{u(a)}\, v(a).$$

It may be verified that the inner product satisfies the following properties:

1. Linearity in the second argument: $\langle u, \alpha v + \beta w \rangle = \alpha \langle u, v \rangle + \beta \langle u, w \rangle$ for all $u, v, w \in \mathbb{C}^\Sigma$ and $\alpha, \beta \in \mathbb{C}$.
2. Conjugate symmetry: $\langle u, v \rangle = \overline{\langle v, u \rangle}$ for all $u, v \in \mathbb{C}^\Sigma$.
3. Positive definiteness: $\langle u, u \rangle \geq 0$ for all $u \in \mathbb{C}^\Sigma$, with $\langle u, u \rangle = 0$ if and only if $u = 0$.

One typically refers to any function satisfying these three properties as an inner product, but this is the only inner product for vectors in complex Euclidean spaces that is considered in this course.

The *Euclidean norm* of a vector $u \in \mathbb{C}^\Sigma$ is defined as

$$\|u\| = \sqrt{\langle u, u \rangle} = \sqrt{\sum_{a \in \Sigma} |u(a)|^2}.$$

The Euclidean norm satisfies the following properties, which are the defining properties of any function that is called a norm:

1. Positive definiteness: $\|u\| \geq 0$ for all $u \in \mathbb{C}^\Sigma$, with $\|u\| = 0$ if and only if $u = 0$.
2. Positive scalability: $\|\alpha u\| = |\alpha| \, \|u\|$ for all $u \in \mathbb{C}^\Sigma$ and $\alpha \in \mathbb{C}$.
3. The triangle inequality: $\|u + v\| \leq \|u\| + \|v\|$ for all $u, v \in \mathbb{C}^\Sigma$.

The Euclidean norm corresponds to the special case $p = 2$ of the class of *p-norms*, defined for each $u \in \mathbb{C}^\Sigma$ as

$$\|u\|_p = \left(\sum_{a \in \Sigma} |u(a)|^p \right)^{1/p}$$

for $1 \leq p < \infty$, and

$$\|u\|_\infty = \max\{|u(a)| \,:\, a \in \Sigma\}.$$

The above three norm properties (positive definiteness, positive scalability, and the triangle inequality) hold for $\|\cdot\|$ replaced by $\|\cdot\|_p$ for any choice of $p \in [1, \infty]$.

The *Cauchy-Schwarz inequality* states that

$$|\langle u, v \rangle| \leq \|u\| \, \|v\|$$

for all $u, v \in \mathbb{C}^\Sigma$, with equality if and only if u and v are linearly dependent. The Cauchy-Schwarz inequality is generalized by *Hölder's inequality*, which states that

$$|\langle u, v \rangle| \leq \|u\|_p \, \|v\|_q$$

for all $u, v \in \mathbb{C}^\Sigma$, provided $p, q \in [1, \infty]$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$ (with the interpretation $\frac{1}{\infty} = 0$).

1.1.3 Orthogonal and orthonormal sets

A collection of vectors

$$\{u_a : a \in \Gamma\} \subset \mathbb{C}^\Sigma,$$

indexed by a given finite, nonempty set Γ, is said to be an *orthogonal set* if it holds that $\langle u_a, u_b \rangle = 0$ for all choices of $a, b \in \Gamma$ with $a \neq b$. Such a set is necessarily linearly independent, provided it does not include the zero vector.

An orthogonal set of *unit* vectors is called an *orthonormal set*, and when such a set forms a basis it is called an orthonormal basis. It holds that an orthonormal set $\{u_a : a \in \Gamma\} \subseteq \mathbb{C}^\Sigma$ is an orthonormal basis of $\mathbb{C}^\Sigma$ if and only if $|\Gamma| = |\Sigma|$.

The *standard basis* of $\mathbb{C}^\Sigma$ is the orthonormal basis given by $\{e_a : a \in \Sigma\}$, where

$$e_a(b) = \begin{cases} 1 & \text{if } a = b \\ 0 & \text{if } a \neq b \end{cases}$$

for all $a, b \in \Sigma$.

Remark 1.1. When using the *Dirac notation*, one writes $|a\rangle$ rather than e_a when referring to standard basis elements; and for arbitrary vectors one writes $|u\rangle$ rather than u (although ϕ, ψ, and other Greek letters are much more commonly used to name vectors). We will generally not use Dirac notation in this course, because it tends to complicate the sorts of expressions we will encounter. One exception is the use of Dirac notation for the presentation of simple examples, where it seems to increase clarity.

1.1.4 Real Euclidean spaces

Real Euclidean spaces are defined in a similar way to complex Euclidean spaces, except that the field of complex numbers $\mathbb{C}$ is replaced by the field of real numbers $\mathbb{R}$ in each of the definitions and concepts in which it arises. Naturally, complex conjugation acts trivially in the real case, and may therefore be omitted.

Although complex Euclidean spaces will play a much more prominent role than real Euclidean spaces in this course, we will restrict our attention to real Euclidean spaces in the context of convexity theory. This will not limit the applicability of these concepts: they will generally be applied to the real Euclidean space consisting of all Hermitian operators acting on a given complex Euclidean space. Such spaces will be discussed later in this lecture.

1.2 Linear operators

Given complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, one writes $\mathrm{L}(\mathcal{X}, \mathcal{Y})$ to refer to the collection of all linear mappings of the form

$$A : \mathcal{X} \to \mathcal{Y}. \tag{1.1}$$

Such mappings will be referred to as *linear operators*, or simply *operators*, from $\mathcal{X}$ to $\mathcal{Y}$ in this course. Parentheses are typically omitted when expressing the action of linear operators on vectors when there is little chance of confusion in doing so. For instance, one typically writes Au rather than $A(u)$ to denote the vector resulting from the application of an operator $A \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$ to a vector $u \in \mathcal{X}$.

