EXERCISES
NOTES
IMPORTANT
PROOF

Lecture 1

Alexandru Petrescu[1,*]

[1] *LPENS, Département de physique, Ecole normale supérieure, Centre Automatique et Systèmes (CAS), MINES Paris, Université PSL, Sorbonne Université, CNRS, Inria, 75005 Paris, France*

(Dated: September 12, 2023. Lecture of September 13, 2022; September 12, 2023.)

The first part of the course will cover basic notions of quantum mechanics, with the aim of providing the essential concepts necessary for practical calculations. One important prerequisite to quantum mechanics is linear algebra. We will begin from this, with the aim of covering the postulates of quantum mechanics by the end of this lecture. In this first part of the course, we draw heavily upon Chapter 2 of Nielsen and Chuang, *Quantum computation and quantum information*, Cambridge University Press (2010).

I. LINEAR ALGEBRA

The basic objects of linear algebra are *vector spaces*. For quantum mechanics, the vector space $\mathbb{C}^n$ is of interest. This is made up of n-tuples, or *vectors*, of complex numbers, or *c-numbers*, $(z_1, \ldots, z_n)$. The following properties define the vector space:

- Addition takes a pair of vectors to another vector in the vector space, which is determined by

$$\begin{pmatrix} z_1 \\ z_2 \\ . \\ . \\ . \\ z_n \end{pmatrix} + \begin{pmatrix} z_1' \\ z_2' \\ . \\ . \\ . \\ z_n' \end{pmatrix} = \begin{pmatrix} z_1 + z_1' \\ z_2 + z_2' \\ . \\ . \\ . \\ z_n + z_n' \end{pmatrix}. \tag{1}$$

  The vector space is said to be closed under addition.

* alexandru.petrescu@inria.fr

- Multiplying a vector by a scalar (i.e. a c-number) gives another vector

$$z \begin{pmatrix} z_1 \\ z_2 \\ . \\ . \\ . \\ z_n \end{pmatrix} = \begin{pmatrix} zz_1 \\ zz_2 \\ . \\ . \\ . \\ zz_n \end{pmatrix}. \tag{2}$$

  The vector space is closed under multiplication by a scalar.

- There is a zero vector,

$$\begin{pmatrix} z_1 \\ z_2 \\ . \\ . \\ . \\ z_n \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ . \\ . \\ . \\ 0 \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \\ . \\ . \\ . \\ z_n \end{pmatrix}. \tag{3}$$

A *vector subspace* of a vector space V is a subset W of V such that W is a also a vector space, i.e. it is closed under addition and multiplication by a scalar.

We may now revert to Dirac notation, where a vector v is denoted by the ket $|v\rangle$.

Bases and linear independence. A *spanning set* for a vector space is a set of vectors $|v_1\rangle, \ldots, |v_n\rangle$ such that any vector $|v\rangle$ in the vector space can be expressed as a linear combination $|v\rangle = \sum_{i=1}^{n} a_i |v_i\rangle$ for some scalars a_i, $i = 1, \ldots, n$. For example, the vector space $\mathbb{C}^2$ is of interest in quantum mechanics because it is used to represent the state of a spin-$\frac{1}{2}$. One can easily check that both $|v_1\rangle = \begin{pmatrix} 1 & 0 \end{pmatrix}^T$ and $|v_2\rangle = \begin{pmatrix} 0 & 1 \end{pmatrix}^T / \sqrt{2}$ and $|v_1'\rangle = \begin{pmatrix} 1 & 1 \end{pmatrix}^T$ and $|v_2'\rangle = \begin{pmatrix} 1 & -1 \end{pmatrix}^T / \sqrt{2}$ are spanning sets. A set of non-zero vectors $|v_1\rangle, \ldots, |v_n\rangle$ is *linearly dependent* if there exist non-zero scalars $a_1, \ldots, a_n$ such that $a_1 |v_1\rangle + \ldots + a_n |v_n\rangle = 0$. A *basis* for the vector space is a *linearly independent spanning set*.

As an exercise, show that any two linearly independent spanning sets of vector space V contain the same number of elements. This common number of elements is the *dimension* of V. In this course we will be dealing with both *finite* (for example, spins) and *infinite* Hilbert spaces (for example, simple harmonic oscillator).

Linear operators and their matrix representations. A linear operator between vector space V and vector space W is defined to be any function A linear in its inputs, that is $A(\sum_i a_i |v_i\rangle) = \sum_i a_i A(|v_i\rangle)$. Note that $|v_i\rangle$ are completely arbitrary vectors, and in particular they *do not* need to be basis elements. In practice we will omit the parenthesis, and simply write $A(|v\rangle) = A|v\rangle$, in line with Dirac notation. [1] However, if the action of A is known on the elements of some basis $|w_1\rangle, \ldots |w_n\rangle$, then its action is known on any input, by linearity and the definition of the basis. Moreover, compositions of linear operators are linear, that is, given linear operators $A : V \to W$ and $B : W \to X$, then $BA : V \to X$ is also a linear operator.

