Phys 251A Problem Set 1 Solutions

1. In this problem we will prove some properties of the commutator $[A, B] = AB - BA$ and anticommutator $\{A, B\} = AB + BA$, where A, B are operators on some Hilbert space. The notation $\mathrm{ad}_A^{\pm}(B) = [A, B]_{\pm} = AB \mp BA$, such that $[A, B]_+$ is the commutator and $[A, B]_-$ is the anticommutator, can be useful.

   (a) Show that $[A, BC]_+ = [A, B]_{\pm} C \pm B[A, C]_{\pm}$. Convince yourself of the similarity to the "product rule" for derivatives: $\mathrm{ad}_A^+(BC) = \mathrm{ad}_A^{\pm}(B)C \pm B\,\mathrm{ad}_A^{\pm}(C)$.

   We will add and subtract BAC to show

$$
\begin{aligned}
ABC - BCA &= ABC - BAC + BAC - BCA \\
&= [A, B]C + B[A, C] \\
&= \{A, B\}C - B\{A, C\} \\
&= \mathrm{ad}_A^{\pm}(B)C \pm B\,\mathrm{ad}_A^{\pm}(C)
\end{aligned}
$$

   where in the second line we grouped the first two terms and the last two terms, and in the third line we grouped the first and third terms, and the second and fourth terms.

   (b) Show, preferably using the previous identity, that

$$
[AB, CD] = A\{B, C\}D - \{A, C\}BD + CA\{B, D\} - C\{A, D\}B
$$

   We will use the identity in part (a), first with the positive sign and second with the negative sign. We will also use the antisymmetry of the commutator and the symmetry of the anticommutator respectively: $[A, B]_{\pm} = \mp[B, A]$.

$$
\begin{aligned}
[AB, CD] &= [AB, C]D + C[AB, D] \\
&= -[C, AB]D - C[D, AB] \\
&= -\{C, A\}BD + A\{C, B\}D - C\{D, A\}B + CA\{D, B\} \\
&= A\{B, C\}D - \{A, C\}BD + CA\{B, D\} - C\{A, D\}B.
\end{aligned}
$$

2. Consider a matrix X, that is not necessarily Hermitian or Unitary, written as

$$
X = a_0 + \boldsymbol{a} \cdot \boldsymbol{\sigma}
$$

   where $a_{0,1,2,3}$ are numbers, and $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices

$$
\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
$$

   (a) Relate a_0 and a_k to $\mathrm{tr}\, X$ and $\mathrm{tr}(X\sigma_k)$.

   (b) Write a_0 and a_k in terms of the matrix elements X_{ij}.

   Explicit calculation leads to

$$
X = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} = \begin{pmatrix} a_0 + a_3 & a_1 - ia_2 \\ a_2 + ia_3 & a_0 - a_3 \end{pmatrix}
$$

and

$$\begin{aligned}
\operatorname{tr} X &= X_{11} + X_{22} = 2a_0 \\
\operatorname{tr} X\sigma_x &= X_{12} + X_{21} = 2a_1 \\
\operatorname{tr} X\sigma_y &= iX_{12} - iX_{21} = 2a_2 \\
\operatorname{tr} X\sigma_z &= X_{11} - X_{22} = 2a_3
\end{aligned}$$

There is more abstract way to do this computation however, if one notes that $\{\sigma_i, \sigma_j\} = 2\delta_{ij}$ and $\operatorname{tr}\sigma_k = 0$.

These properties together imply $\operatorname{tr}\sigma_\mu\sigma_\nu = 2\delta_{\mu\nu}$ where $\sigma_\mu = I$ for $\mu = 0$ and $\sigma_\mu = \sigma_k$ for $\mu = k = 1, 2, 3$. This immediately implies $\operatorname{tr} X\sigma_\mu = 2a_\mu$ where $X = \sum_\mu a_\mu\sigma_\mu$.

3. Using bra-ket notation prove or evaluate the following

(a) $\operatorname{tr}(XY) = \operatorname{tr}(YX)$ for operators X and Y

We insert and remove resolution of the identity $1 = \sum_j |j\rangle\langle j|$ in a basis labeled by j to obtain

$$\operatorname{tr}(XY) = \sum_i \langle i| XY |i\rangle = \sum_{i,j}\sum_i \langle i| X |j\rangle \langle j| Y |i\rangle = \sum_{i,j} \langle j| Y |i\rangle \langle i| X |j\rangle = \sum_j \langle j| YX |j\rangle = \operatorname{tr}(YX)$$

(b) $(XY)^\dagger = Y^\dagger X^\dagger$ for operators X and Y.

