Homework One Solutions

Keith Fratus

June 28, 2011

1 Problem One

1.1 Part a

In this problem, we'll assume the fact that the sum of two complex numbers is another complex number, and also that the product of two complex numbers is another complex number.

Whenever we define a vector space, we need three things. The first thing is the "field" over which the vector space is defined, which is a fancy way of saying we need to say what the scalar multiples are (either real numbers, complex numbers, or something weirder). The second thing we need are the vectors themselves, and the third thing is an addition rule. In this problem it is explicitly stated that the vectors are complex 2 by 2 matrices, while it is implicit that the rule of addition is usual matrix addition, and the scalar multiples are complex numbers.

Let's first check that the set of 2 by 2 matrices is closed under this operation (this isn't one of the usual 8 axioms, but it's always important, often times implied, since the whole notion of addition on the vector space wouldn't make sense in the first place if we could add two 2 by 2 complex matrices and get something that wasn't a complex 2 by 2 matrix). If a, b, c, d, e, f, g, and h are arbitrary complex numbers, then we see that, using the usual rules of matrix addition,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix}. \tag{1}$$

Since the sum of two complex numbers is another complex number, we see that both sides of this equation are elements of this set, and so the set of 2 by 2 complex matrices are indeed closed under this operation.

Now, to show the 8 axioms of a vector space, we start by checking the associativity of addition, which is to say that for any three matrices A, B, and C, we require

$$A + (B + C) = (A + B) + C. \tag{2}$$

Again, writing these as arbitrary complex matrices, we see that we have.

$$\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} \right) + \begin{bmatrix} i & j \\ k & l \end{bmatrix} = \begin{bmatrix} (a+e)+i & (b+f)+j \\ (c+g)+k & (d+h)+l \end{bmatrix}, \tag{3}$$

while it is also true that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \left(\begin{bmatrix} e & f \\ g & h \end{bmatrix} + \begin{bmatrix} i & j \\ k & l \end{bmatrix} \right) = \begin{bmatrix} a+(e+i) & b+(f+j) \\ c+(g+k) & d+(h+l) \end{bmatrix}. \tag{4}$$

Because addition of individual complex numbers is associative, we see that the right side of the first equation is the same as the right side of the second equation, showing that the two sums are equal. Thus, associativity of matrix addition is verified.

Likewise, the equation

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix} = \begin{bmatrix} e+a & f+b \\ g+c & h+d \end{bmatrix} = \begin{bmatrix} e & f \\ g & h \end{bmatrix} + \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{5}$$

verifies the commutativity of matrix addition.

Next, we see that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{6}$$

verifies the existence of a zero vector, which is just the matrix where every element is zero.

To see that every element has an additive inverse, we notice that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} -a & -b \\ -c & -d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \tag{7}$$

which verifies that by negating every element of a matrix, we can arrive at its additive inverse, thus proving that every matrix has an additive inverse.

Now, if λ is an arbitrary complex number, notice that

$$\lambda \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} \right) = \lambda \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix} = \begin{bmatrix} \lambda(a+e) & \lambda(b+f) \\ \lambda(c+g) & \lambda(d+h) \end{bmatrix}. \tag{8}$$

Because scalar multiplication distributes over individual complex numbers, we see that this is equal to

$$\begin{bmatrix} \lambda a + \lambda e & \lambda b + \lambda f \\ \lambda c + \lambda g & \lambda d + \lambda h \end{bmatrix} = \lambda \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \lambda \begin{bmatrix} e & f \\ g & h \end{bmatrix} \tag{9}$$

which verifies that scalar multiplication distributes over addition of matrices.

To show that scalar multiplication distributes over the addition of scalar multiples, first notice that, for complex numbers λ and γ, we have

$$(\lambda + \gamma) \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} (\lambda+\gamma)\, a & (\lambda+\gamma)\, b \\ (\lambda+\gamma)\, c & (\lambda+\gamma)\, d \end{bmatrix} = \begin{bmatrix} \lambda a + \gamma a & \lambda b + \gamma b \\ \lambda c + \gamma c & \lambda d + \gamma d \end{bmatrix}, \tag{10}$$

which can be separated to yield

$$\lambda \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \gamma \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{11}$$

which verifies that scalar multiplication distributes over the addition of complex numbers.