The set $\mathrm{L}(\mathcal{X}, \mathcal{Y})$ forms a vector space, where addition and scalar multiplication are defined as follows:

1. Addition: given $A, B \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$, the operator $A + B \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$ is defined by the equation

   $$(A + B)u = Au + Bu$$

   for all $u \in \mathcal{X}$.
2. Scalar multiplication: given $A \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$ and $\alpha \in \mathbb{C}$, the operator $\alpha A \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$ is defined by the equation

   $$(\alpha A)u = \alpha\, Au$$

   for all $u \in \mathcal{X}$.

The dimension of this vector space is given by $\dim(\mathrm{L}(\mathcal{X}, \mathcal{Y})) = \dim(\mathcal{X}) \dim(\mathcal{Y})$.

The *kernel* of an operator $A \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$ is the subspace of $\mathcal{X}$ defined as

$$\ker(A) = \{u \in \mathcal{X} \,:\, Au = 0\},$$

while the *image* of A is the subspace of $\mathcal{Y}$ defined as

$$\mathrm{im}(A) = \{Au \,:\, u \in \mathcal{X}\}.$$

The *rank* of A, denoted $\mathrm{rank}(A)$, is the dimension of the subspace $\mathrm{im}(A)$. For every operator $A \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$ it holds that

$$\dim(\ker(A)) + \mathrm{rank}(A) = \dim(\mathcal{X}).$$

1.2.1 Matrices and their association with operators

A *matrix* over the complex numbers is a mapping of the form

$$M : \Gamma \times \Sigma \to \mathbb{C}$$

for finite, nonempty sets Σ and Γ. The collection of all matrices of this form is denoted $\mathcal{M}_{\Gamma,\Sigma}(\mathbb{C})$. For $a \in \Gamma$ and $b \in \Sigma$ the value $M(a, b)$ is called the (a, b) *entry* of M, and the elements a and b are referred to as *indices* in this context: a is the *row index* and b is the *column index* of the entry $M(a, b)$.

The set $\mathcal{M}_{\Gamma,\Sigma}(\mathbb{C})$ is a vector space with respect to vector addition and scalar multiplication defined in the following way:

1. Addition: given $M, K \in \mathcal{M}_{\Gamma,\Sigma}(\mathbb{C})$, the matrix $M + K \in \mathcal{M}_{\Gamma,\Sigma}(\mathbb{C})$ is defined by the equation

   $$(M + K)(a, b) = M(a, b) + K(a, b)$$

   for all $a \in \Gamma$ and $b \in \Sigma$.

2. Scalar multiplication: given $M \in \mathcal{M}_{\Gamma,\Sigma}(\mathbb{C})$ and $\alpha \in \mathbb{C}$, the matrix $\alpha M \in \mathcal{M}_{\Gamma,\Sigma}(\mathbb{C})$ is defined by the equation
$$(\alpha M)(a,b) = \alpha M(a,b)$$
for all $a \in \Gamma$ and $b \in \Sigma$.

As a vector space, $\mathcal{M}_{\Gamma,\Sigma}(\mathbb{C})$ is therefore equivalent to the complex Euclidean space $\mathbb{C}^{\Gamma\times\Sigma}$.

Multiplication of matrices is defined in the following standard way. Given matrices $M \in \mathcal{M}_{\Gamma,\Delta}(\mathbb{C})$ and $K \in \mathcal{M}_{\Delta,\Sigma}(\mathbb{C})$, for finite nonempty sets Γ, Δ, and Σ, the matrix $MK \in \mathcal{M}_{\Gamma,\Sigma}(\mathbb{C})$ is defined as
$$(MK)(a,b) = \sum_{c\in\Delta} M(a,c)K(c,b)$$
for all $a \in \Gamma$ and $b \in \Sigma$.

Linear operators from one complex Euclidean space to another are naturally represented by matrices. For $\mathcal{X} = \mathbb{C}^{\Sigma}$ and $\mathcal{Y} = \mathbb{C}^{\Gamma}$, one associates with each operator $A \in \mathrm{L}(\mathcal{X},\mathcal{Y})$ a matrix $M_A \in \mathcal{M}_{\Gamma,\Sigma}(\mathbb{C})$ defined as
$$M_A(a,b) = \langle e_a, Ae_b\rangle$$
for each $a \in \Gamma$ and $b \in \Sigma$. Conversely, to each matrix $M \in \mathcal{M}_{\Gamma,\Sigma}(\mathbb{C})$ one associates a linear operator $A_M \in \mathrm{L}(\mathcal{X},\mathcal{Y})$ defined by
$$(A_M u)(a) = \sum_{b\in\Sigma} M(a,b)u(b) \tag{1.2}$$
for each $a \in \Gamma$. The mappings $A \mapsto M_A$ and $M \mapsto A_M$ are linear and inverse to one other, and compositions of linear operators are represented by matrix multiplications: $M_{AB} = M_A M_B$ whenever $A \in \mathrm{L}(\mathcal{Y},\mathcal{Z})$, $B \in \mathrm{L}(\mathcal{X},\mathcal{Y})$ and $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ are complex Euclidean spaces. Equivalently, $A_{MK} = A_M A_K$ for any choice of matrices $M \in \mathcal{M}_{\Gamma,\Delta}(\mathbb{C})$ and $K \in \mathcal{M}_{\Delta,\Sigma}(\mathbb{C})$ for finite nonempty sets Σ, Δ, and Γ.