For arbitrary vectors ψ and ϕ,

$$\langle\phi|XY\psi\rangle = \langle(XY)^\dagger\phi|\psi\rangle$$

by definition of the Hermitian conjugate, but we can also apply the definition sequentially for X and Y such that

$$\langle\phi|XY\psi\rangle = \langle X^\dagger\phi|Y\psi\rangle = \langle Y^\dagger X^\dagger\phi|\psi\rangle\,.$$

Since ϕ and ψ are arbitrary, we conclude $Y^\dagger X^\dagger = (XY)^\dagger$.

(c) $\exp(if(A))$ in ket-bra form in terms of the eigenvalues and eigenstates of A, for a function f that can be expanded in a power series if you wish.

We will work in an eigenbasis

$$A = \sum_a \lambda_a |a\rangle\langle a|\,, \qquad A^n = \sum_a \lambda_a^n |a\rangle\langle a|$$

such that, for $f(x) = \sum_n f_n x^n$, we have

$$\begin{aligned}
e^{if(A)} &= \sum_n \frac{i^n}{n!} f(A)^n = \sum_n \frac{i^n}{n!}\left(\sum_m f_m A^m\right)^n = \sum_n \frac{i^n}{n!}\left(\sum_m f_m \lambda_a^m \sum_a |a\rangle\langle a|\right)^n \\
&= \sum_a |a\rangle\langle a| \sum_n \frac{i^n}{n!}\left(\sum_m f_m\lambda_a^m\right)^n = \sum_a e^{if(\lambda_a)}|a\rangle\langle a|
\end{aligned}$$

(d) $\sum_n \overline{\psi_n(x_1)}\psi_n(x_2)$ where n labels a complete set of states and $\psi_n(x) = \langle x|n\rangle$.

$$\sum_n \overline{\psi_n(x_1)}\psi_n(x_2) = \sum_n \langle n|x_1\rangle \langle x_2|n\rangle = \langle x_2| \left(\sum_n |n\rangle \langle n|\right) |x_1\rangle = \langle x_2|x_1\rangle$$

4. Suppose $|m\rangle$ and $|n\rangle$ are eigenstates of a Hermitian operator A. Find a general condition under which we can conclude that $|m\rangle + |n\rangle$ is an eigenstate of A.

Suppose $|m\rangle$ and $|n\rangle$ have eigenvalues λ_m and λ_n respectively. If we evaluate

$$A(|m\rangle + |n\rangle) = \lambda_m |m\rangle + \lambda_n |n\rangle$$

we see that $|m\rangle + |n\rangle$ is an eigenstate if $\lambda_m = \lambda_n$.

Conversely, if $\lambda_m \neq \lambda_n$ we see that $|m\rangle + |n\rangle$ cannot be an eigenstate because, acting to the left we obtain $\langle m| A(|m\rangle + |n\rangle) = \lambda_m \langle m|m\rangle$ and $\langle n| A(|m\rangle + |n\rangle) = \lambda_n \langle n|n\rangle$. Acting to the right, and assuming $|m\rangle + |n\rangle$ is an eigenstate with some eigenvalue $\tilde{\lambda}$, then leads to $\tilde{\lambda} = \lambda_m = \lambda_n$ which is a contradiction.

5. Consider a Hermitian operator A with non-degenerate eigenvalues a_n labeled by n, and, using the eigenbasis of A,

   (a) Prove that $\prod_n (A - a_n) = 0$

   In the eigenbasis of A we have

   $$\sum_m \prod_n (a_m - a_n) |m\rangle \langle m|$$

   which vanishes since for each m the term with $n = m$ in the product is zero.

   (b) Explain the action of $X_m = \prod_{n\neq m} \frac{A-a_n}{a_m-a_n}$.

   Acting on an eigenstate of A, say $|k\rangle$ with eigenvalue a_k, we have

   $$X_m |k\rangle = \prod_{n\neq m} \frac{A - a_n}{a_m - a_n} |k\rangle = \prod_{n\neq m} \frac{a_k - a_n}{a_m - a_n} |k\rangle =$$

   The product vanishes if $k \neq m$ due to the factor in the numerator with $n = k$. If $k = m$, however, this factor is cancelled by the $(a_m - a_k)$ factor in the denominator, and in fact the whole product is a product of ones. We therefore have $X_m |k\rangle = \delta_{m,k} |k\rangle$, such that X_m is the orthogonal projector onto the eigenstate $|m\rangle$.