The easiest way to understand linear operators is through their matrix representations. The map from $\mathbb{C}^n \to \mathbb{C}^m$ defined by multiplication from the left by a matrix A of size $m \times n$ is clearly linear. Conversely, every linear operator can be given a matrix representation. Suppose $A : V \to W$ is a linear operator, and pick $\{|v_1\rangle, \ldots |v_n\rangle\}$ and $\{|w_1\rangle, \ldots |w_m\rangle\}$ bases for V and W, respectively. Then, by completeness, $A|v_j\rangle = \sum_i A_{ij}|w_i\rangle$, for some c-numbers A_{ij}, which provide the matrix representation. For a matrix representation, it is therefore necessary to specify the bases of the input and output vector spaces.

Exercise: Bit-flip linear operator. a) Take V a vector space with basis $|0\rangle, |1\rangle$ and a linear operator $A : V \to V$ such that $A|0\rangle = |1\rangle$ and $A|1\rangle = |0\rangle$. Give a matrix representation of A with respect to the basis given above for the input and output vector space V. b) Change input and output bases to get a different matrix representation of A. *Answer:* a) $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. b) Change basis to $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. Then $A|0\rangle = |1\rangle$ and $A|1\rangle = |0\rangle$ rewrite as $A(|+\rangle \pm |-\rangle)/\sqrt{2} = (|+\rangle \mp |-\rangle)/\sqrt{2}$ so $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Exercise: Show that the matrix representation of an operator product BA with $A : V \to W$ and $B : W \to X$ is the matrix product of the matrix representations of B and A, with appropriate bases.

Exercise: Show that the identity operator on a vector space, taken with identical input and output bases, has the same matrix representation regardless of the chosen basis.

Exercise: We define the set of *Pauli matrices* together with the identity as $\sigma_0 = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\sigma_1 = \sigma_x = X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_2 = \sigma_y = Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, and $\sigma_3 = \sigma_z = Z =$

$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. a) Show that any 2×2 Hermitian matrix, *i.e.* $\begin{pmatrix} a & c+id \\ c-id & b \end{pmatrix}$, with a, b, c, d real, can be expressed as a real linear combination of the Pauli matrices and the identity. b) Show that, if we define the commutator of two linear operators as $[A, B] = AB - BA$, then $[\sigma_i, \sigma_j] = 2i \sum_k \epsilon_{ijk} \sigma_k$ for $i, j, k \in \{1, 2, 3\}$ and ϵ_{ijk} is the Levi-Civita symbol, *i.e.* $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$, $\epsilon_{213} = \epsilon_{321} = \epsilon_{132} = -1$, and zero whenever two of its indices are equal. c) Show that, if we define the anticommutator as $\{A, B\} = AB + BA$, then $\{\sigma_i, \sigma_j\} = 2\delta_{ij} I$. d) Show that, if the trace of a square matrix is defined by the sum of the elements on its diagonal, $\mathrm{Tr}A = A_{11} + \ldots A_{nn}$, then $\mathrm{Tr}\sigma_i = 0$ for $i = 1, 2, 3$, but $\mathrm{Tr}\{\sigma_i \sigma_j\} = 2\delta_{ij}$. e) Letting $\vec{a} = a\hat{n}$ with $\hat{n}$ being a unit vector in $\mathbb{R}^3$, i.e. $|\hat{n}| = 1$, and $a > 0$ a real number, show that $e^{ia\hat{n}\cdot\vec{\sigma}} = I \cos a + i\hat{n} \cdot \vec{\sigma} \sin a$. $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ is a vector whose components are the three Pauli matrices, and $\hat{n} \cdot \vec{\sigma} = \hat{n}_1\sigma_1 + \hat{n}_2\sigma_2 + \hat{n}_3\sigma_3$. f) Show that $R_n(-a)\vec{\sigma}R_n(a) \equiv e^{i\frac{a}{2}\hat{n}\cdot\vec{\sigma}}\vec{\sigma}e^{-i\frac{a}{2}\hat{n}\cdot\vec{\sigma}} = \vec{\sigma}\cos a + \hat{n} \times \vec{\sigma} \sin a + \hat{n}(\hat{n} \cdot \vec{\sigma}) \sin a$. Hint: the right hand side and left hand side of the previous equation are both 3-vectors of 2×2 matrices. The exercise above encompasses all possible single qubit gates.

Inner products. An inner product is a function from $V \times V \to \mathbb{C}$ that satisfies the following three properties:

- $(\cdot, \cdot)$ is linear in the second argument, i.e. $(|u\rangle, |v\rangle + |w\rangle) = (|u\rangle, |v\rangle) + (|u\rangle, |w\rangle)$.
- $(|u\rangle, |v\rangle) = (|v\rangle, |u\rangle)^*$
- $(|v\rangle, |v\rangle) \geq 0$, with equality iff ('if and only if') $|v\rangle = 0$.