This correspondence between linear operators and matrices will hereafter not be mentioned explicitly in these notes: we will freely switch between speaking of operators and speaking of matrices, depending on which is more suitable within the context at hand. A preference will generally be given to speak of operators, and to implicitly associate a given operator's matrix representation with it as necessary. More specifically, for a given choice of complex Euclidean spaces $\mathcal{X} = \mathbb{C}^{\Sigma}$ and $\mathcal{Y} \in \mathbb{C}^{\Gamma}$, and for a given operator $A \in \mathrm{L}(\mathcal{X},\mathcal{Y})$, the matrix $M_A \in \mathcal{M}_{\Gamma,\Sigma}(\mathbb{C})$ will simply be denoted A and its (a,b)-entry as $A(a,b)$.

1.2.2 The entry-wise conjugate, transpose, and adjoint

For every operator $A \in \mathrm{L}(\mathcal{X},\mathcal{Y})$, for complex Euclidean spaces $\mathcal{X} = \mathbb{C}^{\Sigma}$ and $\mathcal{Y} = \mathbb{C}^{\Gamma}$, one defines three additional operators,
$$\overline{A} \in \mathrm{L}(\mathcal{X},\mathcal{Y}) \quad \text{and} \quad A^{\mathsf{T}}, A^{*} \in \mathrm{L}(\mathcal{Y},\mathcal{X}),$$
as follows:

1. The operator $\overline{A} \in \mathrm{L}(\mathcal{X},\mathcal{Y})$ is the operator whose matrix representation has entries that are complex conjugates to the matrix representation of A:
$$\overline{A}(a,b) = \overline{A(a,b)}$$
for all $a \in \Gamma$ and $b \in \Sigma$.

2. The operator $A^{\mathsf{T}} \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ is the operator whose matrix representation is obtained by *transposing* the matrix representation of A:

$$A^{\mathsf{T}}(b, a) = A(a, b)$$

for all $a \in \Gamma$ and $b \in \Sigma$.

3. The operator $A^* \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ is the unique operator that satisfies the equation

$$\langle v, Au \rangle = \langle A^* v, u \rangle$$

for all $u \in \mathcal{X}$ and $v \in \mathcal{Y}$. It may be obtained by performing both of the operations described in items 1 and 2:

$$A^* = \overline{A^{\mathsf{T}}}.$$

The operators $\overline{A}$, A^{T}, and A^* will be called the *entry-wise conjugate*, *transpose*, and *adjoint* operators to A, respectively.

The mappings $A \mapsto \overline{A}$ and $A \mapsto A^*$ are conjugate linear and the mapping $A \mapsto A^{\mathsf{T}}$ is linear:

$$\begin{aligned}
\overline{\alpha A + \beta B} &= \overline{\alpha}\,\overline{A} + \overline{\beta}\,\overline{B}, \\
(\alpha A + \beta B)^* &= \overline{\alpha} A^* + \overline{\beta} B^*, \\
(\alpha A + \beta B)^{\mathsf{T}} &= \alpha A^{\mathsf{T}} + \beta B^{\mathsf{T}},
\end{aligned}$$

for all $A, B \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$ and $\alpha, \beta \in \mathbb{C}$. These mappings are bijections, each being its own inverse.

Every vector $u \in \mathcal{X}$ in a complex Euclidean space $\mathcal{X}$ may be identified with the linear operator in $\mathrm{L}(\mathbb{C}, \mathcal{X})$ that maps $\alpha \mapsto \alpha u$. Through this identification the linear mappings $\overline{u} \in \mathrm{L}(\mathbb{C}, \mathcal{X})$ and $u^{\mathsf{T}}, u^* \in \mathrm{L}(\mathcal{X}, \mathbb{C})$ are defined as above. As an element of $\mathcal{X}$, the vector $\overline{u}$ is of course simply the entry-wise complex conjugate of u, i.e., if $\mathcal{X} = \mathbb{C}^{\Sigma}$ then

$$\overline{u}(a) = \overline{u(a)}$$

for every $a \in \Sigma$. For each vector $u \in \mathcal{X}$ the mapping $u^* \in \mathrm{L}(\mathcal{X}, \mathbb{C})$ satisfies $u^* v = \langle u, v \rangle$ for all $v \in \mathcal{X}$. The space of linear operators $\mathrm{L}(\mathcal{X}, \mathbb{C})$ is called the *dual space* of $\mathcal{X}$, and is often denoted by $\mathcal{X}^*$ rather than $\mathrm{L}(\mathcal{X}, \mathbb{C})$.

Assume that $\mathcal{X} = \mathbb{C}^{\Sigma}$ and $\mathcal{Y} = \mathbb{C}^{\Gamma}$. For each choice of $a \in \Gamma$ and $b \in \Sigma$, the operator $E_{a,b} \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$ is defined as $E_{a,b} = e_a e_b^*$, or equivalently

$$E_{a,b}(c, d) = \begin{cases} 1 & \text{if } (a = c) \text{ and } (b = d) \\ 0 & \text{if } (a \neq c) \text{ or } (b \neq d). \end{cases}$$

The set $\{E_{a,b} : a \in \Gamma,\ b \in \Sigma\}$ is a basis of $\mathrm{L}(\mathcal{X}, \mathcal{Y})$, and will be called the *standard basis* of this space.

1.2.3 Direct sums

The *direct sum* of n complex Euclidean spaces $\mathcal{X}_1 = \mathbb{C}^{\Sigma_1}, \ldots, \mathcal{X}_n = \mathbb{C}^{\Sigma_n}$ is the complex Euclidean space

$$\mathcal{X}_1 \oplus \cdots \oplus \mathcal{X}_n = \mathbb{C}^{\Delta},$$

where

$$\Delta = \{(1, a_1) : a_1 \in \Sigma_1\} \cup \cdots \cup \{(n, a_n) : a_n \in \Sigma_n\}.$$

One may view Δ as the *disjoint union* of $\Sigma_1, \ldots, \Sigma_n$.