To show that scalar multiplication is associative, we first see that

$$(\lambda\gamma) \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} (\lambda\gamma)\, a & (\lambda\gamma)\, b \\ (\lambda\gamma)\, c & (\lambda\gamma)\, d \end{bmatrix}. \tag{12}$$

Because scalar multiplication of complex numbers is associative, we can write this as

$$\begin{bmatrix} \lambda\left(\gamma a\right) & \lambda\left(\gamma b\right) \\ \lambda\left(\gamma c\right) & \lambda\left(\gamma d\right) \end{bmatrix} = \lambda \begin{bmatrix} \gamma a & \gamma b \\ \gamma c & \gamma d \end{bmatrix} = \lambda \left(\gamma \begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) \tag{13}$$

which verifies the associativity of scalar multiplication.

Lastly, we must verify there exists a multiplicative identity. Notice that

$$1 * \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 * a & 1 * b \\ 1 * c & 1 * d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{14}$$

verifies the existence of this object, which is just the usual multiplicative identity in the complex numbers.

1.2 Part b

Our proposed inner product is

$$\langle A|B\rangle = \frac{1}{2} Tr\left(A^\dagger B\right), \tag{15}$$

where we recall that the trace of an N by N square matrix is given by

$$Tr\left(A\right) = \sum_{i=1}^{N} A_{ii}. \tag{16}$$

In other words, the trace is the sum over the diagonal entries. We must verify that this is a valid inner product on S, or in other words, it satisfies the three axioms of an inner product.

First, notice that, for any matrix C, we have,

$$Tr\left(C^\dagger\right) = \left(Tr\left(C\right)\right)^*. \tag{17}$$

To see this, recall that from the definition of the Hermitian conjugate, the diagonal elements of $C^\dagger$ will obey,

$$\left(C^\dagger\right)_{ii} = \left(C_{ii}\right)^* \tag{18}$$

since the diagonal elements are unaffected by the transposition of rows and columns. Thus

$$Tr\left(C^\dagger\right) = \sum_{i=1}^{N} \left(C^\dagger\right)_{ii} = \sum_{i=1}^{N} \left(C_{ii}\right)^* = \left(\sum_{i=1}^{N} C_{ii}\right)^* = \left(Tr\left(C\right)\right)^*, \tag{19}$$

where the third equality follows from the fact that complex conjugation is a linear operation. Now, choosing C = $A^\dagger$B, we see that

$$\langle A|B\rangle = \frac{1}{2} Tr\left(A^\dagger B\right) = \frac{1}{2} Tr\left(\left(B^\dagger A\right)^\dagger\right) = \frac{1}{2} Tr\left(B^\dagger A\right)^* = \langle A|B\rangle^*, \tag{20}$$

where we used the rule for taking the Hermitian conjugate of a product of two matrices (make sure to review this if it's something you're not comfortable with). Thus, the first axiom of the inner product is satisfied.

Next, recall that the trace is a linear operation, since we have,

$$Tr\left(\lambda C+\gamma D\right)=\sum_{i=1}^{N}\left(\lambda C_{ii}+\gamma D_{ii}\right)=\lambda\sum_{i=1}^{N}C_{ii}+\gamma\sum_{i=1}^{N}D_{ii}=\lambda Tr\left(C\right)+\gamma Tr\left(D\right). \tag{21}$$

Therefore

$$\langle A|\lambda B\rangle=\frac{1}{2}Tr\left(A^{\dagger}(\lambda B)\right)=\frac{1}{2}Tr\left(\lambda A^{\dagger}B\right)=\frac{\lambda}{2}Tr\left(A^{\dagger}B\right)=\lambda\langle A|B\rangle, \tag{22}$$

where in the second equality I used the fact that scalar multiplication of matrices commutes through any product. Likewise,

$$\langle A|B+C\rangle=\frac{1}{2}Tr\left(A^{\dagger}(B+C)\right)=\frac{1}{2}Tr\left(A^{\dagger}B\right)+\frac{1}{2}Tr\left(A^{\dagger}C\right)=\langle A|B\rangle+\langle A|C\rangle, \tag{23}$$

and thus the second axiom is verified.