You can prove that the inner product is conjugate-linear in its first argument, by using the first two properties above (*exercise*). For a commonly encountered example, on $\mathbb{C}^n$, the inner product is $((y_1, \ldots, y_n), (z_1, \ldots, z_n)) = \sum_i y_i^* z_i$. A vector space equipped with an inner product is an *inner product space*. For finite dimensional vector spaces, this is exactly the same as *a Hilbert space*, and we stick with this notion hereafter. Back to Dirac notation, we will favor the more compact notation $(|u\rangle, |v\rangle) = \langle u|v\rangle$. We have hereby introduced the *dual*, which is a linear operator from the vector space V to complex numbers $\mathbb{C}$, defined by $\langle u|(|v\rangle) = \langle u|v\rangle$. In Dirac notation, the dual $\langle u|$ is called the 'bra' corresponding to the 'ket' $|u\rangle$.

The *norm* of a vector in Hilbert space is defined by the square root of the inner product of

the vector with itself $|| \left|v\right\rangle || = \sqrt{\left\langle v\middle|v\right\rangle}$. A unit vector or a normalized vector is a vector for which the norm equals 1. A set of vectors $\left|v_1\right\rangle, \ldots \left|v_n\right\rangle$ is called *orthonormal* iff $\left\langle v_i\middle|v_j\right\rangle = \delta_{ij}$, for any $i, j = 1, \ldots n$.

These definitions allow us to introduce an orthonormalization procedure called *Gram-Schmidt procedure*. Given a basis $\left|w_1\right\rangle, \ldots, \left|w_n\right\rangle$ of a Hilbert space V, one can produce an orthonormal basis set $\left|v_1\right\rangle, \ldots, \left|v_n\right\rangle$ for the Hilbert space V as follows: Define the first vector in the new basis to be $\left|v_1\right\rangle = \left|w_1\right\rangle / || \left|w_1\right\rangle ||$. For $1 \leq k \leq n-1$, define inductively the next basis vector as $\left|v_{k+1}\right\rangle \equiv \left[\left|w_{k+1}\right\rangle - \sum_{i=1}^{k} \left\langle v_i\middle|w_{k+1}\right\rangle \left|v_i\right\rangle\right] / || \ldots ||$, where the denominator is just the norm of the numerator, i.e. each newly added vector is normalized. It is straightforward to show that this generates an orthonormal basis (*exercise*). Hereafter we will implicitly assume an orthonormal basis when referring to matrix representations.

The inner product can be given a matrix representation. Given an orthonormal basis $\left|1\right\rangle, \ldots, \left|n\right\rangle$ of a Hilbert space V, and two vectors $\left|v\right\rangle = \sum_i v_i \left|i\right\rangle$ and $\left|w\right\rangle = \sum_i w_i \left|i\right\rangle$, then $\left\langle v\middle|w\right\rangle = \sum_{ij} v_i^* w_j \delta_{ij} = \sum_i v_i^* w_i$, i.e. it is equal to the inner product of the two complex matrix representations of the two vectors. The inner product does not depend on the choice of basis, and a different orthonormal basis $\{\left|i'\right\rangle\}$ would yield the same result (*exercise*).

The *outer product* is defined for two vectors $\left|v\right\rangle \in V$ and $\left|w\right\rangle \in W$ as the linear operator $\left|w\right\rangle \left\langle v\right| : V \to W$ such that $\left|w\right\rangle \left\langle v\right| (\left|v'\right\rangle) = \left|w\right\rangle \left\langle v\middle|v'\right\rangle = \left\langle v\middle|v'\right\rangle \left|w\right\rangle$. Using the outer product we can define an expansion for the identity operator. Given an orthonormal basis $\{\left|i\right\rangle\}$ and an arbitrary vector $\left|v\right\rangle$, we have $\left|v\right\rangle = \sum_i v_i \left|i\right\rangle$ for some complex numbers v_i, and therefore $\left\langle i\middle|v\right\rangle = v_i$. Then one easily finds that $(\sum_i \left|i\right\rangle \left\langle i\right|) \left|v\right\rangle = \left|v\right\rangle$. Given that $\left|v\right\rangle$ was chosen arbitrarily, it must be that the operator before it is the identity $I = \sum_i \left|i\right\rangle \left\langle i\right|$. This equation is sometimes called *resolution of identity* or *completeness*. This relationship allows us to write down easily matrix representations of linear operators (over orthonormal bases, which, as mentioned above, we tacitly assume hereafter). Using completeness, given an arbitrary linear operator $A : V \to W$, we may write $I_W A I_V = \sum_{ij} (\left|w_j\right\rangle \left\langle w_j\right|) A (\left|v_i\right\rangle \left\langle v_i\right|) = \sum_{ij} \left\langle w_j\right| A \left|v_i\right\rangle \left|w_j\right\rangle \left\langle v_i\right|$, provided that $\{\left|v_i\right\rangle\}$ and $\{\left|w_j\right\rangle\}$ were orthonormal bases of Hilbert spaces V and W, respectively.