For vectors $u_1 \in \mathcal{X}_1, \ldots, u_n \in \mathcal{X}_n$, the notation $u_1 \oplus \cdots \oplus u_n \in \mathcal{X}_1 \oplus \cdots \oplus \mathcal{X}_n$ refers to the vector for which

$$(u_1 \oplus \cdots \oplus u_n)(j, a_j) = u_j(a_j),$$

for each $j \in \{1, \ldots, n\}$ and $a_j \in \Sigma_j$. If each vector u_j is viewed as a column vector of dimension $\left|\Sigma_j\right|$, the vector $u_1 \oplus \cdots \oplus u_n$ may be viewed as a (block) column vector

$$\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$$

having dimension $|\Sigma_1| + \cdots + |\Sigma_n|$. Every element of the space $\mathcal{X}_1 \oplus \cdots \oplus \mathcal{X}_n$ can be written as $u_1 \oplus \cdots \oplus u_n$ for a unique choice of vectors $u_1, \ldots, u_n$. The following identities hold for every choice of $u_1, v_1 \in \mathcal{X}_1, \ldots, u_n, v_n \in \mathcal{X}_n$, and $\alpha \in \mathbb{C}$:

$$u_1 \oplus \cdots \oplus u_n + v_1 \oplus \cdots \oplus v_n = (u_1 + v_1) \oplus \cdots \oplus (u_n + v_n)$$

$$\alpha(u_1 \oplus \cdots \oplus u_n) = (\alpha u_1) \oplus \cdots \oplus (\alpha u_n)$$

$$\langle u_1 \oplus \cdots \oplus u_n, v_1 \oplus \cdots \oplus v_n \rangle = \langle u_1, v_1 \rangle + \cdots + \langle u_n, v_n \rangle .$$

Now suppose that $\mathcal{X}_1 = \mathbb{C}^{\Sigma_1}, \ldots, \mathcal{X}_n = \mathbb{C}^{\Sigma_n}$ and $\mathcal{Y}_1 = \mathbb{C}^{\Gamma_1}, \ldots, \mathcal{Y}_m = \mathbb{C}^{\Gamma_m}$ for positive integers n and m, and finite, nonempty sets $\Sigma_1, \ldots, \Sigma_n$ and $\Gamma_1, \ldots, \Gamma_m$. The matrix associated with a given operators of the form $A \in \mathrm{L}\left(\mathcal{X}_1 \oplus \cdots \oplus \mathcal{X}_n, \mathcal{Y}_1 \oplus \cdots \oplus \mathcal{Y}_m\right)$ may be identified with a block matrix

$$A = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & \ddots & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix},$$

where $A_{j,k} \in \mathrm{L}\left(\mathcal{X}_k, \mathcal{Y}_j\right)$ for each $j \in \{1, \ldots, m\}$ and $k \in \{1, \ldots, n\}$. These are the uniquely determined operators for which it holds that

$$A(u_1 \oplus \cdots \oplus u_n) = v_1 \oplus \cdots \oplus v_m,$$

for $v_1 \in \mathcal{Y}_1, \ldots, v_m \in \mathcal{Y}_m$ defined as

$$v_j = (A_{j,1} u_1) + \cdots + (A_{j,n} u_n)$$

for each $j \in \{1, \ldots, m\}$.

1.2.4 Tensor products

The *tensor product* of $\mathcal{X}_1 = \mathbb{C}^{\Sigma_1}, \ldots, \mathcal{X}_n = \mathbb{C}^{\Sigma_n}$ is the complex Euclidean space

$$\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n = \mathbb{C}^{\Sigma_1 \times \cdots \times \Sigma_n}.$$

For vectors $u_1 \in \mathcal{X}_1, \ldots, u_n \in \mathcal{X}_n$, the vector $u_1 \otimes \cdots \otimes u_n \in \mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n$ is defined as

$$(u_1 \otimes \cdots \otimes u_n)(a_1, \ldots, a_n) = u_1(a_1) \cdots u_n(a_n).$$

Vectors of the form $u_1 \otimes \cdots \otimes u_n$ are called *elementary tensors*. They span the space $\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n$, but not every element of $\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n$ is an elementary tensor.

The following identities hold for every choice of $u_1, v_1 \in \mathcal{X}_1, \ldots, u_n, v_n \in \mathcal{X}_n$, $\alpha \in \mathbb{C}$, and $k \in \{1, \ldots, n\}$:

$$\begin{aligned} u_1 \otimes \cdots \otimes u_{k-1} \otimes (u_k + v_k) \otimes u_{k+1} \otimes \cdots \otimes u_n & \\ = u_1 \otimes \cdots \otimes u_{k-1} \otimes u_k \otimes u_{k+1} \otimes \cdots \otimes u_n & \\ + u_1 \otimes \cdots \otimes u_{k-1} \otimes v_k \otimes u_{k+1} \otimes \cdots \otimes u_n & \end{aligned}$$

$$\alpha(u_1 \otimes \cdots \otimes u_n) = (\alpha u_1) \otimes u_2 \otimes \cdots \otimes u_n = \cdots = u_1 \otimes u_2 \otimes \cdots \otimes u_{n-1} \otimes (\alpha u_n)$$

$$\langle u_1 \otimes \cdots \otimes u_n, v_1 \otimes \cdots \otimes v_n \rangle = \langle u_1, v_1 \rangle \cdots \langle u_n, v_n \rangle .$$

It is worthwhile to note that the definition of tensor products just presented is a concrete definition that is sometimes known as the *Kronecker product*. In contrast, tensor products are often defined in a more abstract way that stresses their close connection to *multilinear functions*. There is valuable intuition to be drawn from this connection, but for our purposes it will suffice that we take note of the following fact.