Lastly, recall that the product of two matrices can be written in component form as

$$(CD)_{ij}=\sum_{k=1}^{N}C_{ik}D_{kj}. \tag{24}$$

As a result, we can write

$$\langle A|A\rangle=\frac{1}{2}Tr\left(A^{\dagger}A\right)=\sum_{i=1}^{N}\left(A^{\dagger}A\right)_{ii}=\sum_{i=1}^{N}\sum_{k=1}^{N}\left(A^{\dagger}\right)_{ik}A_{ki}=\sum_{i=1}^{N}\sum_{k=1}^{N}|A_{ki}|^{2}, \tag{25}$$

where in the last equality I used

$$\left(A^{\dagger}\right)_{ik}A_{ki}=\left(A_{ki}\right)^{*}A_{ki}=|A_{ki}|^{2}. \tag{26}$$

Now, this double sum runs over every single element in the matrix. Because the norm of a complex number is never negative, we can see that this sum will never be negative, and the only way it can be zero is for every single term in the sum, and thus every single element in the matrix, to be zero, which would imply that A is the zero matrix. Thus, we have verified the third axiom, which is that the inner product is positive definite.

Notice that all three of these axioms could be proven with a brute-force approach, by writing down two arbitrary 2 by 2 complex matrices and working out all of the matrix multiplications. However, this approach will become virtually impossible when the matrices become large enough, and so it is good to have some familiarity with how to do these things in other ways.

1.3 Part c

We need to verify two things about this collection of matrices. The first is that they are orthonormal (in the sense of the inner product given in part b), and the second is that they form a basis for all complex 2 by 2 matrices.

Verifying that they are all orthonormal is straight-forward. We simply compute the inner product of each pair of matrices. First, notice that the Pauli matrices are "Hermitian", which means that they are equal to their own Hermitian conjugate (please check this for yourself if you are unfamiliar with this idea). So we can effectively ignore the Hermitian conjugation in the definition of the inner product.

To see that the identity is orthonormal under this inner product, we compute

$$\langle I|I\rangle = \frac{1}{2}Tr\left(I * I\right) = \frac{1}{2}Tr\left(I\right) = \frac{1}{2} * (1+1) = 1, \tag{27}$$

where we have used the fact that the square of the identity matrix is again the identity matrix.

Furthermore, there is a property of the Pauli matrices which is that the square of any Pauli matrix is the identity. That is,

$$\sigma_i * \sigma_i = I \tag{28}$$

where the index i can represent either x, y, or z. Thus, for any of the Pauli matrices, we have

$$\langle \sigma_i|\sigma_i\rangle = \frac{1}{2}Tr\left(\sigma_i * \sigma_i\right) = \frac{1}{2}Tr\left(I\right) = \frac{1}{2} * (1+1) = 1, \tag{29}$$

which verifies that all of the Pauli matrices are normalized under this inner product.

To see that all of the Pauli matrices are orthogonal to each other, we first notice that the product of any two Pauli matrices, up to a factor of ± i, is another Pauli matrix. That is,

$$\sigma_x * \sigma_y = i\sigma_z, \ \ \sigma_y * \sigma_x = -i\sigma_z, \tag{30}$$

$$\sigma_y * \sigma_z = i\sigma_x, \ \ \sigma_z * \sigma_y = -i\sigma_x \tag{31}$$

$$\sigma_z * \sigma_x = i\sigma_y, \ \ \sigma_x * \sigma_z = -i\sigma_y, \tag{32}$$

which can be verified by explicitly computing the above matrix products. Thus, if i, j, and k are all different, we have

$$\langle \sigma_i|\sigma_j\rangle = \frac{1}{2}Tr\left(\sigma_i * \sigma_j\right) = \frac{1}{2}Tr\left(\pm i\sigma_k\right) = \frac{\pm i}{2}Tr\left(\sigma_k\right). \tag{33}$$

However, it turns out that all of the Pauli matrices have *zero trace* (which is something you should verify for yourself if you haven't seen it before). Thus, we arrive at the conclusion

$$\langle \sigma_i|\sigma_j\rangle = \frac{\pm i}{2}Tr\left(\sigma_k\right) = 0, \tag{34}$$

and so the Pauli matrices are indeed orthonormal under this inner product.