Using completeness we may derive a geometric fact about Hilbert space, called the *Cauchy-Schwarz inequality*. Given arbitrary $\left|v\right\rangle$ and $\left|w\right\rangle$ in Hilbert space V, $|\left\langle v\middle|w\right\rangle|^2 \leq \left\langle v\middle|v\right\rangle \left\langle w\middle|w\right\rangle$, with equality iff the two vectors are collinear $\left|w\right\rangle \propto \left|v\right\rangle$ (symbol means the

two vectors differ only by a complex proportionality constant). Proof: Use Gram-Schmidt to make an orthonormal basis such that the first element is $|1\rangle = |w\rangle / \sqrt{\langle w|w\rangle}$. Then one can rewrite the rhs ('right hand side') of the Cauchy-Schwarz inequality using completeness as $\langle v|v\rangle\langle w|w\rangle = \sum_i \langle v|i\rangle\langle i|v\rangle\langle w|w\rangle$. Now we may split the sum into its first term corresponding to the first basis vector $i = 1$ and the rest, and get $\langle v|v\rangle\langle w|w\rangle = \langle v|w\rangle\langle w|v\rangle + \sum_{i\geq 2}\langle v|i\rangle\langle i|v\rangle\langle w|w\rangle \geq |\langle v|w\rangle|^2$, which concludes the proof. Assuming equality, necessarily $\sum_{i\geq 2}\langle v|i\rangle\langle i|v\rangle\langle w|w\rangle = 0$ so $|\langle v|i\rangle|^2 = 0$ for all $i \geq 2$, and the proportionality of $|v\rangle$ and $|w\rangle$ follows immediately. *Alternative, shorter proof:* For any complex λ, $(\langle v| + \lambda^* \langle u|)(|v\rangle + \lambda |u\rangle) \geq 0$. Setting $\lambda = -\langle u|v\rangle / \langle u|u\rangle$ yields the result.

Eigenvectors and eigenvalues. An eigenvector of a linear operator A is a vector $|v\rangle$ such that for some complex number v, called the eigenvalue, $A|v\rangle = v|v\rangle$. The eigenvalues are the roots of the characteristic polynomial $c(X) = \det(A - \lambda I)$. By the fundamental theorem of algebra, this polynomial has at least one complex root. An *eigenspace* is a Hilbert subspace spanned by the set of eigenvectors sharing an eigenvalue v. The eigenvectors are said to be *degenerate*. The *diagonal representation* of a linear operator A is $A = \sum_i \lambda_i |i\rangle\langle i|$, where $|i\rangle$ are the eigenvectors, and λ_i are the corresponding eigenvalues of A (not necessarily distinct). For example, the Pauli matrix Z defined above has diagonal representation $Z = |1\rangle\langle 1| - |0\rangle\langle 0|$. As an *exercise*, find the eigenvectors and eigenvalues of the Pauli matrices.

Adjoints and Hermitian operators. For any linear operator A on a Hilbert space V, there exists unique linear operator $A^\dagger$, called the adjoint of A, such that for any $|u\rangle, |v\rangle$ in V, $(|u\rangle, A|v\rangle) = (A^\dagger|u\rangle, |v\rangle)$. The adjoint has the following properties:

$$\begin{aligned}
(AB)^\dagger &= B^\dagger A^\dagger, \\
|v\rangle^\dagger &= \langle v|, \\
(A|v\rangle)^\dagger &= \langle v| A^\dagger, \\
(|w\rangle\langle v|)^\dagger &= |v\rangle\langle w|, \\
\left(\sum_i a_i A_i\right)^\dagger &= \sum_i a_i^* A_i^\dagger, \\
(A^\dagger)^\dagger &= A.
\end{aligned} \tag{4}$$

For A a matrix representation of a linear operator, the adjoint is the conjugate of the transpose $A^\dagger = (A^*)^T = (A^T)^*$. A *Hermitian operator*, or *self-adjoint operator*, is an operator A such that $A = A^\dagger$.

Projectors are an important class of Hermitian operators. Consider a Hilbert subspace $W \subset V$ of dimension $k < d$, where d is the dimension of V. Using Gram-Schmidt, one can always construct an orthonormal basis of V $|1\rangle, \ldots, |k\rangle, \ldots |d\rangle$, where the first k vectors are an orthonormal basis of W. Then let $P \equiv \sum_{i=1}^{k} |i\rangle \langle i|$. By construction $P = P^\dagger$. Defining also the complement $Q = I - P = \sum_{i=k+1}^{d} |i\rangle \langle i|$, we also have $Q = Q^\dagger$. Then $QP = PQ = 0$, but $Q^2 = Q$ and $P^2 = P$.

A *normal operator* is a linear operator that commutes with its adjoint, that is $AA^\dagger = A^\dagger A$. Clearly, all Hermitian operators are also normal. However, a normal matrix is also Hermitian iff it has real eigenvalues (*prove*). A normal operator with real eigenvalues is Hermitian (*prove*).

