

Name: **Huan Q. Bui**
 Course: **8.321 - Quantum Theory I**
 Problem set: **#2**

1. Let A be a skew-Hermitian operator, i.e., $A^\dagger = -A$.

(a) Let λ and $|\lambda\rangle$ be an eigenvalue and eigenvector of A , respectively. Then we have

$$A|\lambda\rangle = \lambda|\lambda\rangle \implies \lambda\langle\lambda|\lambda\rangle = \langle\lambda|A|\lambda\rangle = -\langle\lambda|A|\lambda\rangle = \langle\lambda|A^*|\lambda\rangle = \lambda^*\langle\lambda|\lambda\rangle \implies -\lambda = \lambda^*.$$

Since $\lambda \in \mathbb{C}$, the only solution is $\lambda = 0$. Thus, the only real eigenvalue of A (up to multiplicity/degeneracy) is 0.

(b) Let A, B be Hermitian operators. Then

$$[A, B] = AB - BA = A^\dagger B^\dagger - B^\dagger A^\dagger = (BA - AB)^\dagger = -(AB - BA)^\dagger = -[A, B]^\dagger.$$

Thus $[A, B]$ is skew-Hermitian.

2. Let H, K be Hermitian operators with non-negative eigenvalues and assume that the trace defined throughout this problem. Since H, K are Hermitian operators we may assume that there exist complete orthonormal (eigen)bases $\{|h_i\rangle\}$ and $\{|k_i\rangle\}$ for H, K respectively with $H|h_i\rangle = h_i|h_i\rangle$ and $K|k_i\rangle = k_i|k_i\rangle$, and $h_i, k_i \geq 0$ for all i . Then we can spectral-decompose H, K in their product as follows

$$HK = \sum_n h_n |h_n\rangle\langle h_n| \sum_m k_m |k_m\rangle\langle k_m| = \sum_{n,m} h_n k_m |h_n\rangle\langle h_n|k_m\rangle\langle k_m|.$$

Since $\text{tr}(A) = \sum_i \langle\phi_i|A|\phi_i\rangle$ for any orthonormal basis $\{\phi_i\}$, we have

$$\begin{aligned} \text{tr}(HK) &= \sum_j \langle h_j | \left[\sum_{n,m} h_n k_m |h_n\rangle\langle h_n|k_m\rangle\langle k_m| \right] | h_j \rangle \\ &= \sum_{n,m} h_n k_m \langle h_n | k_m \rangle \langle k_m | h_n \rangle, \quad \text{by orthonormality} \\ &= \sum_{n,m} h_n k_m |\langle h_n | k_m \rangle|^2. \end{aligned}$$

Since $h_i, k_i \geq 0$ for all i , and the modulus square is always nonnegative, we see that $\text{tr}(HK) \geq 0$, as desired.

Suppose $\text{tr}(HK) = 0$, then by nonnegativity we must have $h_n k_m |\langle h_n | k_m \rangle|^2 = 0$ for all n, m , or equivalently $h_n k_m \langle h_n | k_m \rangle = 0$ for all n, m . In view of the first equation for HK , we see that $HK = 0$.

3. Let a Hermitian operator H be given with positive spectrum and a complete orthonormal basis.

(a) We want to prove that for any two vectors $|\alpha\rangle, |\beta\rangle$

$$|\langle\alpha|H|\beta\rangle|^2 \leq \langle\alpha|H|\alpha\rangle \langle\beta|H|\beta\rangle.$$

There are two ways to go about this proof, in which both approaches are actually the same and only differ by appearance. I will present the notationally “light” version first. This goes as follows: Since H is Hermitian with positive spectrum, we may find a complete orthonormal basis in which H is diagonal. The transformation between H and its diagonalization D is given by a unitary operator U as $H = U^\dagger D U$. Since D is diagonal with positive entries, we can define its square root \sqrt{D} . From here, we can also define the square root of H , denoted \sqrt{H} by $U^\dagger \sqrt{D} U$. We can check:

$$\sqrt{H}\sqrt{H} = U^\dagger \sqrt{D} U U^\dagger \sqrt{D} U = U^\dagger \sqrt{D} \sqrt{D} U = U^\dagger D U = H.$$