To show that these matrices form a basis for all complex 2 by 2 matrices, we need to verify that these matrices are a set of linearly independent matrices from which any other 2 by 2 complex matrix can be written. The fact that they are linearly independent immediately follows from the fact that they are orthonormal (if any of the vectors could be written as a linear combination of the others, then their inner products amongst themselves could not be zero, and hence they would not be orthonormal). The fact that any complex 2 by 2 matrix can be written as a linear combination of them can be seen by explicitly finding the coefficients in the expansion. If a, b, c, and d are arbitrary complex numbers, and x, y, z, and w are the proposed coefficients in the linear expansion, then we want to try to solve the equation

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = x\ \sigma_x + y\ \sigma_y + z\ \sigma_z + w * I. \tag{35}$$

If we use the explicit form of the Pauli matrices, a short exercise in matrix algebra reduces the above expression to

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} z+w & x-iy \\ x+iy & -z+w \end{bmatrix}. \tag{36}$$

This matrix equation is really a set of four simultaneous equations, which we can easily solve to find

$$x = \frac{1}{2}(b+c),\ y = \frac{i}{2}(b-c) \tag{37}$$

$$z = \frac{1}{2}(a-d),\ w = \frac{1}{2}(a+d). \tag{38}$$

Because I have explicitly found these four coefficients for any arbitrary 2 by 2 matrix that I can think of, I have proven that it is indeed always possible to do this. Notice that I had four equations in four unknowns. If the number of matrices in my proposed basis was either more or less, then these equations would either be under-determined or overdetermined, respectively.

2 Problem Two

2.1 Part a

We can show that this operation is not a linear operation through a simple counter example. For an operator T to be linear, we must have

$$T(\lambda x) = \lambda T(x), \tag{39}$$

where λ is a scalar multiple and x is an element in the space of objects being acted on by T (which may or may not be a vector space, depending upon the context in which T is being introduced). But in our case, the objects in question are positive real numbers, and the operator T is the square root, and so our proposed relation is

$$\sqrt{\lambda x} = \lambda \sqrt{x} \tag{40}$$

for positive real numbers x and scalar multiples λ. But of course we know this is not true at all, and thus the square root is not a linear operation.

2.2 Part b

This is also not a linear operation, which we can again see by a counter-example. We have

$$\ln(1+1) = \ln(2), \tag{41}$$

while we also have

$$\ln(1) + \ln(1) = 2\ln(1) = \ln\left(1^2\right) = \ln(1) \neq \ln(2), \tag{42}$$

and so we see that linearity does not hold.

2.3 Part c

First, recall that the derivative is a linear operation, which is to say that for any functions f and g, and any constant multiples a and b, we have

$$\frac{d}{dx}(a * f(x) + b * g(x)) = a * \frac{df}{dx} + b * \frac{dg}{dx}. \tag{43}$$

Next, notice that for two linear operators T and U, the result

$$TU(a*f+b*g) = T(U(a*f+b*g)) = T(a*(Uf)+b*(Ug)) = a*(TUf)+b*(TUf) \tag{44}$$

implies that the product of two linear operators is another linear operator. Thus, because the first derivative is a linear operator, and because any repeated application of this operator will yield another linear operator, we see that the nth-derivative is indeed a linear operator.

2.4 Part d

If T and U are linear operators, a, b, c, and d are constant multiples, and f and g are vectors, then notice that we have

$$(cT + dU)\ (a * f + b * g) = cT(a * f) + cT(b * g) + dU(a * f) + dU(b * g). \tag{45}$$

Using the linearity property of T and U, we can pull out the constant multiples a and b, and then rearrange this expression to find that this is equal to

$$a(cT + dU)\ (f) + b(cT + dU)\ (g) \tag{46}$$

which shows that the linear combination of operators, cT+dU, is also another linear operator itself. Now, because we know that the nth-derivative and the identity are linear operator, and that linear combinations of linear operators are also linear operators themselves, we see that the operator

$$\alpha \frac{d^2}{dx^2} + \beta \frac{d}{dx} + \gamma \hat{I} \tag{47}$$

is indeed a linear operator.

2.5 Part e

Notice that we have

$$(I + I)^{-1} = (2I)^{-1} = \frac{1}{2} I \tag{48}$$

where I is the identity operator. However,

$$I^{-1} + I^{-1} = I + I = 2I, \tag{49}$$

and so we see that the matrix inverse cannot be a linear operation, since it fails in this case.

3 Problem Three

3.1 Part a

The logic for this problem proceeds in the same way as the proof given in class; however, we want to use bra-ket notation.