Spectral decomposition theorem. Any normal linear operator M on vector space V is diagonal with respect to some orthonormal basis on V. Conversely, any diagonalizable operator is normal. *Proof:* For the converse, if M is diagonalizable one may write $M = \sum_i \lambda_i |i\rangle \langle i|$ where $\lambda_i, |i\rangle$ are the eigenvalues and eigenvectors of M. Then clearly $[M, M^\dagger] = 0$. For the direct result, the proof can be done by induction on d, the dimension of V. If $d = 1$, the result is trivial. Let λ be an eigenvalue of M, and let P be the projector onto the $\lambda-$subspace, with $Q = I - P$ the projector onto its complement subspace. Then $M = (P+Q)M(P+Q) = PMP + QMQ + PMQ + QMP$. By construction, $PMP = \lambda P$. $QMP = 0$ since $QMP = Q\lambda P = \lambda QP = 0$. Moreover, $PMQ = 0$. To show this, take $|v\rangle$ to be an element of the subspace P. $MM^\dagger |v\rangle = M^\dagger M |v\rangle = \lambda M^\dagger |v\rangle$, so $M^\dagger |v\rangle$ is also in the subspace P, that is $QM^\dagger P = 0$. Taking the adjoint, and using the Hermiticity of projectors, $PMQ = 0$. Hence we find $M = PMP + QMQ$. It will be sufficient to show that QMQ is normal. To show this, we use two preliminary facts: $QM = QM(P+Q) = QMQ$ and $QM^\dagger = QM^\dagger(P+Q) = QM^\dagger Q$. Then, to show that QMQ is normal, we evaluate $QMQQM^\dagger Q = QMQM^\dagger Q = QMM^\dagger Q = QM^\dagger MQ = QM^\dagger QQMQ$. Lastly, assuming the direct statement was proved for the Q subspace, adding the P subspace, on which M is diagonal in the basis of eigenvectors of eigenvalue λ, retains the property.

Unitary operators and matrices have the property that their adjoint is their inverse $UU^\dagger = U^\dagger U = I$. A linear operator is unitary iff any of its matrix representations is unitary. Any unitary operator is also normal. It therefore has a spectral decomposition. Unitary operators are *norm-preserving*, that is $(U|v\rangle, U|u\rangle) = (U^\dagger U |u\rangle, |v\rangle) = (|u\rangle, |v\rangle)$. This is evident in Dirac notation, since $(U|u\rangle)^\dagger = \langle u| U^\dagger$, and hence $(\langle u| U^\dagger)(U|v\rangle) = \langle u| (U^\dagger U) |v\rangle = \langle u|v\rangle$.

Unitary operators encode therefore a *basis change.* For any two orthonormal bases $\{|v_i\rangle\}$ and $\{|w_i\rangle\}$ of a Hilbert space V, the linear operator $U = \sum_i |u\rangle_i \langle w|_i$ is unitary.

Exercises: a) Prove that all eigenvalues of a unitary operator have norm 1, and therefore can be written as $e^{i\theta}$ for some real number θ. b) Show that all Pauli matrices are unitary. c) Prove that two eigenvectors of a Hermitian operator with different eigenvalues are necessarily orthogonal. d) Prove that any eigenvalue of a projector P can be either 0 or 1.

A linear operator A is called *positive* iff $\langle v| A |v\rangle \geq 0$ for all $|v\rangle$ in the Hilbert space V. It is *positive-definite* if the inequality above is strict, i.e. $\langle v| A |v\rangle > 0$ for all $|v\rangle$ in the Hilbert space V. Any positive linear operator is Hermitian. To show this, write the operator as $A = B + iC$, where B, C are Hermitian. Since A is positive, for any $|v\rangle$, $\langle v| B |v\rangle + i \langle v| C |v\rangle \geq 0$. So it must be that $\langle v| C |v\rangle$ is imaginary. However, by the Hermiticity of C, $\langle v| C |v\rangle = \langle v| C^\dagger |v\rangle = \langle v| C |v\rangle^*$, and hence $\langle v| C |v\rangle = 0$ for all $|v\rangle$. So $C = 0$. Moreover, for any linear operator A, $A^\dagger A$ is positive. This is proven simply, as $\langle v| A^\dagger A |v\rangle$ is the norm of $A |v\rangle$, and that is non-negative by definition.