   (c) Write out the previous two parts for $A = S^z = \frac{\hbar}{2}\sigma_z$, the spin operator in the z direction for a spin $\frac{1}{2}$ particle.

   For the first part we have

   $$(S^z + \frac{\hbar}{2})(S^z - \frac{\hbar}{2}) = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = 0$$

   While for the second part

   $$X_\pm = \frac{S^z - (\mp\hbar/2)}{\pm\hbar/2 - (\mp\hbar/2)} = \frac{\pm S^z + \hbar/2}{\hbar} = \begin{pmatrix} \pm\frac{1}{2} + \frac{1}{2} & 0 \\ 0 & \mp\frac{1}{2} + \frac{1}{2} \end{pmatrix}$$

which projects onto the top component for the $+$ sign and the down component for the $-$ sign respectively.

6. Consider a Hamiltonian of a two state system

$$H = \varepsilon(|1\rangle\langle 1| - |2\rangle\langle 2| + |1\rangle\langle 2| + |2\rangle\langle 1|),$$

where ε has dimensions of energy. Find the energy eigenvalues and eigenvalues of H (in terms of $|1\rangle$ and $|2\rangle$).

The question should have specified that $|1\rangle$ and $|2\rangle$ are orthonormal. Let us write a trial eigenstate as $\alpha|1\rangle + \beta|2\rangle$. Acting with H we obtain

$$H(\alpha|1\rangle + \beta|2\rangle = \varepsilon((\alpha+\beta)|1\rangle + (\alpha-\beta)|2\rangle$$

We must solve $\alpha + \beta = \lambda\alpha$ and $\alpha - \beta = \lambda\beta$. Dividing the equations we have

$$\frac{\alpha+\beta}{\alpha-\beta} = \frac{\alpha}{\beta} \implies \frac{\alpha}{\beta} = 1 \pm \sqrt{2}.$$

So that

$$\lambda = \frac{\alpha}{\beta} - 1 = \pm\sqrt{2}$$

and the eigenvectors are

$$|\pm\rangle = \frac{1}{\sqrt{(4 \pm 2\sqrt{2})}}\Big((1 \pm \sqrt{2})|1\rangle + |2\rangle\Big)$$

7. Prove the Cauchy-Schwarz identity $|\langle a|b\rangle| \leq \|a\|\|b\|$ by observing

$$\big(\langle a| + \overline{\lambda}\langle b|\big)\left(|a\rangle + \lambda|b\rangle\right) \geq 0,$$

for all complex numbers λ, and then choosing λ appropriately.

The given inequality corresponds to the positive definiteness of the inner product $\langle\psi|\psi\rangle$ where $|\psi\rangle = |a\rangle + \lambda|b\rangle$. Expanding out the inner product we obtain

$$\|a\|^2 + |\lambda|^2\|b\|^2 + \overline{\lambda}\langle b|a\rangle + \lambda\langle a|b\rangle \geq 0.$$

In order to get the absolute value $|\langle a|b\rangle|^2$ we will choose $\lambda = c\langle b|a\rangle$ for some real constant c. With this choice we have

$$\|a\|^2 + c^2|\langle a|b\rangle|^2\|b\|^2 + 2c|\langle a|b\rangle|^2 \geq 0$$

To complete the proof we choose $c = -1/\|b\|^2$, so that the final two terms have the same scaling with b, and after multiplying through by $\|b\|^2$ the first term becomes $\|a\|^2\|b\|^2$:

$$\|a\|^2\|b\|^2 + |\langle a|b\rangle|^2 - 2|\langle a|b\rangle|^2 \geq 0 \implies |\langle a|b\rangle|^2 \leq \|a\|^2\|b\|^2$$

8. Suppose two Hermitian operators anticommute: $\{A, B\} = 0$. Is it possible to have a simultaneous eigenstate of both A and B? Justify your answer

It is possible, but only if the eigenvalue is zero for either A or B. A trivial way to see it is possible is to replace A or B with the zero matrix, but in the solutions here we will prove the general statement and provide a "generic" example.