Proposition 1.2. *Let $\mathcal{X}_1, \ldots, \mathcal{X}_n$ and $\mathcal{Y}$ be complex Euclidean spaces, and let $\phi : \mathcal{X}_1 \times \cdots \times \mathcal{X}_n \to \mathcal{Y}$ be a multilinear function (i.e., a function for which the mapping $u_j \mapsto \phi(u_1, \ldots, u_n)$ is linear for all $j \in \{1, \ldots, n\}$ and all choices of $u_1, \ldots, u_{j-1}, u_{j+1}, \ldots, u_n$. It holds that there exists an operator $A \in \mathrm{L}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n, \mathcal{Y})$ for which*

$$\phi(u_1, \ldots, u_n) = A(u_1 \otimes \cdots \otimes u_n).$$

1.3 Algebras of operators

For every complex Euclidean space $\mathcal{X}$, the notation $\mathrm{L}(\mathcal{X})$ is understood to be a shorthand for $\mathrm{L}(\mathcal{X}, \mathcal{X})$. The space $\mathrm{L}(\mathcal{X})$ has special algebraic properties that are worthy of note. In particular, $\mathrm{L}(\mathcal{X})$ is an *associative algebra*; it is a vector space, and the composition of operators is associative and bilinear:

$$\begin{aligned} (AB)C &= A(BC), \\ C(\alpha A + \beta B) &= \alpha CA + \beta CB, \\ (\alpha A + \beta B)C &= \alpha AC + \beta BC, \end{aligned}$$

for every choice of $A, B, C \in \mathrm{L}(\mathcal{X})$ and $\alpha, \beta \in \mathbb{C}$.

The identity operator $\mathbb{1} \in \mathrm{L}(\mathcal{X})$ is the operator defined as $\mathbb{1}u = u$ for all $u \in \mathcal{X}$, and is denoted $\mathbb{1}_{\mathcal{X}}$ when it is helpful to indicate explicitly that it acts on $\mathcal{X}$. An operator $A \in \mathrm{L}(\mathcal{X})$ is *invertible* if there exists an operator $B \in \mathrm{L}(\mathcal{X})$ such that $BA = \mathbb{1}$. When such an operator B exists it is necessarily unique, also satisfies $AB = \mathbb{1}$, and is denoted A^{-1}. The collection of all invertible operators in $\mathrm{L}(\mathcal{X})$ is denoted $\mathrm{GL}(\mathcal{X})$, and is called the *general linear group* of $\mathcal{X}$.

For every pair of operators $A, B \in \mathrm{L}(\mathcal{X})$, the *Lie bracket* $[A, B] \in \mathrm{L}(\mathcal{X})$ is defined as $[A, B] = AB - BA$.

1.3.1 Trace and determinant

Operators in the algebra $\mathrm{L}(\mathcal{X})$ are represented by *square* matrices, which means that their rows and columns are indexed by the same set. We define two important functions from $\mathrm{L}(\mathcal{X})$ to $\mathbb{C}$, the *trace* and the *determinant*, based on matrix representations of operators as follows:

1. The *trace* of an operator $A \in \mathrm{L}(\mathcal{X})$, for $\mathcal{X} = \mathbb{C}^{\Sigma}$, is defined as
$$\mathrm{Tr}(A) = \sum_{a \in \Sigma} A(a, a).$$
2. The *determinant* of an operator $A \in \mathrm{L}(\mathcal{X})$, for $\mathcal{X} = \mathbb{C}^{\Sigma}$, is defined by the equation
$$\mathrm{Det}(A) = \sum_{\pi \in Sym(\Sigma)} \mathrm{sign}(\pi) \prod_{a \in \Sigma} A(a, \pi(a)),$$
where $Sym(\Sigma)$ is the group of permutations on the set Σ and $\mathrm{sign}(\pi)$ is the *sign* of the permutation π (which is $+1$ if π is expressible as a product of an even number of transpositions of elements of the set Σ, and -1 if π is expressible as a product of an odd number of transpositions).

The trace is a linear function, and possesses the property that
$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$$
for any choice of operators $A \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$ and $B \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$, for arbitrary complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$.

By means of the trace, one defines an inner product on the space $\mathrm{L}(\mathcal{X}, \mathcal{Y})$, for any choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, as
$$\langle A, B \rangle = \mathrm{Tr}\left(A^* B\right)$$
for all $A, B \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$. It may be verified that this inner product satisfies the requisite properties of being an inner product:

1. Linearity in the second argument:
$$\langle A, \alpha B + \beta C \rangle = \alpha \langle A, B \rangle + \beta \langle A, C \rangle$$
for all $A, B, C \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$ and $\alpha, \beta \in \mathbb{C}$.
2. Conjugate symmetry: $\langle A, B \rangle = \overline{\langle B, A \rangle}$ for all $A, B \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$.
3. Positive definiteness: $\langle A, A \rangle \geq 0$ for all $A \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$, with $\langle A, A \rangle = 0$ if and only if $A = 0$.

This inner product is sometimes called the *Hilbert–Schmidt inner product*.

The determinant is multiplicative,
$$\mathrm{Det}(AB) = \mathrm{Det}(A)\,\mathrm{Det}(B)$$
for all $A, B \in \mathrm{L}(\mathcal{X})$, and its value is nonzero if and only if its argument is invertible.