It is easy to show that \sqrt{H} is also Hermitian:

$$\sqrt{H}^\dagger = (U^\dagger \sqrt{D} U)^\dagger = U^\dagger \sqrt{D}^\dagger U = U^\dagger \sqrt{D} U = \sqrt{H},$$

where we have used the fact that \sqrt{D} is strictly diagonal and positive, thus Hermitian. The rest of the proof is now a simple application of the Cauchy-Schwarz inequality for inner products:

$$\begin{aligned} |\langle \alpha | H | \beta \rangle|^2 &= \left| \langle \alpha | \sqrt{H} \sqrt{H} | \beta \rangle \right|^2 = \left| \langle \alpha | \sqrt{H}^\dagger \sqrt{H} | \beta \rangle \right|^2 = \left| \langle \alpha \sqrt{H}^\dagger | \sqrt{H} \beta \rangle \right|^2 \\ &\leq \langle \sqrt{H} \alpha | \sqrt{H} \alpha \rangle \langle \sqrt{H} \beta | \sqrt{H} \beta \rangle \\ &= \langle \alpha | \sqrt{H}^\dagger \sqrt{H} | \alpha \rangle \langle \beta | \sqrt{H}^\dagger \sqrt{H} | \beta \rangle = \langle \alpha | H | \alpha \rangle \langle \beta | H | \beta \rangle \end{aligned}$$

as desired.

The more notationally heavy approach is to consider a complete orthonormal eigenbasis for H , which we may call $\{|\lambda_i\rangle\}$ where $\{\lambda_i\}$ are the eigenvalues of H . Under this basis, we have

$$|\alpha\rangle = \sum_i a_i |\lambda_i\rangle \quad |\beta\rangle = \sum_i b_i |\lambda_i\rangle$$

and so

$$|\langle \alpha | H | \beta \rangle|^2 = \left| \sum_i a_i^* \langle \lambda_i | \lambda_j b_j | \lambda_j \rangle \right|^2 = \left| \sum_i a_i^* \lambda_i b_i \right|^2 = \left| \sum_i (a_i \sqrt{\lambda_i})^\dagger (b_i \sqrt{\lambda_i}) \right|^2.$$

Note that $\sqrt{\lambda_i} \in \mathbb{R}^+$, which is possible because $\lambda_i > 0$. Now, call

$$|\alpha'\rangle = \sum_i a_i \sqrt{\lambda_i} |\lambda_i\rangle \quad |\beta'\rangle = \sum_i b_i \sqrt{\lambda_i} |\lambda_i\rangle.$$

It is clear that

$$|\langle \alpha | H | \beta \rangle|^2 = |\langle \alpha' | \beta' \rangle|^2.$$

On the other hand, we have

$$\begin{aligned} \langle \alpha | H | \alpha \rangle &= \sum_{i,j} a_i^* a_j \lambda_j \langle \lambda_i | \lambda_j \rangle = \sum_i |a_i|^2 \lambda_i = \langle \alpha' | \alpha' \rangle \\ \langle \beta | H | \beta \rangle &= \sum_{i,j} b_i^* b_j \lambda_j \langle \lambda_i | \lambda_j \rangle = \sum_i |b_i|^2 \lambda_i = \langle \beta' | \beta' \rangle. \end{aligned}$$

Applying the Cauchy-Schwarz inequality,

$$|\langle \alpha | H | \beta \rangle|^2 = |\langle \alpha' | \beta' \rangle|^2 \leq \langle \alpha' | \alpha' \rangle \langle \beta' | \beta' \rangle = \langle \alpha | H | \alpha \rangle \langle \beta | H | \beta \rangle$$

we successfully proved the desired result.