To begin the proof, we recall that for any vector $|w\rangle$, we have

$$\langle w|w\rangle \geq 0. \tag{50}$$

Since this can be any vector, we choose

$$|w\rangle = |\alpha\rangle + \lambda|\beta\rangle \tag{51}$$

where $|\alpha\rangle$ and $|\beta\rangle$ are vectors, and λ is a complex scalar multiple. Now, the dual vector of this, in bra-ket notation, will be

$$\langle w| = \langle\alpha| + \lambda^*\langle\beta|, \tag{52}$$

so we can write the inner product in bra-ket notation as

$$(\langle\ \alpha| + \lambda^*\langle\beta|) * (|\alpha\rangle + \lambda|\beta\rangle) \geq 0, \tag{53}$$

which we can expand out to get

$$\langle\alpha|\alpha\rangle + \lambda\langle\alpha|\beta\rangle + \lambda^*\langle\beta|\alpha\rangle + |\lambda|^2\langle\beta|\beta\rangle \geq 0. \tag{54}$$

Notice that in bra-ket notation, the scalar multiples are "already pulled out," in the sense that the complex conjugation of the scalar multiples is already taken care of when we write down the form of the dual vector. Now, since λ can be anything, we are free to choose

$$\lambda = -\frac{\langle\beta|\alpha\rangle}{\langle\beta|\beta\rangle}. \tag{55}$$

If we substitute this value of λ into the previous expression, the simplified result becomes

$$\langle\alpha|\alpha\rangle\langle\beta|\beta\rangle - \langle\alpha|\beta\rangle\langle\beta|\alpha\rangle \geq 0. \tag{56}$$

Now, if we recall that

$$\langle\beta|\alpha\rangle = \langle\alpha|\beta\rangle^*, \tag{57}$$

then the previous result can be written as

$$\langle\alpha|\alpha\rangle\langle\beta|\beta\rangle - |\langle\alpha|\beta\rangle|^2 \geq 0 \Rightarrow \langle\alpha|\alpha\rangle\langle\beta|\beta\rangle \geq |\langle\alpha|\beta\rangle|^2, \tag{58}$$

and since the norm is an inherently non-negative quantity, we can take the square root of both sides to get

$$\sqrt{\langle\alpha|\alpha\rangle\langle\beta|\beta\rangle} = |\alpha||\beta| \geq |\langle\alpha|\beta\rangle| \tag{59}$$

which is the Cauchy-Schwarz inequality.

3.2 Part b

The position operator is defined by

$$\hat{X}_i\ \psi(x) = x_i\ \psi(x), \tag{60}$$

which is to say that the action of the operator is to multiply the function $\psi(x)$ by a factor of x_i, where x_i can be either one of the three usual position coordinates, x, y, and z (corresponding to i = 1, 2, and 3). Likewise, the momentum operator is defined by

$$\hat{P}_i\ \psi(x) = -i\hbar\frac{\partial}{\partial x_i}\psi(x). \tag{61}$$

When evaluating these commutators, the easiest way to avoid mistakes is to use a "test function." For any two linear operators T and U, the commutator, which is defined by

$$[T, U] = TU - UT, \tag{62}$$

is a linear operator itself, since sums and products of linear operators produce other linear operators. So what we want to compute is

$$[T, U]f = TUf - UTf, \tag{63}$$

where f is some arbitrary function which we will throw away at the end of the calculation. The reason we do this is because these commutator calculations often include using various rules of differentiation, and attempting to do this in the abstract without acting on an actual function can lead to mistakes.

To begin, let's compute the commutator

$$[\hat{X}_i, \hat{X}_j] = \hat{X}_i \hat{X}_j - \hat{X}_j \hat{X}_i. \tag{64}$$

Using a test function, we see that this becomes

$$[\hat{X}_i, \hat{X}_j] f = \hat{X}_i \left(\hat{X}_j f \right) - \hat{X}_j \left(\hat{X}_i f \right) = x_i (x_j f) - x_j (x_i f). \tag{65}$$

Now, because x_i and x_j are just coordinates, which are merely numbers, their multiplication is associative and commutative. So we can write the above expression as

$$[\hat{X}_i, \hat{X}_j] f = x_i x_j f - x_j x_i f = x_i x_j f - x_i x_j f = 0. \tag{66}$$

Now, because the action of the commutator on *any* arbitrary function is zero, then the commutator itself must be the zero operator, and so we have

$$[\hat{X}_i, \hat{X}_j] = 0 \tag{67}$$

as expected.