1.3.2 Eigenvectors and eigenvalues

If $A \in \mathrm{L}(\mathcal{X})$ and $u \in \mathcal{X}$ is a nonzero vector such that $Au = \lambda u$ for some choice of $\lambda \in \mathbb{C}$, then u is said to be an *eigenvector* of A and λ is its corresponding *eigenvalue*.

For every operator $A \in \mathrm{L}(\mathcal{X})$, one has that

$$p_A(z) = \mathrm{Det}(z\mathbb{1}_{\mathcal{X}} - A)$$

is a monic polynomial in z having degree $\dim(\mathcal{X})$. This polynomial is the *characteristic polynomial* of A. The *spectrum* of A, denoted $\mathrm{spec}(A)$, is the multiset containing the roots of the polynomial $p_A(z)$, with each root appearing a number of times equal to its multiplicity. As p_A is monic, it holds that

$$p_A(z) = \prod_{\lambda \in \mathrm{spec}(A)} (z - \lambda)$$

Each element $\lambda \in \mathrm{spec}(A)$ is an eigenvalue of A.

The trace and determinant may be expressed in terms of the spectrum as follows:

$$\mathrm{Tr}(A) = \sum_{\lambda \in \mathrm{spec}(A)} \lambda$$

and

$$\mathrm{Det}(A) = \prod_{\lambda \in \mathrm{spec}(A)} \lambda$$

for every $A \in \mathrm{L}(\mathcal{X})$.

1.4 Important classes of operators

A collection of classes of operators that have importance in quantum information are discussed in this section.

1.4.1 Normal operators

An operator $A \in \mathrm{L}(\mathcal{X})$ is *normal* if and only if it commutes with its adjoint: $[A, A^*] = 0$, or equivalently $AA^* = A^*A$. The importance of this collection of operators, for the purposes of this course, is mainly derived from two facts: (1) the normal operators are those for which the spectral theorem (discussed later in Section 1.5) holds, and (2) most of the special classes of operators that are discussed below are subsets of the normal operators.

1.4.2 Hermitian operators

An operator $A \in \mathrm{L}(\mathcal{X})$ is *Hermitian* if $A = A^*$. The set of Hermitian operators acting on a given complex Euclidean space $\mathcal{X}$ will hereafter be denoted $\mathrm{Herm}(\mathcal{X})$ in this course:

$$\mathrm{Herm}(\mathcal{X}) = \{A \in \mathrm{L}(\mathcal{X}) \,:\, A = A^*\}.$$

Every Hermitian operator is obviously a normal operator.

The eigenvalues of every Hermitian operator are necessarily real numbers, and can therefore be ordered from largest to smallest. Under the assumption that $A \in \mathrm{Herm}(\mathcal{X})$ for $\mathcal{X}$ an n-dimensional complex Euclidean space, one denotes the k-th largest eigenvalue of A by $\lambda_k(A)$. Equivalently, the vector

$$\lambda(A) = (\lambda_1(A), \lambda_2(A), \ldots, \lambda_n(A)) \in \mathbb{R}^n$$

is defined so that

$$\mathrm{spec}(A) = \{\lambda_1(A), \lambda_2(A), \ldots, \lambda_n(A)\}$$

and

$$\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_n(A).$$

The sum of two Hermitian operators is obviously Hermitian, as is any real scalar multiple of a Hermitian operator. This means that the set $\mathrm{Herm}(\mathcal{X})$ forms a vector space over the real numbers. The inner product of two Hermitian operators is real as well, $\langle A, B\rangle \in \mathbb{R}$ for all $A, B \in \mathrm{Herm}(\mathcal{X})$, so this space is in fact a real inner product space.

We can, in fact, go a little bit further along these lines. Assuming that $\mathcal{X} = \mathbb{C}^\Sigma$, and that the elements of Σ are ordered in some fixed way, let us define a Hermitian operator $H_{a,b} \in \mathrm{Herm}(\mathcal{X})$, for each choice of $a, b \in \Sigma$, as follows:

$$H_{a,b} = \begin{cases} E_{a,a} & \text{if } a = b \\ \frac{1}{\sqrt{2}}(E_{a,b} + E_{b,a}) & \text{if } a < b \\ \frac{1}{\sqrt{2}}(iE_{a,b} - iE_{b,a}) & \text{if } a > b. \end{cases}$$

The collection $\{H_{a,b} : a, b \in \Sigma\}$ is orthonormal (with respect to the inner product defined on $\mathrm{L}(\mathcal{X})$), and every Hermitian operator $A \in \mathrm{Herm}(\mathcal{X})$ can be expressed as a real linear combination of matrices in this collection. It follows that $\mathrm{Herm}(\mathcal{X})$ is a vector space of dimension $|\Sigma|^2$ over the real numbers, and that there exists an isometric isomorphism between $\mathrm{Herm}(\mathcal{X})$ and $\mathbb{R}^{\Sigma\times\Sigma}$. This fact will allow us to apply facts about convex analysis, which typically hold for real Euclidean spaces, to $\mathrm{Herm}(\mathcal{X})$ (as will be discussed in the next lecture).