- (b) The trace of H is simply the sum of its eigenvalues, so $\text{tr}(H) > 0$. To show explicitly, we use the orthonormal basis introduced in Part (a). Since $\lambda_i > 0$ for all i , we have

$$\text{tr}(H) = \sum_i \langle \lambda_i | H | \lambda_i \rangle = \sum_i \lambda_i \langle \lambda_i | \lambda_i \rangle = \sum_i \lambda_i > 0.$$

4. Let a unitary operator U be given which satisfies the eigenvalue equation $U |\lambda\rangle = \lambda |\lambda\rangle$.

(a) Since $\langle \lambda | \lambda \rangle \neq 0$ (because $|\lambda\rangle$ is an eigenvector), we have

$$\langle \lambda | \lambda \rangle = \langle \lambda | U^\dagger U | \lambda \rangle = |\lambda|^2 \langle \lambda | \lambda \rangle \implies |\lambda|^2 = 1.$$

Since $\lambda \in \mathbb{C}$, it must be of the form $\lambda = e^{i\theta}$ where $\theta \in \mathbb{R}$.

(b) Let distinct eigenvectors $|\mu\rangle$ and $|\lambda\rangle$ be given with corresponding (distinct) eigenvalues $e^{i\theta_\mu}$ and $e^{i\theta_\lambda}$. We have

$$\langle \mu | \lambda \rangle = \langle \mu | U^\dagger U | \lambda \rangle = e^{-i\theta_\mu} e^{i\theta_\lambda} \langle \mu | \lambda \rangle.$$

Since the eigenvalues are not the same, we have that $e^{-i\theta_\mu} e^{i\theta_\lambda} \neq 1$ (i.e., that the complex conjugate of one is not the complex conjugate of the other). Thus, equality holds only if $\langle \mu | \lambda \rangle = 0$.

5.

(a) First, we will show that the set of $N \times N$ complex matrices form a vector space (over the complex numbers).

- The zero matrix O is the identity for vector (matrix) addition.
- For every matrix A , the matrix $-A$ exists and $A + (-A) = O$, so every matrix has an additive inverse.
- Matrix addition is associative.
- Matrix addition is commutative.
- Scalar multiplication: For $a, b \in \mathbb{C}$ and a matrix A , we have $a(bA) = (ab)A = (ab)A$, as usual.
- The number $1 \in \mathbb{C}$ is the identity for scalar multiplication.
- Scalar multiplication is distributive with respect to matrix addition. Given $a \in \mathbb{C}$ and matrices A, B we have $a(A + B) = aA + aB$.
- Finally, for $a, b \in \mathbb{C}$ and a matrix A , we have $(a + b)A = aA + bA$.

Basically, the rules for matrix addition show that the set of $N \times N$ complex matrices form a vector space. To show that the dimension of this space is N^2 , we consider the following set of N^2 matrices $\{M(ij)\}_{i,j=1}^N$ where each $M(ij)$ is an $N \times N$ matrix whose entries are all zeros except for a 1 in the ij position. It is clear that there exists no non-trivial linear combination of the $M(ij)$'s that gives the zero matrix. Thus, $\{M(ij)\}_{i,j=1}^N$ is a linearly independent set. Moreover, it is also obvious that any $N \times N$ matrix can be written as a linear combination of the $M(ij)$ matrices (i.e., given a matrix $A = [a_{ij}]$ we have $A = \sum a_{ij} M(ij)$). Therefore, the vector space of $N \times N$ complex matrices is N^2 -dimensional.

(b) Let $(A, B) = \text{Tr}(A^\dagger B)$. We will show that (\cdot, \cdot) defines an inner product over the vector space \mathcal{V} above.

- Positive semidefinite: Given $A \in \mathcal{V}$. Then

$$\text{Tr}(A^\dagger A) = (A^\dagger A)_{ii} = A_{ij}^\dagger A_{ji} = A_{ji}^* A_{ji} = \sum_{i,j=1}^N |A_{ij}|^2 \geq 0,$$

with equality occurring if and only if $A_{ij} = 0$ for all i, j , i.e., $A = 0$.