To verify the second commutation relation, we again use a test function, and compute

$$[\hat{P}_i, \hat{P}_j] f = \hat{P}_i \left(\hat{P}_j f \right) - \hat{P}_j \left(\hat{P}_i f \right) = -i\hbar \left(\frac{\partial}{\partial x_i} \left(-i\hbar \frac{\partial}{\partial x_j} f \right) - \frac{\partial}{\partial x_j} \left(-i\hbar \frac{\partial}{\partial x_i} f \right) \right). \tag{68}$$

If we pull through the numerical factors, this becomes

$$[\hat{P}_i, \hat{P}_j] f = -\hbar^2 \left(\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} f - \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_i} f \right). \tag{69}$$

Now, if a function is at least twice continuously differentiable (which means that you can at least take its second derivative and that that second derivative is continuous), then we know that the mixed partial derivatives of that function will all be equal. We are told that these operators act on functions which are members of C_∞, which is to say that they can be differentiated an infinite number of times, and thus they more than meet the requirement to have their partial derivatives be equal. Because of this, we have

$$\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} f = \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_i} f \Rightarrow [\hat{P}_i, \hat{P}_j] = 0 \tag{70}$$

as expected.

To prove the final commutator expression, we begin by writing

$$[\hat{X}_i, \hat{P}_j] f = \hat{X}_i \left(\hat{P}_j f \right) - \hat{P}_j \left(\hat{X}_i f \right) = -i\hbar \left(x_i \cdot \frac{\partial f}{\partial x_j} - \frac{\partial}{\partial x_j} (x_i \cdot f) \right). \tag{71}$$

Now, if i $\neq$ j, which means that x_i and x_j are different coordinates, then we have

$$\frac{\partial}{\partial x_j}(x_i \cdot f) = x_i \cdot \frac{\partial f}{\partial x_j} \tag{72}$$

since x_i is just a constant with respect to the differentiation. Thus, the commutator becomes,

$$[\hat{X}_i, \hat{P}_j]f = -i\hbar\left(x_i \cdot \frac{\partial f}{\partial x_j} - x_i \cdot \frac{\partial f}{\partial x_j}\right) = 0, \tag{73}$$

so that

$$[\hat{X}_i, \hat{P}_j] = 0, \tag{74}$$

as expected. Now, if i = j, then we have instead

$$\frac{\partial}{\partial x_j}(x_i \cdot f) = \frac{\partial}{\partial x_i}(x_i) \cdot f + x_i \cdot \frac{\partial}{\partial x_i}(f) = f + x_i \cdot \frac{\partial f}{\partial x_i}, \tag{75}$$

which implies that the commutator becomes

$$[\hat{X}_i, \hat{P}_j]f = -i\hbar\left(x_i \cdot \frac{\partial f}{\partial x_j} - f - x_i \cdot \frac{\partial f}{\partial x_i}\right) = i\hbar f, \tag{76}$$

which is to say that the action of the commutator on any arbitrary function is to multiply it by i$\hbar$, or,

$$[\hat{X}_i, \hat{P}_j] = i\hbar. \tag{77}$$

Combining the two cases, we can write

$$[\hat{X}_i, \hat{P}_j] = i\hbar\delta_{ij}, \tag{78}$$

as expected.

4 Problem Four

The statement of the triangle inequality is, for any two vectors x and y,

$$||x + y|| \leq ||x|| + ||y||, \tag{79}$$

or, in bra-ket notation,

$$\sqrt{(\langle x| + \langle y|)(|x\rangle + |y\rangle)} \leq \sqrt{\langle x|x\rangle} \; + \; \sqrt{\langle y|y\rangle}. \tag{80}$$

To prove this inequality, we begin by writing

$$(\langle x| + \langle y|)(|x\rangle + |y\rangle) = \langle x|x\rangle + \langle x|y\rangle + \langle y|x\rangle + \langle y|y\rangle. \tag{81}$$