1.4.3 Positive semidefinite operators

An operator $A \in \mathrm{L}(\mathcal{X})$ is *positive semidefinite* if and only if it holds that $A = B^*B$ for some operator $B \in \mathrm{L}(\mathcal{X})$. Hereafter, when it is reasonable to do so, a convention to use the symbols P, Q and R to denote general positive semidefinite matrices will be followed. The collection of positive semidefinite operators acting on $\mathcal{X}$ is denoted $\mathrm{Pos}(\mathcal{X})$, so that

$$\mathrm{Pos}(\mathcal{X}) = \{B^*B : B \in \mathrm{L}(\mathcal{X})\}.$$

There are alternate ways to describe positive semidefinite operators that are useful in different situations. In particular, the following items are equivalent for a given operator $P \in \mathrm{L}(\mathcal{X})$:

1. P is positive semidefinite.
2. $P = B^*B$ for some choice of a complex Euclidean space $\mathcal{Y}$ and an operator $B \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$.
3. u^*Pu is a nonnegative real number for every choice of $u \in \mathcal{X}$.
4. $\langle Q, P\rangle$ is a nonnegative real number for every $Q \in \mathrm{Pos}(\mathcal{X})$.

5. P is Hermitian and every eigenvalue of P is nonnegative.
6. There exists a complex Euclidean space $\mathcal{Y}$ and a collection of vectors $\{u_a : a \in \Sigma\} \subset \mathcal{Y}$, such that $P(a,b) = \langle u_a, u_b \rangle$.

Item 6 remains valid if the additional constraint $\dim(\mathcal{Y}) = \dim(\mathcal{X})$ is imposed.

The notation $P \geq 0$ is also used to mean that P is positive semidefinite, while $A \geq B$ means that $A - B$ is positive semidefinite. (This notation is only used when A and B are both Hermitian.)

1.4.4 Positive definite operators

A positive semidefinite operator $P \in \mathrm{Pos}(\mathcal{X})$ is said to be *positive definite* if, in addition to being positive semidefinite, it is invertible. The notation

$$\mathrm{Pd}(\mathcal{X}) = \{P \in \mathrm{Pos}(\mathcal{X}) : \mathrm{Det}(P) \neq 0\}$$

will be used to denote the set of such operators for a given complex Euclidean space $\mathcal{X}$. The following items are equivalent for a given operator $P \in \mathrm{L}(\mathcal{X})$:

1. P is positive definite.
2. $\langle u, Pu \rangle$ is a positive real number for every choice of a nonzero vector $u \in \mathcal{X}$.
3. P is Hermitian, and every eigenvalue of P is positive.
4. P is Hermitian, and there exists a positive real number $\varepsilon > 0$ such that $P \geq \varepsilon \mathbb{1}$.

1.4.5 Density operators

Positive semidefinite operators having trace equal to 1 are called *density operators*, and it is conventional to use lowercase Greek letters such as ρ, ξ, and σ to denote such operators. The notation

$$\mathrm{D}(\mathcal{X}) = \{\rho \in \mathrm{Pos}(\mathcal{X}) : \mathrm{Tr}(\rho) = 1\}$$

is used to denote the collection of density operators acting on a given complex Euclidean space.

1.4.6 Orthogonal projections

A positive semidefinite operator $P \in \mathrm{Pos}(\mathcal{X})$ is an *orthogonal projection* if, in addition to being positive semidefinite, it satisfies $P^2 = P$. Equivalently, an orthogonal projection is any Hermitian operator whose only eigenvalues are 0 and 1. For each subspace $\mathcal{V} \subseteq \mathcal{X}$, we write $\Pi_{\mathcal{V}}$ to denote the unique orthogonal projection whose image is equal to the subspace $\mathcal{V}$.

It is typically that the term *projection* refers to an operator $A \in \mathrm{L}(\mathcal{X})$ that satisfies $A^2 = A$, but which might not be Hermitian. Given that there is no discussion of such operators in this course, we will use the term *projection* to mean *orthogonal projection*.

1.4.7 Linear isometries and unitary operators

An operator $A \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$ is a *linear isometry* if it preserves the Euclidean norm—meaning that $\|Au\| = \|u\|$ for all $u \in \mathcal{X}$. The condition that $\|Au\| = \|u\|$ for all $u \in \mathcal{X}$ is equivalent to $A^*A = \mathbb{1}_{\mathcal{X}}$. The notation

$$\mathrm{U}(\mathcal{X}, \mathcal{Y}) = \{A \in \mathrm{L}(\mathcal{X}, \mathcal{Y}) : A^*A = \mathbb{1}_{\mathcal{X}}\}$$

is used throughout this course. Every linear isometry preserves not only the Euclidean norm, but inner products as well: $\langle Au, Av \rangle = \langle u, v \rangle$ for all $u, v \in \mathcal{X}$.

The set of linear isometries mapping $\mathcal{X}$ to itself is denoted $\mathrm{U}(\mathcal{X})$, and operators in this set are called *unitary operators*. The letters U, V, and W are conventionally used to refer to unitary operators. Every unitary operator $U \in \mathrm{U}(\mathcal{X})$ is invertible and satisfies $UU^* = U^*U = \mathbb{1}_{\mathcal{X}}$, which implies that every unitary operator is normal.

1.5 The spectral theorem

The *spectral theorem* establishes that every *normal* operator can be expressed as a linear combination of projections onto pairwise orthogonal subspaces. The spectral theorem is so-named, and the resulting expressions are called spectral decompositions, because the coefficients of the projections are determined by the spectrum of the operator being considered.

1.5.1 Statement of the spectral theorem and related facts

A formal statement of the spectral theorem follows.

Theorem 1.3 (Spectral theorem)**.** *Let $\mathcal{X}$ be a complex Euclidean space, let $A \in \mathrm{L}(\mathcal{X})$ be a normal operator, and assume that the distinct eigenvalues of A are $\lambda_1, \ldots, \lambda_k$. There exists a unique choice of orthogonal projection operators $P_1, \ldots, P_k \in \mathrm{Pos}(\mathcal{X})$, with $P_1 + \cdots + P_k = \mathbb{1}_{\mathcal{X}}$ and $P_i P_j = 0$ for $i \neq j$, such that*

$$A = \sum_{i=1}^{k} \lambda_i P_i. \tag{1.3}$$

For each $i \in \{1, \ldots, k\}$, it holds that the rank of P_i is equal to the multiplicity of λ_i as an eigenvalue of A.