- Linear in the second argument: For $\beta \in \mathbb{C}$ and $A, B \in \mathcal{V}$, we have, by the linearity of the trace function, $\text{Tr}(A^\dagger \beta B) = \beta \text{Tr}(A^\dagger B)$. Moreover, given $C \in \mathcal{V}$, we have

$$\text{Tr}(A^\dagger (B + C)) = \text{Tr}(A^\dagger B + A^\dagger C) = \text{Tr}(A^\dagger B) + \text{Tr}(A^\dagger C).$$

- Conjugate-linear in the first argument (optional since the previous condition suffices): For $\alpha \in \mathbb{C}$ and $A, B \in \mathcal{V}$, we have, by the linearity of the trace function, $\text{Tr}((\alpha A)^\dagger B) = \beta \text{Tr}(a^* A^\dagger B) = a^* \text{Tr}(A^\dagger B)$. Similarly, given $C \in \mathcal{V}$,

$$\text{Tr}((A + C)^\dagger B) = \text{Tr}(A^\dagger B + C^\dagger B) = \text{Tr}(A^\dagger B) + \text{Tr}(C^\dagger B).$$

- Conjugate symmetry: Given $A, B \in \mathcal{V}$, we have $\text{Tr}(B^\dagger A) = \text{Tr}((A^\dagger B)^\dagger) = \overline{\text{Tr}((A^\dagger B)^\top)} = \overline{\text{Tr}(A^\dagger B)}$, using the fact that $\text{Tr}(X) = \text{Tr}(X^\top)$ for a square matrix X .

(c) By inspection, we can see that the collection $\{\mathbb{I}, \sigma_1, \sigma_2, \sigma_3\}$ is linearly independent. It remains to show that it spans the space of 2×2 complex matrices. To this end, let a 2×2 matrix A be given.

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

We may write A as a linear combination of $\{\mathbb{I}, \sigma_1, \sigma_2, \sigma_3\}$ as

$$A = \left(\frac{a+d}{2}\right)\mathbb{I} + \left(\frac{b+c}{2}\right)\sigma_1 + \left(\frac{c-b}{2i}\right)\sigma_3 + \left(\frac{a-d}{2}\right)\sigma_3.$$

Therefore, $\{\mathbb{I}, \sigma_1, \sigma_2, \sigma_3\}$ also spans the space. Thus, $\{\mathbb{I}, \sigma_1, \sigma_2, \sigma_3\}$ is a basis for this space. Now, we claim that this basis, under the normalization factor $1/\sqrt{2}$:

$$B = \left\{ \frac{1}{\sqrt{2}}\mathbb{I}, \frac{1}{\sqrt{2}}\sigma_1, \frac{1}{\sqrt{2}}\sigma_2, \frac{1}{\sqrt{2}}\sigma_3 \right\}$$

is orthonormal with respect to the inner product defined in Part (b). To see this, we observe that each element of the basis B is already Hermitian and that $\sigma_i^2 = \mathbb{I} = \mathbb{I}^2$ for $i = 1, 2, 3$. So, we have that $\text{Tr}\left((\sigma_i^\dagger/\sqrt{2})(\sigma_i/\sqrt{2})\right) = \text{Tr}(\sigma_i^2)/2 = \text{Tr}(\mathbb{I})/2 = \text{Tr}(\mathbb{I}^2)/2 = 1$, for each element of B has unit norm. Moreover, since each of $\sigma_1, \sigma_2, \sigma_3$ is traceless, and that

$$\sigma_1\sigma_2 = i\sigma_3 \quad \sigma_2\sigma_3 = i\sigma_1 \quad \sigma_3\sigma_1 = i\sigma_2$$

all of which are traceless, we have $\text{Tr}(\sigma_i^\dagger\sigma_j) = \text{Tr}(\sigma_i\sigma_j) \propto \text{Tr}(\sigma_k) = \text{Tr}(\mathbb{I}\sigma_k) = 0$ for all $i \neq j$. Therefore, B is mutually orthogonal collection. In view of the previous result, B is an orthonormal basis.