Tensor products provide a way to put together Hilbert spaces to form larger Hilbert spaces. If V and W are two Hilbert spaces of dimensions m and n, respectively, $V \otimes W$ is a tensor product Hilbert space of dimension mn. If $|v\rangle \in V$ and $|w\rangle \in W$, then $|v\rangle \otimes |w\rangle \in V \otimes W$. If $\{|v_i\rangle\}$ is an orthonormal basis of V, and $\{|w_j\rangle\}$ is an orthonormal basis of W, then $\{|v_i\rangle \otimes |w_j\rangle\}$ is an orthonormal basis of $V \otimes W$. Tensor products have the following properties

- For any scalar z, $|v\rangle \in V$ and $|w\rangle \in W$, $z(|v\rangle \otimes |w\rangle) = (z |v\rangle) \otimes |w\rangle = |v\rangle \otimes (z |w\rangle)$;
- they are linear in the first argument;
- they are linear in the second argument;
- $(A \otimes B)(|v\rangle \otimes |w\rangle) = A |v\rangle \otimes B |w\rangle$
- $(A \otimes B) \sum_i a_i |v_i\rangle \otimes |w_i\rangle = \sum_i a_i A |v_i\rangle \otimes B |w_i\rangle$, i.e. $A \otimes B$ is a well-defined linear operator on $V \otimes W$.

The *inner product on a tensor-product space* inherits its definition from the inner products on the constituent spaces, that is $\left(\sum_i a_i |v_i\rangle \otimes |w_i\rangle , \sum_j b_j \left|v'_j\right\rangle \otimes \left|w'_j\right\rangle\right) = \sum_{ij} a_i^* b_j \langle v_i | v'_j \rangle \langle w_i | w'_j \rangle$. If $|v_i\rangle = |v'_i\rangle$ and $|w_j\rangle = \left|w'_j\right\rangle$, then the latter equals $\sum_i a_i^* b_i$.

The *matrix representation* of the tensor product is the *Kronecker product*. The Kronecker product is defined as follows

$$A \otimes B = \begin{pmatrix} A_{11}B & \dots & A_{1n}B \\ \vdots & & \vdots \\ A_{m1}B & \dots & A_{mn}B \end{pmatrix}. \tag{5}$$

The resulting matrix has dimensions $mp \times nq$ if the original matrices A and B had dimensions $m \times n$ and $p \times q$, respectively.

The tensor product is *not commutative*. Transpose, complex conjugation, and adjoint distribute over the tensor product, that is $(A \otimes B)^* = A^* \otimes B^*, (A \otimes B)^T = A^T \otimes B^T$, and $(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger$ (*prove these*). The tensor product of two unitary/Hermitian/positive/projector operators is a unitary/Hermitian/positive/projector operator (*prove*).

~~*Exercise:* Show that $\exp(A_1 \otimes I_2 + I_1 \otimes A_2) = \exp(A_1) \otimes \exp(A_2)$.~~

Exercise: Prove that $|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ cannot be expressed as a tensor product.

Exercise: Show that $\exp(A_1 \otimes I_2 + I_1 \otimes A_2) = \exp(A_1) \otimes \exp(A_2)$ for some linear operators $A_{1,2}$ acting on Hilbert spaces $V_{1,2}$, and $I_{1,2}$ the respective identity operators.

Operator functions. $A = \sum_a a \, |a\rangle \langle a|$ is the spectral decomposition of a normal operator A. One can uniquely define $f(A) \equiv \sum_a f(a) \, |a\rangle \langle a|$.

Exercises: $e^{i\theta Z} = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$. Find the logarithm of the matrix $\begin{pmatrix} 2 & 3/2 \\ 3/2 & 2 \end{pmatrix}$. Show that $\exp(i\theta \vec{v} \cdot \vec{\sigma}) = \cos(\theta) I + i \sin(\theta) \vec{v} \cdot \vec{\sigma}$, with the usual notations and for $\vec{v}$ a unit vector.

The *trace of a square matrix* is defined as the sum of its diagonal entries $\mathrm{Tr}(A) = \sum_i A_{ii}$. It is (proofs of these properties left as *exercise*)

- cyclic: $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$,
- linear: $\mathrm{Tr}(A + B) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$,
- invariant under similarity transformations: $\mathrm{Tr}(UAU^{-1}) = \mathrm{Tr}(A)$.

The *trace of an operator* is the trace of any one of its matrix representations.

The trace allows us to define an inner product for operators, the *Hilbert-Schmidt product*. Let $\mathcal{L}(V)$ be the set of linear operators on a Hilbert space V. It is a vector space, since the sum of two linear operators is also a linear operator; zA is a linear operator if A is a linear

operator and z a scalar; and there is a zero element, the linear operator 0, that takes every vector to the null vector. The Hilbert-Schmidt inner product is defined as $(A, B) = \mathrm{Tr}(A^\dagger B)$.

An important operation on tensor-product spaces is the *partial trace*. Let

$$A = \sum_{nn'mm'} A_{nn'mm'} |nn'\rangle \langle mm'| , \tag{6}$$

where $|nn'\rangle = |n\rangle \otimes |n'\rangle$ is an orthonormal basis for $V \otimes V'$. Then the partial trace over Hilbert space V' is defined as

$$\mathrm{Tr}_{V'}(A) = \sum_{nn'm} A_{nn'mn'} |n\rangle \langle m| . \tag{7}$$

As an *exercise*, show that this definition is independent of the choice of orthonormal basis.