Let us begin with an eigenstate of A, $|a\rangle$, with eigenvalue λ_a. Then we verify

$$AB\,|a\rangle = -BA\,|a\rangle = -\lambda_a(B\,|a\rangle)$$

We see that $B\,|a\rangle$ is also an eigenvector of A with eigenvalue $-\lambda_a$. This appears inconsistent with $|a\rangle$ being an eigenvector of B, and it is unless $\lambda_a = 0$ or $B\,|a\rangle = 0$; i.e. the associated eigenvalue of either A or B is zero, as claimed.

An example is given by

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

where $\begin{pmatrix} 0 & 1 & 0 \end{pmatrix}^T$ has eigenvalue 0 under A and eigenvalue 1 under B.

9. True or false questions. Explain your reasoning.

(a) The commutator of two Hermitian operators is Hermitian

False. The commutator of two Hermitian operators is anti-Hermitian.

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

(b) When two Hermitian operators A and B commute, any eigenvector of A is also an eigenvector of B.

False. Two commuting Hermitian operators can be diagonalized simultaneously by one single unitary transformation, so the same set of eigenvectors can be found for both operators. However, if one operator A has degenerate eigenvectors, then the linear combination of these vectors is still an eigenvector for A, but not for operator B if the eigenvalues in B are not degenerate.

As an example, take

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

where $\begin{pmatrix} 1 & 1 \end{pmatrix}^T$ is an eigenvector for A but not for B.

(c) An operator that is both Unitary and Hermitian must square to the identity.

True. For unitary operators: $U^\dagger U = I$, and for Hermitian operators: $U = U^\dagger$, so we have $U^2 = I$

Bonus Let us suppose that the silver atoms in the Stern-Gerlach experiment have classical magnetic moments $\boldsymbol{\mu}$ that are random in direction; the direction is uniformly distributed over the sphere of all possible directions. Furthermore suppose their magnitude $\mu_0 = |\boldsymbol{\mu}|$ is drawn from a continuous probability distribution $p(\mu_0)$. Compute the distribution of μ_z values that would be seen in this case, in terms of the function $p(\mu_0)$.

We will use the general result on conditional probabilities

$$P(X) = \sum_Y P(Y)P(X|Y)$$

where $P(X|Y)$ is the probability of X given that Y is known to have occured. We will substitute the probability $P(Y) = p(\mu_0)d\mu_0$ and the sum for a corresponding integral. We must also evaluate $p(\mu_z|\mu_0)$. Note that $\mu_z = \mu_0 \cos\theta$, where θ is the angle that the direction of $\boldsymbol{\mu}$ makes with the z axis. We note that the probability density over the sphere is homogeneous implying that probabilities are proportional to the area element as $p(\theta, \phi)d\theta d\phi = \frac{1}{4\pi}dA = \frac{1}{4\pi}\sin\theta d\theta d\phi$, where ϕ is the azimuthal angle and the factor of $1/4\pi$ is included so that the probability distribution integrates to 1. Since we are only interested on θ, we can integrate over ϕ to get the probability distribution over θ

$$p(\theta)d\theta = \frac{1}{4\pi}\sin\theta d\theta \int_0^{2\pi} d\phi = \frac{1}{2}\sin\theta d\theta.$$

Because the values of μ_z and θ are in one-to-one correspondence, the associated probabilities (not the probability densities) are equal:

$$|p(\mu_z|\mu_0)d\mu_z| = |p(\theta)d\theta| = \frac{1}{2}\sin\theta|d\theta| = \frac{1}{2}|d\cos(\theta)| = \frac{1}{2\mu_0}|d\mu_z|.$$

We can then conclude $p(\mu_z|\mu_0) = \frac{1}{2\mu_0}$, for $\mu_z \in [-\mu_0, \mu_0]$ and zero otherwise. Note that we have introduced absolute values above so that the probability densities are positive (we are not interested in the negative sign from $d\mu_z/d\theta < 0$ because this will be corrected by the fact that a positive-orientation integration range for θ, from 0 to π, corresponds to a negative one for μ_z (from μ_0 to $-\mu_0$).

The total probability distribution is then

$$p(\mu_z) = \int_0^\infty p(\mu_0)p(\mu_z|\mu_0)d\mu_0 = \int_0^\infty \frac{1}{2\mu_0}\theta(\mu_0 - |\mu_z|)d\mu_0 = \int_{|\mu_z|}^\infty \frac{p(\mu_0)}{2\mu_0},$$

where we have used the Heaviside theta function $\theta(x)$, which is zero when $x < 0$ and one when $x > 0$.