(d) Let $\Sigma(\cdot)$ denote the spectrum and \mathcal{E} the set of eigenvectors for each matrix. Note: to avoid confusion, we use the capital Σ rather than the lowercase. Except for the case of \mathbb{I} , the characteristic polynomial for each of $\sigma_1, \sigma_2, \sigma_3$ is $\lambda^2 = 1$, so $\lambda = \pm 1$.

$$\begin{aligned} \mathbb{I} : \Sigma(\mathbb{I}) &= \{1, 1\} & \mathcal{E}(\mathbb{I}) &= \{\vec{v} : \vec{v} \in \mathbb{C}^2\} \\ \sigma_1 : \Sigma(\sigma_1) &= \{1, -1\} & \mathcal{E}(\sigma_1) &= \{(1 \ 1)^\top, (1 \ -1)^\top\} \\ \sigma_2 : \Sigma(\sigma_2) &= \{1, -1\} & \mathcal{E}(\sigma_2) &= \{(1 \ i)^\top, (1 \ -i)^\top\} \\ \sigma_3 : \Sigma(\sigma_3) &= \{1, -1\} & \mathcal{E}(\sigma_3) &= \{(1 \ 0)^\top, (0 \ 1)^\top\}, \end{aligned}$$

where the eigenvectors are ordered to match their corresponding eigenvalues.

(e) Let $\mathbf{A} = (A_1, A_2, A_3)$ and $\mathbf{B} = (B_1, B_2, B_3)$ be given such that $[A_j, \sigma_i] = 0$ and $[B_n, \sigma_m] = 0$ for all i, j, m, n . Then, using the following identities

$$\sigma_1\sigma_2 = i\sigma_3, \quad \sigma_2\sigma_3 = i\sigma_1, \quad \sigma_3\sigma_1 = i\sigma_2, \quad \sigma_2\sigma_1 = -i\sigma_3, \quad \sigma_3\sigma_2 = -i\sigma_1, \quad \sigma_1\sigma_3 = -i\sigma_2, \quad \sigma_i^2 = \mathbb{I},$$

we find

$$\begin{aligned} (\sigma \cdot \mathbf{A})(\sigma \cdot \mathbf{B}) &= (\sigma_1 A_1 + \sigma_2 A_2 + \sigma_3 A_3)(\sigma_1 B_1 + \sigma_2 B_2 + \sigma_3 B_3) \\ &= (A_1 B_1 + A_2 B_2 + A_3 B_3) + i\sigma_1(A_2 B_3 - A_3 B_2) + i\sigma_2(A_3 B_1 - A_1 B_3) + i\sigma_3(A_1 B_2 - A_2 B_1) \\ &= (\mathbf{A} \cdot \mathbf{B})\mathbb{I} + i\sigma \cdot (\mathbf{A} \times \mathbf{B}) \end{aligned}$$

as desired.

(f) We claim:

$$\exp(i\theta \sigma \cdot \mathbf{n}) = \cos \theta \mathbb{I} + i \sigma \cdot \mathbf{n} \sin \theta.$$

In view of Part (e), we observe that $[\sigma \cdot \mathbf{n}]^{2n} = [(\mathbf{n} \cdot \mathbf{n})\mathbb{I}]^n = \mathbb{I}$ and thus $[\sigma \cdot \mathbf{n}]^{2n+1} = \sigma \cdot \mathbf{n}$. This will help with simplifying the power series expansion of $\exp(i\theta \sigma \cdot \mathbf{n})$ below by splitting up the odd-powered and even-powered terms:

$$\begin{aligned} \exp(i\theta \sigma \cdot \mathbf{n}) &= \sum_{n=0}^{\infty} \frac{(i\theta)^n}{n!} [\sigma \cdot \mathbf{n}]^n \\ &= \sum_{k=0}^{\infty} \frac{(i\theta)^{2k}}{(2k)!} \mathbb{I} + [\sigma \cdot \mathbf{n}] \sum_{j=0}^{\infty} \frac{(i\theta)^{2j+1}}{(2j+1)!} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k \theta^{2k}}{(2k)!} \mathbb{I} + i[\sigma \cdot \mathbf{n}] \sum_{j=0}^{\infty} \frac{(-1)^j \theta^{2j+1}}{(2j+1)!} \\ &= \cos \theta, \mathbb{I} + i \sigma \cdot \mathbf{n} \sin \theta. \end{aligned}$$

And we're done with the proof.