Now, notice that

$$\langle x|y\rangle + \langle y|x\rangle = \langle x|y\rangle + \langle x|y\rangle^* = 2Re[\langle x|y\rangle], \tag{82}$$

and because the real part of a complex number is always less than or equal to the absolute value of that complex number (which you should verify for yourself), we have

$$2Re[\langle x|y\rangle] \leq 2|\langle x|y\rangle|, \tag{83}$$

which implies

$$(\langle x| + \langle y|)(|x\rangle + |y\rangle) \leq \langle x|x\rangle + 2|\langle x|y\rangle| + \langle y|y\rangle. \tag{84}$$

Now, by the Cauchy-Schwarz inequality, we have

$$|\langle x|y\rangle| \leq \sqrt{\langle x|x\rangle}\sqrt{\langle y|y\rangle}, \tag{85}$$

so we can further write

$$(\langle x| + \langle y|)(|x\rangle + |y\rangle) \leq \langle x|x\rangle + 2\sqrt{\langle x|x\rangle}\sqrt{\langle y|y\rangle} + \langle y|y\rangle. \tag{86}$$

Finally, noticing that

$$\langle x|x\rangle + 2\sqrt{\langle x|x\rangle}\sqrt{\langle y|y\rangle} + \langle y|y\rangle = \left(\sqrt{\langle x|x\rangle} + \sqrt{\langle y|y\rangle}\right)^2, \tag{87}$$

we ultimately get

$$(\langle x| + \langle y|)(|x\rangle + |y\rangle) \leq \left(\sqrt{\langle x|x\rangle} + \sqrt{\langle y|y\rangle}\right)^2, \tag{88}$$

which gives the desired result when we take the square root of both sides.

The triangle inequality shows up all the time when calculating limits from the definition, and also has an obvious geometric interpretation in a variety of physical contexts.

5 Problem Five

Let's begin by normalizing the first vector in our space, the function f(x) = 1. Remember that for a norm derived from an inner product, we have

$$||f|| = \sqrt{\langle f|f\rangle}, \tag{89}$$

which in our case implies,

$$||f|| = \sqrt{\int_{-\infty}^{\infty} e^{-x^2} f^*(x)f(x)\ dx} = \sqrt{\int_{-\infty}^{\infty} e^{-x^2}|f(x)|^2\ dx}. \tag{90}$$

Thus, the norm of f(x) = 1 is

$$||1|| = \sqrt{\int_{-\infty}^{\infty} e^{-x^2}|1|^2\ dx} = \sqrt{\int_{-\infty}^{\infty} e^{-x^2}\ dx} = \sqrt{\sqrt{\pi}} = \pi^{\frac{1}{4}} \tag{91}$$

where the value of the integral can be found in a table or computed with some program like Mathematica. Now that we have the norm of f(x) = 1, we can take our first basis vector to be

$$\hat{e}_1 = \frac{1}{||1||} = \pi^{-\frac{1}{4}} \tag{92}$$

To find the second orthonormal basis vector, we want to take the original second vector, f(x) = x, and subtract off the portion of it which is parallel to the first basis vector. This will leave a vector which is orthogonal to the first basis vector, which we can then normalize to get the second basis vector. The portion of the original second vector which is parallel to the first basis vector is given by

$$|x\rangle_{||} = \langle e_1|x\rangle \ \hat{e_1}, \tag{93}$$

which is the component along the first orthonormal basis vector, times the first basis vector. Notice that if $\hat{e_1}$ were not already normalized, we would need to divide the above expression by the norm of $\hat{e_1}$, which is the more general expression for the projection along a given direction. Now, the above inner product is given by

$$\langle e_1|x\rangle = \int_{-\infty}^{\infty} e^{-x^2}\pi^{-\frac{1}{4}} \ x \ dx = 0, \tag{94}$$

which follows from the fact that we are integrating an odd function over a symmetric interval. So the original second vector is already orthogonal to the first basis vector. Thus, we merely need to normalized the original second vector. This norm is given by

$$||x|| = \sqrt{\int_{-\infty}^{\infty} e^{-x^2}|x|^2 \ dx} = \sqrt{\int_{-\infty}^{\infty} e^{-x^2}x^2 \ dx} = \sqrt{\frac{\sqrt{\pi}}{2}} = \frac{1}{\sqrt{2}}\pi^{\frac{1}{4}}, \tag{95}$$

so that we can take our second basis vector to be

$$\hat{e}_2 = \frac{x}{||x||} = \sqrt{2} \ \pi^{-\frac{1}{4}} \ x. \tag{96}$$

To find the third basis vector, we will start with the original third vector, and then subtract off the portion which is parallel to $\hat{e_1}$, and also the portion which is parallel to $\hat{e_2}$. Formally, this is

$$|x^2\rangle_{\perp} = |x^2\rangle - |x^2\rangle_{||} = |x^2\rangle - \langle e_1|x^2\rangle \ \hat{e_1} - \langle e_2|x^2\rangle \ \hat{e_2}. \tag{97}$$