The *commutator* of two linear operators is defined as $[A, B] = AB - BA$.

Theorem. Suppose A, B are Hermitian. Then $[A, B] = 0$ iff there exists an orthonormal basis such that A and B are diagonal with respect to that basis. We say that A and B are *simultaneously diagonalizable*. *Proof:* The $\leftarrow$ direction is trivial. A and B are normal and they share eigenvectors. Then writing their spectral decomposition shows that they commute. For the $\rightarrow$, let $\{|a, j\rangle\}$ be the orthonormal basis of the eigenspace corresponding to eigenvalue a of A, V_a, with j labeling possible degenerate eigenvectors. Then $AB |a, j\rangle = BA |a, j\rangle = aB |a, j\rangle$. So $B |a, j\rangle \in V_a$. Let P_a be the projector onto V_a. Define the restriction of B on V_a, $B_a = P_a B P_a$. B_a is Hermitian, so it is normal, so it has a spectral decomposition in terms of a set of eigenvectors which span V_a, $\{|a, b, k\rangle\}$, where a denotes the eigenvalue of A, b denotes the corresponding eigenvalue of B_a, and k is introduced for possible multiplicity of the b eigenvalue. From definitions, $P_a |a, b, k\rangle = |a, b, k\rangle$, $B |a, b, k\rangle = P_a B |a, b, k\rangle$. So $B |a, b, k\rangle = P_a B P_a |a, b, k\rangle = B_a |a, b, k\rangle = b |a, b, k\rangle$, whence the conclusion of the proof follows easily.

II. POSTULATES OF QUANTUM MECHANICS

Postulate 1. Associated to any isolated physical system is a complex vector space with inner product, i.e. a Hilbert space, known as the *state space* of the system. The system is completely defined by its *state vector*, which is a unit vector in *state space*.

Postulate 2. The evolution of a closed quantum system is described by a *unitary transformation*. That is, the state $|\psi_1\rangle$ of the system at time t_1 is related to the state $|\psi_2\rangle$ at

time t_2 by a unitary operator U that depends only on the times $t_{1,2}$,

$$|\psi_2\rangle = U(t_2, t_1)\,|\psi_1\rangle\,. \tag{8}$$

Alternatively, we may formulate *postulate 2* as follows: the time evolution of the state of a closed system is described by the *Schrödinger equation,*

$$i\frac{d\,|\psi\rangle}{dt} = \frac{H}{\hbar}\,|\psi\rangle\,, \tag{9}$$

where $\hbar$ is a physical constant to be determined experimentally. H is a Hermitian operator known as the *Hamiltonian* of the system. As a consequence, suppose H is time independent. Then H being normal, it has a spectral decomposition $H = \sum_E E\,|E\rangle\,\langle E|$, with eigenvalues E called *energies* and eigenvectors $|E\rangle$ called energy eigenstates. Under time evolution, $|E\rangle \to e^{-iEt/\hbar}\,|E\rangle$. Back to the first formulation of the second postulate above, $|\psi(t_2)\rangle = e^{-i\frac{H(t_2-t_1)}{\hbar}}\,|\psi(t_1)\rangle \equiv U(t_2, t_1)\,|\psi(t_1)\rangle$. The operator U that we have just defined is unitary because H is Hermitian.

Exercise: Solve Schrödinger's equation for $H = \hbar\omega X$.

Postulate 3. Quantum measurements are described by a collection of measurement operators $\{M_m\}$ acting on the state space of the system to be measured. The index m refers to the measurement outcome. If the state of the system is $|\psi\rangle$ immediately before the measurement, then the probability to obtain result m upon measurement is

$$p(m) = \langle\psi|\,M_m^\dagger M_m\,|\psi\rangle\,, \tag{10}$$

and the state of the system right after measuring outcome m is

$$\frac{M_m\,|\psi\rangle}{\sqrt{\langle\psi|\,M_m^\dagger M_m\,|\psi\rangle}} = \frac{M_m\,|\psi\rangle}{\sqrt{p(m)}}. \tag{11}$$

The measurmeent operators satisfy the *completeness relation*

$$\sum_m M_m^\dagger M_m = I, \tag{12}$$

from which one deduces that the probabilities of all outcomes sum to 1 for any state $|\psi\rangle$: $1 = \sum_m p(m) = \sum_m \langle\psi|\,M_m^\dagger M_m\,|\psi\rangle$.