6.

- (a) To see that $R := (1/\sqrt{2})(\mathbb{I} + i\sigma_x)$ is a rotation by $-\pi/2$ around the x -axis, it suffices to show that (1) R keeps the σ_x eigenstates invariant and (2) R rotates clockwise the σ_z eigenstates into the σ_y eigenstates.

It is clear from the definition of R that $R|\pm, x\rangle = |\pm, x\rangle$ (since $|\pm, x\rangle$ is a simultaneous eigenket of both \mathbb{I} and σ_x). So R is an identity function when restricted to the x -axis. Now, the matrix representation of this operator in the z basis is

$$R = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$$

Consider $|+, z\rangle = (1 \ 0)^T$. We see that $R|+, z\rangle = (1/\sqrt{2})(1 \ i)^T = |+, y\rangle$. Applying R one more time, we find $R|+, y\rangle = R^2|+, z\rangle = i\sigma_x|+, z\rangle = i|-, z\rangle \equiv |-, z\rangle$. The total effect is that $+z$ gets rotated into $+y$ and $+y$ gets rotated into $-z$, all with x fixed. As a result, R is a rotation by $-\pi/2$ about the x -axis.

To see this even more clearly, plot the yz -plane with the x -axis pointing out of the paper. Let the state $|+, z\rangle$ represent the $+z$ direction and $|+, y\rangle$ represent the $+y$ direction. Because $R|+, z\rangle = |+, y\rangle$, $R|+, y\rangle \equiv |-, z\rangle$ and $R|\pm, x\rangle = |\pm, x\rangle$, we have that $+z$ gets sent to $+y$, and $+y$ gets sent to $-z$. So, the yz -plane gets rotated by $-\pi/2$ about the x -axis.

- (b) We will set $\hbar/2 \equiv 1$ for convenience. The matrix elements of S_z in the y -basis are given by $\langle y_i | S_z | y_j \rangle$. So, in the y -basis, S_z is

$$S_z|_y = \begin{pmatrix} \langle +, y | S_z | +, y \rangle & \langle +, y | S_z | -, y \rangle \\ \langle -, y | S_z | +, y \rangle & \langle -, y | S_z | -, y \rangle \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

where we have used the fact that $S_z|+, y\rangle = |-, y\rangle$ and $S_z|-, y\rangle = |+, y\rangle$. Alternatively, we could calculate the matrix elements exclusively using known results in the z -basis (see Part (d)).

7. We want to construct a matrix which connects the z -basis to the x -basis. To do this, we must know how $|\pm, z\rangle$ appears in the x -basis:

$$\begin{aligned} |+, z\rangle &= \frac{1}{\sqrt{2}} |+, x\rangle + \frac{1}{\sqrt{2}} |-, x\rangle \\ |-, z\rangle &= \frac{1}{\sqrt{2}} |+, x\rangle - \frac{1}{\sqrt{2}} |-, x\rangle. \end{aligned}$$

So, to see what vectors in the z -basis look like in the x -basis, we apply the following matrix to those vectors:

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Compare this with the formula:

$$\begin{aligned} U &= \sum_r |x_r\rangle \langle z_r| \\ &= |+,x\rangle \langle +,z| + |-,x\rangle \langle -,z| \\ &= |+,x\rangle \left(\frac{1}{\sqrt{2}} \langle +,x| + \frac{1}{\sqrt{2}} \langle -,x| \right) + |-,x\rangle \left(\frac{1}{\sqrt{2}} \langle +,x| - \frac{1}{\sqrt{2}} \langle -,x| \right) \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \end{aligned}$$

which is consistent with what we found before.