Now, we have

$$\langle e_1|x^2\rangle = \int_{-\infty}^{\infty} e^{-x^2}\pi^{-\frac{1}{4}} \ x^2 \ dx = \frac{1}{2}\pi^{\frac{1}{4}}, \tag{98}$$

while we also have

$$\langle e_2|x^2\rangle = \int_{-\infty}^{\infty} e^{-x^2}\sqrt{2} \ \pi^{-\frac{1}{4}} \ x \cdot x^2 = 0 \tag{99}$$

which is a result of the integrand being an odd function. So we see that we have

$$|x^2\rangle_\perp = |x^2\rangle - \frac{1}{2}\pi^{\frac{1}{4}}\ \hat{e_1} = x^2 - \frac{1}{2}. \tag{100}$$

Finally, to normalize this vector, we have,

$$||x_\perp^2|| = \sqrt{{}_\perp\langle x^2|x^2\rangle_\perp} = \sqrt{\int_{-\infty}^{\infty} e^{-x^2}\ |\ x^2 - \frac{1}{2}\ |^2\ dx} = \sqrt{\frac{\sqrt{\pi}}{2}} = \frac{\pi^{\frac{1}{4}}}{\sqrt{2}} \tag{101}$$

so that our third basis vector becomes

$$\hat{e}_3 = \frac{|x^2\rangle_\perp}{||x_\perp^2||} = \sqrt{2}\ \pi^{-\frac{1}{4}}\ (x^2 - \frac{1}{2}). \tag{102}$$

Finally, to find the fourth vector in our basis, we need to find the portion of the fourth original vector which is perpendicular to the three basis vectors we have so far. This looks like

$$|x^3\rangle_\perp = |x^3\rangle - |x^3\rangle_{||} = |x^3\rangle - \langle e_1|x^3\rangle\ \hat{e_1} - \langle e_2|x^3\rangle\ \hat{e_2} - \langle e_3|x^3\rangle\ \hat{e_3} \tag{103}$$

Using Mathematica, or a table, or some other means of computing the integrals, we find

$$\langle e_1|x^3\rangle = \int_{-\infty}^{\infty} e^{-x^2}\pi^{-\frac{1}{4}}\ x^3\ dx = 0, \tag{104}$$

$$\langle e_2|x^3\rangle = \int_{-\infty}^{\infty} e^{-x^2}\sqrt{2}\ \pi^{-\frac{1}{4}}\ x\cdot x^3\ dx = \frac{3\pi^{\frac{1}{4}}}{2\sqrt{2}}, \tag{105}$$

$$\langle e_3|x^3\rangle = \int_{-\infty}^{\infty} e^{-x^2}\sqrt{2}\ \pi^{-\frac{1}{4}}\ (x^2 - \frac{1}{2})\cdot x^3\ dx = 0. \tag{106}$$

Thus, we have,

$$|x^3\rangle_\perp = |x^3\rangle - \frac{3\pi^{\frac{1}{4}}}{2\sqrt{2}}\ \hat{e_2} = x^3 - \frac{3}{2}x \tag{107}$$

The norm of this perpendicular portion is found to be

$$||x_\perp^3|| = \sqrt{{}_\perp\langle x^3|x^3\rangle_\perp} = \sqrt{\int_{-\infty}^{\infty} e^{-x^2}\ |\ x^3 - \frac{3}{2}x\ |^2\ dx} = \sqrt{\frac{3\sqrt{\pi}}{4}} = \frac{\sqrt{3}}{2}\ \pi^{\frac{1}{4}}, \tag{108}$$

which finally lets us write

$$\hat{e}_3 = \frac{|x^3\rangle_\perp}{||x_\perp^3||} = \frac{2\pi^{-\frac{1}{4}}}{\sqrt{3}}\ (x^3 - \frac{3}{2}x). \tag{109}$$