As suggested above, the expression of a normal operator A in the form of the above equation (1.3) is called a *spectral decomposition* of A.

A simple corollary of the spectral theorem follows. It expresses essentially the same fact as the spectral theorem, but in a slightly different form that will be useful to refer to later in the course.

Corollary 1.4. *Let $\mathcal{X}$ be a complex Euclidean space, let $A \in \mathrm{L}(\mathcal{X})$ be a normal operator, and assume that* $\mathrm{spec}(A) = \{\lambda_1, \ldots, \lambda_n\}$*. There exists an orthonormal basis $\{x_1, \ldots, x_n\}$ of $\mathcal{X}$ such that*

$$A = \sum_{i=1}^{n} \lambda_i x_i x_i^*. \tag{1.4}$$

It is clear from the expression (1.4), along with the requirement that the set $\{x_1, \ldots, x_n\}$ is an orthonormal basis, that each x_i is an eigenvector of A whose corresponding eigenvalue is λ_i. It is also clear that any operator A that is expressible in such a form as (1.4) is normal—implying that the condition of normality is equivalent to the existence of an orthonormal basis of eigenvectors.

We will often refer to expressions of operators in the form (1.4) as *spectral decompositions*, despite the fact that it differs slightly from the form (1.3). It must be noted that unlike the form (1.3), the form (1.4) is generally not unique (unless each eigenvalue of A has multiplicity one, in which case the expression is unique up to scalar multiples of the vectors $\{x_1, \ldots, x_n\}$).

Finally, let us mention one more important theorem regarding spectral decompositions of normal operators, which states that the same orthonormal basis of eigenvectors $\{x_1, \ldots, x_n\}$ may be chosen for any two normal operators, provided that they commute.

Theorem 1.5. *Let $\mathcal{X}$ be a complex Euclidean space and let $A, B \in \mathrm{L}(\mathcal{X})$ be normal operators for which $[A, B] = 0$. There exists an orthonormal basis $\{x_1, \ldots, x_n\}$ of $\mathcal{X}$ such that*

$$A = \sum_{i=1}^{n} \lambda_i x_i x_i^* \qquad \text{and} \qquad B = \sum_{i=1}^{n} \mu_i x_i x_i^*$$

are spectral decompositions of A and B, respectively.

1.5.2 Functions of normal operators

Every function of the form $f : \mathbb{C} \to \mathbb{C}$ may be extended to the set of normal operators in $\mathrm{L}(\mathcal{X})$, for a given complex Euclidean space $\mathcal{X}$, by means of the spectral theorem. In particular, if $A \in \mathrm{L}(\mathcal{X})$ is normal and has the spectral decomposition (1.3), then one defines

$$f(A) = \sum_{i=1}^{k} f(\lambda_i) P_i.$$

Naturally, functions defined only on subsets of scalars may be extended to normal operators whose eigenvalues are restricted accordingly. A few examples of scalar functions extended to operators that will be important later in the course follow.

The exponential function of an operator

The exponential function $\alpha \mapsto \exp(\alpha)$ is defined for all $\alpha \in \mathbb{C}$, and may therefore be extended to a function $A \mapsto \exp(A)$ for any normal operator $A \in \mathrm{L}(\mathcal{X})$ by defining

$$\exp(A) = \sum_{i=1}^{k} \exp(\lambda_i) P_i,$$

assuming that the spectral decomposition of A is given by (1.3).

The exponential function may, in fact, be defined for all operators $A \in \mathrm{L}(\mathcal{X})$ by considering its usual Taylor series. In particular, the series

$$\exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

can be shown to converge for all operators $A \in \mathrm{L}(\mathcal{X})$, and agrees with the above notion based on the spectral decomposition in the case that A is normal.

Non-integer powers of operators

For $r > 0$ the function $\lambda \mapsto \lambda^r$ is defined for nonnegative real values $\lambda \in [0, \infty)$. For a given positive semidefinite operator $Q \in \mathrm{Pos}(\mathcal{X})$ having spectral decomposition (1.3), for which we necessarily have that $\lambda_i \geq 0$ for $1 \leq i \leq k$, we may therefore define

$$Q^r = \sum_{i=1}^{k} \lambda_i^r P_i.$$

For integer values of r, it is clear that Q^r coincides with the usual meaning of this expression given by the multiplication of operators. The case that $r = 1/2$ is particularly common, and in

this case we also write $\sqrt{Q}$ to denote $Q^{1/2}$. The operator $\sqrt{Q}$ is the unique positive semidefinite operator that satisfies

$$\sqrt{Q}\sqrt{Q} = Q.$$

Along similar lines, for any real number $r < 0$, the function $\lambda \mapsto \lambda^r$ is defined for positive real values $\lambda \in (0, \infty)$. For a given positive definite operator $Q \in \mathrm{Pd}(\mathcal{X})$, one defines Q^r in a similar way to above.

The logarithm of an operator

The function $\lambda \mapsto \log(\lambda)$ is defined for every positive real number $\lambda \in (0, \infty)$. For a given positive definite operator $Q \in \mathrm{Pd}(\mathcal{X})$, having a spectral decomposition (1.3) as above, one defines

$$\log(Q) = \sum_{i=1}^{k} \log(\lambda_i) P_i.$$

Logarithms of operators will be important during our discussion of von Neumann entropy.