The postulate 3 is best illustrated by a simple *example.* The measurement operators for a qubit can be written as $M_i = |i\rangle\,\langle i|$ for $i = 0, 1$. Note that $M_i = M_i^\dagger$ and $M_i^2 = M_i$. They

are projectors onto complementary subspaces, $M_0^\dagger M_0 + M_1^\dagger M_1 = |0\rangle\langle 0| + |1\rangle\langle 1| = I$. Now taking arbitrary $|\psi\rangle = a_0|0\rangle + a_1|1\rangle$, with $|a_0|^2 + |a_1|^2 = 1$, the probabilities of measurement outcomes are $p(i) = \langle\psi| M_i^\dagger M_i |\psi\rangle = |a_i|^2$. The state immediately after measurement of outcome $i = 0, 1$ is $M_i|\psi\rangle / |a_i| = a_i/|a_i| \, |i\rangle$. Note that these are just the kets $|i\rangle$, up to phase factors. Phase factors are irrelevant, since a phase rotation does not change the measurement probability $p(i)$. This was an example of a *projective measurement*.

Projective measurements are described by an observable M, a Hermitian operator on the state space of the system being observed. It then has a spectral decomposition $M = \sum_m mP_m$, where m is the possible outcome, and P_m is the projector onto the eigenspace of M with eigenvalue m. Upon measuring a system known to be in state $|\psi\rangle$, the probability of obtaining outcome m is $p(m) = \langle\psi| P_m |\psi\rangle$, and if the outcome m is recorded, then the state of the system right after the measurement is $P_m|\psi\rangle / \sqrt{p(m)}$. We have just reformulated Postulate 3 with one extra requirement, that M_m be orthogonal projectors. From the above, we can write down the average value of a measurement

$$E(M) = \sum_m mp(m) = \sum_m m\langle\psi| P_m |\psi\rangle = \langle\psi| \sum_m mP_m |\psi\rangle = \langle\psi| M |\psi\rangle . \tag{13}$$

This is to be interpreted as the average value of the measurement outcome upon performing a large number of measurements and storing the corresponding outcomes in an array. The variance is defined as

$$\Delta(M)^2 = E(M^2) - E(M)^2 = \langle\psi| M^2 |\psi\rangle - \langle\psi| M |\psi\rangle^2 . \tag{14}$$

These definitions allow us to formulate and prove the Heisenberg uncertainty principle at the end of this lecture.

Exercises: a) You have a qubit in $|\psi\rangle = |0\rangle$ and you measure X. What is $E(X)$ and what is the variance? b) Show that $\vec{v}\cdot\vec{\sigma}$ has eigenvalues ± 1 if $||v\rangle| = 1$, and that the projectors onto the corresponding eigenvectors are $P_\pm = \frac{I \pm \vec{v}\cdot\vec{\sigma}}{2}$. Calculate the probability of obtaining outcome $+1$ when measuring $\vec{v}\cdot\vec{\sigma}$, given that the state prior to measurement is $|0\rangle$. What is the state of the system right after measuring outcome $+1$?

Postulate 4. The state space of a composite physical system is the tensor product of the state spaces of the component physical systems. Moreover, if systems $1, \ldots, n$ are prepared in state $|\psi_1\rangle, \ldots, |\psi_n\rangle$, then the composite system is in state $|\psi_1\rangle \otimes \ldots \otimes |\psi_n\rangle$.

We now prove the *Heisenberg uncertainty principle*. Consider A, B Hermitian acting on state space V. Let $\langle\psi| AB |\psi\rangle = x + iy$ with x, y real. Then $\langle\psi| [A, B] |\psi\rangle = 2iy$ whereas

$\langle\psi|\{A,B\}|\psi\rangle = 2x$, and therefore $|\langle\psi|[A,B]|\psi\rangle|^2 + |\langle\psi|\{A,B\}|\psi\rangle|^2 = 4|\langle\psi|AB|\psi\rangle|^2$. Applying Cauchy-Schwarz inequality to the two state vectors $A|\psi\rangle$ and $B|\psi\rangle$, one obtains $|\langle\psi|AB|\psi\rangle|^2 \leq \langle\psi|A^2|\psi\rangle\langle\psi|B^2|\psi\rangle$. The two relations above give $\frac{1}{4}|\langle\psi|[A,B]|\psi\rangle|^2 \leq \langle\psi|A^2|\psi\rangle\langle\psi|B^2|\psi\rangle$. As a last step, let $A \to A - \langle A\rangle$, and $B \to B - \langle B\rangle$, to get

$$\Delta(A)\Delta(B) \geq \frac{|\langle\psi|[A,B]|\psi\rangle|}{2}, \tag{15}$$

for all $|\psi\rangle$. The physical interpretation is as follows. Upon performing a large number of measurements of A, then a large number of measurements of B, the variances of the two sets containing the outcomes will obey the above inequality.

Example. $[X,Y] = 2iZ$ implies $\Delta(X)\Delta(Y) \geq \langle Z\rangle$, so for $|\psi\rangle = |0\rangle$, $\Delta(X)\Delta(Y) \geq 1$.

Example. $[x,p] = i\hbar$ implies $\Delta(x)\Delta(p) \geq \hbar/2$ for all $|\psi\rangle$. The right hand side is free of expectation value, it is the same for all state vectors.

---

[1] In the literature, operators are sometimes expressed with a hat $\hat{A}$, but we will avoid this notation for simplicity, at least until it becomes strictly necessary.